# Table of Contents

## CH. 1: ANALYZING CHOICE UNDER ALTERNATIVE RULES

### A. INTRODUCTORY CONSIDERATIONS

### B. ELEMENTARY GAME THEORY

1. **THE PRISONER’S DILEMMA**
   a. **THE STUDENT’S DILEMMA: UNIVERSITY v. EAGER**
   b. **THE PRISONER’S DILEMMA AS A MATRIX MODEL**
      - Exhibit 1.1: PD Matrix Model
   c. **THE “PRISONER’S DILEMMA” DEFINED**
      - Exhibit 1.1: PD Matrix Model

2. **ANOTHER IMPORTANT, CLOSELY RELATED GAME: “CHICKEN”**

3. **OTHER USEFUL JARGON: “EXTERNALITIES”**

4. **ETHICAL AND OTHER EXTRALEGAL BEHAVIORAL REGULATIONS**

5. **EXTERNAL BENEFITS, EXCLUDABILITY AND “FREE RIDERS”**

### C. THE LAW AND ECONOMICS OF PROPERTY RIGHTS

1. **ROAD MAINTENANCE PROBLEM**
   - Exhibit 1.3 Benefits of Road Repair Levels

2. **JOINTNESS AND EXCLUDABILITY AS ELEMENTS OF PROPERTY**

3. **CHICKEN vs. PRISONER’S DILEMMA**
   - Exhibit 1.4 Road Problem As A “Chicken” Game
   - Exhibit 1.5 Road Problem As A Prisoner’s Dilemma

4. **GAMING ABOUT INFORMATION**
   - UNTRUTHS CAN IMPROVE BUSINESS-SCHOOL GRADE; PEER PRESSURE CLASHES WITH ETHICS

5. **“CHICKEN” IN ACTION: LAND ASSEMBLY**
   - HOMEOWNER BICKERING ENDS $100,000 OFFER FOR HOUSES

6. **WHAT GENERALIZATIONS CAN BE MADE?**

### D. TRANSFERABILITY AS AN ELEMENT OF PROPERTY

1. **ASSIGNABILITY OF CAUSES OF ACTION**
   - THE CASE OF NIKE v. LOBEL

2. **OTHER KINDS OF PROPERTY RIGHTS IN LEGAL PROCESS**
   - Mutuality of Estoppel
CH 2: RIVALROUS AND RISKY DECISIONS

A. BEHAVIOR INVOLVING UNCERTAINTY

1. PROSPECTS AND CERTAINTY EQUIVALENTS

2. EXPECTED VALUES
   Exhibit 2.1 Computing Expected Values

3. RISK PREFERENCE
   Exhibit 2.2 Money and Utility

B. APPLICATIONS

1. LAW OF THE LEANING TREE

2. THE CASE OF MRS. CRISPY’S CHICKEN

3. EX ANTE ANALYSIS OF PUNISHMENTS
   IN RE JOHN LYNCH
   RICHARDS v. ALLSTATE INSURANCE CO.
   CRAIG v. BOREN

4. THE HUNTING OF A TAKEOVER

5. AGENT-PRINCIPAL CONFLICTS OF INTEREST
   a. Using Agents To Sell Real Estate
      Exhibit 2.3 - Real Estate Market Information
   b. Agent-Principal Conflicts in Legal Contracts
      HEINZMAN v. FINE, FINE, LEGUM & FINE

6. “CURVE” MODELS: MARGINAL AND TOTAL EFFECTS
   Exhibit 2.4 – Marginal Effect As Slope of Total Curve
   Exhibit 2.5 - Area Under A Marginal Curve
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.</td>
<td>ATTORNEY- CLIENT RELATIONSHIPS</td>
<td>71</td>
</tr>
<tr>
<td>8.</td>
<td>SHOULD SUBROGEES KEEP “EXCESS” RECOVERIES?</td>
<td>72</td>
</tr>
<tr>
<td>C.</td>
<td>LEAST-COST RISK BEARING</td>
<td>73</td>
</tr>
<tr>
<td>1.</td>
<td>WHAT IS “LEAST-COST” RISK BEARING?</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Exhibit 2.6 – Binomial Distribution</td>
<td>75</td>
</tr>
<tr>
<td>2.</td>
<td>SOURCES OF LEAST COST RISK-BEARING</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Exhibit 2.7 – Pool of 5 Shared $1 Risks with 40% Probability</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Exhibit 2.8 – Pool of 10 Shared $1 Risks with 40% Probability</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Exhibit 2.9 – Pool of 20 Shared $1 Risks with 40% Probability</td>
<td>78</td>
</tr>
<tr>
<td>3.</td>
<td>PORTFOLIO ANALYSIS AND THE PRUDENCE STANDARD</td>
<td>79</td>
</tr>
<tr>
<td>D.</td>
<td>OPTIMAL INSURANCE AND RISK ADJUSTMENT</td>
<td>80</td>
</tr>
<tr>
<td>1.</td>
<td>“INSURANCE” AGAINST CONTRACT BREACH</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Exhibit 2.10 (top panel) and Exhibit 2.11 (bottom panel)</td>
<td>82</td>
</tr>
<tr>
<td>2.</td>
<td>AFTER-THE-FACT INDEMNIFICATION: REALLY INSURANCE?</td>
<td>83</td>
</tr>
<tr>
<td>3.</td>
<td>“RISK DISTRIBUTION” INSURANCE AGAINST HARMFUL EVENTS</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>NATIONAL STEEL SERVICE CENTER v. GIBBONS</td>
<td>85</td>
</tr>
<tr>
<td>E.</td>
<td>JUSTICE AND FAIRNESS</td>
<td>88</td>
</tr>
<tr>
<td>1.</td>
<td>A Contractarian Theory of Justice as Fairness</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Exhibit 2.12 – Social Choice Matrix</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Exhibit 2.13 – Maxi-Min Social Choice Matrix</td>
<td>91</td>
</tr>
<tr>
<td>2.</td>
<td>FAIR DIVISION SCHEMES: DIVIDE AND CHOOSE</td>
<td>91</td>
</tr>
<tr>
<td>3.</td>
<td>EQUITABLE DISTRIBUTION OF VOTING POWER</td>
<td>92</td>
</tr>
<tr>
<td>a.</td>
<td>Pivotal-Voter Measure of Political Power</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>Exhibit 2.14: Power of a Voter Via Representation</td>
<td>94</td>
</tr>
<tr>
<td>b.</td>
<td>An Equilibrium-Displacement Measure of Voter Power</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Exhibit 2.15: Potential Displacement Via Single Voter</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Exhibit 2.16: Electoral College Power Under Alternative Theories</td>
<td>96</td>
</tr>
<tr>
<td>4.</td>
<td>EFFICIENCY ASPECTS OF FAIRNESS</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>STATE OF IOWA v. HENRY PARRISH</td>
<td>96</td>
</tr>
<tr>
<td>5.</td>
<td>“PATERNALISM” IN THE LAW</td>
<td>100</td>
</tr>
<tr>
<td>CH. 3: INTERTEMPORAL COSTS AND BENEFITS</td>
<td>103</td>
<td></td>
</tr>
<tr>
<td>A.</td>
<td>INTEREST RATES AND “PRESENT VALUES”</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>Exhibit 3.1: ((1+r)^n) Compound Interest Growth Factors</td>
<td>105</td>
</tr>
</tbody>
</table>
B. VALUING THE BENEFITS OF INTERESTS IN LAND 108

1. DETERMINING THE VALUE OF A LEASEHOLD 108

2. VALUING A PROPERTY ENCUMBERED BY A LEASE 108

C. INTEREST ON DAMAGES AND PENALTIES 109

1. PREJUDGMENT INTEREST ON DAMAGES 109
   BUSIK v. LEVINE 109

2. “BACK DOOR” PREJUDGMENT INTEREST? 113
   UNITED TELECOMMUNICATIONS, INC. v. AMERICAN TELEVISION AND COMMUNICATIONS CORP. 113

3. PREJUDGMENT INTEREST ON PENALTIES 115

D. EVALUATING LOST FUTURE INCOME AS DAMAGES 116

   Exhibit 3.3: Lost Income Damages, Assumption Set #1 117
   Exhibit 3.4: Lost Income Damages, Assumption Set #2 118
   BEAULIEU v. ELLIOTT 119
   DOCA v. MARINA MERCANTE NICARAGUENSE, S.A. 123

E. COMBINED UNCERTAINTY AND DISCOUNTING 129

1. Introductory Considerations 129
   a. Tintd Hypothetical 130
      Exhibit 3.5: Tintds’ Probability of Survival to Various Birthdays 130
   b. Annuity Hypothetical 130
   c. Life Tenancy and Remainder Rights 131

2. AVERAGE vs. actuarily expected LIFE earnings 132
   Exhibit 3.6: Average Lifespan vs. Actuarial Calculations 133

3. RISK SELECTION AND RISK COMPENSATION 134
   Exhibit 3.7: Analysis of Rentals 135

4. DEFAULT COST UNDER “BACKDOOR” INTEREST SCHEMES 136
   Exhibit 3.8: The “Time Price” Interest Subterfuge 136

F. TAX TREATMENT OF TORT DAMAGES 139
   DeLUCCA v. UNITED STATES 139

G. FAIRNESS IN EMPLOYEE BENEFITS 142
# Table of Contents

1. Contributions to Pension Plans
   LOS ANGELES DEPT. OF WATER AND POWER v. MANHART
   142

2. CONSTRUING THE AGE DISCRIMINATION IN EMPLOYMENT ACT
   145

CH. 4: MAPPINGS AND “DEMAND” DECISIONS
   147

A. TOPOGRAPHICAL MAP READING
   147

1. Decisions Based on Spatial Locations
   Exhibit 4.1 – Topographic Map Showing Relevant Features
   147

2. INDIFFERENCE CURVE MODELS AND MARKET DECISIONS
   Exhibit 4.2: Economic Topography and Constraints
   149

B. GROUP DECISIONS UNDER MAJORITY RULE
   151

1. “Rational” Group Preferences
   151

2. MODELING OF VOTING DECISIONS
   152
   a. Topographic Model of 3-Person Majority Rule Decisions
      Exhibit 4.3 – Topographic Map of Budgetary Decisionmaking
      152
   b. A MATRIX MODEL OF VOTE TRADING
      Exhibit 4.4 – Voters’ Benefit Changes For Measures A & B
      153

C. DECISIONMAKING IN THE COURTS
   STATE OF MISSOURI v. BARTON
   HARRIS v. RIVERA
   155

D. EDGEWORTH BOX ANALYSIS OF CONTRACTS
   CONTRACTUAL REALLOCATION THROUGH “EFFICIENT BREACH”
   Exhibit 4.5 – Indifference Curve Model of Efficient Breach
   158

E. CONSUMER CHOICE: MODELING PROMISSORY RELIANCE
   Exhibit 4.6 – Intertemporal Model of Promissory Reliance
   161

F. PERMISSIBLE AND IMPERMISSIBLE EFFECTS OF GOVERNMENT GRANTS
   164
   MATTHEWS et al. v. MASSELL
   TILTON v. RICHARDSON
   Hypothetical: Christian Friars Wine Sales to Military
   164

G. BRIBES VERSUS THREATS
   UNITED STATES v. BUTLER
   STEWARD MACHINE CO. v. DAVIS
   CARROTS vs. STICKS
   172
### H. MEASURING RESPONSIVENESS: ELASTICITIES

HARRIS v. McRAE  
UNITED STATES v. E. I. DU PONT DE NEMOURS & CO.

### 3. LEGAL APPLICATIONS FROM DEMAND THEORY

CH. 5: CHANNELING BEHAVIOR VIA LEGAL INCENTIVES

### A. TORT LIABILITY MODELS

1. THE LEARNED HAND RULE: A MARGINAL COST-BENEFIT FORMULATION  
   Exhibit 5.1 – Marginal Cost and Marginal Benefit of Precautions

2. THE HYPOTHETICAL MODEL UNDERLYING THE DIAGRAMS  
   Exhibit 5.2: Effect of Recognizing Only 75% of Actual Harm

3. ERRORS IN THE APPLICATION OF TORT LAW  
   a. Incorrect Assessment Of Damages

4. Mistakes In Applying The Liability Standard  
   Exhibit 5.3: Probability of Jury Requiring Care Level  
   Exhibit 5.4: Behavior Effects of Errors Regarding Care Standard

### B. MULTI-PARTY HARM PRODUCTION FUNCTIONS

1. Precautions That Minimize Net Social Cost of Accidents  
   Exhibit 5.5 – Multiparty Social Cost Function #1

2. Comparative Negligence  
   Exhibit 5.6 – Multiparty Social Cost Function #2  
   BENTZLER v. BRAUN  
   VICTOR SCHWARTZ, COMPARATIVE NEGLIGENCE, p. 297  
   STATE v. KAATZ

### C. THE “LEVEL” OR QUANTITATIVE PROBLEM

GREENLAWN NEIGHBORHOOD ASSOCIATION v. FIRST GUARANTY BANK

### D. TAKING COST-MINIMIZATION SERIOUSLY(?)

The Case Of Innocente v. Vinho  
Drunks Won’t Let The Preacher Alone

### E. PROMISE-BREAKING ANALYZED AS A TORT

1. Economically Optimal Damages For Broken Promises

2. Sec. 90, Restatement Of Contracts
# CH. 6: SUPPLY AND SYSTEMIC INTERACTIONS

## A. PRODUCTION COSTS AND CONTRACT BREACH

1. **Cost Conditions For A Firm Producing Standprods**
   - Exhibit 6.1: Costs at Alternative Outputs

2. **Cases About Damages for Breaches of Duty**
   - JERICHO SASH AND DOOR CO. v. BUILDING ERECTORS, INC.
   - NOBS CHEMICAL, U.S.A., INC. v. KOPPERS CO., INC.
   - CURT’S TRUCKING CO. v. CITY OF ANCHORAGE
   - LUND v. COMMONWEALTH

## B. ADJUSTMENT OF A COMPETITIVE INDUSTRY

1. **A Simons Model of Competitive Equilibrium**
   a. Part I. Initial Equilibrium State
   - Exhibit 6.2: Graph of Firm’s Cost Curves
   - Exhibit 6.3: Short-run Equilibrium of Supply and Demand
   b. Partial Displacement And Short-Run Equilibrium Adjustment
   - Exhibit 6.4: Costs to Firm Paying ASD Royalty
   c. Progress Toward New Equilibrium
   d. Additional Implications
   e. A Simple Mathematical Model [Optional]
   - Exhibit 6.4b: Solutions of Market Model

2. **Bankruptcy Laws And Fixed Costs**

## C. INCIDENCE ANALYSIS: SHIFTING AND PASSING ON

- HANOVER SHOE, INC. v. UNITED SHOE MACHINERY CORP.
- ILLINOIS BRICK CO. v. ILLINOIS

## D. ANTICOMPETITIVE SUPPLY RESTRICTION

- Exhibit 6.5: Expansion versus Market Spoilage
- CLARK v. UNIVERSAL BUILDERS, INC.
- STATEMENT OF U.S. DEPT. OF TRANSPORTATION BEFORE THE INTERSTATE COMMERCE COMMISSION, WESTERN COAL INVESTIGATION-GUIDELINES FOR RAILROAD RATE STRUCTURE
- GHEN v. RICH
- CONSTRUCTION INDUSTRY ASSOCIATION OF SONOMA COUNTY v. CITY OF PETALUMA
<table>
<thead>
<tr>
<th>Exhibit</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exhibit 6.2a</td>
<td>Correctly Graphed Cost Curves</td>
<td>289</td>
</tr>
<tr>
<td>Exhibit 6.3a</td>
<td>Completed Graph of Demand and Shifting Supply Curves</td>
<td>289</td>
</tr>
</tbody>
</table>
CH. 1: ANALYZING CHOICE UNDER ALTERNATIVE RULES

A. INTRODUCTORY CONSIDERATIONS

There can be little doubt that one primary purpose of legal systems is to modify human behavior, to induce at least some people to act in ways that they would not choose but for the pressure of legal incentives or disincentives. If this behavior-modification or “channeling” function of law is truly important, then it also becomes imperative to understand the predictable behavioral implications of alternative legal doctrines and policies. When a particular rule change is enacted, exactly what will happen? Or, if we cannot predict “exactly” what will happen, can we make any useful generalizations at all about the results of a proposed modification in law? And if it is indeed impossible to make useful predictions about consequences, then on what grounds are institutional changes to be advocated?

These observations suggest that the study of Law can be thought of as having a close affinity to the social sciences. Jurisprudence is, after all, inevitably concerned with predicting and describing the behavior of human beings under alternative institutional arrangements. Increased awareness of this is precisely why the nexus between Law and traditional social sciences, especially Economics, has become the focus of intensified interest in recent years. Methodological tools and concepts of the social sciences are increasingly applied in legal scholarship because they provide insights about the interaction of human beings judges, parties litigant, potential tortfeasors, etc.-as they confront alternative legal rules.

The use of economics in legal analysis is sometimes viewed with alarm and alleged to be infected with ideological connotations. Of course, any “approach” to legal analysis can be, and frequently is, bent to serve ideological purposes. Just bear in mind that economic analysis is no more and no less subject to this danger than many another tool of intellectual inquiry. Whatever the goals of its users, the power of economic analysis to predict and describe many facets of human behavior has no necessary link with value judgments about what conduct or institutions should be classified as good, bad, or indifferent. One can, after all, describe in essentially neutral fashion the objective implications of alternative rule systems without sitting in judgment on the results themselves. Occasionally, the mere laying bare of policy implications seems to lead rather directly to value judgments, but this is not an inherent characteristic of the analysis itself; rather, it reflects what is frequently the natural reaction of an observer to any revelatory process.

Economists use special terms of art to distinguish descriptive-predictive analysis from prescriptive-judgmental statements, labeling the two analytical modes as “positive economics” and “normative economics,” respectively. Inevitably, objective economic analysis and subjective opinion tend to become intertwined in the hands of many practitioners. Nevertheless, an expert in positive social science analysis does not, merely by virtue of his technical expertise, warrant having any greater-than-ordinary deference paid to whatever purely normative opinions he may from time to time intermingle with his analysis.

On the other hand, a reasonable ability to describe the implications of alternatives does seem to be a necessary precondition for competent formation of value judgments, if forming value judgments is what one ultimately is interested in doing. After all, unless the objective consequences of alternatives are first correctly understood, how can subjective value weightings be rationally and intelligently attached? One useful result of analysis therefore is to expose counterintuitive effects or ill-considered indirect consequences of a policy. Indeed, it is not uncommon to hear economic reasoning used to suggest that a legal policy adopted in pursuit of some goal is unintentionally counterproductive. Hence, although prediction is not itself an inherently normative act, it serves as an essential input into the process wherein people ultimately arrive at reasoned value judgments about law-or any other public policies, for that matter. From this point of view, predictions about the economic consequences of legal decisions are a valid and useful part of the “means to ends” debate in a legislative or judicial proceeding. Of course, in an adversary proc-
ess, it is perhaps unsurprising that the predictive inputs are frequently selective and self-serving. Still, such inputs are no more and no less than specialized and occasionally quite persuasive formal tools of legal argumentation. Potentially powerful tools of legal argumentation can be ignored only at one’s peril. In fact, the probability of having to deal with economic reasoning in law seems to be growing, whether in the form of affirmative arguments to be advanced and supported or as allegations that someone else raises and which, therefore, must be probed, criticized, and rebutted.

Advocacy is not the only use of economic-type behavioral analysis, however. Sometimes the application is quite simply educative and explanatory, an attempt to understand something or reconcile it with apparently conflicting information. For instance, a pervasive concern in legal studies is the role of coercion. Behavioral strictures rise above the status of mere exhortations and become “laws” precisely because they are backed by a sanction or coercive force. Yet, in a democratic society, laws supposedly arise out of the consent of the governed. To many, this smacks of mystery, or at least paradox. Why would people consent to-and, presumably, benefit from—a system of coercion? Some of the early analysis in this chapter is directed precisely at the question of why self-imposed coercion is a plausible and voluntarily chosen response to a wide class of societal problems.

Behavioral analysis frequently reveals that results which appear to have been chosen freely and voluntarily are, on closer inspection, not so greatly desired as they superficially seem. This is a very important realization because the fact that a decisionmaker freely chooses an option is often offered as an allegedly persuasive empirical proof that the chosen option must be “better” than the other available choices. Otherwise, why did the person choose what he did rather than something else? In fact, whereas there is indeed a kernel of truth in that somewhat Panglossian argument, it is at best a prima facie argument or presumption, subject to qualifications that are of great relevance in precisely the factual circumstances that typify many legal applications. Even in this initial chapter, we shall see that circumstances frequently induce people to choose results that they would certainly regard as “inferior” to other available results.

Another recognizable objective of economic analysis in Law is that of mere explanation. Under this rubric, one may hypothesize about why people did historically make a particular policy choice without also necessarily endorsing the values of the original decisionmakers. This involves a process of working backwards, of discovering a plausible rationale for observed behavior, of ferreting out what may have been the motivation for a law. One is then free to take the additional step of approving or disapproving what are perceived as the underlying objectives of those who actually made the choices. Becoming aware of an underlying systematic basis for some class of phenomena can powerfully assist one in learning a body of doctrine, manipulating it, appreciating its nuances, and even predicting its evolution in response to changing conditions.

This chapter introduces some elementary concepts that will be useful in reasoning about legal phenomena in terms of behavioral science concepts. Initially, the behavioral science concepts employed are not notably “economic” in character. Although the vocabulary of the economist’s conceptual bag of tools has become increasingly important in legal applications, the application of economic terminology should not imply any claim that the phenomena being analyzed are necessarily economic in the narrow sense of the term. In fact, it just happens that certain conceptual tools created by economists for the analysis of explicitly economic transactions can usefully be adapted to the legal environment. Legal issues frequently do have an explicitly economic component, of course. Nonetheless, if the methodology presented below is properly understood, it can legitimately be applied in many areas where mere economic motivation would be a gross oversimplification. In many cases, the language of economics represents merely the application of a specialized tool of formal logic, a deducing of conclusions from premises in a way somewhat akin to the still more abstract methods of symbolic logic and mathematics upon which, in turn, economic theory itself draws. It is perhaps true that the more formal and elegant a mode of analysis, the more an unwary person runs the risk of being overly dazzled. On the other hand, formal reasoning is valuable precisely
because it permits a careful observer to scrutinize critically both the premises and the logical links that are alleged to justify certain conclusions.

Finally, one should be sensitive to the limits of economic argumentation in legal contexts, even when “economic” is construed in its most expansive and generous sense. In some circumstances, arguments based on economic reasoning will have undeniable relevance and potent persuasive force. In other situations, economic factors may be of relatively trivial weight as compared to amorphous-yet perfectly valid-ethical, moral or even viscerally instinctive notions of what is right and just. As with any other source of legal arguments, economic factors must ultimately be evaluated through each individual’s own views of their relevance and weight. Economic analysis is not a single great searchlight that will penetrate and illuminate every nook and cranny of the law, but neither is any other “approach,” whether it be rooted in ethics, sociology, legal history, or some other discipline that can be brought to bear on legal problems. Since this is a book about economic reasoning in law, practical constraints will impose what may occasionally resemble methodological tunnel vision. With only that caveat, the reader is credited with the good sense to integrate, where necessary, the narrowly focused materials that follow into a more fully articulated intellectual framework.

The material in this book is not intended to be elegant or highly technical economics. An attempt is made to develop and apply only that limited set of economic-type constructs that is most useful in legal reasoning. Even so, there is relatively short shrift given to many traditional economic topics that are adequately treated elsewhere, in standard economic textbooks at an elementary or intermediate level. The analytic concepts are introduced here in a carefully phased process wherein the applications grow progressively more sophisticated. Considerable cohesiveness and intrinsic interest has been designed into these legal applications, since the intent is for this to be genuinely a law book rather than intermediate economic theory thinly cloaked with market-linked legal issues. But the reader should be warned that communication of underlying analytic principles is the real goal and the organizing principle of these pages is the economics, not the law. Accordingly, communication of economic content is frequently pursued in ways detrimental to a potentially more informative treatment of the substantive law involved.

Several of the first few concepts introduced below are originally drawn from game theory. Why give such a primacy of place to game theory? A standard dictionary definition of a game is: “a contest, physical or mental, according to set rules, undertaken for amusement or for a stake.” Two elements of that definition should be especially noted. One is that games are explicitly defined as behavioral situations governed by rules. The second is that game theory deals neither exclusively nor even primarily with contests undertaken for amusement; rather, it analyzes the behavior of “players” in serious conflict situations, frequently for heavy stakes. Conflict within a set of rules is what a great deal of the law is really about. Some of the “games” affected are played out in a marketplace where buyers and sellers compete for resources subject to a set of property and trading rules. This is the traditional domain of economic analysis. Legal proceedings can themselves be conceptualized as games, governed by their own peculiar system of rules, entitlements, costs, etc. Many of the underlying conflict situations giving rise to legal proceedings can also be understood as games. In any case, students of the Law ought to be fundamentally interested in how the strategies of players change and how the predicted outcomes vary as the applicable rules of various law-related “games” are modified.

QUESTIONS

1. There are many situations in which some agenda of societal goals and the trade-offs between them has already been fully articulated and predetermined. A legal analyst’s role, therefore, is merely to offer counsel to decisionmakers regarding the advisability of alternative laws as means to these ends.

   a. What real-world jobs for a lawyer fit the above scenario? In what respects is predictive competence necessary?
b. Drug addicts are responsible for a large amount of crime. This crime is thought to be motivated by the addicts’ eagerness to “earn” the income necessary to support an expensive drug habit. In 1973, the State of New York passed tough new mandatory sentencing provisions which increased the penalties for drug distribution. It is plausible to assume that the legislators were attempting to reduce the availability of drugs and the attendant crime problem. Are tougher drug laws an effective means to that end? Does your answer depend on empirical data?

c. Supply some specific examples wherein it can persuasively be argued that a statute or common law rule adopted as a means to an identifiable end is actually detrimental to that end. Can you supply another example of a law that has unintended “byproduct” consequences that substantially vitiate the benefits of what was originally viewed as its direct or primary objective?

2. You are a basketball fan. A proposal is made to penalize a team two points per inch or fraction thereof for the number of inches by which the average height of its players used in a game exceeds 76.

a. Could a basketball coach help you in describing what effect this rule change would have on how the game is played? What kinds of things would the “expert” be able to tell you that you couldn’t figure out for yourself? Would you expect any surprises? (E.g., can you advance any argument that suggests that having one or more real giants on a team would become more important than ever before?)

b. Once he had expounded on the consequences to the game, to what extent would you be willing further to defer to the expert’s opinion as to whether the resultant changes would make the game better or worse? If possible, provide an example of a sports rule wherein you suspect that the preferences of coaches and fans are antagonistic.

c. Are the “experts” always right in predicting consequences? NCAA rules allowed conferences to use several kinds of “experimental” basketball rules-such as the three-point shot and the time clock-for the first time during the 1982-83 college season. If you do not remember yourself, ask a friend who is a real fan. When are experts most likely to be right or to be wrong?

3. Richard Posner and others have argued that common law legal rules implement goals of “economic efficiency.”

The rules assigning property rights and determining liability, the procedures for resolving legal disputes, the constraints on law enforcers, methods of computing damages and determining the availability of injunctive relief-these and other important elements of the legal system can best be understood as attempts, though rarely acknowledged as such, to promote an efficient allocation of resources.¹

This position is, as might be imagined, a controversial one. In fact, the very sense in which the term economic efficiency should be understood is itself subject to some debate.

a. Suppose that economic efficiency were defined as a situation in which potential economic gains were fully exploited or, what is pretty much the same thing, “wasteful” conduct were discouraged

and avoided. How plausible does it seem to you that a large part of common law doctrines might be explicable in terms of economic efficiency goals?

b. One seldom finds explicit use of economic analysis in judicial decisions, a fact which does not seem to support the assertion that economic considerations are determinants of judicial policy. The following remark by Prof. Neil Komesar may be interpreted as casting some doubt on the strength of that evidence:

Traditional legal analysis teaches that the reasons articulated by the decisionmaker are seldom sufficient-and are sometimes irrelevant-as indicators of the actual determinants of decisions. Judicial opinions are more often observations to be explained than sources of explanation. They yield insights only to one who can approach them systematically.²

What do you think? To what extent are written opinions themselves probative as to the underlying motivation of judicial decisions?

c. One objection to the efficiency hypothesis about common law formation is that many jurists never have and never will know very much about economics. To what extent would a Darwinian-type “survival of the fittest” explanation meet that objection?³

d. Assume that the common law efficiency hypothesis were persuasively established as a more or less accurate reflection of historical fact, but you are a Marxist and regard the pursuit of economic efficiency as a vicious capitalistic excuse for grinding the faces of the poor. Would knowing the economic efficiency implications of various legal rules have any value to you as a lawyer?

4. The following rule is proposed: when property subject to a security interest is claimed by a creditor and sold, the debtor must be credited with the actual proceeds of the resale or the fair retail value, whichever is higher.

a. What are the implications of such a rule if adopted? Are such implications likely to be important in determining the decision? If your answer would be different depending on the forum (court, legislature, regulatory commission, etc.), indicate why.

b. Would you know how to cross-examine an economist testifying as to these implications? If the forum were a court, would expert testimony of that type be allowed?

B. ELEMENTARY GAME THEORY

1. THE PRISONER’S DILEMMA

The following scenario illustrates a famous game-theoretic model that is attributed to the mathematician A. W. Tucker. Widely known as the “Prisoner’s Dilemma,” this situation, together with several closely related fact patterns, provides powerful insights into behavioral problems that have importance in the law.⁴ The scenario presented here is embellished slightly for pedagogical purposes but, as the title suggests, closely follows Tucker’s original anecdote.

---

³ For instance, Rubin, supra, posits an evolutionary mechanism guided by the decisions of the litigants rather than the judges. The evolutionary approach is criticized in Cooter and Kornhauser, “Can Litigation Improve the Law Without the Help of Judges,” 9 J.Legal Stud. 235 (1980).
⁴ There is an extensive literature dealing with this famous game-theoretic concept. For general discussions at a reasonably accessible level, see R. Luce and H. Raiffa, Games and Decisions, 94-102 (1967); A Rapoport and A. Chamah, Prisoner's Dilemma (1965); A. Rapoport, N-Person Game Theory (1970).
Imagine that you are an urban clerical worker who stops off in a downtown bar after working late one night. You encounter a kindred soul, hitherto a perfect stranger. In the course of having several drinks, the two of you commiserate over your mutual senses of extreme financial deprivation. One thing leads to another and, after suitably cautious circumlocutions, this new friend finally suggests that there might be a very nice gain indeed from knocking over a certain downtown jewelry store. In true Hollywood fashion, the fadeout of one scene depicts the larcenous proposal. Then, cut to the next scene wherein the two of you are braced against a building wall being frisked during an arrest by the Metropolitan Police. The jewelry store alarm bell is ringing and a few nice rope-of-gold necklaces are drooping accusingly out of your pocket.

QUESTIONS

1. First a predictive question. If it were necessary to complete the “script” of the above scenario, what is a plausible ending to the story? How sure are you?

   c. First, assume that an agreement between the two prisoners not to testify is unenforceable by any sanction, whether legal or retributive. Would both prisoners testify? Neither one? How would they behave and why? How would you behave under similar circumstances? Does it matter whether or not the prisoners are allowed ample opportunity to communicate with each other?

   d. If an agreement not to talk is unenforceable legally, one or both parties may nonetheless attempt to convince the other that an agreement not to testify will be enforced by subsequent retribution against the party who “squeals.” Assess the credibility of this threat. Would the prisoners be better off if they both belonged to the Mafia?

   e. Suppose that a court of law would provide traditional damage remedies for any agreement between the prisoners. Specifically, the prisoners could make a legally binding contract and the party who breached the contract would be held liable for damages suffered by the breachee. This possibility suggests another predictive question: Exactly what contract would you expect them to make?

   f. How can the prisoner’s dilemma concept be applied more generally to the whole field of executory contracts? What relationship is there to the traditional legal concept of “reliance”?  

2. Now some “normative” considerations.

   a. “Should” such a contract between prisoners be enforceable? Why?

   b. A corollary question. Is it appropriate that prosecutors be permitted to place prisoners in conditions that produce a Prisoner’s Dilemma?

3. Now a conceptualization. Assume that you are not able to use the above scenario as an illustration. How would you explain or define a Prisoner’s Dilemma to someone else?

4. A policy application. It has sometimes been suggested that the same attorney should not be permitted to represent two co-defendants being tried in connection with the same criminal incident. The thesis is that conflicts of interest inevitably arise between two defendants such that no common position can fairly represent the interests of both clients.

   a. What relevance does the Prisoner’s Dilemma have to this proposed ban against joint representation? In the text above, a “dilemma” was created by the prosecutor’s structuring of a possible deal. Does joint representation address difficulties that may arise merely out of the facts of the case and the applicable law?

---


b. Prof. Gary Lowenthal observes the following:

To the extent that counsel do not cooperate with each other, it reasonably can be inferred that the interests of the clients conflict. When clients have conflicting interests, one attorney representing all of the defendants can present a unified strategy only at the expense of prejudicing one or more of the clients.

Lowenthal, supra, at 986. Do you agree?

5. Finally, an attempt to generalize or analogize: describe some other fact situation wherein the participants seem to be confronted with a behavioral problem similar to the Prisoner’s Dilemma.

a. THE STUDENT’S DILEMMA: UNIVERSITY v. EAGER

The following facts describe a situation that arose in a science course at a large state university:
The University has a rule that penalizes cheating by dismissal and the notation “Dismissed for Cheating” on the offender’s transcript.
The mechanics of the examination process are such that it is very easy to cheat. There is a negligible risk of any penalty; the probability of being caught is very close to zero, but not zero.
Grades are distributed strictly on a required curve: A to the highest 10%, B+ to the next 20%, etc.
The course is a key subject required for pre-med students and is known to be an important consideration in the process of medical school admissions reviews. Hence, the benefits from high grades and the penalties for low ones are exceptionally great.

Lightning strikes! While the class is taking the final examination, the proctoring system manifests one of its rare, seemingly spasmodic surges into action. Fourth-year student Edward L. Eager is caught in flagrante delicto. The evidence that he did in fact cheat is absolutely conclusive; consequently, Ed faces dismissal.

Chris Pensive, another student in the class, is troubled by Ed’s predicament and was overheard to make the following comment.

“Something is wrong, all out of proportion, with the way Ed is being treated. Maybe what he did wasn’t right, but he’s being punished almost as if he were a criminal. Under the circumstances, his behavior just doesn’t seem that bad to me. Ed was almost under duress, almost compelled to do what he did. I don’t blame him. I blame the school for holding exams like that.”

Ed now faces an administrative hearing prior to the imposition of the penalty pursuant to University rules.

QUESTIONS

1. Is a Prisoner’s Dilemma among the students involved here? Or is the situation “something like” a PD?
2. How would you defend Ed if assigned as his counsel at the hearing? Is the University arguably at fault? Is Ed subjected to an irresistible temptation? Is this a case of “self defense”?
3. Chris apparently finds the penalty disproportionate. How should an appropriate penalty ideally be determined?
4. Bribery of foreign officials by the agents of U. S. corporations operating abroad is subject to sanction under U. S. law. Does this bear any similarity to the Ed Eager scenario?
5. What are some real-world analogues of University v. Eager? What legal or policy issues are suggested?
b. The Prisoner’s Dilemma as a Matrix Model

The use of models is almost ubiquitous in social science. In one sense, the term “model” has almost its vernacular meaning of a representation. When used in a theoretical context, however, a model transcends mere representation; indeed, it may to some extent deliberately sacrifice accuracy of representation for a more abstract and general description that highlights the most important elements in the phenomenon being “modeled.” Models may even be caricatures, having the same goal as a deft cartoonist who, with a few expertly chosen but highly oversimplified lines, seeks to convey a meaning more clearly than through a photographic likeness. It must already be apparent that the Prisoner’s Dilemma is a conceptual model, a paradigm or pattern with potentially wide application. In short, the insights of the PD can be analogized to many other interesting situations, as the Ed Eager scenario in the last section exemplifies. It is interesting to see, however, that the paradigm of the Prisoner’s Dilemma can itself be modeled in several different ways, varying greatly in their degree of formality. One approach is the one just used: simply to tell the story, employing only ordinary English words similar to those used in the text above. Alternative descriptions, or models, of the same situation may convey additional insights, or at least conceptualize the situation with greater clarity. This possibility will probably be apparent as we now construct a more formal and compact rendition of the same Prisoner’s Dilemma conflict.

The Prisoner’s Dilemma is frequently presented in the form of a matrix model of the type shown in Exhibit 1.1. In this simple form of model, the rows of the matrix represent the possible choices open to a particular prisoner: either Testify or Silence. The columns represent the exogenously controlled circumstances. Hence, the “other” prisoner will either have chosen Testify or Silence. In creating these matrix models, the same general convention is usually followed: choices or “strategies” under the decisionmaker’s own control appear on the rows while the externally-controlled “state of nature” is represented by the columns. Each “cell” formed by the intersections of the rows and columns therefore corresponds to a unique combination of the individual’s own decision and the external conditions imposed on him exogenously, in this case by the other person’s choice.

Do the row and column labels in Exhibit 1.1 encompass all of the relevant possibilities both for the choosing individual’s own choice and the other individual’s choice? If so, then the resulting four cells properly represent the four possible outcomes of the Prisoner’s Dilemma. Such a representation of the possible choices and outcomes is an important first step in modeling. In more formal language, this process is frequently called identification of the “choice space” or “opportunity set.”

Once the possible outcomes are thus identified, the implications of each outcome can also be indicated. In Exhibit 1.1, this is accomplished by placing a number in each cell to tell the outcome or “payoff” in terms of years spent in jail. Other kinds of numbers might be more appropriate for certain purposes. (Can you think of any drawbacks of using years as opposed to something else?)

In any event, the resulting matrix model does give a concise kind of “Picture” of the situation faced by any decisionmaker. It is now easy to see that the Testify strategy has what game theorists call the quality of “dominance”: no matter which column is assumed, the Testify strategy row always results in the best payoff, i.e., the lowest jail sentence. [Place a circle around the lowest jail sentence in each column of Exhibit 1.1.] In other words, dominance exists when, no matter what the external situation is (as represented by the columns), the “best” strategy for the decisionmaker is always the same row. This presence
of dominance is why we might reasonably predict that a prisoner who understands the situation will choose to Testify.

But the other prisoner faces the same kind of matrix—a condition known as “symmetry” of the payoff matrix for the two individuals and hence can also be predicted to Testify. That enables us to advance a step further and hazard a prediction about the particular outcome of the game that we are most likely to see: because Testify is a dominant strategy for both players, the expected “solution” of the game will be the lower right-hand cell. This, of course, is precisely what the Prosecutor’s office hopes for and is the reason why it has structured the “rules” of the game the way they are. A classical Prisoner’s Dilemma, then, falls into the wider category of games with a predictable outcome or “solution.” In economic models, such a solution will be called an “equilibrium” outcome.

Not all models have determinate solutions or equilibrium outcomes. [The “Chicken” game described later in this chapter will be one that does not.] Even when there does exist what seems to be a determinate theoretical solution, one should be wary of the fact that models are almost always abstractions and simplifications of reality. Still, within appropriate limits, models are powerful analytic tools because they help clarify what forces are at work and how these forces interact. Hence, it is properly cautious to think of models as formalized descriptions of “what tends to emerge” under specified behavioral conditions and factual assumptions.

As we have already seen, models may also be articulated in mere vernacular speech rather than through formal constructions. Lawyers, after all, traditionally pride themselves on their ability to distinguish, deduce, and generally massage and manipulate the English language during any process of legal analysis. One thesis of these pages, however, is that there is frequently some advantage in going through a more formal modeling process, of “writing down” the behavioral conditions in slightly more technical fashion. For instance, do you find that Exhibit 1.1 conveys the essence of the Prisoner’s Dilemma game more clearly than mere words? Most, though not all, people feel that it does. The construction just used, the matrix model, is a relatively simple one. As subsequent situations are analyzed, the modeling will require many additional building-blocks of a conceptual type. These building-block concepts are labels and analytical devices that provide tools of thought in much the same way that having more brushes and colors enables an artist to paint better pictures. Equally important, many of the concepts are concise and convenient means of communicating about complex situations with others who “speak the language,” just as terms such as negligence, consideration, and reliance are key concepts that facilitate communication among students of the law. That is why it is important to have the patience to learn a certain amount of behavioral science jargon and terminology.

At the same time, a caveat should be issued as we undertake to add to the arsenal of technical terms and concepts: beware of using technical jargon in an inappropriate forum. Economic jargon is frequently not only incommunicative but, worse yet, runs the risk of arousing negative reactions from those—including senior partners and powerful judges—who have not themselves assimilated the language. After working out the implications of a situation in more rigorous terms, therefore, the legal practitioner is often confronted by the necessity to translate arguments and conclusions into more traditional legal terms. Although the translation process may require care and sensitivity, what truly makes sense in one language ought to be susceptible to communication in another form, even if less elegantly. In sum, rigorous terms of art may be ideal in working out one’s own personal analysis, but ultimately the language of argument and persuasion must be carefully adapted to its audience.

c. THE “PRISONER’S DILEMMA” DEFINED

Labeling various behavioral conflicts with names such as Prisoner’s Dilemma is merely a technique of classifying situations in a way that may lead to generalizable insights; one should not take the strict requirements of the classic “models” or paradigms too seriously. Nevertheless, a significant advantage of
formalizing definitions is that distinctions between otherwise similar situations become apparent. Note that, although we have used the concept and discussed it at some length, a Prisoner’s Dilemma has not yet been formally defined. Although you may think that you more or less know what one is, could you really distinguish a classic PD-type situation from just a garden-variety “tough spot”? A rigorous formal definition should clarify just what are the important elements of the phenomenon being discussed. Once these elements of the definition are isolated, one can more easily recognize relevant distinctions among closely related phenomena. Looking ahead, a formal definition will suggest why it is worth distinguishing Prisoner’s Dilemma from one of its close lineal descendants in the game theory family, the game of Chicken.

In order to construct an appropriate definition, the structure of the Prisoner’s Dilemma game will once more be presented in its now-familiar matrix form. In Exhibit 1.2, the standard conventions are used whereby the rows designate the strategies, the things the player does control, while the columns reflect the “states of nature” that are externally controlled. “Cooperate” is the strategy that involves cooperation among the people in the Dilemma and “Defect” is any departure from this mutually beneficial behavior. By substituting abstract symbols for the numbers in the cells, however, a great deal of additional generality can be derived from the model. Hence, the algebraic symbols a, b, c, and d are now used to represent the payoffs associated with each box or “cell” of the matrix. Note that these symbols may still signify years in jail, as postulated in the original scenario. They are also consistent, however, with payoffs denominated in other kinds of units: money, an index of “satisfaction,” etc. The payoff matrix is assumed to be “symmetric,” that is the other player is looking at a payoff matrix that has the same structure.

Whereas we have previously described the Prisoner’s Dilemma only in rather imprecise terms, use of symbols in the payoff matrix permits a formal definition that is quite generalizable. In order to qualify technically as a Prisoner’s Dilemma, there must exist a particular set of relationships among the relative magnitudes of payoffs a, b, c, and d in Exhibit 1.3.

Exhibit 1.1: PD Matrix Model

<table>
<thead>
<tr>
<th>Your Own Behavior</th>
<th>Behavior by the Other “Player”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cooperates With You</td>
</tr>
<tr>
<td>Cooperate With Other</td>
<td>a</td>
</tr>
<tr>
<td>Defect</td>
<td>b</td>
</tr>
</tbody>
</table>

Specifically, the payoff matrix of any PD situation must satisfy these three conditions:
1. \( b > a \) and \( d > c \), so that Defect is a “dominant” strategy, one which always leads to the highest payoff no matter what behavior is expected of others.
2. \( a > c \) and \( b > d \), so that more cooperation by the other player always increases one’s own payoff.
3. \( (a + a) \) exceeds the similar summation of the two players’ payoffs for any other outcome, so that the cooperative solution is really “best” from the standpoint of the players.

The function of conditions 1 and 2 is probably easy to see: the former makes it advantageous for each party to defect, and the latter makes that defection costly to the other party. Condition 3 is a bit more subtle, and at this point its function need be indicated only briefly. Suppose that the sum of the payoffs in a Cooperate/ Defect outcome were higher than in a Cooperate/ Cooperate, i.e., that \( (b + c) > (a + a) \). Assume, then, that a proposed “deal” involves one player being randomly selected to cooperate and the other to defect. Why might this proposed deal be regarded as possibly superior to a Cooperate/Cooperate solution? (If you don’t see at least intuitively why, return to this question after the treatment of “expected values” in Chapter II below.) In short, condition 3 is there merely to avoid the possibility that Cooperate/ Cooperate is not really the desired solution.
In addition, some social scientists attribute a behavioral significance to the within-column payoff variations \((b - a)\) and \((d - c)\). These are termed “temptation differentials” since they show the incentive to profit by abandoning a cooperative mode of behavior. The larger these are, the less cooperation is predicted, ceteris paribus. Similarly, the diagonal difference \((a - d)\) is called the “cooperation differential” since it is an index of the incentive to achieve the cooperative solution.

2. ANOTHER IMPORTANT, CLOSELY RELATED GAME: “CHICKEN”

One use of the formal PD definition just derived is to distinguish the Prisoner’s Dilemma from other similar “dilemma”-type situations. For instance, another interesting game, called “Chicken,” can be created by altering assumption 1 above so that if the other party is sufficiently uncooperative then cooperation becomes the optimal strategy. Specifically, assume \(b > a\) but \(d < c\). If you were sure that the other person were not going to cooperate, now it would actually be better to cooperate than to remain noncooperative. Since the other player would react similarly if convinced that you would not cooperate, the circumstances are such that it pays to convince the other party of one’s stubbornness and obduracy. Note that \(b\) is the highest payoff achievable (you defect, other cooperates). This can be gained by successfully bluffing your opponent.

As will become apparent in subsequent sections, bluffing problems are frequently involved in situations of interest to lawyers and that is why Chicken is an important concept. A Chicken Game will be exemplified later in this chapter in the form of the Road Problem, dealing with negotiating agreement on road repair levels. After reflecting on the Road Problem in its purely verbal form, it will be profitable to fit the words into the sort of formal matrix model just laid out. With that forewarning about the role of the Chicken model, we shall develop some additional terminology and then turn to the factual scenario.

3. OTHER USEFUL JARGON: “EXTERNALITIES”

The Prisoner’s Dilemma has been chosen as a starting point for several reasons. One is that--as will become clearer and clearer as the applications multiply--the situation is an interesting one per se; it has a more or less direct bearing on many interesting legal phenomena. A second reason, however, is that it helps to edge one into a whole set of terminology that is helpful in conceptualizing and distinguishing other situations.

A fundamental characteristic of the Prisoner’s Dilemma situation is the presence of some act whose benefits exceed its costs (if any) for a single individual but whose aggregate costs to all affected parties exceed its benefits. Condition 3 in the definition of the PD given above implies that this is true; since the \((a + a)\) aggregate payoffs to the two individuals at the Cooperate/Cooperate solution are larger than the summed payoffs at any other solution, then any defection from that solution necessarily involves a lesser magnitude of gain to the defector than the magnitude of the loss imposed on the other player. In pursuing their individual benefits, the players in the Prisoner’s Dilemma game will tend to impose on each other what may be a very considerable amount of costs that overbalance any attendant benefits. It does not help the situation that each individual can recognize the ultimately unsatisfactory outcome, the mutually self-defeating nature of the behavior involved. This point deserves some emphasis. It is a simple but dismayingly common error to believe that the reciprocal nature of the damaging behavior either “cancels out,” and can therefore be neglected, or that their own recognition of the reciprocity of the damages will cause the “players” to desist from the behavior.

The Prisoner’s Dilemma anecdote sets the stage for some additional concepts dealing with costs and benefits. Specifically, there are several economic terms of art worth introducing at this point, both because they are useful in further articulating the Prisoner’s Dilemma type of problem and also because they have important applications in other contexts. The terms in question deal with what economists call “externalities.”
Economists classify the effects of any act as being “internal” to the extent that they are felt only by the actor and “external” to the extent that they affect third parties. Hence, the Prisoner’s Dilemma falls in a class of situations where the internal benefits selfishly pursued by any individual are exceeded by the external costs imposed by others engaged in similar activity. The regulation of many types of external effects, both costs and benefits, will turn out to be a very pervasive theme in Law and Economics problems. Unfortunately, a variety of terminology is sometimes used to describe external effects. Sometimes they are referred to simply as “externalities” or “spillover” effects. Also, the words “private” and “social” are substituted for internal and external, respectively. These alternative sets of terminology are essentially interchangeable. The key question involved is whether the impact of an act is confined to the person who performs it and, if not, whether others are affected beneficially or detrimentally.

The Prisoner’s Dilemma illustrates a situation where the external effect is a cost, a harmful effect. Consider what happens to the “other” prisoner if one prisoner decides to testify rather than remain silent: the second prisoner suffers an increase in the years that he will spend in jail. Thus, the testifying creates an internal benefit for the prisoner who squeals, but it levies an external cost on the second prisoner. The exact magnitude of the external cost depends on whether the second prisoner is already cooperating with the prosecution or not.

There is a rule of thumb about external effects that is worth remembering: activities that produce external costs tend to generate the perception of a “problem” involving “too much” of the activity in question. The PD game, of course, is a classical illustration of the imposition of heavy external costs because of (as viewed by the prisoners) “too much” testifying. We shall shortly be dealing with problems caused when the external effect is a benefit rather than a cost. Such a situation is generally regarded as producing a problem of exactly the opposite sort: a tendency for there to be “too little” of the activity that produces the external benefit.

QUESTIONS

1. Under the facts of the original PD scenario presented above, what are these magnitudes of external cost that the prisoners impose on one another? What are the associated internal benefits to the prisoner who squeals? (Answer as specifically as possible, e.g., in terms of years or some other quantity.)

2. Suppose that a prisoner who squealed were penalized by having to serve additional time equal to the external costs imposed on the other prisoner. What would the outcome be in terms of the amount of cooperation with the prosecution that you would then expect?

3. Now suppose that a prisoner had to compensate the other for any external costs suffered, but that the compensation could be in the form of a monetary payment. (This would be the equivalent of treating the squealing as a tort.) Would either prisoner testify? Indicate any additional factual assumptions necessary to answer this question. [Hint: Do we have to know anything about the money value of years of freedom to each prisoner?]

4. One of the first applications of the Prisoner’s Dilemma model in the economic literature was in relation to the decisions of property owners whether or not to invest in maintenance and improvements. See Davis and Whinston, “Economics of Urban Renewal,” 26 J.Law and Contemp.Prob. 105 (1961). Do such activities produce “externalities”? What would you guess was the structure and content of the PD matrix used by Davis and Whinston? Why would they apply this model to explain such things as building codes and “urban renewal” through government acquisition and redevelopment of large tracts?

4. ETHICAL AND OTHER EXTRALEGAL BEHAVIORAL REGULATIONS

The simple game-theoretic notions introduced above obviously provide one possible rationale for the coercive function of a legal system in removing individuals from predictably mutually-destructive
“dilemmas.” The role of “liability” or punishment in channeling behavior will be an important theme in much of the analysis below. One should not, however, suppose that extralegal institutions have an unimportant role in the social problems addressed by laws. Extralegal behavioral influences may supplement law in important ways or even obviate its need entirely. On the other hand, the breakdown of extralegal pressures may supply the impetus for creation of new or modified legal rules.

This general subject is, for instance, dealt with by ethician J.L. Mackie in terms of exactly the game-theoretic methodology introduced above. He analyzes the following anecdote as a variant of the Prisoners’ Dilemma. [Ethics: Inventing Right and Wrong, 115-121 (1977).] Tom and Dan are soldiers assigned to adjacent fortified posts in the path of an enemy attack. If both remain at these posts, they have a reasonable chance of holding out until a relief column arrives, and so of both surviving. If they both desert their posts, the enemy will break through immediately and the chance of either of them surviving is markedly reduced. On the other hand, if one stays at his post while the other flees, the coward will have a much better chance of survival than if both soldiers remain. In this latter case, the one who stays will have a worse chance of survival than if they both deserted their posts.

Under the circumstances, it would be rational for both men to agree to be literally chained to their posts. Each soldier would benefit from this loss of his own freedom of choice, provided that his comrade’s freedom was similarly constrained. An effective alternative would be some external discipline that confronted a deserter with the prospect of a sufficiently severe punishment, such as execution. But Mackie stresses that there can be psychological substitutes for physical chains and external penalties. Military traditions of honor and loyalty to comrades can serve as invisible chains, while the stigma of cowardice, with its attendant disgrace and shame, can substitute for external sanctions. Indeed, given the hypothetical situation, one would actually prefer to belong to a group whose members were encumbered with an appropriate set of psychological fetters.

Mackie also points out that if Tom and Dan have a general tradition of keeping agreements, they will be able to pledge that each will persevere at his post, and the agreement-keeping tradition will tend to hold each man there. In practice, however, a general agreement-keeping tradition is likely to be rather less effective in situations of extreme stress (such as our hypothetical one) than more focused, situation-specific mechanisms as exemplified by military traditions of honor and loyalty. On the other hand, a general ethic of agreement-keeping has the advantage of being more flexible and widely applicable than special-purpose ethical constraints. A tradition of observing agreements can support the making and keeping of all sorts of useful bargains. Thus, observes Mackie, Hume was quite right in saying that a man is the more useful, both to himself and to others, the greater degree of probity and honor he is endowed with.

The particular example used to illustrate this form of two-person game is both dramatic and realistic, but it has the disadvantage that it does not lend itself to repeated trials by the same two players. That is, it deals with a single-trial game rather than an “iterated” one. Hence, Mackie considers another example where even if Tom, say, comes off badly at the first trial he will still survive to play with Dan again. The following assumptions are made: (1) that each man has only a weak agreement-keeping tendency; (2) that neither can see, on any one occasion, whether the other is keeping the bargain until he himself is committed either to keeping it or to breaking it; (3) that if both men keep the agreement on one occasion, each is more likely to keep it next time, whereas if either or both men break the agreement on one occasion, each is less likely to keep it next time; (4) that all these tendencies are known to both men; and (5) that each time Tom and Dan play this game they know that they will have to play it again with one another.

These assumptions are said to alter the form of the game in the following way:

[If], on any one occasion, Dan is going to keep the agreement, it will be to Tom’s selfish advantage, with a view to the future, to do so too, though if Dan is going to break the agreement this time, it will be to Tom’s advantage to break it. And of course the situation is still symmetrical. Self-interest no longer unambiguously urges each man to break the agreement on any one occasion:
consequently only a fairly weak agreement-keeping tendency will be needed to tip the balance. Fairly obvious and natural assumptions lead to a similar conclusion if we extend the game in another direction as well, and assume that there are more than two players.\(^7\)

Mackie also considers the possibility of outcomes that are not symmetrical for the parties involved:

Even if Tom and Dan are initially placed alike, there may be several possible agreements between them, each of which is better for each man singly than the results of failure to agree or of failure to keep the agreement, but some of which are in various degrees more advantageous to Tom than to Dan, and vice versa. In these circumstances the man who is, or gives the appearance of being, the more reluctant to make, or to adhere to, an agreement is likely to get more advantageous terms. Though complete intransigence in either party is disastrous for both, incomplete relative intransigence is differentially advantageous to its possessor. This holds, as I have said, even if the initial situation is symmetrical; but if one party has less to lose by failure to agree, or less to gain from a stable agreement, further possibilities of unequal agreements arise.\(^8\)

Finally, the conclusion of this ethician’s section on “Game Theory” provides an appropriate combination both of endorsement of the game-theoretic conceptualization and of caveat about the limitations of any general model:

There can be no doubt that many real-life situations contain, as at least part of their causally relevant structure, patterns of relationship of which various simple ‘games’ are an illuminating description. An international arms race is one obvious example: another is the situation where inflation can be slowed down only if different trade unions can agree to limit their demands for wage increases. One merit of such simplified analyses is that they show dramatically how the combined outcome of several intentional actions, even of well-informed and rational agents, may be something that no one of the agents involved has intended or would intend. * * * The main moral is the practical value of the notion of obligation, of an invisible and indeed fictitious tie or bond, whether this takes the form of a general requirement to keep whatever agreements one makes or of various specific duties like those of military honour or of loyalty to comrades or to an organization. * * * The real weakness of the Hobbesian solution lies not in anything that the game theory models show, but in what, just by being models, they leave out. Real situations always incorporate, along with the skeletal structure of some fairly simple game, other forces and tendencies whose strength varies through time.\(^9\)

**QUESTIONS**

5. Provide an illustration of a situation wherein an ethical, moral or other extralegal compulsion seems adequately to regulate a form of behavior that might otherwise necessitate some more formal sanction system. Exemplify, if possible, a situation wherein the progressive breakdown of a formerly effective extralegal behavioral convention has already caused, or seems likely to cause, recourse to a formal coercive regulation.

6. Exactly how does a factor such as “guilt” or “approbation” change the PD matrices? What changes in the facts of the original PD scenario in the text above would make it plausible that these factors would eliminate the dilemma for the prisoners?

7. What circumstances determine whether extralegal rules are preferable to formal legal sanctions? Are the most relevant considerations on the “cost” side or the “benefit” side? Are relative “flexibility” advantages of interest?

---

\(^7\) J.L. Mackie, Ethics: Inventing Right and Wrong, (1977) at 118.

\(^8\) Ibid., 118-19.

\(^9\) Ibid., 120.
8. Even though there are fines for traffic infractions and legal liabilities imposed for driving-related
torts, there is arguably a certain gaming aspect to careful driving, an activity that produces “external
benefits.” This exchange in F. Scott Fitzgerald’s The Great Gatsby is instructive:

“You’re a rotten driver,” I protested. “Either you ought to be more careful or you oughtn’t to
drive at all.”
“I am careful.”
“No, you’re not.”
“Well, other people are,” she said lightly.
“What’s that got to do with it?”
“They’ll keep out of my way,” she insisted. “It takes two to make an accident.”
“Suppose you met somebody just as careless as yourself.”
“I hope I never will,” she answered. “I hate careless people. That’s why I like you.”

In your experience, do informal codes of “driving courtesy” differ substantially as one travels from
place to place? If so, how and why? Would you expect their degree of observance or “strength” to vary
predictably in accordance with any particular circumstances?

5. Re-read and critique the paragraph quoted above on the iterated game as played under Mackie’s
five assumptions. What is your own analysis of the difference it makes if the game is an iterated one
rather than a single-play game?

6. Suppose that a Prisoners Dilemma game is to be played in iterated form, but the number of trials is
known to be equal to n, where n is any positive number. What can you say about the results of this
game?10

7. Do you understand the paragraph quoted in the text above that deals with asymmetrical results and
the advantages of intransigence? If you do not understand it, return to this question after completing the
“Road Problem” in the next section below.

5. EXTERNAL BENEFITS, EXCLUDABILITY AND “FREE RIDERS”

As indicated earlier, external effects may be produced in the form of benefits rather than the harms
exemplified in the original Prisoner’s Dilemma situation. We turn now to a specific consideration of such
external benefit production. The “road problem” scenario provided below deals with external benefits and
the difficulties that occur when such benefits are potentially available at no cost. In the original Prisoner’s
Dilemma presented above, it is at least arguable that the prisoners “should” be kept in their dilemma.
Indeed, one function of Law may be to keep certain types of people in a dilemma-like environment
because, although this is detrimental to the “prisoners,” it is regarded as producing benefits to other parts
of society. By contrast, in the road scenario below, the sympathy or normative feeling that one has for the
players will be less ambiguous; most people would agree that the situation is somehow troubling and that
the “players” should be removed from the dilemma if possible.

10 Although logic may seem to suggest that repeated plays of the same game should not produce any change in
behavior, a number of game theorists argue that altered behavior will in fact occur. See, e.g., R. Luce & H. Raiffa,
Games and Decisions, at 102-04. A stronger case for unaltered behavior can be made when the number of iterations
(the value n in the text) is known. The nth iteration becomes, in effect, a non-iterated game and the non-cooperative
strategy is therefore clearly dominant. But, it is reasoned, if this is known, then there is no reason to cooperate on the
(n-1)th play. A similar logic can then, of course, be extended backward without limit.
C. THE LAW AND ECONOMICS OF PROPERTY RIGHTS

1. ROAD MAINTENANCE PROBLEM

Adams and Braun own lots A and B, respectively, of Section 236, bordering on State Route 116. The Adams lot is adjacent to the state road and the Braun lot is immediately behind it. Access to 236B is provided for in an easement under which the owner of 236A must allow free passage to occupants of all inner lots. Adams already has a house on lot A and has resided there for five years. Upon constructing the house, Adams also caused a road to be prepared, running from Route 116 to the back border of his lot. The road is not hard-surfaced, but is a type of “oiled gravel” surface common to private roads and driveways in rural areas. This kind of surface is subject to erosion and general deterioration over time. However, Pitts Gravel will come and provide any number of “doses” of road repair at a price of $25.

[This concept of a “dose” of an input is a pedagogical simplification that will be utilized a number of times below. It will be used when it is expositionally desirable to collapse a heterogeneous, multi-dimensional package of possible adjustments into a single dimension. Hence, “road repair” could be a mix of many different services in varying proportions, but no useful purpose is served by explicitly considering multidimensional adjustments. Through the simplification provided by the “dose” concept, we can focus on a straightforward quantitative adjustment.]

Fortunately, we just happen to be in possession of Adams’ true subjective evaluations of how much different levels of road repair would be “worth” to him. Column 2 of Exhibit 1.3 shows the total value of the services received from the road, assuming different levels of maintenance, i.e., different qualities of road. Note that these values are “gross” of any costs incurred in actually purchasing the maintenance.

<table>
<thead>
<tr>
<th>Number of Doses Provided</th>
<th>Total Benefits To One Person</th>
<th>Marginal Benefits To One Person</th>
<th>Marginal Benefits To Two Persons</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$50</td>
<td>-NA-</td>
<td>-NA-</td>
</tr>
<tr>
<td>1</td>
<td>$90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$120</td>
<td>$30</td>
<td>$30</td>
</tr>
<tr>
<td>3</td>
<td>$140</td>
<td>$20</td>
<td>$20</td>
</tr>
<tr>
<td>4</td>
<td>$150</td>
<td>$10</td>
<td>$10</td>
</tr>
<tr>
<td>5</td>
<td>$150</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

QUESTIONS

9. Fill in Column 3 of Exhibit 1.3 with the incremental values that each dose of road repair adds to Adams’ evaluation of the benefits from the road. In economists’ jargon, these values are called “marginal” evaluations. [Strictly speaking, the word “marginal” should be associated only with infinitesimal changes, but it is commonly used loosely to refer, as in the present case, to discrete changes.]

10. If only Adams were on the scene, how much road repair do you predict would be purchased? Why? Try to explain your answer in terms of marginal benefits and marginal costs.

11. Suppose Braun now builds a house on lot B. Braun has, by remarkable coincidence, the same preferences about this road repair as does Adams. Enter in Column 4 the aggregate 2-person marginal benefits now that both Adams and Braun benefit in precisely the same way from each unit of road
repair. (Since the frequency of their trips on the road is low, the possibility of detrimental mutual interference because of attempted simultaneous use can be regarded as negligible.)

12. How much road repair do you think will be purchased now? Who will pay what? Describe the negotiating process.

13. Enter Chernak, the owner of lot C. Mirabile dictu, Chernak also has the same preferences. What happens now? How do things change as Donatelli, Eglin, etc. also arrive on the scene?

14. In what sense might the people on the road regard the situation as “unsatisfactory”? Would they like to be coerced? To do what?

15. What difference, if any, did it make in this problem that the preferences of the individuals were assumed to be identical? That is, was it a mere “simplifying assumption”? Or did it affect the results in some significant way?

16. Can you explain this situation as due to a “defect in the applicable property rights”? Does the creation of patent rights fit this model in any way? Can you think of any other examples?

17. One way of providing roads would be for all roads to be privately-owned, excludable facilities, i.e., toll roads. Why is this not done? Is it because of the cost of exclusion? Other disadvantages?

18. In practice, what policies does society adopt in response to the impracticality of providing roads that are privately (excludably) owned?

2. JOINTNESS AND EXCLUDABILITY AS ELEMENTS OF PROPERTY

The road scenario illustrates that the concepts of ownership and property are sometimes elusive. Because of the way the legal rights regarding road use are structured, nobody can effectively be excluded from the benefits of any road maintenance once it is provided. Economists apply the terms “excludability” and “appropriability” in this context. If a good lacks excludability, then the one who provides it cannot appropriate from its users a return on his investment in its provision; that is, the provider cannot charge for or sell what he provides. This may be true even if the provider nominally “owns” what he has provided.

“Jointness” in consumption is the characteristic whereby some goods, such as the road, can be consumed by many people at once. The road case was, for simplicity, presented as a case of perfect jointness in consumption, i.e., where the value of the good to any single user is the same regardless of how many co-consumers there are. In practice, this is seldom true outside of narrow limits. But the fact that a good “congests” and begins to degrade in value beyond a certain point does not radically alter the kind of problem epitomized in the road anecdote. [Ask yourself what difference it would have made to the analysis if each co-user resulted in, say, a 5% degradation of the road’s prior value to the previous users.]

In combination, these characteristics of jointness and non-excludability present the opportunity for someone to be a “Free Rider,” one who receives the external benefits of another’s act and pays nothing in return. In the hypothetical presented above, there is the literal prospect of a free ride on the road, but the phenomenon is also widely generalized as the so-called “Free Rider Problem.”

In the hypothetical, there was nothing inherently preventing the excludability of the private road, so that the underlying problem could be remedied by a reformulation of the legal rights. Thus, you should be sensitive to instances in which courts can attack such problems by “creating” excludability in the form of new kinds of property rights. Sometimes, however, a mere legal fiat is meaningless, either because the enforcement of the exclusion right is prohibitively expensive, because enforcement interferes with the flow of benefits, or for some other practical consideration. Can you think of examples of such cases?

In sum, the rectification of a property right “defect” may or may not suffice to cure a problem of the type outlined above. Alternatively, some form of direct mandate of a result may be adopted. For instance many types of goods having very high excludability costs and jointness in consumption are provided
coercively by a government that dispenses the goods and then taxes individual beneficiaries for some share of the cost. Because jointly consumable, non-excludable goods are so commonly provided by governments, economists frequently refer to such goods as “public goods,” using the terminology of a classic article by Paul Samuelson\(^\text{I}\) on the welfare economics of the public sector. Samuelson derives “optimality” conditions for the level of so-called public good provision: jointly consumable goods should be produced up to the point where the sum of the marginal evaluations of all potential consumers equals the marginal cost of the good. The application of this “Samuelson Summation Condition” for public goods provision is, of course, exemplified in the road hypothetical where each resident is a potential beneficiary. Note, however, that excludability may exist in various degrees, so that private provision of “public” goods may be quite feasible. In some situations, considerable excludability can be enforced, even if not against all users. In other cases, a good or activity provides distinguishable types of benefits, some of which are excludable and some not. As long as some excludability is practical, private provision is a distinct possibility. The Samuelson Summation Condition can then be used as a benchmark to determine how close the expectable private level of provision comes to the “optimal” level. (Remember, however, that optimal is defined in a special technical sense.) An important point is that the “public goods” terminology is somewhat of a misnomer because public provision for such goods is not the only, nor even necessarily the most desirable, institutional arrangement for the provision of such goods.

**QUESTIONS**

1. “The doctrine of adverse possession boils down to this: when, because of expiration of the statute of limitations, a landowner loses the right of excludability, he also loses title.” Comment.

2. The legal right of excludability may be more apparent than real. Describe a situation in which an asset is protected by a damage or injunctive remedy for unauthorized use, but enforcement of this nominal excludability is impractical.

3. **CHICKEN VS. PRISONER’S DILEMMA**

   Much of what we have been discussing above deals with the circumstances in which individuals are likely, absent some external impetus, to fail to engage in behavior that would ordinarily be termed cooperative or mutually beneficial. The failure to exploit available mutual gains is, roughly speaking, what economists mean by “inefficient” results. Failure to cooperate usually implies an inefficient result, but it does not necessarily mean that nothing will be done at all. The distinction between Prisoner’s Dilemma and the closely related Chicken Game makes this apparent.

   According to its original set of facts, the road scenario yields a Chicken game rather than a Prisoner’s, Dilemma. As an exercise, you should attempt to confirm this by filling in the cells of the matrix in Exhibit 1.4. This matrix is a little more complicated than the original ones constructed above. Those were (2 x 2) or 4-cell matrices because there were fewer possibilities. Now, the description of the opportunity set has to allow for varying levels of cooperation. Note, however, the implicit assumption that doses of road maintenance can only be bought in discrete units. Relaxation of this assumption will be a desirable feature of other types of models that we will develop shortly.

   For each cell in Exhibit 1.4, consult the corresponding row and column headings in order to determine the total amount of road maintenance provided. That amount is the sum of the quantity provided by you (the row label) and the quantity provided by others (the column label). Then check Exhibit 1.3 to find out

---

the total value of the road benefits at that level of maintenance. Remember, though, that Exhibit 1.3 presents only gross benefits whereas the Exhibit 1.4 entries should be in net terms. Deduct your contribution (units you provided times their unit cost), if any, and enter the resulting net payoff in the appropriate cell of the matrix. There is little or no need to complete the entries for total provisions of more than 4 units of maintenance. (Why?)

In Exhibit 1.4, circle the cell that contains the best payoff in each column, i.e., the best strategy for each alternative assumption about what Others will provide. Note that this circled cell corresponds to different row strategies, depending on what the Others provision is. Do you think that there is a “dominant” strategy in this game? (Remember that dominance requires that there exist some best strategy regardless of what the other party is expected to do.) Is there a determinate or predictable outcome? In particular, how likely do you think it is that no road repairs will be undertaken, even if one is as pessimistic as possible about “cooperation”?

### Exhibit 1.4 Road Problem As A “Chicken” Game

<table>
<thead>
<tr>
<th>Doses Provided By Others</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>By You</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
</tr>
<tr>
<td>3</td>
<td>XXXXX</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>XXXXX</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Now change the facts by assuming that a dose of road maintenance costs $55.00 rather than $25.00. Complete Exhibit 1.5 by using the same procedure as for Exhibit 1.4. The resultant matrix should satisfy the requirements of a classic Prisoner’s Dilemma. Is cooperative behavior by one player likely in the absence of cooperative behavior by the other? Can you explain in “common-sense” terms why changing the cost alters the game from Chicken to Prisoner’s Dilemma?

### Exhibit 1.5 Road Problem As A Prisoner’s Dilemma

<table>
<thead>
<tr>
<th>Doses Provided By Others</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>By You</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
<td>XXXXX</td>
</tr>
<tr>
<td>3</td>
<td>XXXXX</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>XXXXX</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Although Chicken shares some of the same strategic characteristics as Prisoner’s Dilemma, it dramatically demonstrates the potential importance of bluffing and hold-out behavior when bargaining or trying to influence the behavior of others. By contrast, PD is not really a bluffing game, since the maximizing strategy is unaffected by what one’s opponent does or is expected to do. A simple form of Chicken bluffing is merely to adopt a very stubborn, uncooperative stance, even if it seems to involve “cutting off your nose to spite your face.” In the road example, for instance, Adams might threaten to provide no road maintenance, although Braun knows that Adams would find two units worthwhile even without any cost contribution from anyone else. But matters usually are much worse in the real world where, instead of the preferences of the parties being known, their respective valuations of the alternatives can only be guessed at or based on information interchange. Unfortunately, under these circumstances, there is a very real loss of ability of the bargainers to communicate with each other. The reason is that the parties have reason to offer false information to each other, to convince each other that there is less possibility for mutual gain.
than really exists. This is the essence of the bargaining bluff. Unfortunately, the parties then lose much of their ability to discriminate between “real” information and the kind of disinformation commonly produced as a self-interested bargaining ploy.

The production of misinformation during bargaining is easily understood if one contemplates the typical process of buying a used car. Assume that you have gone to the lot and see a car that fits your purposes and which, if necessary, you would be willing to acquire at a price as high as $1000. Unbeknownst to you, the used car dealer would be happy to sell at any price over $600. Since the car is worth $1000 and $600, respectively, to the potential buyer and seller, there is a range of “gains from trade” equal to $400. The bargaining is, in essence, about who gets what part of that $400 potential gain. During the bargaining process, each party can benefit by attempting to convince the other that his own “fullback position” or “best offer” is tougher than it really is. What kinds of things do the buyer and seller typically say to each other in an attempt to mislead about their willingness to sell. Is this deceptive process costly? And do you think that beneficial deals are sometimes lost because one party perceives the other party’s real fallback as merely a tough bargaining position?

In sum, our earlier models reflected the implicit assumption that each party had “full information,” i.e., substantially complete knowledge of the gains and losses attaching to the relevant alternatives, both for himself and also for the “opponent” in the game. That assumption is now relaxed since, unfortunately, many real world situations are characterized by so-called “limited information” about the implications of the alternatives. In such circumstances, the flow of information itself becomes part of the gaming process and is affected by strategic considerations.

4. GAMING ABOUT INFORMATION

Persons who have not engaged in tough bargaining are frequently naive about the possible benefits-sometimes even the outright necessity for survival-of what may euphemistically be called “dissimulation” during negotiating processes. The following story, based on a March, 1980 Wall Street Journal article, makes a sobering point.

TELLING LIES JUST A MATTER OF COURSE

UNTRUTHS CAN IMPROVE BUSINESS-SCHOOL GRADE; PEER PRESSURE CLASHES WITH ETHICS

BOSTON. A Harvard Business School student won the highest grade in part of his Competitive Decision Making course because “I was willing to lie to get a better score.”

That reaction is just fine with Prof. Howard R. Raiffa, whose course is aimed at teaching would-be business executives to negotiate in the real world. Like it or not, Prof. Raiffa says, lying or “strategic misrepresentation,” as he calls it, is sometimes resorted to in business negotiations.

Each week, Prof. Raiffa and his students play a game. He pairs them off and assigns them roles in a negotiation. One week a big-city mayor and a police-union leader bargain over a contract. Another week, a plaintiff and an insurance company try to reach a settlement. Next time, one executive tries to buy a company from another. The students negotiate outside the classroom during the week, then report the results.

The results determine the grade. The mayor who held the police union to the smallest wage-and-benefit package gets the best grade among the mayors; the police union chief who negotiated with him gets the worst grade among the union leaders.

Students find that hiding certain facts, bluffing or even outright lying often gains them a better deal. But the objective isn’t necessarily to teach them to lie. Rather, Prof. Raiffa says, it is to teach them they may be lied to. Raiffa thinks the students become “much more aware.” “They’re very naive when they start,” he claims.

One-third of the course grade is based on success in the negotiating games. For ambitious, aggressive students, the pressure to win is intense, and the course evokes strong reactions. During one class discussion of a game, a woman burst into
tears. She had discovered that the man she negotiated with, who repeatedly assured her he opposed any misrepresentation, had in fact lied blatantly. Another student, who has worked for Arabian American Oil Co. in Saudi Arabia negotiating construction contracts, said he found the students here less reasonable than the people he dealt with in the business world.

Raiffa doesn’t tell his students how to negotiate in any particular game. The students develop their own strategies and methods as they go along. Part of the course is theoretical; students learn how to analyze competitive situations. But to Raiffa and most of his students, the actual negotiations are the heart of the course.

“You learn a lot about negotiating and a lot about yourself,” says a lawyer who put in a year as counsel to a congressional subcommittee before taking the course.

Raiffa says he structures the negotiations so that in some of the early games, “the truth teller is at an extreme disadvantage” against someone who lies or bluffs. In later games, liars may lose a chance to reach a profitable settlement because their opponent is outraged and becomes more stubborn. “People have to learn to understand the nature of the game,” Raiffa says, “and understand how they are vulnerable.”

Many students are surprised at the amount of lying, claiming that while some people never misrepresent their position, others do it pathologically. Another student, who was a research assistant at Harvard before coming to the business school, says that attitudes toward lying changed during the semester. “There was a period when it seemed as if everyone was lying. It wasn’t bluffing; it was outright lying,” she says. “I did it too.”

This student says the experience taught her that peer pressure can override personal ethics. Since she doesn’t want to lie, she plans to avoid fields where she thinks dishonesty is commonplace.

Another member of the class, who had worked for the National Park Service, says she preserved her ethical standards, but only at the cost of losing in several negotiations with people who lied. Deeply disturbed, she went to Prof. Raiffa to discuss the course and her future in business. “I concluded there are businesses I’d better not go into,” she says. “I’m not willing to compromise my principles to the point of boldfaced lying.”

Some students say that lying in the course is acceptable while that lying outside isn’t. But most students feel that the way they play the games does reflect what techniques they’ll be using in their careers.

To what extent is misdirection during negotiation a mutually destructive process? Would all (or most?) parties be better off if only truthful information could be exchanged? Ask yourself whether any legal mechanisms seem to address these questions or, indeed, whether society attempts to provide extralegal “psychological” incentives not to dissimulate. Finally, bear in mind the possibility that dissimulation may not be an optimal long-term policy in an iterated or repetitious negotiating game, especially where parties have some control over whom one “plays” with.

QUESTIONS

1. Do most bilateral negotiations or trades broadly fit the paradigm of the Chicken game? Illustrate your answer with respect to the used car sale example in the text above.

2. What application of the models in this chapter is there to negotiations designed to reach out-of-court settlements?

3. Do threats of suit frequently fall into the mold of the Chicken paradigm? Can you give some (other) examples from the legal environment?

4. Is it desirable to be able to “rely” on what someone else tells you? What costs are incurred if it is impossible to rely on information proffered? Are there any legal mechanisms designed to ensure the reliability of information used in arriving at bargains?

5. How do bargainers invest resources in making their bargaining stances credible (even when they are fundamentally untrue)? Is this an “unprofitable” process in some overall sense?
6. What is the distinction between what the Harvard professor called “strategic misrepresentation” and what the law calls fraud?

5. “CHICKEN” IN ACTION: LAND ASSEMBLY

Misinformation is not the only obstacle to bargaining. One should not become overly optimistic merely because the “real” payoffs to the parties are known. When the parties possess the relevant facts about the options available to them, the situation is termed by economists as involving “full information.” One can debate whether the real world example supplied immediately below was literally one of full information, but it is certainly plausible that the parties are substantially informed about the respective benefits and costs and that dissimulation was not the major obstacle to the bargaining process. In order to further concretize the example, however, assume that the real estate developer could have made a gross profit of $11,000,000, before any deductions for land acquisition costs, if he could have assembled the entire tract as a casino site. Also, assume that each homeowner could move to a comparable house in another location and be just as well off with a compensation payment of anything over $30,000.

HOMEOWNER BICKERING ENDS $100,000 OFFER FOR HOUSES

[Adapted from newspaper accounts in 1979]

ATLANTIC CITY, N.J. An offer of $100,000 apiece for 72 homes in a Boardwalk neighborhood has fallen through because greedy residents ruined the deal, according to the real estate agent who tried to assemble the land.

The tract, which contains mostly small homes that sold for about $15,000 three years ago, is adjacent to Playboy’s casino hotel, and was desired as a site for a proposed casino.

Many homeowners on the Boardwalk block between Bellevue and Texas Avenues, reportedly are bitter over the attempt to assemble the building site.

“We went through a year of aggravation on this block. And for what? Nothing,” said one disgruntled Bellevue Avenue homeowner. “Most of these people didn’t come out with one red cent.”

“I’m glad the whole stir is done with,” said Sam Glaspert, a longtime Texas Avenue resident. “But next time, the people around here are going to be a lot more careful, money up front and the rest. If you didn’t learn a lesson from this deal, you never will.”

“The pressure was unbelievable. I had two neighbors come into my living room and pressure me. They wouldn’t leave until I agreed to sell.”

Thomas Blossom, the agent who last year offered $100,000 to each homeowner if all agreed to sell, says many owners “suddenly came back and argued their land was worth much more.”

He said only 20 of the 68 homeowners who agreed to sell out in March were willing to close the deal in June.

“Getting the whole tract in one shot was the only way we felt we could make the deal go,” Blossom said. “Now some of these holdouts are in for an unpleasant surprise. If that area can’t become a casino site immediately, their homes won’t be worth even $30,000.”

QUESTIONS

1. Suppose that the developer had managed to purchase 71 of the required 72 pieces of property for an average price of $50,000. If you owned the 72nd house, how much could you hope to get for it? (Remember the facts on above indicate that ownership of the entire tract would allow the developer to generate a profit of $11,000,000.)
2. In a similar situation, it is proposed to use a contingent contract scheme. An offer is made to each homeowner at a price of $550,000. Acceptance by the homeowner is irrevocable, but the contract is subject to the condition that all 72 accept. Is this technique likely to produce the required amount of agreement?

3. Suppose the 72 landowners could be compelled to accept offers of $50,000 apiece. Would they find a mechanism for such coercion to be desirable?

4. What is the amount of the loss if no agreement is reached? Who loses and by how much? Would you characterize this loss as “waste”?

5. Does the land assembly bargaining problem fit the paradigm of either the Prisoner’s Dilemma or the Chicken game?

6. WHAT GENERALIZATIONS CAN BE MADE?

   The last several sections should serve to suggest that the kinds of formal models represented by Exhibits 1.4 and 1.5 above can be extended either exactly or by analogy to a wide range of situations, many of which have a rather obvious relationship to the legal system. Two specific comments about the generalization of these models should be made. One is that the models can handle more complex behavioral choices than the binary Cooperate-Defect types of decisions originally considered. But it is easy to see that matrix models become very cumbersome as the number of row and column alternatives increase. Indeed, even in the road maintenance case, the notion of an arbitrarily-sized “dose” of maintenance is an extreme simplification. What if it were possible to purchase half-doses, quarter-doses, etc.? Hence, it will come as no surprise that other techniques of modeling will be urgently required when the behavior adjustments are potentially very complex, even continuously variable.

   Another conclusion is that Exhibits 1.4 and 1.5 can be applied to many-person situations as well as two-person games; the column headings are ambiguous as to whether the “other” provision comes from one or several persons. Therefore, the game concepts developed in the simple 2-person context have many-person analogues. The term of art for “large” games is “n-person game” where n stands for any number greater than two. In sum, the important insights gained from two-person, two-strategy games can be applied in more complicated real-world situations.

QUESTIONS

1. Is a coercive settlement better for people in a Prisoner’s Dilemma, as opposed to letting them work things out for themselves? (Sometimes, always, or never?) Is your answer the same for parties caught in a Chicken game?

2. “If there is no impediment to enforceable contracts, then the ‘dilemma’ problem is not generally a valid rationale for externally imposed coercion in the two-person cases; in n-person cases, coercion may well be prescribed.” True?

D. TRANSFERABILITY AS AN ELEMENT OF PROPERTY

   When a good or activity is characterized by perfect jointness in consumption, multiple beneficiaries do not interfere with each other. Consequently, more than one person can truly and completely own the entire property and its stream of benefits--if not in the legal sense, at least in a more fundamental underlying sense. When jointness does not apply, persons cannot fully share “ownership” in precisely the same way. In our legal system there are many forms of what are commonly thought of as shared ownership: tenants in common; life tenant and remainderman; lessor and lessee; common stockholder with voting rights and non-voting stockholder; owner of fee simple and owner of mineral rights; etc. In general, these are really divisions of ownership rather than sharing, since in the absence of perfect jointness one interest bars a simultaneous participation in precisely the same part of the benefits flow. Thus, for nonjoint goods, appropriability requires not only the minimal prerequisite of excludability, but also the potential assignment or transferal of part of the benefits to someone else.
In legalese, what one hopes to get in return for an assignment of benefits is a “consideration.” Trade takes place because the original owner of property finds an opportunity to transfer away all or part of his rights for a consideration that is deemed more valuable than the rights given up. In contrast, the giver of the consideration, the purchaser, must regard the rights as more valuable than the consideration paid for them. In sum, for trade to be possible, the trading partners must disagree about the relative values of the things being exchanged. Because of this disagreement, both parties can benefit by converting a lesser-valued thing into a higher-valued thing.

What does this all have to do with property? Well, a bit of reflection should indicate that “ownership” without the privilege of transferability is only a limited form of ownership. The owner can use the good himself, but loses the possibility to exact a charge for what is at least potentially the even higher use-value that someone else applies to the good. This “appropriability” (i.e., ability to “make one’s own”) the benefits of something to a non-owner user may take place through transfer of full and permanent rights, as through conferral of fee simple title. Alternatively, the exercise of appropriability may involve transfer merely of an entitlement to some part of the fruits of the property, limited either in manner or in time, as by a lease, entitlement to a “share,” etc. Legal rules occasionally inhibit or completely bar the transfer of certain kinds of property. The question is: Why?

Rather than deal with commonplace kinds of property, the following pages first examine the implications of excludability and transferability for some legal phenomena not commonly thought of as property. These phenomena involve actual or potential property rights in some aspect of legal process itself. Partially, this is done to drive home the fact that property is what the legal system makes it, by conferring or withholding excludability and transferability. Also, quite frankly, the notions of potential “property” about to be considered would be regarded by many traditional students of jurisprudence as somewhat bizarre. All the better! Some of the most insightful uses of formal analysis arise from being compelled to view familiar phenomena from an unaccustomed angle.

After dealing with somewhat unorthodox notions of property in legal process, we shall take up a line of more traditional “nuisance” cases.

By posing the questions in the next few sections, the intent is to engage in analytical exercises rather than to advocate policy changes. Inevitably, the factual or positive economics implications are likely to lead to normative judgments. Still, you should attempt to analyze the situations dispassionately without any visceral preconceptions. Whatever “bottom line” normative judgments that emerge in your own mind, be alert for objective implications that surprise or are counterintuitive. In particular, try to be sensitive to how, if at all, your normative views of the questions are influenced.

1. ASSIGNABILITY OF CAUSES OF ACTION

This section deals with legal causes of action as a species of intangible property. Preliminarily, we should note some general facts about the body of American law that governs the assignability of a cause of action by a plaintiff to a third party. Although courts were originally hostile to assignment of virtually all causes of actions, many types of claims are now transferable. Claims arising out of contract are, in general, fully assignable and are, indeed, actively exchanged in commerce. Many tort claims, e.g., those for damage to property, nuisance, etc., are also generally assignable. We focus here, however, on personal injury claims, one type of cause of action that is normally not assignable. Although personal injury claims are not directly transferable, they may be de facto transferable, at least in part, through certain assign-
ments of interests, perhaps most notably through subrogation by an insured to the insurance carrier.\^ Strangely, some jurisdictions have permitted assignment of the proceeds of personal injury claims although the assignment of the claim itself is barred.\^  

**THE CASE OF NIKE v. LOBEL**  

Victoria Nike is a graduate student in aerospace engineering and, until a fateful day last Fall, an avid competitive runner for her local track club. On the day in question, Victoria was running on one of the streets adjacent to her university’s campus. Suddenly her dreams of a medal-winning performance in the annual Lynchboro 10-Miler were rudely shattered by an “encounter of the worst kind” with an automobile driven by a local businessman, George Lobel. Lobel was cited by the investigating police officer for driving under the influence of alcohol. However, it has subsequently become apparent that, if Nike sues, Lobel’s attorney will vigorously raise a defense of contributory negligence. Nike was, he will contend, running on a heavily traveled street at dusk and failed to wear a reflective vest or any highly visible kind of clothing. Also, the damage calculations are likely to be the subject of fierce dispute since the extent and permanence of Nike’s injuries are not perfectly clear. Specifically, Nike’s ability to pursue her athletic career in the future is highly debatable.

Nike’s roommate, Carla Konomos, is a first-year student in the economics Ph.D. program. Konomos has a low tolerance for Nike’s frequent wailing about the painful uncertainties of her legal position. Indeed, Konomos has recently been heard to mutter under her breath: “Idiot! Why doesn’t she just peddle her claim for a little upfront money?! Just sell out and avoid all the months and months of grief over this.”

**QUESTIONS**

1. Suppose that Nike did want to sell her claim in exchange for some advance compensation. finder present law in the U.S., important restrictions exist as to her ability to transfer rights to a cause of action in exchange for financial or other consideration. Whom do you understand to be the intended beneficiary of these restrictions on transfer? Are there any general public policy purposes that the restrictions are designed to serve?

2. Absent any legal restraints, why might Nike want to sell her claim rather than litigate it herself? Enumerate as many conceptually distinguishable reasons as possible.

3. Suppose that Nike had a disability policy under which her insurance carrier will pay her immediately for various bodily injuries under a fixed-payment schedule. To what extent is it possible for an insurance company to recover from Lobel under Nike’s personal injury claim?

4. What problems, if any, do you see with legal restrictions against free assignment of tort claims to third parties? Are there any areas in which you would be prepared to expand on the present degree of assignability?

5. Does membership in the Bar constitute a good reason why a potential claim purchaser might not be permitted to engage in such transactions?\^p


\^ Although New York Statutory and common law both appeared to bar assignment of personal injury claims, in Richard v. National Transportation Co., 158 Misc. 324, 285 N.Y.S. 870 (1936) an injured party’s assignment to a hospital of the proceeds from such a claim was held valid. But see Southern Farm Bureau Casualty Insurance Co. v. Wright Oil Co., 248 Ark. 803, 454 S.W.2d 69 (1970) where the distinction between the cause of action and a contingent claim on proceeds is rejected.

\^p For instance, New York law bars an attorney from acquiring an interest in any suit where the primary, and not merely incidental, purpose of the assignment is to bring a suit. See N.Y. Judiciary Law Sec. 488 (McKinney 1968).
6. Now assume the same basic facts as in the Nike situation described above, except that all legal restrictions on the assignability of legal claims have been abolished. After careful consideration of her situation, Nike feels, that compared to litigating the claim herself, it would be preferable to sell the claim if she can get any price over $7200. Six lawyers have approached Nike and expressed an interest in buying her claim. Although Nike of course has no way of knowing it, the maximum bids that the lawyers are willing to pay are $6800, $7300, $7550, $7650, $7900, and $8200, respectively.

7. What would explain why the various parties place different values on Nike’s claim?

8. What prediction would you make about the price that Nike is likely to receive?

9. Suppose that only lawyers #2 and #5 above exist. They realize that, by concerted action, they probably could get Nike to sell for $7201. What is the concerted action? Use a Prisoner’s Dilemma analysis to assess the prospect of their being able to collude effectively against Nike. How would the presence of additional lawyers affect matters?

10. Back to all six lawyers again. In addition, assume that, by remarkable coincidence, there are four other cases exactly like Nike’s. Each lawyer is only capable of handling at most one case. How would the existence of the four other cases affect the price that Nike could expect?

11. “Laymen tend to have a fundamental misconception about the behavior of markets. They believe that buyers compete against sellers and sellers against buyers. Actually, it is the buyers who are their own worst opponents, and likewise for the sellers.” Comment on this statement.

SOME COMMENTS ON THE “ANSWERS”

WARNING: If you read this section without having first tried to work out for yourself the answers to the questions above, the author cannot take responsibility for resultant defects in your learning!

Here are a few thoughts on what you might have gotten out of the preceding questions. Not all of the items are dealt with, and you should work out the details of some of these general answers, if you have not already done so.

Most people initially feel that an underlying aim of the restrictions on transferability is protection of the unsophisticated victim from exploitation. One important lesson of the Nike hypothetical is the troubling implication that the restrictions on assignability tend actually to disadvantage plaintiffs. The favored parties appear to be (a) defendants, who effectively buy the cause of action in what we call “settlement;” and (b) lawyers and insurance companies to whom a modified interest in the cause of action can be given through contingency fees and subrogation, respectively. These parties are not obliged to compete against possibly higher bids for claims that might otherwise be forthcoming from persons to whom full or partial assignment is legally impermissible.

The gain to defendants is perfectly clear. However, it might be objected that, if the number of lawyers or insurance companies is large, then competition will tend to eliminate any net gains to the lawyers and insurance companies: the consideration given for the contingency fee or subrogation will tend to equal the value of the interest surrendered by the victim. For instance, market pressure will force the price of insurance policies to be lowered by the prospective value of the insurance company’s potential recovery under the subrogation privilege. Similarly, lawyers will have to give “value received” in services for the contingency fees they get. Thus, these “favored” parties would not really gain at the expense of the Nikes of the world. Is this argument correct? There is a subtle aspect to the answer that requires analytical results treated in Chapter II. By way of preview, the answer is: Yes and no; market pressure does guarantee some return to assignments of interests through the “back door” means of contingency fees and subrogations, but the value received in this way is likely to be less than it would be if the claims were capable of outright sale. Can you guess why at this point?

The small-scale competition among several would-be purchasers is designed to show something of the process whereby freer transferability might ensure the victim of receiving a fair value for the claim. Allowing a competitive market generates information about the claim’s value. Unlike the two-person
Chicken game analyzed above, bluffing and misinformation is not likely to work when the number of parties increases; one’s bluff is likely to be in effect “called” by a competitor. Thus, Nike does not have to be a sophisticated evaluator of her own claim, so long as the potential purchasers are themselves sophisticated and are compelled to compete against one another.

The answer to the part of the problem dealing with price prediction is that the potential purchaser who places the highest value on the claim will get the claim. This bidder must at least outbid the second-highest valuation. Where between that minimum outbid and the maximum valuation will the actual payment made to Nike fall? That ultimate result is somewhat indeterminate, a product of the bargaining process between Nike and the high-value lawyer.

The key to the remainder of the question is this: if four claims are now to be (bought) sold instead of one, the price must be low enough to induce the fourth most reluctant buyer to buy and, at the same time, high enough to induce the fourth most reluctant seller to sell. What range of prices meets these criteria? The price predictions just suggested involved an area of indeterminacy, a bargaining range.

What do you think happens to the range of indeterminacy as the number of buyers and sellers increases? If you can work out the answers to these questions, you are well on your way to understanding the process of price formation in a competitive market.

An objection frequently raised against transferability is that it will result in cases being litigated that otherwise would have been settled. Now, it is not so clear that this would be such a bad thing even if true. Bias against plaintiffs may now cause cases not to be litigated that should be litigated. More surprisingly, it is not even so clear that the incidence of settlement would not be increased rather than decreased.

Cases tend to be settled when the opposing parties have similar estimates of the expected value of the recovery. For instance, suppose that plaintiff and defendant have identical estimates of the outcome of a case if it is litigated to final judgment: they both estimate that the prospect of judgment to the plaintiff is worth about $55,000. [In the next chapter, we will get a little bit more sophisticated about how prospects are evaluated when they are to any extent uncertain.] Suppose further that litigation costs are $4000 for plaintiff and $3000 for defendant. In total, the parties can save $7000 in costs by not going to trial. Hence, they will both be better off at any settlement figure between $51,000 and $58,000.

The above example can be generalized. Estimates of the outcome need not be identical in order for mutual gains from settlement to exist; the divergence in views must merely be less than the cost to both parties of proceeding to final judgment. Do you see why this is true?

Now for the relevance of assignability to all of this. A plausible argument might be that assignability would produce a more objective valuation of claims by people who are specialists in such evaluations. After all, if the present holder of a claim placed an unduly low value on it, there would be money to be made from detecting this fact and purchasing the claim for a figure closer to its real value. Indeed, a market-like valuation of one’s claim might induce a downward revaluation by one who overvalues his cause of action. After all, if I see that nobody else values my asset as highly as I do, it may produce some reassessment in my mind as to the accuracy of my original subjective valuation. In sum, transferability might well produce more uniformity in the valuation of claims. If so, an increase rather than a decrease in pre-trial settlement is a defensible prediction. Do you find this a persuasive argument? Why?

**2. OTHER KINDS OF PROPERTY RIGHTS IN LEGAL PROCESS**

Here are several shorter questions, all on the general topic of property rights in legal process, but drawing on elements from almost all of the preceding sections of the chapter.

**d. Mutuality of Estoppel**

The doctrine of mutuality of estoppel, or issue preclusion, traditionally maintained that one may not invoke a judgment in his own favor unless an adverse judgment could have been used against him. This rule has increasingly been subject to attack as possibly requiring wasteful use of resources in the
repetitive relitigation of identical issues. To make matters even worse, this relitigation may also result in inconsistent resolutions of the issues involved.

**QUESTIONS**

1. Party A has been adjudged guilty of negligence in an accident that caused injuries to B. The same act that caused B’s injuries also led to damages sustained by C. What legal considerations determine whether an adverse party can invoke B’s judgment against A?

2. Can mutuality of estoppel be defended as a response to a possible Free Rider problem? Or a Chicken problem? Or a Prisoner’s Dilemma? Devices such as class actions and compulsory joinder are sometimes regarded as mechanisms to overcome Free Rider problems among litigants. How effective do you think they are? For an informative discussion of the legal devices to deal with multiparty actions, see McCoid, “A Single Package for Multiparty Disputes,” 28 Stanford L. Rev. 707 (1976).

3. The rules restricting estoppel confer a species of excludability on a judgment, but they do not create full appropriability because a determination is not saleable. What would be wrong with making a determination transferable for consideration? For instance, multiple plaintiffs in an airline disaster could either “produce” proof of liability against the airline themselves or they could “purchase” the favorable judgment secured by an early litigant.

4. “One way of explaining the present situation is to say that the law treats ‘precedents’ differently, depending on whether they are precedents as to law or as to facts. We usually don’t call the latter ‘precedents,’ but they are.” What does the quoted statement mean? Is there any good reason why judgments as to facts ought to be treated differently from judgments as to law?

**e. Discovery of Computerized Litigation Files**

Presumably, the ideal behind the discovery process is to facilitate the uncovering of truth by making available as many facts and “investigative resources” as possible. A very real impediment to understanding certain complex cases is the overwhelming mass of documents and facts unearthed during discovery. At some point, the truth-seeking process almost grinds to a halt under the sheer weight of the accumulated evidence. Computerized litigation support systems are emerging as an important tool in coping with complex cases. Documents are coded, classified, summarized, and sometimes entered in “full text” form into computer-accessible files. The ability to quickly and accurately retrieve and process information is thought to be very useful by litigators. However, putting documents into computer-accessible form can be a very expensive process, often running into the hundreds of thousands of dollars.

Assume that there is a computer-readable file containing the full text of all documents discovered in a pending suit. The file has been prepared by litigant A at a cost of $650,000. The file contains nothing of what would ordinarily be called attorney work product; it consists only of the full text of the original documents, now reduced to computer-readable form. Litigant R seeks discovery of the computerized file.

a. What are the pros and cons of permitting discovery by party B?

b. The cost to A of duplicating the tapes for B is approximately $1000 dollars. If B is granted access to the file, what should B have to pay to A?

c. What could A have done to improve the possibility that a computerized document file would be held non-discoverable? Is it fair to consider the extra cost of this as an investment in “staking a claim” to property?

**BROADCAST MUSIC, INC. v. MOOR-LAW, INC.**

527 F. Supp. 758 (D. Delaware, 1981)

[Almost all licensing of U.S. performing rights to musical compositions is done by two large organizations, Broadcast Music, Inc. (BMI) and the American Society of Composers, Authors & Publishers (AS-}
CAP), each of whom represent thousands of individual copyright owners. Licensees fall into two categories: (1) broadcast licensees, who include television and radio stations or networks; and (2) non-broadcast licensees, who include hotels, concert halls, and “GLAs”. GLAs are small establishments, such as nightclubs and bars, that provide live music under a blanket General Licensing Agreement. In 1377, BMI filed a copyright infringement action against Moor-Law, Inc., doing business as the Triple Nickel Saloon. Moor-Law raised the affirmative defense of copyright misuse and counterclaimed that BMI’s licensing practices to GLAs violated Secs. 1 and 2 of the Sherman Act and Sec. 3 of the Clayton Act.

The relevant market in this case is the licensing of musical performing rights to GLA licensees. An examination of the characteristics of this unique market is essential to an evaluation of the Triple Nickel’s antitrust and copyright misuse claims. Although the parties dispute the significance of some market characteristics, they are in basic agreement about many features of the market, including the applicability of the economic concepts of natural monopoly and public goods.

Both parties’ experts agreed that this market has natural monopoly characteristics. Because there are thousands of individual copyright “sellers” seeking to deal with thousands of GLA buyers, the potential transaction costs are very high. Economies of scale exist as sellers band together to spread transaction costs of identical transactions over a larger group. Thus, some pooling of copyrights by individual copyright holders is a necessity in order to take advantage of the natural monopoly characteristics of the market.

In addition, both parties’ experts agree that the goods in this market—the performing rights to the musical compositions—have the characteristics of “public goods”. Public goods have two salient characteristics which operate in this market. First, unlike private goods (e.g. apples), one can use a public good without leaving any less for others to consume. Once a musical composition is created, the marginal cost of additional consumption is zero. The second characteristic of public goods is that it is difficult to exclude persons who do not pay from using the good. The owners of private goods can withhold their goods from the market and release them only in return for payment; but, once a composer’s song becomes known, he or she finds it difficult to prevent that good from being “stolen” by users. The enforcement problem resulting from this public good characteristic manifests itself in the GLA market through users who don’t pay any licensing fee. During the course of this litigation, this has been labeled the “free rider” problem.

Because the high transaction costs derived from natural monopoly characteristics are increased by the public good enforcement problem, very large performing rights organizations, like BMI and ASCAP, in which individual copyright holders pool their rights are necessary to achieve efficiency. The larger the organization, the more efficient it will be in reducing transaction costs; indeed, Triple Nickel’s expert, Dr. Cirace, advocated one combined licensing operation in this market. The necessity for these large licensing organizations makes competition in the sense of many sellers competing against each other in the GLA market unrealistic.

The parties are also in agreement that the nature of the GLA market makes some kind of blanket license a necessity. As the Supreme Court observed in CBS IV:

Individual sales transactions in this industry are quite expensive, as would be individual monitoring and enforcement, especially in light of the resources of single composers. Indeed, as both the Court of Appeals and CBS recognize the costs are prohibitive for licenses with individual radio stations, nightclubs, and restaurants, * * * and it was in that milieu that the blanket license arose.

CBS IV, 441 U.S. at 20.

Moreover, the Supreme Court’s recognition in the CBS case, Id. at 20, that most users want “unplanned, rapid and indemnified access” to a wide range of compositions, is particularly apt in the GLA market. Testimony at trial made clear that GLA users typically do not know in advance what compositions will be performed nightly in their establishments and yet want the right to perform them instantaneously. The blanket license provides instantaneous access to any composition desired.
A corollary of the conclusion that the blanket license is a necessity in the GLA market is that the alternatives required by BMI’s consent decree of direct licensing with individual copyright owners or of per piece licensing are unfeasible in this market. Again, the parties seem to agree on this point. Unlike the situation in CBS where large networks were interested in a relatively small number of compositions known in advance of performance, GLAs like the Triple Nickel are small establishments which lack the resources or the advance notice to contract copyright owners individually on a large scale. Likewise, because GLA owners rarely know in advance of performance the songs a band intends to play and because GLA bands often take audience requests, a prospective per piece license is unrealistic.

Finally, although the parties disagree over appropriate methods of pricing, they seem to agree that the natural market forces of supply and demand do not operate normally on pricing in this market. Because of the public good characteristic that the marginal cost of using a musical composition is zero, normal cost-based pricing is not feasible. The parties seem to agree that some form of pricing based on benefit conferred is appropriate. But, since as a practical matter a GLA needs a license from both ASCAP and BMI, the normal constraint on benefit pricing—alternative supply—does not operate in this market.

Thus, although large performing rights organizations are a necessity in this market, the result in the current market is that BMI can exercise substantial monopoly power over price. This monopoly power of the seller is particularly strong in a negotiating situation where there is not corresponding power on the buyer’s side. Unlike the television network market where buyers like CBS exercise some monopsony power of their own, the buyers in the GLA market are weak and diffuse.

While normal competitive forces do not operate in this market, it is not true that BMI’s price for its GLA license is unconstrained. Testimony at trial convinced me that the free rider problem does provide a significant constraint on the price BMI charges. The higher the price it charges, the greater the resistance of GLA users is likely to be, and, conversely, the lower the price, the lower the resistance will be. Since the free rider problem tends to make BMI’s enforcement costs high and can, indeed, cause increased costs to more than consume increased revenue from a higher price, BMI considers this problem when setting a price.

QUESTIONS

1. In the case excerpted above, the court is sympathetic, based on a free-rider argument, to a pricing arrangement that might otherwise be deemed a violation of the antitrust laws. Summarize in your own words how a free-rider situation is alleged to result and why the court might view this as bad. Are you persuaded by the argument? Can you suggest other industries that might plausibly make similar claims?

2. A free-rider argument may be made to explain certain restrictions on price competition imposed by manufacturers on their dealers. Briefly, this thesis suggests that manufacturers may want to induce dealers to substitute non-price competition—in the form of service, advertisement, parts availability, etc.—for price competition. In principle, the restrictions may be beneficial to all parties, i.e., manufacturer, dealers, and customers. Indicate how you think this argument might be developed and what might be the countervailing anti-trust risks.8

3. COASE AND COSTS, “OPPORTUNITY” AND OTHERWISE

The facts of the nuisance cases in this part of the chapter can be used to comment briefly on a disarmingly simple but extremely powerful proposition frequently applied in economic analysis of Law: the so-called “Coase Theorem.” The Coase Theorem had its genesis in a classic article by Ronald Coase, “The Problem of Social Cost,” 3 J. Law & Econ. 1 (1960). In this article, which was an important determinant of Coase’s receipt of a Nobel Prize in Economics, the author discussed a series of specific fact situations

that link common law tort and nuisance doctrines with the economic theory of externalities. Since Coase himself never indicated explicitly what generalizations were to be drawn from his specific examples, the so-called “Theorem” bearing his name has been the subject of slightly different formulations, but its general flavor can be gleaned from the situation faced by the Fontainebleau and Eden Roc Hotels.

### a. Allocation of Rights in the Small-Number Context

**FONTAINEBLEAU HOTEL CORP. v. FORTY-FIVE TWENTY-FIVE, INC.**

114 So.2d 357 (Fla. 1959)

* * * In this action, plaintiff-appellee sought to enjoin the defendants-appellants from proceeding with the construction of an addition to the Fontainebleau, alleging that the construction would interfere with the light and air on the beach in front of the Eden Roc and cast a shadow of such size as to render the beach wholly unfitted for the use and enjoyment of its guests, to the irreparable injury of the plaintiff; further, that the construction of such addition on the north side of defendants’ property, rather than the south side, was actuated by malice and ill will on the part of the defendants’ president toward the plaintiff’s president; and that the construction was in violation of a building ordinance requiring a 100-foot setback from the ocean. It was also alleged that the construction would interfere with the easements of light and air enjoyed by plaintiff and its predecessors in title for more than twenty years and “impliedly granted by virtue of the acts of the plaintiff’s predecessors in title, as well as under the common law and the express recognition of such rights by virtue of Chapter 9837, Laws of Florida 1923. * * * “ Some attempt was also made to allege an easement by implication in favor of the plaintiff’s property, as the dominant, and against the defendants’ property, as the servient, tenement.

* * *

The chancellor * * * entered a temporary injunction restraining the defendants from continuing with the construction of the addition. His reason for so doing was stated by him as follows:

* * * The ruling is not based on alleged presumptive title nor prescriptive right of the plaintiff to light and air nor is it based on any deed restrictions nor recorded plats in the title of the plaintiff nor of the defendant nor of any plat of record. It is not based on any zoning ordinance nor any provision of the building code of the City of Miami Beach nor on the decision of any court, nisi pruis or appellate. It is based solely on the proposition that no one has a right to use his property to the injury of another and that the intended use by the Fontainebleau will materially damage the Eden Roc. There is evidence indicating that the construction of the proposed annex by the Fontainebleau is malicious or deliberate for the purpose of injuring the Eden Roc, but it is scarcely sufficient, standing alone, to afford a basis for equitable relief.

This is indeed a novel application of the maxim *sic utere tuo ut alienum non laedas*. This maxim does not mean that one must never use his own property in such a way as to do any injury to his neighbor. It means only that one must use his property so as not to injure the lawful rights of another. *Cason v. Florida Power Co.*, 74 Fla. 1, 76 So. 535. In *Reaver v. Martin Theatres* (Fla.1951) 52 So.2d 682, 683, under this maxim, it was stated that “it is well settled that a property owner may put his own property to any reasonable and lawful use, so long as he does not thereby deprive the adjoining landowner of any right of enjoyment of his property which is recognized arid protected by law, and so long as his use is not such a ones as the law refill pronounce a nuisance.” [Emphasis supplied.]

No American decision has been cited, and independent research has revealed none, in which it has been held that--in the absence of some contractual or statutory obligation--a landowner has a legal right to the free flow of light and air across the adjoining land of his neighbor. Even at common law, the landowner had no legal right, in the absence of an easement or uninterrupted use and enjoyment for a period of 20 years, to unobstructed light and air from the adjoining land. *Blumberg v. Weiss*, (N.J.1941), 17 A.2d 823.
* * * There being, then, no legal right to the free flow of light and air from the adjoining land, it is universally held that where a structure serves a useful and beneficial purpose, it does not give rise to a cause of action, either for damages or for an injunction under the maxim sic utere tuo ut alienum non laedas, even though it causes injury to another by cutting off the light and air and interfering with the view that would otherwise be available over adjoining land in its natural state, regardless of the fact that the structure may have been erected partly for spite. See the cases collected in the annotation in 133 A.L.R. at pp. 701 et seq.; 1 Am.Jur., Adjoining Landowners, Sec. 54, p. 536. * * *

We see no reason for departing from this universal rule. If, as contended on behalf of plaintiff, public policy demands that a landowner in the Miami Beach area refrain from constructing buildings on his premises that will cast a shadow on the adjoining premises, an amendment of its comprehensive planning and zoning ordinance, applicable to the public as a whole, is the means by which such purpose should be achieved.

* * *

Since it affirmatively appears that the plaintiff has not established a cause of action against the defendants by reason of the structure here in question, the order granting a temporary injunction should be and it is hereby reversal with directions to dismiss the complaint.

QUESTIONS

1. If the Florida courts had ruled in favor of the Eden Roc owners, what do you think would subsequently have happened?

2. To whom do you think what are, in effect, the “solar property rights” were most valuable before the Fontainebleau commenced construction of the tower? If the law had at that time been known to be clearly against the highest-valued user, what option would have been available? Is the option a practical one?

3. “Solar rights constitute an area where the legal system need only define and enforce a clear system of rights through injunctive remedies. Then these property rights, like any other, can be efficiently reallocated via transfers, easements, etc.” Comment.

Are the problems of allocating wind rights essentially the same as those involved with solar rights? What about aesthetic rights?

b. Applicability of the Coase Theorem

In determining whether blockage of the sunlight is actionable or not, what the court really does is to determine who “owns” the relevant aspect of the local environment. If the Eden Roc is denied any remedy against the blockage of the sunlight, then the Fontainebleau has effectively been declared the owner of a valuable right, the use of the sunlight that would be impeded by Fontainebleau’s new building. If, conversely, the Fontainebleau is restrained by the law through the issuance of an injunction or damage remedies, then the Eden Roc effectively owns the property right to that sunlight. Coase’s insight was to point out that the story does not necessarily stop after this initial determination of ownership. Let us consider why.

Assume that the value of the right to put the new building in a position that blocks the Eden Roc’s sunlight is worth $X to the Fontainebleau while the right to have the light unobstructed is worth $Y to the Eden Roc. If the law assigns the right to the Fontainebleau, how much does it cost the Fontainebleau to use that right? It is tempting to say that it costs nothing, but that answer is clearly wrong. While no expenditure is required in order for the Fontainebleau to use a right that it already owns, that use necessitates what legal scholars have traditionally called a “forebearance.” In our present example, the Fontainebleau must turn down the offer of up to $Y that the Eden Roc would be willing to pay to acquire the right. Following a similar reasoning, modern economists correctly insist that the cost of something is what one gives up in order to have it, and there is a very relevant sense in which it makes no difference whether the
lost opportunity derives from an outlay or a foresworn receipt. This doctrine is known as the “opportunity cost” concept, where the relevant opportunity is the highest-valued alternative that is lost. Determining opportunity cost requires that one ask: What am I giving up in order to do this? Sometimes the answer is in the form of a thing (e.g., money) that is given up in the form of a conventional “outlay” cost. Other times, however, the opportunity cost takes the form of an option foresworn, i.e., a forebearance. Further, it is important to remember that opportunity cost is always the highest-valued thing that is given up. Economists are very insistent—and rightly so—that “cost” must always be reckoned as opportunity cost, not merely conventional outlays. We shall see, in several different contexts, both here and in much later topics, why being a consistent opportunity-coster is important.

This opportunity cost concept leads directly to the main thrust of what has come to be known as the “Coase Theorem.” It suggests that, in order for any right to be “used” by a party, that party must place a higher value on the right than any other bidder for the same right. If this were not so, the owner would be exercising a right that is worth (to him) less than its opportunity cost. Hence, who ultimately winds up exercising the right will tend to be determined only by the relative magnitudes of the numbers that correspond to X and Y, i.e., the values of the alternative uses. Note that the same logic applies if we now shift our assumption and suppose that the court had been willing to enjoin the Fontainebleau, thus making the Eden Roc the initial owner of the legal entitlement.

In brief, the thesis is that in an environment where there are no obstacles to transacting, legal rights will tend ultimately to be allocated, through trade if necessary, to the party that values them most highly, regardless of their initial assignment. The emphasized proviso in the previous sentence is critically important. We already have reason to believe, for instance, that the proviso is not met if the transferability or excludability of rights is somehow defective, and other such situations will appear below. Subject to its very emphatic qualifier, however, the Coase Theorem still points out an important truth: to the extent that transactions can reasonably be counted upon to reallocate rights, if necessary, to the “highest valued” use, then courts need not worry about the “waste” that might occur if low-valued uses of rights impeded higher-valued ones.

In sum, the Coase Theorem may be interpreted as an identification of circumstances in which the assignment of legal rights “makes no difference” to the achievement of an efficient use of resources because bargaining will reallocate and correct any initial misassignment. But the theorem can easily be misinterpreted as saying additionally that there is no difference in a much wider sense. Plainly that is not true. The initial allocation of rights ownership does make a difference because it affects the distribution of income, thus effectively making some people richer and some poorer. The Coase Theorem states only that no significant economic waste consequences are thereby implicated. Properly understood, it does not deny that the court’s decision will make one party or the other better or worse off by the amounts that the relevant rights are “worth,” but merely asserts that an initial ownership decision does not by itself mandate the final result as to how rights are used and lay whom. The latter questions are still the focus of a cost-benefit calculus ultimately determined by the parties themselves.

The theorem has sometimes also been interpreted as claiming that the same allocation of resources will result regardless of the initial assignment. This claim is plainly erroneous, although perhaps precipitated by the misleadingly simple examples commonly used to describe the theorem. For instance, if X is greater than Y, then it is asserted that the Fontainebleau will also emerge as the end-user, whichever way the court initially decided. This thesis, however, fails to consider the possible effect of the rights assignment on the values that the parties attach to the rights after an initial allocation has been determined. In general, a so-called “wealth effect” causes people to attach higher or lower values to things--i.e., change the numbers corresponding to X and Y--when their wealth level varied. Thus, through its original assignment of the wealth represented by the transfer value of the right, a legal system may create an “income effect” on the magnitudes of X and Y, the values placed by different parties on some particular right people. Reversal of the relative magnitudes of X and Y perhaps requires what is usually an implausibly large wealth effect, but the point should be made that, where the valuations of the parties are close, even comparatively small wealth effects caused by legal assignments can tip the balance as to who will place the
highest value on the right when post-assignment bargaining becomes possible. Under those conditions, altering the original assignment of rights does result in different final resource allocations (because the wealth effect actually determines who is the highest-value user of the right) but those different final results are equally consistent with the Coase Theorem’s key prediction that, regardless of the legal rule, final allocations will be highest-valued uses.

Of course, the optimistic prediction of the Coase Theorem will frequently be inapplicable because there are substantial impediments to the reallocation of rights through bargaining. In general, for instance, the Coase Theorem has greater applicability in Contracts, where the parties are already involved in a rights-and-duties trading process, than in Torts, where the parties may even be unknown to each other before the precipitating event. Hence, great caution must be exercised to invoke the Coase Theorem only under appropriate fact circumstances.

Although it presented an important common-sense insight, the Coase Theorem can facilitate a certain brand of myopia in legal analysis if one over-concentrates on the environment of extremely low bargaining costs. Whatever the narrowness of vision of some of those who invoke his Theorem, Coase himself had an admirable perspective:

I would not wish to conclude without observing that, while consideration of what would happen in a world of zero transactions costs can give us valuable insights, these insights are, in my view, without value except as steps on the way to analysis of the real world of positive transactions costs. We do not do well to devote ourselves to a detailed study of the world of zero transactions costs, like augurs divining the future by the minute inspection of the entrails of a goose.

It has just been suggested that the famous Coase Theorem is really a straightforward corollary of the opportunity cost concept. Discussion of a few additional cost concepts will also be useful for purposes of possible application to legal reallocation. Legal process costs, including filing, preparation, litigation, etc., will be referred to as “process costs.” Eventually, it will be useful for certain purposes to further subdivide transaction and process costs into other components such as “monitoring” and “precautionary” costs.

Finally, the notion of “error costs” should also be introduced. In brief, error costs are the foregone gains due to not achieving the “ideal” result, i.e., the costs of “getting it wrong.” For example, in the road maintenance scenario above, it would be the increase in net benefits, if any, derivable by producing the optimal level of road repairs rather than the existing one. Paradoxically enough, error costs are a benchmark concept and do not necessarily represent “true” costs in the opportunity cost sense. For instance, it may not be technologically or practically possible to achieve the ideal results, in which case the “opportunity” to have perfection is not a real alternative. More importantly, however, the concept of error costs is used below to focus attention on the fact that it is seldom desirable in the real world to attempt to achieve “perfect” or “ideal” outcomes. In general, any reallocation yielding lower error costs can only be accomplished at the expense of additional transaction or process costs. It is the sum of these various costs that presumably should be minimized, usually by trading off among categories. Frequently, therefore, it is sensible to accept substantial error costs because the transactions or processes available to bring about superior reallocations are not cost-effective. Zero error costs are the benchmark result that would hold only if there were an omniscient, omnipotent, and benevolent entity prepared to effectuate and enforce “optimal” results.

---

Chapter 1  ANALYZING HUMAN CHOICE

c. Cases Involving Transactions-Cost Problems

BOOMER v. ATLANTIC CEMENT CO.

257 N.E.2d 870 (Court of Appeals of New York, 1970)

BERGAN, Judge. Defendant operates a large cement plant near Albany. These are actions for injunction and damages by neighboring land owners alleging injury to property from dirt, smoke and vibration emanating from the plant. A nuisance has been found after trial, temporary damages have been allowed; but an injunction has been denied.

The cement making operations of defendant have been found by the court at Special Term to have damaged the nearby properties of plaintiffs. That court, as it has been noted, accordingly found defendant maintained a nuisance and this has been affirmed at the Appellate Division. The total damage to plaintiffs’ properties is, however, relatively small in comparison with the value of defendant’s operation and with the consequences of the injunction which plaintiffs seek.

The ground for the denial of injunction, notwithstanding the finding both that there is a nuisance and that plaintiffs have been damaged substantially, is the large disparity in economic consequences of the nuisance and of the injunction. This theory cannot, however, be sustained without overruling a doctrine which has been consistently reaffirmed in several leading cases in this court and which has never been disavowed here, namely, that where a nuisance has been found and where there has been any substantial damage shown by the party complaining an injunction will be granted.

The rule in New York has been that such a nuisance will be enjoined although marked disparity be shown in economic consequence between the effect of the injunction and the effect of the nuisance.

The problem of disparity in economic consequence was sharply in focus in Whalen v. Union Bag & Paper Co., 208 N.Y. 1, 101 N.E. 805. A pulp mill entailing an investment of more than a million dollars polluted a stream in which plaintiff, who owned a farm, was “a lower riparian owner”. The economic loss to plaintiff from this pollution was small. This court, reversing the Appellate Division, reinstated the injunction granted by the Special Term against the argument of the mill owner that in view of “the slight advantage to plaintiff and the great loss that will be inflicted on defendant” an injunction should not be granted (p. 2, 101 N.E. p. 805). “Such a balancing of injuries cannot be justified by the circumstances of this case”, Judge Werner noted (p. 4, 101 N.E. p. 806). * * * The rule laid down in that case, then, is that whenever the damage resulting from a nuisance is found not “unsubstantial”, viz., $100 a year, injunction would follow. This states a rule that had been followed in this court with marked consistency.

***

This result at Special Term and at the Appellate Division is [thus] a departure from a rule that has become settled; but to follow the rule literally in these cases would be to close down the plant at once. This court is fully agreed to avoid that immediately drastic remedy; the difference in view is how best to avoid it.

One alternative is to grant the injunction but postpone its effect to a specified future date to give opportunity for technical advances to permit defendant to eliminate the nuisance; another is to grant the injunction conditioned on the payment of permanent damages to plaintiffs which would compensate them for the total economic loss to their property present and future caused by defendant’s operations. For reasons which will be developed, the court chooses the latter alternative.

If the injunction were to be granted unless within a short period, e.g., 18 months, the nuisance be abated by improved methods, there would be no assurance that any significant technical improvement would occur.

The parties could settle this private litigation at any time if defendant paid enough money and the imminent threat of closing the plant would build up the pressure on defendant. If there were no improved techniques found, there would inevitably be applications to the court at Special Term for extensions of time to perform on showing of good faith efforts to find such techniques. * * * Moreover, techniques to eliminate
dust and other annoying by-products of cement making are unlikely to be developed by any research the defendant can undertake within any short period.

On the other hand, to grant the injunction unless defendant pays plaintiffs such permanent damages as may be fixed by the court seems to do justice between the contending parties. All of the attributions of economic loss to the properties on which plaintiffs’ complaints are based will have been redressed.

The nuisance complained of by these plaintiffs may have other public or private consequences, but these particular parties are the only ones who have sought remedies and the judgment proposed will fully redress them. The limitation of relief granted is a limitation only within the four corners of these actions and does not foreclose public health or other public agencies from seeking proper relief in a proper court.

It seems reasonable to think that the risk of being required to pay permanent damages to injured property owners by cement plant owners would itself be a reasonably effective spur to research for improved techniques to minimize nuisance.

***

The orders should be reversed, without costs, and the cases remitted to Supreme Court, Albany County to grant an injunction which shall be vacated upon payment by defendant of such amounts of permanent damage to the respective plaintiffs as shall for this purpose be determined by the court.

JASEN, Judge (dissenting).

***

I see grave dangers in overruling our long-established rule of granting an injunction where a nuisance results in substantial continuing damage. In permitting the injunction to become inoperative upon the payment of permanent damages, the majority is, in effect, licensing a continuing wrong. It is the same as saying to the cement company, you may continue to do harm to your neighbors so long as you pay a fee for it. Furthermore, once such permanent damages are assessed and paid, the incentive to alleviate the wrong would be eliminated, thereby continuing air pollution of an area without abatement.

It is true that some courts have sanctioned the remedy here proposed by the majority in a number of cases, but none of the authorities relied upon by the majority are analogous to the situation before us. In those cases, the courts, in denying an injunction and awarding money damages, grounded their decision on a showing that the use to which the property was intended to be put was primarily for the public benefit. Here, on the other hand, it is clearly established that the cement company is creating a continuing air pollution nuisance primarily for its own private interest with no public benefit.

This kind of inverse condemnation (Ferguson v. Village of Hamburg, 272 N.Y. 234, 5 N.E.2d 801) may not be invoked by a private person or corporation for private gain or advantage. Inverse condemnation should only be permitted when the public is primarily served in the taking or impairment of property. The promotion of the interests of the polluting cement company has, in my opinion, no public use or benefit.

Nor is it constitutionally permissible to impose servitude on land, without consent of the owner, by payment of permanent damages where the continuing impairment of the land is for a private use. (See Fifth Ave. Coach Lines v. City of New York, 183 N.E.2d 684, 686; Walker v. City of Hutchinson, 352 U.S. 112.) This is made clear by the State Constitution (art. I, Section 7) which provides that “[p]rivate property shall not be taken for public use without just compensation” (emphasis added). It is, of course, significant that the section makes no mention of taking for a private use.

In sum, then, by constitutional mandate as well as by judicial pronouncement, the permanent impairment of private property for private purposes is not authorized in the absence of clearly demonstrated public benefit and use.

I would enjoin the defendant cement company from continuing the discharge of dust particles upon its neighbors’ properties unless, within 18 months, the cement company abated this nuisance.

***
Consider the following facts for the hypothetical case of Bomber et al. v. Particular Gypsum Co. Sections 38 and 39 are adjacent 100-acre parcels of land in the northwestern part of Clementine County. Section 38 is owned by the Particular Gypsum Co. and is the site of its main production plant. Section 39 has been subdivided into 100 one-acre lots, almost all of which contain owner-occupied residences. Plaintiff Bomber is one of a class of individual property owners in Section 39 and is seeking to enjoin Particular from emitting dust and other particulate matter whose airborne passage onto the property in Section 39 allegedly constitutes a nuisance.

Expert testimony establishes that the losses to Section 39 property values resulting from Particular’s emissions are approximately $1,000 per acre. In turn, Particular offers unrebutted evidence that (1) it is not feasible to produce at all in the present location without causing an objectionable level of emissions and (2) relocation of the plant would result in a cost to it of $12,437,000.

QUESTIONS

1. Counsel for Particular suggests that the precedent of Boomer v. Atlantic Cement Co. be applied. Do you agree that the Boomer case should be interpreted as teaching the following rule: an injunction should not issue when it can be shown that the value of a nuisance greatly exceeds the damages that it is imposing? If not, what?

2. Would it make any difference to your treatment of the case if Section 39 were an apartment complex owned in its entirety by Bomber Development Corporation?

3. Suppose that the Particular plant long antedated the residential development. Particular’s counsel urges that no damages are owed, that the homeowners have already been compensated: when they bought the land from the original developer, each parcel sold for $1000 less than it otherwise would have because the purchasers knew about the dust emissions and took this into consideration in establishing the market price of the property. What is your reaction to this argument?

4. Assume that a strong tradition of no damages for a party who “comes to the nuisance” exists in your jurisdiction. Also, time has been turned back so that neither the factory nor the housing development presently exist; the developmental conditions of the area are not yet really ripe for such activities. Will the “comes to the nuisance” rule affect the pace and pattern of development? Is there anything objectionable about the behavior potentially motivated by a “comes to the nuisance” rule? [Hint: The rule in question sets up a “race” for valuable legal rights. Is this race costless? Might society arguably be “better off” if the rights were allocated in some other way? When is it worth paying the costs of having the race?]

5. This case and several that follow involve the notion that property can be defined in terms of legal rules that have varying degrees of potency. It can make a great deal of difference whether the “ownership” of a right is protected by an injunctive remedy or a damages remedy. In the legal literature, the distinction was emphasized by Calabresi and Melamed in “Property Rules, Liability Rules, and Inalienability,” 85 Harv.L.Rev. 1089 (1972). Do you understand the implications of using “property rules” rather than “liability rules”? For an extended discussion, see Polinsky, “Resolving Nuisance Disputes: The Simple Economics of Injunctive and Damage Remedies,” 32 Stan.L.Rev. 1075 (1980).

J. WEINGARTEN, INC. v. NORTHGATE MALL, INC.
404 So.2d 896 (Supreme Court of Louisiana, 1981)

DENNIS, Justice.

We are called upon to decide whether, under the circumstances of this case, a court should specifically enforce a lease by ordering the destruction of the major part of a $4 million building which a shopping center developer erected in an area reserved to its tenant for customer parking. The trial court refused to
order the building razed, but the court of appeal reversed, requiring that approximately 60% of the building be torn down and removed within six months of the effective date of its judgment.

The defendant, Northgate Mall, Inc., developed an enclosed shopping mall on approximately .35 acres of land. In 1968, the defendant subleased space in the mall to the plaintiff, J. Weingarten, Inc., for the operation of a grocery store.

In early 1978, the defendant began planning to renovate and expand the mall to counter expected competition from a new mall to be built in Lafayette. The defendant’s right to expand was limited to roughly 400,000 square feet, but it proposed to add slightly over 100,000 square feet. Therefore, consent of both parties was necessary under the lease to permit expansion. However, no written modification ever resulted from the parties’ preliminary negotiations.

In February, 1973, the defendant erected a construction fence and began moving construction equipment and material onto the job site. The plaintiff alerted the defendant that it considered the defendant’s activities a breach of contract and that injunctive relief would be sought if such activities did not cease. Last minute negotiations between the parties were unsuccessful, and the plaintiff filed suit on March 1, 197.9 seeking preliminary and permanent injunctive relief.

Pursuant to plaintiff’s petition, the trial court issued a temporary restraining order prohibiting further construction activities and continued the order until March 19, 1979 when a hearing was held on plaintiff’s motion to dissolve the restraining order. The trial court denied the preliminary injunction and dissolved the temporary restraining order based on its finding that Weingarten failed to show that it would sustain irreparable damage without injunctive relief. The trial on the merits of the petition for a permanent injunction was not held until October 25 and 26, 1979. By this time the $4 million expansion project was virtually complete and the new stores therein were open for business.

After the trial on the permanent injunction, the trial court reaffirmed its earlier finding that the plaintiff had not demonstrated that it would be irreparably harmed and could be adequately compensated monetarily. The court refused to enforce specifically a provision in the lease stipulating that an injunction to enforce Weingarten’s rights to egress and passage over the parking area occupied by the new building could be obtained without the necessity of showing irreparable harm or the inadequacy of damages. The trial court concluded the agreed remedy provision was against public policy relying on Termplan Arabi, Inc. v. Carollo, 299 So.2d 831 (La.App. 4th Cir.).

The court of appeal reversed the trial court, holding that the plaintiff was entitled to permanent injunctive relief because the agreed remedy provision was valid and because the plaintiff was irreparably harmed. We granted defendant’s application to consider the appropriateness of specific performance in this case.

Northgate agreed not to erect any additional buildings in the parking area of the shopping center except within the space shown on a plat attached to the lease designated as sites for future department store and a proposed theater. The defendant also promised to maintain a ratio of six car parking spaces for each 1,000 feet of floor space in the shopping center. It further agreed that its tenant, Weingarten, would have the right to an “irrevocable non-exclusive easement” over all parking areas shown on the plat attached to the lease. In connection with the “easement,” the lease provides that Weingarten “shall have the right to obtain an injunction specifically enforcing such rights and interests without the necessity of proving inadequacy of legal remedies or irreparable harm.”

We concur in the court of appeal’s findings that Northgate breached each of the contractual provisions. However, we disagree with the court of appeal finding that the plaintiff had shown that it would be irreparably injured by the breaches of contract. The evidence fully supports the trial judge’s determination that plaintiff failed to demonstrate that its injury would be irreparable or insusceptible to adequate compensation.
The decisive issue presented by the breaches of contract is whether Weingarten is entitled to the substantive right of specific performance under the circumstances of this case. * * * Plaintiff Weingarten has asked specific performance of defendant Northgate’s obligation not to do something—namely, not to infringe on plaintiff’s contractual rights over areas reserved for parking by the lease. Civil Code articles 1926 through 1929 govern the enforcement of obligations to do, or not to do.

Articles 1926 through 1929 Provide:

Art. 1926. On the breach of any obligation to do, or not to do, the obliger is entitled either to damages, or, in cases which permit it, to a specific performance of the contract, at his option, or he may require the dissolution of the contract, and in all these cases damages may be given where they have accrued, according to the rules established in the following section.

Art. 1927. In ordinary cases, the breach of such a contract entitles the party aggrieved only to damages, but where this would be an inadequate compensation, and the party has the power of performing the contract, he may be constrained to a specific performance by means prescribed in the laws which regulate the practice of the courts.

Art. 1928. The obliger may require that any thing which has been done in violation of a contract, may be undone, if the nature of the cause will permit, and that things be restored to the situation in which they were before the act complained of was done, and the court may order this to be effected by its officers, or authorize the injured party to do it himself at the expense of the other, and may also add damages, if the justice of the case require it.

Art. 1929. If the obligation be not to do, the obligee may also demand that theobliger be restrained from doing any thing in contravention of it, in cases where he proves an attempt to do the act covenanted against.

Since they are not models of clarity, the articles are susceptible to more than one reasonable interpretation.

A literal reading of the codal language may lead one to conclude that the general rule, i.e., the one to be applied in ordinary cases, is that the breach of a contract entitles the party aggrieved only to damages; and specific performance may only be obtained where damages would be an inadequate compensation. Under this review, an obligee could only require a thing done in violation of the contract to be undone if the nature of the cause would permit, viz., in the exceptional case in which damages are an inadequate remedy. Such an interpretation is suspect, however, because of its common law overtones.

Eminent Louisiana civilian commentators have argued persuasively to the contrary that these articles of the code were designed to enhance the rank of specific performance over damages as a remedy for the nonperformance of obligations to do or not to do in a manner not expressed in the comparable French articles.

We agree with the distinguished doctrinal writers that Articles 1926 through 1929 were intended to give first rank to the obligee’s right to performance in specific form, consistently with other provisions of the Code. Therefore, we reject the common law view that the obligee must first clear the inadequacy of damage-irreparable injury hurdle before invoking the remedy.

A reading of the articles as a whole, however, implies that courts are empowered to withhold specific performance in some exceptional cases even when specific performance is possible. The phrases, “if the nature of the cause will permit” and “in ordinary cases,” suggest a reference to a traditional civilian concept. The civil law systems, i.e., those descended from Roman law, have by and large proceeded on the premise that specific redress should be ordered whenever possible, unless disadvantages of the remedy outweigh its advantages. The main reservations have been for cases where specific relief is impossible, would involve disproportionate cost, would introduce compulsion into close personal relationships or compel the expression of special forms of artistic or intellectual creativity. Dawson, Specific Performance in France and Germany, 57 Mich.L.Rev. 495, 520(1959). Professor Litvinoff has described the French jurisprudence as allowing courts to deny specific performance whenever the inconvenience of such forced
execution would exceed the advantage, as when the cost of performing in kind is disproportionate to the actual damage caused, or when it would have a negative effect upon the interest of third parties.

For the foregoing reasons, we conclude that the legislative aim of the redactors of the code was to institute the right to specific performance as an obligee’s remedy for breach of contract except when it is impossible, greatly disproportionate in cost to the actual damage caused, no longer in the creditor’s interest, or of substantial negative effect upon the interests of third parties. Applying this interpretation to the present case, we conclude that its nature will not permit specific performance.

The evidence clearly reflects that the cost of tearing down most of the $4 million building and doing incalculable damage to the remainder of the shopping center greatly outweighs any actual damage caused to the plaintiff. It is dubious that the devastation of such a building in a shopping center built with the hope of competing with a neighboring market complex is in Weingarten’s real interest. It is evident that the third persons not party to the contract or to this lawsuit would be negatively affected. Shopping center tenants in the new building would lose their store space and risk losing their investments. H.J. Wilson Company, which occupies over 63,000 square feet of the building, would be particularly disadvantaged. The owners of the land and John Hancock Mutual Life Insurance Company, whose long term loan is secured by a mortgage on the shopping center and pledges of the leases of the five major tenants, would suffer a substantial negative effect. Although perhaps not controlling, additional considerations weighing against specific performance and its consequent destruction of a major commercial building are the potential negative effects upon the community in the form of economic waste, energy dissipation, and possible urban blight.

Although the contractual provisions involved here cannot be specifically enforced under the circumstances of this case, we hold that they could have been under different conditions. Indeed, unless exceptional conditions prevail as in this case anything which has been done in violation of the contract may be undone, including the destruction of a building. Moreover, plaintiff is not without a remedy because it is entitled to be compensated fully in damages for any loss it sustains as a result of the breach of contract. Because the record presently does not provide an adequate basis for the assessment of damages, however, the case will be remanded to the trial court for that purpose.

Reversed and Remanded.

QUESTIONS

1. Assume that, in addition to Weingarten, there were originally 20 other lessees who held similar contractual rights. Could this case have been decided on the same principles as Boomer? Would your answer to this question change if the 20 others had all settled amicably through out of court negotiations before Weingarten filed suit?

2. Does the court’s announced principle for resolving the case arguably result in bad behavioral incentives for future parties who are in the position of Northgate?

SPUR INDUSTRIES, INC. v. DEL E. WEBB DEVELOPMENT CO.

494 P.2d 700, (Supreme Ct. of Ariz., In Bane, 1972)

CAMERON, Vice Chief Justice.

From a judgment permanently enjoining the defendant, Spur Industries, Inc. from operating a cattle feedlot near the plaintiff Del. E. Webb Development Company’s Sun City, Spur appeals. Webb cross appeals. Although numerous issues are raised, we feel that it is necessary to answer only two questions. They are:

---

6 Even when a valid agreed remedy that provision such as stipulated damages is included in the parties’ contract, the court retains some discretion to modify that provision when circumstances such as partial performance make the agreed the remedy inappropriate or unjust.
1. Where the operation of a business, such as a cattle feedlot is lawful in the first instance, but becomes a nuisance by reason of a nearby residential area, may the feedlot operation be enjoined in an action brought by the developer of the residential area?

2. Assuming that the nuisance may be enjoined, may the developer of a completely new town or urban area in a previously agricultural area be required to indemnify the operator of the feedlot who must move or cease operation because of the presence of the residential area created by the developer?

[The area in question is located in Maricopa County, Arizona, 14 to 15 miles west of Phoenix, on Grand Avenue. About two miles south of Grand Avenue is Olive Avenue which runs east and west. Farming started in this area about 1911. By 1950, the only urban areas in the vicinity were agriculturally related communities. The retirement community of Youngstown, between Grand and Olive Avenues, was commenced in 1954.

In 1956 Spur’s predecessors in interest developed feedlots about one half mile south of Olive Avenue. The area is well suited for cattle feeding, and in 1959, there were 25 cattle feeding pens within a 7 mile radius. In May of 1959 Del Webb began to plan the development of an urban area to be known as Sun City. The price which Del Webb paid for the land was considerably less than the price for land near Phoenix, and this price, along with the success of Youngstown, was a factor influencing the decision to purchase the property in question.

In 1960 Spur began a rebuilding and expansion program and by 1962 had extended its expansion north and south to encompass 114 acres. Del Webb offered its first homes in January 1960. The first unit to be completed was approximately 2.5 miles north of Spur. By May 1960, 450 to 500 houses were completed or under construction. At this time, Del Webb did not consider odors from the Spur feed pens a problem and continued to develop in a southerly direction, until sales resistance became so great that the parcels were difficult if not impossible to sell.

Del Webb filed its original complaint alleging that lots in the southwest portion were unfit for development for sale as residential lots because of the operation of the Spur feedlot. Del Webb’s suit complained that the Spur feeding operation was a public nuisance because of the flies and odor. Despite the admitted good feedlot management and good housekeeping practices by Spur, the odor and flies resulting from the huge amounts of wet manure produced an annoying if not unhealthy situation as far as the senior citizens of southern Sun City were concerned. Some were unable to enjoy the outdoor living which Del Webb had advertised, and Del Webb was faced with sales resistance from prospective purchasers as well as strong and persistent complaints from the people who had purchased homes in that area. It is noted, however, that neither the citizens of Sun City nor Youngstown are represented in this lawsuit and the suit is solely between Del Webb and Spur Industries, Inc.

The difference between a private nuisance and a public nuisance is generally one of degree. A private nuisance is one affecting a single individual or a definite small number of persons in the enjoyment of private rights not common to the public, while a public nuisance is one affecting the rights enjoyed by citizens as a part of the public. To constitute a public nuisance, the nuisance must affect a considerable number of people or an entire community or neighborhood. City of Phoenix v. Johnson, 51 Ariz. 115 (1938).

Where the injury is slight, the remedy for minor inconveniences lies in an action for damages rather than in one for an injunction. Kubby v. Hammond, 68 Ariz. 17 (1948). Moreover some courts have held, in the “balancing of conveniences” cases, that damages may be the sole remedy. See Boomer v. Atlantic Cement Co., 309 N.Y.S.2d 312 (1970), and annotation comments, 40 A.L.R.3d 601.

Thus, it would appear from the admittedly incomplete record as developed in the trial court, that, at most, residents of Youngstown would be entitled to damages rather than injunctive relief.

We have no difficulty, however, in agreeing with the conclusion of the trial court that Spur’s operation was an enjoicable public nuisance as far as the people in the southern portion of Del Webb’s Sun City were concerned.
* * *  
In addition to protecting the public interest, however, courts of equity are concerned with protecting the operator of a lawfully, albeit noxious, business from the result of a knowing and willful encroachment by others near his business.

In the so-called “coming to the nuisance” cases, the courts have held that the residential landowner may not have relief if he knowingly came into a neighborhood reserved for industrial or agricultural endeavors and has been damaged thereby:

“Plaintiffs chose to live in an area uncontrolled by zoning laws or restrictive covenants and remote from urban development. In such an area plaintiffs cannot complain that legitimate agricultural pursuits are being carried on in the vicinity, nor can plaintiffs, having chosen to build in an agricultural area, complain that the agricultural pursuits carried on in the area depreciate the value of their homes. The area being primarily agricultural, any opinion reflecting the value of such property must take this factor into account. The standards affecting the value of residence property in an urban setting, subject to zoning controls and controlled planning techniques, cannot be the standards by which agricultural properties are judged.” *Dill v. Excel Packing Co.*, 331 P.2d 539, 548, 549 (Kan., 1958).

And:

“ * * * a party cannot justly call upon the law to make that place suitable for his residence which was not so when he selected it. * * *.” *Gilbert v. Showerman*, 23 Mich. 448, 2 Brown 15X (1871).

Were Webb the only party injured, we would feel justified in holding that the doctrine of “coming to the nuisance” would have been a bar to the relief asked by Webb, and, on the other hand, had Spur located the feedlot near the outskirts of a city and had the city grown toward the feedlot, Spur would have to suffer the cost of abating the nuisance as to those people locating within the growth slattern elf the expanding city:

“The case affords, perhaps, an example where a business established at a place remote from population is gradually surrounded and becomes part of a populous center, so that a business which formerly was not an interference with the rights of others has become so by the encroachment of the population * * *.” *City of Ft. Smith v. Western Hide & Fur Co.*, 239 S.W. 724, 726 (Ark., 1922).

We agree, however, with the Massachusetts court that:

“The law of nuisance affords no rigid rule to be applied in all instances. It is elastic. It undertakes to require only that which is fair and reasonable under all the circumstances. In a commonwealth like this, which depends for its material prosperity so largely on the continued growth and enlargement of manufacturing of diverse varieties, ‘extreme rights’ cannot be enforced. * * *.” *Stevens v. Rockport Granite Co.*, 104 N.E. 371, 373 (Mass., 1914).

There was no indication in the instant case at the time Spur and its predecessors located in western Maricopa County that a new city would spring up, full-blown, alongside the feeding operation and that the developer of that city would ask the court to order Spur to move because of the new city. Spur is required to move not because of any wrongdoing on the part of Spur, but because of a proper and legitimate regard of the courts for the rights and interests of the public.

Del Webb, on the other hand, is entitled to the relief prayed for (a permanent injunction), not because Webb is blameless, but because of the damage to the people who have been encouraged to purchase homes in Sun City. It does not equitably or legally follow, however, that Webb, being entitled to the injunction, is then free of any liability to Spur if Webb has in fact been the cause of the damage Spur has sustained. It does not seem harsh to require a developer, who has taken advantage of the lesser land values in a rural area as well as the availability of large tracts of land on which to build and develop a new town or city in the area, to indemnify those who are forced to leave as a result.
Having brought people to the nuisance to the foreseeable detriment of Spur, Webb must indemnify Spur for a reasonable amount of the cost of moving or shutting down. It should be noted that this relief to Spur is limited to a case wherein a developer has, with foreseeability, brought into a previously agricultural or industrial area the population which makes necessary the granting of an injunction against a lawful business and for which the business has no adequate relief.

It is therefore the decision of this court that the matter be remanded to the trial court for a hearing upon the damages sustained by the defendant Spur as a reasonable and direct result of the granting of the permanent injunction.

QUESTIONS

1. In the court’s view, which party “owned” the relevant property right? Was that party deprived of the value of any part of its property through the court’s ruling?

2. Why should Del Webb not have been required to bribe Spur to move away?

3. Suppose Del Webb had acquired its tract in 1954 and placed thereupon large signs announcing “Future Site of Sun City Residential Development.” Would this have been likely to change the court’s assessment of Del Webb’s position? If not, what more would Del Webb have had to do in order to avoid risking future legal liability to Spur?

Sprecher v. Adamson Companies et al.

636 P.2d 1121, (Supreme Ct. of Cal., In Bank, 1981)

[Respondent Adamson Companies is a joint owner of a 90-acre parcel of land in Malibu. Opposite the parcel are a number of homes, including that of appellant, Peter Sprecher. Respondents’ tract contains part of an active landslide which extends seaward from the parcel for some 1,700 feet along Malibu Road and beyond the boundaries of respondent’s property. In March 1978, heavy spring rains triggered a major movement of the slide which caused Sprecher’s home to rotate and to press against the home of his neighbor, Gwendolyn Sexton. She filed an action against Sprecher, seeking to enjoin the encroachment of his home upon hers. Appellant cross-complained against Sexton, the County of Los Angeles and respondents. Specifically, appellant sought damages for the harm done to his home by the landslide, alleging that such damage proximately resulted from respondents’ negligent failure to correct or control the landslide condition. Respondents successfully moved for summary judgment, arguing that a possessor of land has no duty to remedy a natural condition of the land in order to prevent harm to property outside his premises.]

* * * Under the common law, the major important limitation upon the responsibility of a possessor of land to those outside his premises concerned the natural condition of the land. (Prosser, Law of Torts (4th Ed. 1971) Sec. 57, p. 354.) While the possessor’s liability for harm caused by artificial conditions was determined in accord with ordinary principles of negligence (id., at p. 355; see Rest. 2d Torts, Secs. 364-370), the common law gave him an absolute immunity from liability for harm caused by conditions considered natural in origin. (Prosser, supra, at p. 354; see Rest. 2d Torts, Sec. 363.) No matter how great the harm threatened to his neighbor, or to one passing by, and no matter how small the effort needed to eliminate it, a possessor of land had no duty to remedy conditions that were natural in origin.

This court has held that it will not depart from the fundamental concept that a person is liable for injuries caused “by his want of ordinary care * * * in the management of his property or person * * * “ except when such a departure is “clearly supported by public policy.” (Rowland v. Christian (1968) 69 Cal.2d 108, 112, 70 Cal. Rptr. 97, 443 P.2d 561.) Accordingly, common law distinctions resulting in wholesale immunities have been struck down when such distinctions could not withstand critical scrutiny. * * * In Rowland, this court stated that “[a] departure from [the] fundamental principle involves the balancing of a number of considerations[,] [T]he major ones are the foreseeability of harm to the plaintiff, the degree of certainty that the plaintiff suffered injury, the closeness of the connection between the defendant’s conduct and the injury suffered, the moral blame attached to the defendant’s conduct, the policy of preventing
future harm, the extent of the burden to the defendant and [the] consequences to the community of imposing a duty to exercise care with resulting liability for breach and the availability, cost, and prevalence of insurance for the risk involved.”

This progression of the law in California mirrors what appears to be a general trend toward rejecting the common law distinction between natural and artificial conditions. Instead, the courts are increasingly using ordinary negligence principles to determine a possessor’s liability for harm caused by a condition of the land. The early case of *Gibson v. Denton* (1896) 4 A.D. 198, 38 N.Y.S. 554 was a precursor of this trend. In *Gibson*, the court held a possessor of land liable for damage caused when a decayed tree on her premises fell on the home of her neighbor during a storm. After noting that the defendant clearly would be liable for the fall of a dilapidated building, or artificial structure, the court observed that it could “see no good reason why she should not be responsible for the fall of a decayed tree, which she allowed to remain on her premises.” (Id., at p. 555.) “[T]he tree was on her lot, and was her property. It was as much under her control as a pole or building in the same position would have been.” (Ibid.) Thus, “[t]he defendant had no more right to keep, maintain, or suffer to remain on her premises an unsound tree * * * than she would have had to keep a dilapidated and unsafe building in the same position.” (Id., at pp. 555-556.)

In more recent years, at least 13 other states and the District of Columbia have begun applying ordinary negligence principles in determining a possessor’s liability for harm caused by a natural condition. * * * Not surprisingly, all these cases involved an injury caused by a fallen tree. However, the principles expressed by these courts are not so limited. For example, the court in *Dudley v. Meadowbrook, Inc.*, 166 A.2d 743, held that a possessor of land has a “duty of common prudence in maintaining his property, including trees thereon, in such a-way as to prevent injury to his neighbor’s property.” (See also *Turner v. Ridley*, 144 A.2d at pp. 270-271 [same court; duty applied to automobile parked on public road].)

The courts are not simply creating an exception to the common law rule of nonliability for damage caused by trees and retaining the rule for other natural conditions of the land. Instead, the courts are moving toward jettisoning the common law rule in its entirety and replacing it with a single duty of reasonable care in the maintenance of property. This development is reflected in the Restatement Second of Torts, which now recognizes that a possessor of land may be subject to liability for harm caused not just by trees but by any natural condition of the land.

* * *

In rejecting the common law rule of nonliability for natural conditions, the courts have recognized the inherent injustice involved in a rule which states that a “landowner may escape all liability for serious damage to his neighbors [or those using a public highway], merely by allowing nature to take its course.” (Prosser, supra, at p. 355.) As one commentator has observed: “[w]here a planted tree has become dangerous to persons on the highway or on adjoining land, and causes harm, the fault lies not in the planting of the tree but in permitting it to remain after it has become unsafe.” (Noel, Nuisances from Land in Its Natural Condition (1943) 56 Harv.L.Rev. 772, 796-797)

Historically, the consideration most frequently invoked to support the rule of nonliability for natural conditions was that it was merely an embodiment of the principle that one should not be obligated to undertake affirmative conduct to aid or protect others [citations]. This doctrine rested on the common law distinction between the infliction of harm and the failure to prevent it, or misfeasance and nonfeasance. Misfeasance was determined to exist when a defendant played some part in the creation of a risk, even if his participation was innocent. Nonfeasance occurred when a defendant had merely failed to intervene in a plaintiff’s behalf. (See generally, Weinrib, The Case for a Duty to Rescue (1980) 90 Yale L.J. 247, 251-258.) Liability for nonfeasance, or the failure to take affirmative action, was ordinarily imposed only where some special relationship between the plaintiff and defendant could be established. (Id., at p. 248 and fn. 7; Rest. 2d Torts, Sec. 314, com. a.)

Proponents of the rule of nonliability for natural conditions argued that a defendant’s failure to prevent a natural condition from causing harm was mere nonfeasance. A natural condition of the land was by definition, they argued, one which no human being had played a part in creating. (See Noel, Nuisances
from Land in Its Natural Condition, supra, 56 Harv.L.Rev., at p. 773.) Therefore, no basis for liability existed because a duty to exercise reasonable care could not arise out of possession alone. (Ibid.) Since there was no special relationship between the possessor of land and persons outside the premises, there could be no liability. (See Rest. 2d Torts, Sec. 314, com. f.) Conversely, a defendant’s failure to prevent an artificial condition from causing harm constituted actionable misfeasance.

Whatever the rule may once have been, it is now clear that a duty to exercise due care can arise out of possession alone. One example is provided by modern cases dealing with the duty of a possessor of land to act affirmatively for the protection of individuals who come upon the premises. In days gone by, a possessor of land was deemed to owe such a duty of care only to invitees. That is, the duty to act affirmatively was grounded in the special relation between the possessor-invitee and the invitee. Rowland held that whether the individual coming upon the land was a trespasser, a licensee or an invitee made no difference as to the duty of reasonable care owed but was to be considered only as to the issue of whether the possessor had exercised reasonable care under all the circumstances.

Modern cases recognize that after Rowland, the duty to take affirmative action for the protection of individuals coming upon the land is grounded in the possession of the premises and the attendant right to control and manage the premises.

* * *

Thus, it becomes clear that the traditional characterization of a defendant’s failure to take affirmative steps to prevent a natural condition from causing harm as nonactionable nonfeasance conflicts sharply with modern perceptions of the obligations which flow from the possession of land. Possession ordinarily brings with it the right of supervision and control. As Justice Mosk has aptly stated, the right of supervision and control “goes to the very heart of the ascription of tortious responsibility. ** * * (Connor v. Great Western Sav. & Loan Assn., 447 P.2d 609 (1968) [dis. opn. of Mosk, J.])

Another deficiency of the historical justification of the rule of nonliability is simply that it proves too much. Under the traditional analysis, a possessor of land should be excused from any duty to prevent harm to persons outside his land whenever he has played no part in the creation of the condition which threatens the harm, be it artificial or natural. However, most courts recognize that the possessor is under an affirmative duty to act with regard to a dangerous artificial condition even though the condition was created solely by some predecessor in title or possession (Rest.2d Torts, Sec. 366; see e.g. Dye v. Burdick 553 S.W.2d 883, 836-837(Ark., 1977)); or by the unauthorized conduct of some other third person (Rest.2d Torts, Sec. 364, subd. (c)). “To impose such a duty is to cross the line from misfeasance to nonfeasance” unless the present possessor somehow aggravates the danger. (See Weinrib, The Case for a Duty to Rescue, supra, 90 Yale L.J. at p. 257.)

Interestingly enough, in the cases holding that a possessor has an affirmative duty to prevent harm by a dangerous artificial condition created solely by another, the liability of the defendant has been predicated upon his possession and control of the artificial condition which caused the harm. (E.g., Dye v. Burdick, supra, 553 S.W.2d 833; cf. Pridgen v. Boston Housing Authority (Mass.,1974) 308 N.E.2d 467, 475-478.) Thus, these cases confirm that mere possession with its attendant right to control conditions on the premises is a sufficient basis for the imposition of an affirmative duty to act. (See generally, 2 Harper & James, supra, at pp. 1523-1525.) In sum, the historical justification for the rule of nonliability for natural conditions has lost whatever validity it may once have had.

* * *

The trend in the law is in the direction of imposing a duty of reasonable care upon the possessor of land with regard to natural conditions of land. The erosion of the doctrinal underpinning of the rule of nonliability is evident from even a cursory review of the case law. Also evident is the lack of congruence between the old common law rule of nonliability and the relevant factors which should determine whether a duty exists. All this leads to but one conclusion. The distinction between artificial and natural conditions should be rejected. “A [person’s] life or limb [or property] does not become less worthy of compensation under the law” because that person has been injured by a natural, as opposed to an artificial, condition.
It must also be emphasized that the liability imposed is for negligence. The question is whether in the management of his property, the possessor of land has acted as a reasonable person under all the circumstances. The likelihood of injury to plaintiff, the probable seriousness of such injury, the burden of reducing or avoiding the risk, the location of the land, and the possessor’s degree of control over the risk-creating condition are among the factors to be considered by the trier of fact in evaluating the reasonableness of a defendant’s conduct.

* * *

Construing appellant’s evidence liberally and respondents’ narrowly, as this court must, a rational inference can be drawn that effective measures for the control of the slide were within respondent’s reach and that such measures would have entailed a substantial expense. Although the cost of implementing only slightly effective measures would not be justified by the benefit to appellant, it can be inferred that the cost of implementing effective measures might (or might not) be justified by the benefit to appellant and to respondents. Although the case is a close one, the evidence does not conclusively establish that no rational inference of negligence can be drawn under the circumstances of this case.

* * *

The judgment of the trial court is reversed and the cause remanded to the trial court for further proceedings consistent with the views expressed in this opinion.

QUESTIONS

1. “This case is a contest about the ownership of a bad rather than a good, a liability rather than an asset. The traditional rule is that the ownership of a bad rests upon the party whom Nature afflicts with its results.” Comment.

2. What is the fundamental difference between the issue in Sprecher on the one hand and the nuisance cases of Boomer and Spur on the other? Is Sprecher analogous to the dispute between the Fontainebleau and Eden Roc?

3. Why does the court seem to find it necessary to overturn the traditional rule on natural immunities? Is its reasoning sound? How would you argue this case on behalf of Adamson Companies?

4. How do you think the rule of liability for mudslides should be formulated? Does your view in any way depend on factual questions? If you would not espouse the same rule for all natural hazards, explain why.

5. Suppose that the owner of a large tract became aware that a volcanic eruption, similar to Mt. St. Helens, was developing under his land. The prospect is that the eruption will cause $1 Billion worth of damages. The eruption can be “plugged” at the cost of $1 Million. Do you think that the Sprecher court would find a duty to plug?
A. BEHAVIOR INVOLVING UNCERTAINTY

Many human decisions are made at a point in time when all or many of their important consequences, whether costs or benefits, lie concealed in the mists of a future that can be known only in terms of probabilities. The dimension of time does not inevitably implicate grave uncertainty about the consequences of a decision; there are sometimes future events about which we can be quite confident. As a practical matter, however, intertemporal choices often involve probability distributions of future results rather than the determinate, single-valued outcomes normally referred to as “certainty.” For instance, even if one knows the costs of filing and litigating a case, the cost-benefit analysis of whether to go forward may require comparing a set of relatively predictable costs with the perhaps wildly variant values of the possible future pay-outs when the case is ultimately decided by a judge or jury. In general, both the costs and benefits that will flow from a decision have at least some elements of uncertainty about them, but different aspects of a choice may implicate the uncertainty issue in different degrees.

How can decision-making in the face of uncertainty be dealt with? Specifically, how can whole constellations of possible results, each having different probability, be reduced to terms that are intelligible for analytic purposes? These materials survey a few of the simple models commonly used to deal explicitly with uncertain future prospects. In many practical applications, risk factors loom as sufficiently important that their effects must be confronted squarely. Bear in mind, however, that we retain one important simplifying assumption: although prospective costs and benefits are adjusted for their uncertainty characteristics, no separate account is taken of timing per se, i.e., the distances into the future at which the events occur. Even at an intuitive level, it is apparent that a dollar of damages paid or received tomorrow is, in an important sense not at all the same magnitude as a dollar whose transfer takes place five years hence. Therefore, a separate issue is that of how costs and benefits that accrue at different points in time must, even after uncertainty has been taken into account, be adjusted to render them comparable for decision-making purposes. Use of an appropriate interest rate to “discount” future costs and benefits is an additional topic, beyond the scope of what we will be doing at this juncture.

1. PROSPECTS AND CERTAINTY EQUIVALENTS

In making present decisions, a person often confronts uncertain future consequences that may conveniently be termed “prospects.” The prospects to be evaluated frequently consist of a range of alternative outcomes, perhaps differing very greatly in value or “payoff.” Moreover, just as the potential consequences of a choice vary in their impact, they may also vary enormously in their likelihood; some may be of very high probability while others have almost negligible chances of occurring.

As we shall illustrate below, it is convenient to express the value of a stream of uncertain future prospects, whether costs or benefits, in terms of a concept called their “certainty equivalent.” A certainty equivalent is the immediate (certain) payment that a person would regard as having a value equal to that of a particular uncertain set of prospects. In other words, a person given a choice should be indifferent between any prospect and its certainty equivalent.

For instance, suppose that your Fairy Godmother is about to toss a normal, six-sided die once. She will pay you $20 if a 2 turns up or $40 if a 4 turns up. You will lose $30 if a 6 turns up, and neither win nor lose if any other number results. Upon reflection about the “odds” involved and the possible payoffs, most people would regard this as an attractive prospect. But before Fairy Godmother tosses the die, Gandalf the Wizard appears and asks what monetary payment you would accept and feel exactly compensated for the fact that he is about to make the Godmother forget her offer. (The wizard also puts you under a spell whereby you are constrained to give the truthful answer, rather than the exaggerated figure that avarice might otherwise suggest.) Your answer would then be your certainty equivalent of the lost prospect.
Of course, a prospect may also have a negative value. If it were the Wicked Witch of the East about to toss the die, the payoffs in the above example could be changed to all negative numbers. They would be alternative sums that Wicked Witch proposes to take away rather than give you. Now the magnitude that might interest Gandalf is what it is worth to you for him to drive the Witch away rather than permit you to be exposed to the risk of her extracting imposing costs on you. In either case, the essential nature of the exercise is really the same: the person facing the prospect formulates its certainty equivalent, whether it be a certain loss that is the equivalent of prospective costs or a certain gain that is the equivalent of uncertain benefits. Obviously, there is also the middle case where the prospects involve a mixture of benefits and costs. Depending on the net predominance of costs or benefits, the certainty value of the prospect will then be either negative or positive.

2. EXPECTED VALUES

Exactly how is a certainty equivalent actually derived from an underlying set of facts? Unfortunately, there are a number of different theories as to how people do this, and many of the theories are extremely complicated. What is presented below therefore is an admittedly oversimplified treatment which, however, will capture the flavor of more complex formulations. The reader interested in more sophisticated and technical theories of decision-making under uncertainty is invited to delve into the mass of sophisticated literature on decision-making under conditions of uncertainty and incomplete information. [See generally, Karl H. Borch, The Economics of Uncertainty, (Princeton Univ. Press, 1968); J. Hirshleifer and J. G. Reilly, “The Analytics of Uncertainty and Information—An Expository Survey,” 17 J. of Econ. Literature 1375 (1979).]

One relatively simple and widely used method of computing a certainty equivalent is the concept of “mathematical expectation,” perhaps more popularly called the “expected value.” In essence, an expected value is the weighted average of all possible payoffs, where the weights used are the probabilities of the different results. Note that, because the probabilities of the events are used in the weighting process, an outcome is given an importance proportional to its likelihood as well as to the magnitude of its effect. For the die-tossing example given above, the expected value would be computed as indicated in Exhibit 2.1.

Observe that the expected value, a $5 gain, is not an actual outcome that would ever be possible in the retrospective or ex post sense; “Win $5” is not an outcome that can occur. The $5 payoff is “expected” only as an average value of the prospect considered ex ante, as opposed to a value that is particularly likely to be observed ex post.

**Exhibit 2.1 Computing Expected Values**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Probability</th>
<th>Calculation</th>
<th>Weighted Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Win $20</td>
<td>1/6</td>
<td>$20 X 1/6 =</td>
<td>$3.33</td>
</tr>
<tr>
<td>Win $40</td>
<td>1/6</td>
<td>$40 X 1/6 =</td>
<td>$6.67</td>
</tr>
<tr>
<td>Lose $30</td>
<td>1/6</td>
<td>-$30 X 1/6 =</td>
<td>-$5.00</td>
</tr>
<tr>
<td>No Change</td>
<td>3/6</td>
<td>$0 X 3/6 =</td>
<td>$0.00</td>
</tr>
<tr>
<td>Expected Value</td>
<td>=</td>
<td></td>
<td>$5.00</td>
</tr>
</tbody>
</table>

The example depicted in Exhibit 2.1 illustrates the general methodology for computing an expected value. The first column identifies all of the possible results and their associated payoffs. The second column is the probability of the outcome, commonly expressed as a decimal fraction but sometimes described as a percent or a common fraction. In any event, the sum of the probabilities in column 2 must be unity, or 100%, if all of the possibilities have indeed been accounted for. The third column is the product of the payoff times its associated probability. The method of computing this third column illustrates how the final results will depend equally on the size of an event’s payoff and on the
magnitude of its probability. Finally, adding up the third column yields the weighted average or “expected value.”

The valuation of future prospects is sometimes depicted as if decision-makers compared prospects exclusively in terms of their expected values. As a rough description, this practice is quite useful and has a considerable measure of plausibility. It does seem sensible, after all, that the alternative consequences should have “weights” approximately equal to their likelihoods. The actual process whereby individuals reduce prospects to certainty equivalents is nonetheless much more complicated. For instance, it is also plausible to think that such a decision-making process would pay some attention to the dispersion of the possible outcomes.

For instance, the same $5 expected value would result if the Fairy Godmother were about to flip a coin and give a person $50,010 for a head but extract $50,000 if a tail comes up:

$$\frac{1}{2} X +$50,010 = +$25,005$$
$$\frac{1}{2} X -$50,000 = -$25,000$$

Expected Value = +$5

The expected value computation above implies that this prospect is a “good” one, that it is worth $5. But would you pay $5 to have this prospect? Would you be indifferent between this prospect and the Fairy Godmother’s die-tossing prospect described above? (Remember, both of these prospects have the same $5 expected value.) Do you think that some people might even pay something to not be exposed to the prospect of flipping a coin for $1000? These questions, when answered in the way most people do, betray the fact that typical decision-makers do not care only about the average results that the expected value describes; they are also concerned with some measure of the degree of inherent uncertainty, the range or dispersion of the results around that average. Since expected values depend exclusively upon an averaging concept and do not indicate the dispersion around that summary measure, it should be apparent that expected values are not perfectly satisfactory descriptions of the way most people compare uncertain prospects and convert them to certainty equivalents. Hence, we shall have to complicate the story a bit in order to formulate even a very simple theory of decision making under uncertainty.

3. **RISK PREFERENCE**

When a decision-making entity cares about the dispersion of prospects in addition to their expected values, “risk preference” is said to be present. The existence of risk preference, its strength, and even its direction are a matter of each individual’s subjective preferences. If a particular person really did not care about the inherent uncertainty or dispersion in his prospects, then expected values would be a totally accurate way of describing that person’s assessment of prospects. Such a person, who would always rank different prospects in the exact order of their expected values, is called “risk neutral.” Many of us, however, would be prepared at times to reject an opportunity that has a higher expected value in favor of one with a lower expected value, provided that the alternative with the higher expected value also had a sufficiently higher level of risk attached. Although higher expected value is a “good,” additional risk in the form of result-dispersion is a “bad,” and we are prepared to trade off losses of expected value in exchange for reductions of risk, at least to some extent. Thus, in arriving at certainty equivalents, one might want to “discount” a prospect’s expected value in order to reflect dislike of its magnitude of risk. Individuals who attach such a negative value to risk are called “risk averse.” [Do not risk provoking horselaughs from the cognoscenti by saying risk adverse rather than risk averse; the proper sense of the term is aversion or avoidance, not opposition.] Finally, there are some people, epitomized by the “gambler” type, who may actually place a positive value on risk and who would adjust expected values upward to reflect a premium for risk. The latter are described as risk-seekers or “risk prone.”

What accounts for the existence of different sets of risk preferences in individuals? One explanation derives from the fact that the payoffs of prospects are commonly denominated in terms of things like money rather than in units of satisfaction. Would a person’s decline in satisfaction from a $1000 loss be
exactly the same magnitude as the increase in satisfaction from a $1000 gift? Suppose that the answer is "no," that the relevant numbers are a -100 and a +60, respectively, as measured on a particular person’s internal pleasure meter. Economists call such satisfaction measures “utility” indexes. For now, we shall assume that we are working with a utility index that is “cardinal,” i.e., whose numbers have a meaning analogous to those on a thermometer, units of utility measuring satisfaction in the same way that degrees on the thermometer measure quantities of heat. (So-called “ordinal” utility indices, which measure directions of preference change but not magnitudes, will be treated in the “indifference curve” models of that we will develop in the “Topographic Models” materials below.) Then, on the above facts, where a $1000 loss implies a 100-unit utility reduction but a $1000 gain produces only a 60-unit utility increase, the coin-flip for $1000 is \textit{not} an attractive prospect in utility terms: (-100 \times .5) + (+60 \times .5) = (-50 + 30) = -20. Even though the pecuniary expected value of the prospect is zero, corresponding to what is sometimes called a “fair bet,” it would not be surprising if a person with the hypothesized characteristics were willing to pay out a small sum of money rather than be forced to play the coin-flip game for those stakes.

The operative principle in the above example is “diminishing marginal utility of money,” i.e., each successive increment of money generates a smaller increase in satisfaction than the one before it. If this condition is true, equal monetary prospects of gains and losses do not counterbalance each other exactly; a dollar’s worth of loss is more important than a dollar’s worth of gain. A rudimentary “utility function” will make this clear. A utility function is just a mathematical way of describing the relationship between money (or some other good or bad) and a person’s utility index. Here is a simple example:

\[ U = 100M^{1/2} \]

In words, this says that the utility index \( U \) rises proportionately to one hundred times the square root of the quantity of money \( M \) that the person has to spend. (A number raised to the 1/2 power is a square root.) In general, utility functions are hypothesized or “made up” in order to exemplify some specific point in much the same way that illustrative facts are supplied in law school hypothetical cases. The example given above deliberately incorporates the principle of declining marginal utility of income. Based on the function specified, one could construct a table that shows various income levels and their resultant utility levels:

\begin{center}
\textbf{Exhibit 2.2 Money and Utility}
\end{center}

<table>
<thead>
<tr>
<th>Money</th>
<th>Utility Index</th>
<th>Marginal Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>387.30</td>
<td>13.13</td>
</tr>
<tr>
<td>16</td>
<td>400.00</td>
<td>____</td>
</tr>
<tr>
<td>17</td>
<td>412.31</td>
<td>____</td>
</tr>
<tr>
<td></td>
<td>____</td>
<td>____</td>
</tr>
<tr>
<td>63</td>
<td>793.73</td>
<td>6.32</td>
</tr>
<tr>
<td>64</td>
<td>____</td>
<td>____</td>
</tr>
<tr>
<td>65</td>
<td>____</td>
<td>____</td>
</tr>
<tr>
<td>66</td>
<td>____</td>
<td>____</td>
</tr>
<tr>
<td></td>
<td>____</td>
<td>____</td>
</tr>
<tr>
<td>120</td>
<td>1095.45</td>
<td>4.57</td>
</tr>
<tr>
<td>121</td>
<td>____</td>
<td>____</td>
</tr>
<tr>
<td>122</td>
<td>____</td>
<td>____</td>
</tr>
<tr>
<td>123</td>
<td>____</td>
<td>____</td>
</tr>
<tr>
<td></td>
<td>____</td>
<td>____</td>
</tr>
</tbody>
</table>
Given the equation $U = 100M^{0.5}$ above, one should be able to fill in, if necessary, the value of $U$ for any value of $M$, even for non-integer values of $M$. For instance, what are the missing entries in column two of Exhibit 2.2? Also, filling in the marginal or incremental utilities in column three confirms the existence of diminishing marginal utility of income.

As exemplified by the equation $U = 100M^{0.5}$ that we have been discussing, a “utility function” is an algebraic method of relating values of a satisfaction index to other values that describe different outcomes. The equation used will often be much more complex than our present example, the reason being that the conditions to be illustrated are themselves more complex. Even the present example may be “generalized” by replacing the 100 and $1/2$, respectively, by the abstract parameters $a$ and $b$; thus yielding $U = aM^b$. It is tempting for lawyers to lose patience with these algebraic expressions, but their advantages become apparent if one compares the utility function to an alternative expression of similar information. Specifically, Exhibit 2.2 conveys some of the same information as is embodied in the utility function, but it does so only incompletely and rather inelegantly. It would, of course, be literally impossible to supply a table that indicates the utility index for all possible values of $M$, even just for all of the integer values. Yet this is exactly what the equation does in a very succinct fashion.

Diminishing marginal utility of income is generally regarded as one plausible rationale for the refusal of persons to be indifferent about prospects that have zero expected money values and even to reject prospects that have positive expected money values. But what would we say about someone who rejects a bet that has zero or positive expected value in utility terms? This phenomenon may be regarded as a form of “pure” risk aversion, the result of a cost being attached to risk per se. It implies that such a person will be willing to trade off some loss in the expected value of utility for a diminution in the dispersion of the possible results.

Risk-neutral behavior is often assumed for expository purposes, despite there being good reason to believe that most people are at least slightly risk averse. The practical problem reflected here is that risk aversity is a subjective phenomenon which differs from person to person; although we know that high-dispersion prospects should probably be subject to some discount, we do not know exactly what magnitude that discount should be in any particular case. The relevant magnitude is likely to be highly idiosyncratic. Hence, the expected values are taken as a simple approximation. Any approximation inevitably introduces some bias and inaccuracy. For many purposes, the distortion costs of this simplification are dwarfed by its expository benefits. Occasionally, however, the existence of so-called “risk preference” must be explicitly taken into consideration because it is an important element in the behavior being modeled.

What is the bottom line here? Well, at minimum, one should be familiar with this terminology and with the way in which expected values are derived. Also, one should be sensitive to those instances in which the assumption of risk neutrality becomes something other than an expository simplification, but rather an element that arguably introduces significant changes in the results derived. Although such cases are the exception, it is worth being alert for them. You should also be aware that the “Expected Utility Model”—the assumption that a decision-maker maximizes the expected value of his utility index—is a behavioral hypothesis that has many complications not discussed in the necessarily simple treatment provided here. For a critical survey of the literature on expected utility decision making, see Schoemaker, “The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations,” 20 J. Econ. Literature 529 (1982).

QUESTIONS

1. Answer the following questions on the assumption that your utility function for money is accurately summarized by the simple square root function $U = 100M^{0.5}$ given in the text above. [Borrow a
calculator from a friend, if necessary! If you are not disposed to calculate actual answers, at least think about how one would go about deriving an exact answer.]

2. Suppose that you already have 100 dollars. You are to be exposed to the prospect of a coin-flip with stakes of $19. Is this an attractive prospect or not? Why?

3. What would be the certainty equivalent of the prospect in the previous question? I.e., what sum would you pay out to avoid the prospect (or avoid losing the prospect, as the case may be).

4. What does the previous question have to do with determining insurance premiums, or the value of warranties?

5. How much would the probability of winning the $19 have to be in order to make the expected value be zero in utility terms? What would the expected value in money terms then be? Can the immediately preceding questions be addressed by solving the following equation for $p$?

$$100(100^{\frac{1}{2}}) = p[100(119^{\frac{1}{2}})] + (1-p)[100(89^{\frac{1}{2}})]$$

If that equation is used, what does it implicitly assume about the presence or absence of “pure” risk aversion? About risk aversion due to diminishing marginal utility of income?

**B. APPLICATIONS**

1. **LAW OF THE LEANING TREE**

   One of the disillusioning aspects of early legal studies is the dawning realization that the law is, alas, not always so certain. An example of this is the question of liability for the effects of a tree that falls and damages property underneath it, or even threatens to do so.

   Abbott and Benson are neighbors. Close to their property line, but with its roots entirely located on Abbott’s land, is a very large and imposing oak tree. Not long ago, a major windstorm almost blew the tree down. The tree now leans across the property line into Benson’s air space, casting a threatening shadow upon a storage room attached to Benson’s house. Now, it is by no means clear that a subsequent windstorm will blow the tree entirely down, but both Abbott and Benson privately assess the probability of a fall as being 60%.

   Benson has approached Abbott with the view of having the tree cut down. “That’s fine with me,” says Abbott, “go right ahead and have it cut down.”

   “No, no. You don’t understand,” Benson replies. “The tree be longs to you and you have to cut it down.”

   “I understand very well that removing the tree will cost $385. If you want it down, do it yourself!” Abbott says.

   In most U.S. jurisdictions, the law governing this situation is a bit muddy. There is little doubt that Benson, at his own expense, has the right to sever that portion of the tree intruding into his own airspace. What is less clear is (1) whether Benson can secure an injunction to have Abbott remove the tree or (2) who pays for the damages if the tree does fall on Benson’s storage room, possibly damaging both the structure itself and the contents therein. In this problem, adopt the initial simplifying assumption that all parties are risk-neutral, basing their assessments of alternative prospects solely on the mathematical expected value. When you are done, reflect on what difference, if any, an assumption of risk aversion would have made.

**QUESTIONS**

1. Do you have any view on what the “right” legal rules should be about the two points described as unclear above? Describe the rationale for your position.
2. Suppose that, if the tree does fall, the damages incurred will be $760. “Should” the tree be cut down now, or would it be “better” to just accept the risk without any attempt at removing the potential danger? By how much would the probability of fall have to change in order for your answer to be different?

3. Assume that local law does not permit Benson to secure an injunction for the tree removal. It is unclear who would have to pay if the tree falls; no similar case has been litigated in this state’s courts. Abbott and Benson have reason to believe that it is a “coin flip” as to liability. What will be done about the tree. Why?

4. Same facts as in the previous question, except that the two neighbors have formed conflicting views as to the likelihood of liability for Abbott. Abbott thinks the odds are 3 to 2 in his favor. Benson feels that the probability that a court will find against Abbott is 60%. What do you predict will happen under these circumstances?

5. Is the characteristic of clarity and certainty a desideratum for legal rules? Why?

6. Suppose that the clear rule in a jurisdiction is that Abbott would not be liable for a tree fall under the circumstances depicted. However, a poll of laymen indicates that 80% of them are under the impression that a person in Abbott’s position would be responsible for the damages suffered by Benson. Does this information affect your assessment of the desirability of the rule itself?

2. **THE CASE OF MRS. CRISPY’S CHICKEN**

Mrs. Crispy is insured by the Faithless Fidelity Co. in the form of a policy with a $10,000 coverage limit. Crispy is sued for $100,000. Liability is disputed under conflicting legal theories and factual contentions. If liability is found, however, both Crispy and her insurance company agree that the judgment is likely to be exactly the $100,000 claimed by the plaintiff. Crispy thinks that the odds of being held liable are about fifty-fifty. The insurance company thinks that there is only a 40% chance of having to pay off if it spends $7,000 on Crispy’s defense.

Plaintiff offers to settle for $10,500. Faithless Fidelity, ignoring Crispy’s offer to chip in $1500 toward the required $10,500, refuses the settlement offer. The suit goes to final judgment and plaintiff is indeed awarded the full $100,000 sought. Mrs. Crispy is understandably miffed at Faithless Fidelity.

Once again, begin your analysis by adopting the simplifying assumption that the parties are risk-neutral, basing their assessments of alternative prospects solely on the mathematical expected value.

**QUESTIONS**

1. Show that the settlement offer was a lesser cost than the expected value of the cost to Faithless Fidelity from going to final judgment.

2. Can Faithless Fidelity’s superficially uneconomic behavior be explained as a Chicken game against Mrs. Crispy? [Hint: What is Mrs. Crispy’s exposure and what can she do about it if Faithless Fidelity refuses to settle?]

3. Even in the light of 20/20 hindsight, it can be argued that Faithless Fidelity’s decision not to settle may have been in its long-term best interests. What is this argument? Does this argument suggest a systematic bargaining advantage to certain categories of litigants? [Hint: Is this a single-shot or an iterated game for Faithless Fidelity? For Mrs. Crispy? For the plaintiff in the original case?]

4. Suppose that the assumption of risk neutrality on the part of the parties were dropped. Could you make any kind of plausible speculation about the relative degrees of risk preference that the insurance company and Mrs. Crispy might have? Would this affect the results in any of the preceding questions?

5. What generic sorts of problems are involved in working out a settlement when at least some degree of joint liability may be involved? See “Settlement Devices with Joint Tort-Feasors,” 25 U.Fla.L.Rev. 762 (1973); “The Mary Carter Agreement—Solving Problems of Collusive Settlements

6. Suppose Crispy sues Faithless Fidelity for having negligently failed to settle. Here is a proposed rule:

* * * It may not be unreasonable for an insured who purchases a policy with limits to believe that a sum of money equal to the limits is available and will be used so as to avoid liability on his part with regard to any covered accident. In view of such an expectation an insurer should not be permitted to further its own interests by rejecting opportunities to settle within the policy limits unless it is also willing to absorb losses which may result from its failure to settle.

7. Crisci v. Security Insurance Co., 426 P.2d 173 (1967). Is this a desirable rule from the standpoint of insurance purchasers? Does the proposed rule change a potential insurance purchaser’s ability to buy a “limited” coverage? Can you suggest a better rule for potential insurance purchasers? [Hint: Does this case in effect give the insured a type of unlimited coverage? Would any of them not want that?]

3. EX ANTE ANALYSIS OF PUNISHMENTS

Social scientists tend to be stubbornly insistent that the behaviorally relevant aspect of rewards and punishments is how they are viewed “ex ante.” After all, a decision-maker is influenced only by subjective expectations of the good or bad consequences of an act; once the act has been done, its consequences may depart radically from the initial expectations. In more familiar legal language, one would discuss the same things as being evaluated “prospectively.” This beforehand assessment is, of course, contrasted with retrospective or “ex post” analysis. The point of view that one adopts in analyzing awards and punishments can affect the results powerfully, especially when the events in question have small probabilities or probabilities that vary widely in a relative sense. The case immediately below provides a dramatic example of how the distinction between ex ante and ex post analysis is relevant in a criminal law context. However, as we shall see, the same principle is relevant for a wide variety of civil and private law examples.

Read Lynch below from the stance of an Assistant Attorney General given the opportunity for reargument of the case on behalf of the State of California. Reflect on how a sensitivity to the probabilistic aspects of punishments might suggest useful rebuttals to the court’s own language in the opinion.

IN RE JOHN LYNCH

503 P.2d 921 (Supreme Ct. of Cal., In Bank, 1972)

MOSK, Justice.

One who commits an act of indecent exposure in California is guilty of a simple misdemeanor and can be punished by no more than a brief jail sentence or a small fine. If he commits the identical act a second time, however, the law declares him guilty of a felony and inflicts on him a punishment of imprisonment in the state prison for the indeterminate period of one year to life. We adjudicate here the question whether the aggravated penalty for second-offense indecent exposure provided by Penal Code Sec. 314 violates the prohibition of the California Constitution against cruel or unusual punishments. We conclude that the penalty offends the Constitution in the respect charged, and petitioner is therefore entitled to relief.

The issue is presented by John Lynch, a state prison inmate. In 1958 he was convicted of misdemeanor indecent exposure in violation of Penal Code Sec. 314. For this offense he spent two years on probation. In 1967 he was again convicted of indecent exposure. The court ruled he was not a mentally disordered sex offender, denied probation, and sentenced him to prison for the indeterminate term provided by Sec. 314 in the case of a second offense.

* * *
Petitioner urges that a life sentence for indecent exposure is cruel or unusual punishment under the California Constitution because it is grossly disproportionate to the offense. No California court has yet held a statutory penalty unconstitutional on the ground it is disproportionate to the crime committed. The rule has been recognized, however, in several opinions considering the constitutionality of the death penalty. Thus in In re Finley (1905) 1 Cal.App. 198, 202 the court stated that a punishment may be denounced as unusual in the constitutional sense only when it “is out of all proportion to the offense; and is beyond question an extraordinary penalty for a crime of ordinary gravity, committed under ordinary circumstances.”

A similar rule has evolved at the federal level in the interpretation of the cruel and unusual punishment clause of the Eighth Amendment. In O’Neil v. Vermont (1892) 144 U.S. 323, the defendant was convicted on multiple counts of unauthorized sale of liquor and sentenced to a fine of over $6,000 or, if he could not pay the fine, to hard labor for more than 54 years. A majority of the United States Supreme Court declined on federal-state grounds to consider whether this sentence constituted cruel and unusual punishment. Dissenting, Justice Field would have held the sentence to be “one which, in its severity, considering the offenses of which [the defendant] was convicted, may justly be termed both ‘unusual and cruel.’” He recognized that the cruel and unusual punishment clause was traditionally thought to prohibit physically torturous methods of punishment such as the rack and the screw, but explained: “The inhibition is directed, not only against punishments of the character mentioned, but against all punishments which by their excessive length or severity are greatly disproportioned to the offenses charged. The whole inhibition is against that which is excessive either in the bail required, or fine imposed, or punishment inflicted.” Less than two decades later Justice Field’s view became law in the landmark case of Weems v. United States (1910), 217 U.S. 349.

* * * The principle was recently reaffirmed in Furman v. Georgia (1972), the United States Supreme Court decision holding the death penalty unconstitutional as applied. In the course of his separate opinion in support of the majority, Justice Brennan gave as his view that “Although the determination that a severe punishment is excessive may be grounded in a judgment that it is disproportionate to the crime, the more significant basis is that the punishment serves no penal purpose more effectively than a less severe punishment.” In turn, Justice Marshall recognized that “a penalty may be cruel and unusual because it is excessive and serves no valid legislative purpose. The decisions [of the United States Supreme Court] are replete with assertions that one of the primary functions of the cruel and unusual punishments clause is to prevent excessive or unnecessary penalties [citations]; these punishments are unconstitutional even though popular sentiment may favor them.”

Finally, the highest courts of our sister states have repeatedly invoked the rule of proportionality in applying their equivalents of our cruel or unusual punishment clause.

* * *

Whether a particular punishment is disproportionate to the offense is, of course, a question of degree. The choice of fitting and proper penalties is not an exact science, but a legislative skill involving an appraisal of the evils to be corrected, the weighing of practical alternatives, consideration of relevant policy factors, and responsiveness to the public will: in appropriate cases, some leeway for experimentation may also be permissible. The judiciary, accordingly, should not interfere in this process unless a statute prescribes a penalty “out of all proportion to the offense” (Robinson v. California (1962), 370 U.S. 660 (concurring opinion of Douglas, J); In re Finley (1905) supra, 81 P. 1041, 1042), i.e., so severe in relation to the crime as to violate the prohibition against cruel or unusual punishment.

We conclude that in California a punishment may violate Article I, Sec. 6, of the Constitution if, although not cruel or unusual in its methods, it is so disproportionate to the crime for which it is inflicted that it shocks the conscience and offends fundamental notions of human dignity. * * * To aid in administering this rule, we point to certain techniques used in the decisions discussed herein.

First, a number of courts have examined the nature of the offense and/or the offender, with particular regard to the degree of danger both present to society. Thus in Anderson we spoke in this connection of
excessive punishment for “ordinary offenses”, and in *Finley* the court referred to an extraordinary penalty for “a crime of ordinary gravity committed under ordinary circumstances” (1 Cal.App. at p. 202, 81 P. at 1042).

The second technique used by the courts is to compare the challenged penalty with the punishments prescribed in the *same jurisdiction for different offenses* which, by the same test, must be deemed more serious. The underlying but unstated assumption appears to be that although isolated excessive penalties may occasionally be enacted, e.g., through “honest zeal” generated in response to transitory public emotion, the Legislature may be depended upon to act with due and deliberate regard for constitutional restraints in prescribing the vast majority of punishments set forth in our statutes. The latter may therefore be deemed illustrative of constitutionally permissible degrees of severity; and if among them are found more serious crimes punished less severely than the offense in question, the challenged penalty is to that extent suspect.

Closely related to the foregoing is the third technique used in this inquiry, i.e., a comparison of the challenged penalty with the punishments prescribed *for the same offense in other jurisdictions* having an identical or similar constitutional provision. Here the assumption is that the vast majority of those jurisdictions will have prescribed punishments for this offense that are within the constitutional limit of severity; and if the challenged penalty is found to exceed the punishments decreed for the offense in a significant number of those jurisdictions, the disparity is a further measure of its excessiveness. At common law indecent exposure was deemed to be no more than a public nuisance, and was punished by a misdemeanor. The penalties were generally a fine or a brief jail sentence. * * * In California a similar pattern prevailed until the present penalty was added in 1952.

The low-key approach of the common law is also that adopted by modern psychiatric science. Clinical studies “support and confirm the traditional legal provisions which have treated this behavior as a social nuisance, as disorderly conduct rather than an offense causing personal injury.” (Gigeroff, Mohr, and Turner, Sex Offenders on Probation: The Exhibitionist (1968) 32 Fed.Prob. (No. 3) 17, 21.)

Turning to the typical offender, we find a similar pattern of nonviolence. “The vast majority of exhibitionists are relatively harmless offenders; mostly they are public nuisances and sources of embarrassments” (Report of Karl M. Bowman, Medical Superintendent of the Langley Porter Clinic, in 2 Assem.J. (1951 Reg.Sess.) p. 2847). They are characterized as “generally passive, inoffensive people” with low self-esteem.

Finally, although indecent exposure is not a “victimless” crime, any harm it may cause appears to be minimal at most. As noted above, the nonviolence of the conduct ensures there is no danger of physical injury to the person who witnesses the exposure. Nor is there any convincing evidence that the person is likely to suffer either long-term or significant psychological damage. Indeed, the statute itself defines the offense as exposure in public or in any place where there are persons who may merely be “offended or annoyed” thereby. Such an “annoyance” is not a sufficiently grave danger to society to warrant the heavy punishment of a life-maximum sentence.

These considerations make a persuasive case for a finding of unconstitutional disproportionality between the offense and the aggravated penalty prescribed by Sec. 314. The case is further strengthened by a comparison of this penalty with the punishments for other crimes in California which are undeniably of far greater seriousness. For example, is it rational to believe that second-offense indecent exposure is a more dangerous crime than the unlawful *killing* of a human being without malice but in the heat of passion? Yet the punishment for manslaughter (up to 15 years) is far less than the life maximum inflicted by Sec. 314. The same is true for such other violent crimes against the person as assault with intent to commit murder (1-14 years), kidnapping (1-25 years), mayhem (up to 14 years), assault with intent to commit mayhem or robbery (1-20 years), assault with caustic chemicals, with intent to injure or disfigure (1-14 years), and assault on a peace officer or fireman engaged in the performance of his duties (up to 2 years, or up to 1 year in jail).

***
We recognize, of course, that an important additional element must be taken into account: Sec. 314 prescribes a life-maximum sentence for indecent exposure only when the offender has previously been convicted of the same crime or of lewd and lascivious acts upon a child. We further recognize that the potential for recidivism is here very real: “exhibitionists are more likely to repeat their offense than other kinds of sex offenders.” (Gigeroff, at 21; accord, Bowman, at 2847.) But this likelihood does not result in a pro tanto repeal of the cruel or unusual punishment clause. Petitioner does not challenge—or do we consider—the validity of our general habitual criminal law or of any other recidivist statute. He is entitled, however, to question whether in the particular context of indecent exposure the phenomenon of recidivism constitutionally justifies the greatly enhanced punishment of Sec. 314. We hold that it does not.

* * *

Finally, we may profitably compare Sec. 314 with other California statutes which prescribe enhanced punishment for recidivism. First, however, we pause to note that in all but three of the above-listed dangerous crimes involving personal violence, sexual assaults, or harm to children, there is no statutory provision increasing the penalty for recidivism. In other words, a man may repeatedly commit manslaughter or mayhem, assault with intent to commit rape or sodomy, child-beating or felony drunk driving, and still be subject each time to a lighter penalty than one who twice exposes his private parts.

Second, of all the statutes which increase the punishment in the case of a second offense, only Sec. 314 and one other compel the enormous single leap from an ordinary misdemeanor to a life-maximum felony. In each of the remaining statutes there is a reasonable relationship between the punishments for the first and subsequent offenses. If the crime is of a minor nature, its enhanced penalty remains proportionately light; if the crime is more serious, its original penalty was proportionately heavy. The theory in each instance is that whatever the response appropriate to the factor of recidivism, the judgment of the Legislature as to the gravity of the act itself should remain relatively constant.

* * *

The last technique to be employed—a comparison of the challenged penalty with the punishments prescribed for the same offense in other jurisdictions—is no less revealing. A study of the indecent exposure statutes of each of our sister states and the District of Columbia reveals only two other states—Michigan and Oklahoma—which permit life-maximum sentences for second-offense exhibitionists. By contrast, 34 states and the District of Columbia do not enhance the punishment for any degree of recidivism; in each, indecent exposure remains a misdemeanor at all times. * * * Thus it is the virtually unanimous judgment of our sister states that indecent exposure, no matter how often it may recur, can be adequately and appropriately controlled by the imposition of a short jail sentence and/or a small fine. In this setting the California penalty of a life-maximum sentence in state prison strikes a discordant note indeed.

* * *

Viewing the total disparity between the life-maximum sentence currently inflicted by Sec. 314 for second-offense indecent exposure and the far lighter penalties in force in California and elsewhere, we conclude with Justice McKenna in Weems that “this contrast shows more than different exercises of legislative judgment. It exhibits a difference between unrestrained power and that which is exercised under the spirit of constitutional limitations formed to establish justice.” (217 U.S. at p. 381.)

* * *

QUESTIONS

1. In this context, what is the difference between the prospect of punishment and the actual punishment provided under the statute? Which of these is more relevant?

2. What would your speculation be about the proportion of convictions to offenses for, respectively, the crimes of indecent exposure and aggravated assault? What implications might a substantial
difference in these proportions have for setting statutory penalties? Does it make a difference whether the objective is deterrence or “punishment”? 

3. If sentences were being adjusted for a “probability of getting away with it” factor, how might this affect the treatment of multiple offenders? (Hint: Suppose one were worried about “over-punishing” a person who committed the crime only once?)

4. Suppose that you were defending the position of the State in this case. What portions of the court’s opinion represent views that might be turned to your advantage?

5. What possible generalization is there of the Lynch case to other areas of law? For instance, do similar considerations explain the trebling of damages in antitrust law? If not, what does? [The anti-trust damages issue is specifically dealt with by Elzinga and W. Breit, The Antitrust Penalties: A Study in Law and Economics, (New Haven, 1976) Ch. 7.] For a general discussion and bibliographical footnotes that provide a guide to applications in many different areas, see A. Polinsky and S. Shavell, “The Optimal Tradeoff Between the Probability and Magnitude of Fines,” 69 Amer.Econ.Rev. 880 (1979).

6. This case deals, of course, with a public law issue. But, what applications are there to private law questions? [Hint: What kinds of damages in tort and contract?] See also, the Richards case infra.

**RICHARDS v. ALLSTATE INSURANCE CO.**

693 F.2d 502. (5th C., 1982)

Before CLARK, Chief Judge, RUBIN and WILLIAMS, Circuit Judges.

Larry Richards filed this diversity suit against Allstate Insurance Company, seeking actual damages for injuries he sustained in a motorcycle accident and punitive damages for Allstate’s bad-faith refusal to honor his claim for those injuries. The jury returned a verdict of $2,500 in compensatory damages and $750,000 in punitive damages. The district judge subsequently ordered a remittitur of $375,000, which Richards accepted, albeit reluctantly. Allstate contends here that the award of punitive damages was without basis in Mississippi law. * * * We affirm.

Mississippi law requires all automobile insurance policies to provide at least $10,000 coverage for personal injuries sustained in accidents caused by motorists who have no insurance. The focal point of this litigation is Exclusion 2 of the policy issued by Allstate to Richards, which attempted to limit such coverage. This provision, which was for many years part of Allstate’s standard Mississippi automobile policy, totally excluded from uninsured motorist coverage injuries suffered by an insured while occupying a vehicle owned by him but not specified as insured under his policy.

Allstate retained Exclusion 2 in its standard policy until 1981 despite the fact that in 1973 the Mississippi Supreme Court struck down a similar exclusion in another company’s policy in Lowery v. State Farm Mutual Automobile Insurance Co., 285 So.2d 767 (Miss.1973). Allstate officials admittedly were aware of the Lowery decision and its fatal effect on Exclusion 2. They did not, however, delete the exclusion from their standard policy. Nor did they make any effort to inform their insureds or their sales agents of the effect of Lowery. Instead, mid-level claims personnel were instructed to honor claims that otherwise would have been denied in reliance on Exclusion 2. Unfortunately, as the result of a series of mistakes, this procedure failed in Mr. Richards’ case and his claim was denied.

The incident that gave rise to Richards’ claim occurred in August of 1977. Early in that month he was injured in a motorcycle accident caused by the negligence of an uninsured motorist. After Richards contacted an Allstate sales agent in an unsuccessful attempt to recover for the damages to his motorcycle, he brought his accident to the attention of Pete Quave, who worked for Richards’ retained counsel. Quave was an ex-employee of Allstate and had handled a similar claim while working there. He contacted Allstate on Richards’ behalf and spoke with Marinella Davis, one of Allstate’s telephone claims handlers, describing the accident and the terms of Richards’ policy, including the fact that the motorcycle was not an insured vehicle. After discussing the case with her supervisor, Margaret Kessler, Davis informed
Quave that Allstate planned to deny coverage in reliance on Exclusion 2. At this point Quave brought the Lowery decision to Davis’ attention and urged her to reconsider. He then talked to Kessler, again emphasizing the effect of Lowery on Richards’ claim. Quave also asked to speak to Cecil Snodgrass, who he knew was aware of Lowery. Because Snodgrass was unavailable, Quave asked Kessler to discuss the case with him. The message somehow got garbled in transmission, and Snodgrass concurred in the denial. After receipt of Allstate’s letter of denial, Richards filed suit.

Allstate contends first that this is not a proper case for an award of punitive damages. The leading Mississippi case on punitive damages in an insurance claim setting is Standard Life Ins. Co. of Indiana v. Veal, 354 So.2d 239 (Miss.1978). There the court recognized that punitive damages are not to be awarded unless the breach of the insurance contract is attended by an intentional wrong or such gross negligence as to amount to an independent tort. But the court allowed punitive damages, finding that the company had denied a legitimate claim for a reason clearly contrary to the express provisions of the policy. The court equated gross negligence with a refusal to pay not based on a legitimate or arguable reason.

This circuit has recognized Veal’s two-part “legitimate or arguable reason” rule to mean that the absence of (1) a justifiable reason, or (2) an arguable basis under Mississippi law, for refusal to pay a valid claim creates a jury issue on punitive damages. Black v. Fidelity & Guaranty Insurance Underwriters, 582 F.2d 984, 990-91 (5th Cir.1978). In this case we conclude that under Veal submission of the punitive damages issue to the jury was compelled by the proof.

At trial Richards alleged that three separate but related actions of Allstate justified a punitive award: (1) retention of the invalid Exclusion 2 in the standard Allstate policy; (2) failure to provide an adequate procedure to prevent erroneous denials; and (3) denial of Richards’ claim itself. The jury was properly instructed on all three. We need go no further than the first allegation to uphold their verdict under the facts in this record.

As we have noted, the Lowery decision in 1973 invalidated Exclusion 2. Although aware of Lowery, Allstate decided not to delete the exclusion from its standard policy. Instead, a procedure was established within the claims department that Allstate thought would be adequate to prevent erroneous denial of claims that were valid but for Exclusion 2. The court instructed the jury that if they found an adequate procedure was established they could not assess punitive damages unless Richards proved that reliance on the procedure amounted to gross negligence.

Allstate urges that its procedure was successful in every case except Richards’. Even accepting this as true, this procedure provided no remedy for those policyholders who read Exclusion 2, assumed their injuries were not covered, and failed to file claims. Failure to delete Exclusion 2 in effect represented a corporate decision by Allstate not to inform its policyholders of undisclosed coverage required by Mississippi law. Whether this conduct was sufficiently culpable to justify an award of punitive damages was a question for the jury. The jury obviously believed that it was.

Allstate contends that retaining Exclusion 2 was justified in spite of Lowery and makes two arguments in support of this claim. Both are wholly without merit. First, Allstate asserts that Exclusion 2 was valid insofar as coverage in excess of that required by law was concerned. But this does not explain the fact that Exclusion 2 eliminated all coverage, not just that above the statutory minimum. In this case, for example, Richards’ claim was within the statutory minimum, but it was denied based on the language of Exclusion 2. Second, Allstate contends that Exclusion 2 was in effect erased by provision 5 of the policy, which stated: “Such terms of this policy as are in conflict with statutes of the state in which this policy is issued are hereby amended to conform.” In Allstate’s view, this provision corrected any deficiency in the policy. Therefore, it contends that deletion of Exclusion 2 was not required. This argument strains credibility. If it were accepted, Allstate could include in its policies any sort of invalid exclusion and then rely on change provision 5 when challenged. This would mean that policyholders, not insurance companies, would bear the burden of keeping abreast of changes in the law. Clearly this is not the intent of Mississippi’s insurance code.
Exclusion 2 as written and as retained in Allstate’s policies from 1973 until well after 1977 was invalid under Lowery. Under the court’s instructions the jury necessarily found that Allstate’s failure to remove Exclusion 2 from its standard automobile policy until after this suit was filed was grossly negligent. This finding is supported by the proof.

Allstate also claims that the punitive award is excessive in light of the actual injuries suffered by Richards. Although the punitive award here, even after remitted to $375,000, may be characterized as unusually large where actual damages are only $2,500, this disproportionality does not mandate reversal.

Punitive damages are awarded not as compensation to the plaintiff but as “punishment for the wrongdoing of the defendant and as an example so that others may be deterred from the commission of similar offenses * * *.” Snowden v. Osborne, 269 So.2d 858, 860 (Miss.1972). Given the jury’s determination that Allstate’s actions were grossly negligent and that Allstate’s financial statement disclosed its latest annual addition to surplus was $390,046,821 and its cumulated policyholders’ surplus is $2,371,135,879, we cannot say that the award of sixteen ten-thousands of that surplus was greater than necessary to serve these twin rationales. See Standard Life Ins. Co. of Indiana v. Veal, 354 So.2d at 249. An additional consideration in determining whether a punitive damage is excessive is whether the plaintiff has performed a public service by bringing the wrongdoer to account. See Fowler Butane Gas Co. v. Varner, 244 Miss. 130, 151, 141 So.2d 226, 233 (1962). Here Richards’ suit has benefitted all Allstate policyholders by directly causing deletion of Exclusion 2 from the standard Mississippi policy. These factors, together with the fact that Mississippi has never articulated any requirement that punitive damages be rationally related to actual damages, convince us that the award was proper.

* * *

For the foregoing reasons, the decision of the district court is Affirmed.

QUESTIONS

1. Did plaintiff Richards receive his bargained-for expectation interest in his insurance contract with Allstate? Is there any sense in which Allstate “broke its promise” to Richards? Is this a contract claim? If not, what?

2. Why did Allstate keep the unenforceable provision in the contract if its contention is true that Richards’ case is a “slip-up” in a general policy of actually honoring claims that would have been barred by the term in question?

3. Suppose that all claims similar to that of Richards had been honored. Would anyone have a cause of action against Allstate? If so, for what damages?

4. The judge apparently thought that $750,000 in punitive damages was excessive on a claim of only $2,500 in compensatory damages. What factors should determine the amount of punitive damages, if any?

5. Most states now have landlord-tenant statutes that restrict the terms that may be written into residential leases. Would it surprise you to find that terms contrary to the statutes are fairly widely used? Why?

6. “Resolved: Anyone who shall induce another to transact on terms known to be contrary to law or public policy shall be liable for punitive damages.” Comment.

7. What is the nexus between this case and the Lynch opinion above?

8.

CRAIG V. BOREN

429 U.S. 190 (1976)

Mr. Justice BRENNAN delivered the opinion of the Court.
The interaction of two sections of an Oklahoma statute prohibits the sale of “nonintoxicating” 3.2% beer to males under the age of 21 and to females under the age of 18. The question to be decided is whether such a gender-based differential constitutes a denial to males 18-20 years of age of the equal protection of the laws in violation of the Fourteenth Amendment.

* * *

* * * Clearly, the protection of public health and safety represents an important function of state and local governments. However, appellees’ statistics in our view cannot support the conclusion that the gender-based distinction closely serves to achieve that objective and therefore the distinction cannot under *Reed* withstand equal protection challenge.

The appellees introduced a variety of statistical surveys. First, an analysis of arrest statistics for 1973 demonstrated that 18-20-year-old male arrests for “driving under the influence” and “drunkenness” substantially exceeded female arrests for that same age period. Similarly, youths aged 17-21 were found to be overrepresented among those killed or injured in traffic accidents, with males again numerically exceeding females in this regard. Third, a random roadside survey in Oklahoma City revealed that young males were more inclined to drive and drink beer than were their female counterparts. Fourth, Federal Bureau of Investigation nationwide statistics exhibited a no-table increase in arrests for “driving under the influence.” Finally, statistical evidence gathered in other jurisdictions, particularly Minnesota and Michigan, was offered to corroborate Oklahoma’s experience by indicating the pervasiveness of youthful participation in motor vehicle accidents following the imbibing of alcohol. Conceding that “the case is not free from doubt,” 399 F.Supp., at 1314, the District Court nonetheless concluded that this statistical showing substantiated “a rational basis for the legislative judgment underlying the challenged classification.”

Even were this statistical evidence accepted as accurate, it nevertheless offers only a weak answer to the equal protection question presented here. The most focused and relevant of the statistical surveys, arrests of 18-20-year-olds for alcohol-related driving offenses, exemplifies the ultimate unpersuasiveness of this evidentiary record. Viewed in terms of the correlation between sex and the actual activity that Oklahoma seeks to regulate—driving while under the influence of alcohol—statistics broadly establish that .18% of females and 2% of males in that age group were arrested for that offense. While such a disparity is not trivial in a statistical sense, it hardly can form the basis for employment of a gender line as a classifying device. Certainly if maleness is to serve as a proxy for drinking and driving, a correlation of 2% must be considered an unduly tenuous “fit”. Indeed, prior cases have consistently rejected the use of sex as a decision making factor even though the statutes in question certainly rested on far more predictive empirical relationships than this.

Moreover, the statistics exhibit a variety of other shortcomings that seriously impugn their value to equal protection analysis. Setting aside the obvious methodological problems,* the surveys do not adequately justify the salient features of Oklahoma’s gender-based traffic-safety law. None purports to measure the use and dangerousness of 3.2% beer as opposed to alcohol generally, a detail that is of particular importance since, in light of its low alcohol level, Oklahoma apparently considers the 3.2% beverage to be “nonintoxicating.” Okla.Stat., Tit. 37, Sec. 163.1 (1958); see *State ex rel. Springer v. Bliss,* 199 Okla. 198, 185 P.2d 220 (1947). Moreover, many of the studies, while graphically documenting the unfortunate increase in driving while under the influence of alcohol, make no effort to relate their findings to age-sex differentials as involved here. Indeed, the only survey that explicitly centered its attention upon young drivers and their use of beer—albeit apparently not of the diluted 3.2% variety—reached results that hardly can be viewed as impressive in justifying either a gender or age classification.

* The very social stereotypes that find reflection in age-differential laws, see *Stanton v. Stanton,* 421 U.S. 7, 14-15 (1975), are likely substantially to distort the accuracy of these comparative statistics. Hence “reckless” young men who drink and drive are transformed into arrest statistics, whereas their female counterparts are chivalrously escorted home. See, e.g. W. Reckless & B. Kay, *The Female Offender* 4, 7, 13, 14-17 (Report to Presidential Commission on Law Enforcement and Administration of Justice, 1967).
There is no reason to belabor this line of analysis. It is unrealistic to expect either members of the judiciary or state officials to be well versed in the rigors of experimental or statistical technique. But this merely illustrates that proving broad sociological propositions by statistics is a dubious business, and one that inevitably is in tension with the normative philosophy that underlies the Equal Protection Clause. Suffice to say that the showing offered by the appellees does not satisfy us that sex represents a legitimate, accurate proxy for the regulation of drinking and driving. In fact, when it is further recognized that Oklahoma’s statute prohibits only the selling of 3.2% beer to young males and not their drinking the beverage once acquired (even after purchase by their 18-20-year-old female companions), the relationship between gender and traffic safety becomes far too tenuous to satisfy Reed’s requirement that the gender-based difference be substantially related to achievement of the statutory objective.

We hold therefore, that under Reed, Oklahoma’s 3.2% beer statute invidiously discriminates against males 18-20 years of age.

* * *

Mr. Justice REHNQUIST, dissenting.

The Court’s disposition of this case is objectionable on two grounds. First is its conclusion that men challenging a gender-based statute which treats them less favorably than women may invoke a more stringent standard of judicial review than pertains to most other types of classifications. Second is the Court’s enunciation of this standard, without citation to any source, as being that “classifications by gender must serve important governmental objectives and must be substantially related to achievement of those objectives.” Ante, at 197 (emphasis added). * * * I think the Oklahoma statute challenged here need pass only the “rational basis” equal protection analysis expounded in cases such as McGowan v. Maryland, 366 U.S. 420 (1961), and Williamson v. Lee Optical Co., 348 U.S. 483 (1955), and I believe that it is constitutional under that analysis.

* * *

Quite apart from * * * alleged methodological deficiencies in the statistical evidence, the Court appears to hold that that evidence, on its face, fails to support the distinction drawn in the statute. The Court notes that only 2% of males (as against .18% of females) in the age group were arrested for drunk driving, and that this very low figure establishes “an unduly tenuous ‘fit’” between maleness and drunk driving in the 18-20-year-old group. On this point the court misconceives the nature of the equal protection inquiry. The rationality of a statutory classification for equal protection purposes does not depend upon the statistical “fit” between the class and the trait within the included class than in the excluded class to justify different treatment. Therefore the present equal protection challenge to this gender-based discrimination poses only the question whether the incidence of drunk driving among young men is sufficiently greater than among young women to justify differential treatment. Notwithstanding the Court’s critique of the statistical evidence, that evidence suggests clear differences between the drinking and driving habits of young men and women. Those differences are grounds enough for the State reasonably to conclude that young males pose by far the greater drunk-driving hazard, both in terms of sheer numbers and in terms of hazard on a per-driver basis. The gender-based difference in treatment in this case is therefore not irrational.

The Court’s argument that a 2% correlation between maleness and drunk driving is constitutionally insufficient therefore does not pose an equal protection issue concerning discrimination between males and females. The clearest demonstration of this is the fact that the precise argument made by the Court would be equally applicable to a flat bar on such purchases by anyone, male or female, in the 18-20 age group; in fact it would apply a fortiori in that cases given the even more “tenuous ‘fit’” between drunk-driving arrests and femaleness. The statistics indicated that about 1% of the age group population as a whole is arrested. What the Court’s argument is relevant to is not equal protection, but due process—whether there are enough persons in the category who drive while drunk to justify a bar against purchases by all members of the group. Cast in those terms, the argument carries little weight, in light of our decisions indicating that such questions call for a balance of the State’s interest against the harm
resulting from any overinclusiveness or underinclusiveness. Vlandis v. Kline, 412 U.S. 441, 448-452 (1973). The personal interest harmed here is very minor—the present legislation implicates only the right to purchase 3.2% beer, certainly a far cry from the important personal interests which have on occasion supported this Court’s invalidation of statutes on similar reasoning. And the state interest involved is significant—the prevention of injury and death on the highways. This is not a case where the classification can only be justified on grounds of administrative convenience. There being no apparent way to single out persons likely to drink and drive, it seems plain that the legislature was faced here with the not atypical legislative problem of legislating in terms of broad categories with regard to the purchase and consumption of alcohol. I trust, especially in light of the Twenty-first Amendment, that there would be no due process violation if no one in this age group were allowed to purchase 3.2% beer. Since males drink and drive at a higher rate than the age group as a whole, I fail to see how a statutory bar with regard only to them can create any due process problem.

The Oklahoma Legislature could have believed that 18-20-year-old males drive substantially more, and tend more often to be intoxicated than their female counterparts; that they prefer beer and admit to drinking and driving at a higher rate than females; and that they suffer traffic injuries out of proportion to the part they make up of the population. Under the appropriate rational-basis test for equal protection, it is neither irrational nor arbitrary to bar them from making purchases of 3.2% beer, which purchases might in many cases be made by a young man who immediately returns to his vehicle with the beverage in his possession. The record does not give any good indication of the true proportion of males in the age group who drink and drive (except that it is no doubt greater than the 2% who are arrested, but whatever it may be I cannot see that the mere purchase right involved could conceivably raise a due process question. There being no violation of either equal protection or due process, the statute should accordingly be upheld.

QUESTIONS

1. Compare this case to In re Lynch, supra (punishment of multiple offender for indecent exposure).
2. The majority finds the different offense rate for the two sexes to be de minimis. Do you find that position persuasive?
3. Explore the hypothesis that this law is in the best interests even of the young males who are “discriminated against.” What are the costs and benefits to them? Are the benefits equal in every year? If not, does some sort of adjustment need to be made for the different time periods in which they accrue? (See Chapter III below.)
4. If it were to be concluded that the law is in the best interests even of young males, how do you explain the suit to overturn it?

4. THE HUNTING OF A TAKEOVER

A corporate executive generally cannot appropriate personally the full gains to the firm from good managerial decisions; only a part of the benefits will accrue to the manager in the form of higher salary, etc., while the remainder goes to stockholders. Normally, the cruel truth is that it is easier for managers, as for most other working people, to do a slipshod job than a very good job. Better management usually requires that the managers themselves shoulder certain personal costs in the form of overtime work, greater intensity of effort, confronting distasteful personal situations, risking embarrassment, etc. To the extent that the executive must bear internal costs that produce what are to him or her substantially external benefits, there exists an incentive to “shirk,” to fail to make some cost justified efforts that would improve the quality of managerial performance and increase net profitability. Losses—i.e., foregone profits—stemming from this conflict of interests between owners and their agents are frequently termed
“agency costs.” The inherent conflict of interest itself, and the incentive response thereto, is explored in greater detail below, e.g., in the “Real Estate Agent” hypothetical. At this point, we shall be concerned with a specific application to the corporate environment.

One response to the existence of suboptimal agency efforts is monitoring by the principal. In the case of corporations, however, the principals are the frequently very numerous shareholders who are themselves implicated in an easily recognizable Free Rider problem: monitoring is a costly activity whose potential benefits are non-excludible vis-a-vis the other stockholders. Moreover, the Free Rider problem is exacerbated by the necessity of any stockholder, once monitoring has detected a suboptimal performance, to persuade a majority of the other stockholders to authorize any disciplinary or remedial action.

One theory of capital markets holds that tender offers are an efficacious alternative method of monitoring the performance of corporate management. If management’s performance is sufficiently suboptimal, the attention of a prospective bidder will tend to be attracted. The outsider who buys up a majority of the shares of such a corporation can install new management, improve the performance of the corporation, and benefit from the subsequent increase in share prices. In principle, all (ownership) parties gain from the takeover. The bidder profits from the increase in value of his purchased majority interest, nontendering shareholders receive a similar price increase, and tendering shareholders enjoy a higher selling price for their stock than if the tender offer had not occurred. Furthermore, the threat of such a takeover bid acts as a spur to managerial efficiency even in firms where a takeover bid never occurs.

The target company’s managers frequently have a substantial personal interest in resisting such takeovers. Since managerial resistance is generally couched in terms of resisting an “inadequate offer” or “looking for a better offer,” it seems difficult to distinguish good faith efforts by management to protect the stockholders’ interests from mere defenses of well feathered nests by slothful birds. In any event, managerial resistance may facilitate an “auction” bidding market for the shares and a resultant higher price to the shareholders. Alternatively, it may defeat the takeover entirely. But this latter eventuality is not necessarily inconsistent with the stockholders’ interests because any determined effort to extract the highest possible price runs the risk of obstructing quick, short-run sales.

When managerial resistance leads to the defeat of a takeover bid, suits by disappointed shareholders against management are sometimes filed. Such suits are almost always unsuccessful because, absent some clear showing of bad faith, the common law “business judgment” rule normally insulates the discretion of managers from judicial review. Assessing the merits of a tender offer is therefore within management’s peculiar discretionary purview and, indeed, courts have held that directors have not only the right but the affirmative duty to resist tender offers deemed not to be in the best interests of shareholders.

NOTES AND QUESTIONS

In a law review article, Easterbrook and Fischel proffer the initially counterintuitive suggestion that the interests of stockholders would best be served by legal rules that restrict or prohibit entirely the ability of a target company’s management to engage in defensive tactics in response to a tender offer, even if the purpose of the resistance is to trigger a bidding contest and higher price. [Easterbrook and Fischel, 94 Harvard L.Rev. 1161 (1981).] Here is a key section of their argument:

* * * The value of any stock can be understood as the sum of two components: the price that will prevail in the market if there is no successful offer (multiplied by the probability that there will be none) and the price that will be paid in a future tender offer (multiplied by the likelihood that some offer will succeed). A shareholder’s welfare is maximized by a legal rule that enables the sum of these two components to reach its highest value. [At 1184.]

5. Can the above paragraph be translated into the equation \( V = (1-p)M + pT \)? What do the different symbols mean? Is \( V \) an “expected value”?

6. Elaborate on the quoted paragraph. What do you think is the rest of the authors’ basic argument? Are “ex post” and “ex ante” distinctions relevant?

7. If the authors’ basic argument above is accepted, does it then also follow that \( V \) is maximized by a strategy of no resistance?

8. On balance, does it seem in the best interests of stockholders as a class that courts impose an affirmative duty of resistance on management? If so, how should that duty be defined?


5. AGENT-PRINCIPAL CONFLICTS OF INTEREST

a. Using Agents To Sell Real Estate

Under the most commonly encountered arrangements, real estate agents are nominally—and legally—in the employ of the potential seller of the property. Their behavior is, of course, meant to be disciplined by the applicable laws of agency. Nevertheless, the complaint is frequently voiced that real estate agents tend in actual practice to act in some ways that are detrimental to the seller and beneficial to the buyer. For instance, an agent will hint to a potential buyer that the seller would really be willing to accept a bid lower than the asking price or will attempt to persuade the seller to accept a lower price than might be warranted. Assume that the agent in the following problem is hired under a simplified version of the standard real estate listing contract. The agent receives 6% of the sales price and bears all selling costs, such as advertising and showing. The agent shows the property to prospective purchasers, each of whom makes a bid that must be reported to the seller.

A considerable amount is known about the market for the property in question. With the simplifying assumption that bids can only be made in multiples of $100, the market information can be summarized in Exhibit 2.3.

### Exhibit 2.3 - Real Estate Market Information

<table>
<thead>
<tr>
<th>Current Bid of Higher Value</th>
<th>Probability From Higher Bid</th>
<th>Expected Value Of Add’l Showing</th>
</tr>
</thead>
<tbody>
<tr>
<td>500000.4760</td>
<td>1134</td>
<td></td>
</tr>
<tr>
<td>502000.4472</td>
<td>1123</td>
<td></td>
</tr>
<tr>
<td>503000.4312</td>
<td>1109</td>
<td></td>
</tr>
<tr>
<td>504000.3893</td>
<td>1083</td>
<td></td>
</tr>
<tr>
<td>505000.3640</td>
<td>1062</td>
<td></td>
</tr>
<tr>
<td>506000.3445</td>
<td>994</td>
<td></td>
</tr>
<tr>
<td>507000.3215</td>
<td>985</td>
<td></td>
</tr>
<tr>
<td>508000.3011</td>
<td>962</td>
<td></td>
</tr>
<tr>
<td>509000.2859</td>
<td>912</td>
<td></td>
</tr>
<tr>
<td>510000.2641</td>
<td>885</td>
<td></td>
</tr>
<tr>
<td>511000.2404</td>
<td>858</td>
<td></td>
</tr>
<tr>
<td>512000.2210</td>
<td>835</td>
<td></td>
</tr>
<tr>
<td>513000.1956</td>
<td>821</td>
<td></td>
</tr>
<tr>
<td>514000.1779</td>
<td>791</td>
<td></td>
</tr>
<tr>
<td>515000.1587</td>
<td>765</td>
<td></td>
</tr>
<tr>
<td>516000.1516</td>
<td>740</td>
<td></td>
</tr>
<tr>
<td>517000.1266</td>
<td>719</td>
<td></td>
</tr>
<tr>
<td>518000.1184</td>
<td>700</td>
<td></td>
</tr>
</tbody>
</table>
The exhibit reflects that fact that the “average” bid will be $50,000, but there will also be considerable dispersion of bids among individuals who will like the property either more or less than the average amount. Neither the seller nor the agent are, of course, concerned with the below-average bidders; the agent is being hired to, in effect, search out the above-average bidders. Bidders are distributed in the classical bell-shaped “Gaussian normal” curve centered around the $50,000 average. That is, the probability of a $50,000 bid is highest and the probabilities of other bids decline as variations from $50,000 increase. As higher and higher bids are found, it becomes more and more difficult to improve matters by searching more. This is true for two reasons: (1) bidders who will exceed the current bid become scarcer; and (2) once found, the probable amount by which a higher bidder exceeds the current bid becomes smaller and smaller.

It costs the real estate agent $10 worth of transportation expense, time, trouble, etc. to bring about each additional “showing.” Hence, this $10 is the relevant opportunity cost to the agent of searching for buyers. Once a bid is received, the agent’s own cost-benefit calculus will determine whether he desires that the owner accept the current bid or would prefer the chance to search for a still higher bid. In sum, the problem is one of deciding whether to settle for the “bird in hand” or to expend additional resources in searching for the fatter bird still lurking in the bush.

**QUESTIONS**

1. Assume that this is an “exclusive listing” with the agent. At approximately what bid level does it behoove the agent to persuade the owner to accept? Why? (If you cannot derive an exact numerical answer, attempt to describe the logic behind the decision process.)

2. If the same agent were selling his own house, what would he think of advice to sell at the price he advised in the previous question? At approximately what price would he be disposed to sell?

3. What difference would it make if the agent’s contract were not for an exclusive listing but involved co-listing with one or more other brokers?

4. How can the underlying cause of the conflict of interest between principal and agent be explained? Can you suggest a revision of the terms in the agent’s contract such that the conflict would be mitigated? Why would it be unsatisfactory for the owner to pay the agent for the number of showings accomplished? Why don’t the agent and seller agree on a sale price and a fixed fee, as opposed to the “sliding scale” of a commission based on an indeterminate price?
5. The facts assumed in this scenario are simplified in order to facilitate the analysis. In the real world, which of the factual assumptions would be unrealistic and to what extent? How much difference, if any, does this make to the results derived?

6. A seller files suit against his agent, claiming that the agent has shirked his duty to show and otherwise promote the sale of a house being marketed by the agent under an exclusive contract. Assuming that there is no difficulty in ascertaining any of the relevant facts, what is the standard of duty to which the agent should be held? How should “shirking” by the agent be understood?

7. If the law could lay some duty on agents to exert more efforts than the profit-maximizing amount for them in each transaction, would agents be worse off? If the answer depends on the nature of the duty, explain what extra duty, if any, might be preferred by agents? Would this be good or bad for principals?

8. Suppose that the “shirking” by an agent can neither be eliminated by monitoring nor shown to be a legally provable breach of contract. (What difficulties would be involved in either of those remedies?) The principal is then “stuck” with less-than-satisfactory performance. One possibility is that a sophisticated principal will have anticipated the problem and adjusted downward his return promise or fee in consideration of the lower level of expected performance. More naive principals may, however, suffer a lesser performance than they thought they bargained for. To what extent will extralegal market sanctions penalize shirking? Will the word “get around” that Shirke Realty represents a seller’s interests poorly, so that sellers will either not deal with them or require a “discounted” fee? Does the nature of the “product” make a difference? (Specifically, is real estate a likely or unlikely market environment for this reputational effect?) See Klein and Leffler, “The Role of Market Forces in Assuring Contractual Performance,” 89 J. of Polit.Econ. 615 (1981).

9. “Sharecropping” is an arrangement wherein a tenant farmer receives the use of farmland in exchange for a rental payment calibrated in terms of a share of the crop. Is this analogous to the real estate agent situation? What problems would you predict and how, if at all, can they be mitigated? See Cheung, “Private Property Rights and Sharecropping,” 76 J.Law and Econ. 1107 (1968); Hsiao, “The Theory of Share Tenancy Revisited,” 83 J.Pol.Econ. 1023 (1975).

b. Agent-Principal Conflicts in Legal Contracts

As it probably apparent, a large part of the difficulty in the real estate principal-agent relationship described in the last section is the practical inability to specify the obligations of the agent in clear, definitive terms. Attorney-client contingency fee contracts implicate similar problems because they also involve a situation in which the precise duties of the agent—in this case, the attorney—are difficult to specify at the time when the contract is formed. As the case below indicates, many states have a special rule that is applicable to breach of an attorney’s contract by the client. See whether you can explain the underlying rationale of this rule in a more compelling way that that expressed by the court in the case below.

**HEINZMAN v. FINE, FINE, LEGUM & FINE**

234 S.E.2d 282 (Supreme Court of Virginia, 1977)

[Heinzman, a plaintiff in a personal injury suit, engaged Legum’s law firm of Fine, Fine, Legum & Fine to represent him in a suit against defendant Mouser. Heinzman signed a written “power of attorney” promising Legum a fee of “1/3 of the gross amount recovered” by way of “compromise,” “trial,” or “appeal.” About three months later, Heinzman discharged Legum and retained attorney Aninios who, on the day before trial, settled the case for $10,000. Upon suit by Legum against Heinzman, the trial court found that Legum was entitled to $3,333.33 in satisfaction for breach of the contingency fee agreement. Heinzman appealed.]

The question remains whether, as Legum contends, an attorney employed under a contingent fee contract and discharged without just cause is entitled to recover from the proceeds of a later settlement the full amount of the contractual fee.

Courts which make no distinction between contracts for legal services and other contracts answer this question in the affirmative. The decisions rest upon various rationales: courts cannot convert bilateral contracts into unilateral contracts and relieve one of the parties of the burdens of his bargain, Friedman v. Mindlin, 155 N.Y.S. 295 (1915); a client “who has wrongfully broken a contract should not be permitted to reap advantage from his own wrong,” Dolph v. Speckart, 94 Or. 550, 186 P. 32, 35 (1920); when the client wrongfully frustrates performance, the attorney is entitled to the presumption that he would have performed, McElhinney v. Kline, 6 Mo.App. 94 (1878).

We reject such rationales, for we believe the premise upon which they rest is wrong. Contracts for legal services are not the same as other contracts.

“[I]t is a misconception to attempt to force an agreement between an attorney and his client into the conventional modes of commercial contracts. While such a contract may have similar attributes, the agreement is, essentially, in a classification peculiar to itself. Such an agreement is permeated with the paramount relationship of attorney and client which necessarily affects the rights and duties of each.” Krippner v. Matz, 287 N.W. 19, 24 (Minn., 1939).

Seldom does a client stand on an equal footing with an attorney in the bargaining process. Necessarily, the layman must rely upon the knowledge, experience, skill, and good faith of the professional. Only the attorney can make an informed judgment as to the merit of the client’s legal rights and obligations, the prospects of success or failure, and the value of the time and talent which he must invest in the undertaking. Once fairly negotiated, the contract creates a relationship unique in the law. The attorney-client relationship is founded upon trust and confidence, and when the foundation fails, the relationship may be, indeed should be, terminated.

Legum does not dispute the fact that a client has an absolute right to discharge his attorney. He contends, however, that when the discharge is without just cause, the act constitutes a breach of contract and that the measure of damages is the fee fixed in the contract. We agree with the courts that hold otherwise.

“That the client may at any time for any reason discharge his attorney is a firmly established rule * * *

* [T]he attorney may recover the reasonable value of the services which he has rendered but he cannot recover for damages for the breach of contract. The discharge of the attorney by his client does not constitute a breach of the contract because it is a term of such contract, implied from the peculiar relationship which the contract calls into existence, that the client may terminate the contract at any time with or without cause.” Martin v. Camp, 114 N.E. 46, 48 (N.Y., 1916).

Overruling its prior decisions, the court in Fracasce v. Brent, 494 P.2d 9, 12-13 (Cal., 1972), applied the same rule and explained the reasons for the rule:

“The right to discharge is of little value if the client must risk paying the full contract price for services not rendered upon a determination by a court that the discharge was without legal cause. The client may frequently be forced to choose between continuing the employment of an attorney in which he has lost faith, or risking the payment of double contingent fees equal to the greater portion of any account eventually recovered. * * * Unless a rule is adopted allowing an attorney as full compensation the reasonable value of services rendered to the time of discharge, clients will often feel required to continue in their service attorneys in whose integrity, judgment or capacity they have lost confidence.”

We share the conclusion that quantum meruit is the most functional and equitable measure of recovery. The law does not favor recoveries based upon conjecture. In awarding the contractual fee, the trial court presupposed that Legum, had he not been discharged, would have achieved a recovery exactly equivalent
to that achieved by Anninos. But this is entirely speculative. Legum might have effected a greater settlement or won a larger verdict; considering the vagaries of trial, he might have recovered nothing.

Having in mind the special nature of a contract for legal services, we hold that when, as here, an attorney employed under a contingent fee contract is discharged without just cause and the client employs another attorney who effects a recovery, the discharged attorney is entitled to a fee based upon *quantum meruit* for services rendered prior to discharge and, as security for such fee, to the lien granted by Code 54-70.

**QUESTIONS**

1. The court applies a special treatment to contracts for legal services as opposed to contracts for other goods and services. Might the rationale of this special treatment have anything to do with agency costs, shirking, and monitoring?

2. Are lawyers better off in states that, unlike Virginia, treat attorney-client relationships as “ordinary” contracts? Does your answer to this question depend in any way on whether the typical client knows about this special rule governing contracts with lawyers?

3. Even if the attorney is discharged without cause and in a manner that normally would constitute a breach by the client, the remedy awarded in Heinzman is “only” quantum meruit rather than the usual contractual measure of damages, which would be the lawyer’s lost profits. Under what circumstances could the recovery of damages based on quantum meruit actually amount to more than the lost profits awarded under an expectation theory of damages? What do you think a court would say if these circumstances were shown to hold under facts otherwise similar to Heinzman?

6. **“CURVE” MODELS: MARGINAL AND TOTAL EFFECTS**

In the last chapter, we used a matrix model to translate the Prisoner’s dilemma from words to a possibly more useful form for analytical purposes. Up to this point, we have been modeling the uncertainty problems in tabular form, as exemplified by the real estate agent situation just discussed. In the next section, we shall once again perform a translation from words into a more formal model, this time introducing the type of “curve” diagrammatic model that is almost ubiquitous in economic analysis. First, however, it will be useful to review certain fundamentals about the kinds of relationships that will appear in these diagrammatic models.

Not surprisingly, many of our models will be concerned with the effect that a particular variable has on something else. Think for a moment of the causal factor as an “input” and the thing that it affects as an “output.” Exhibit 2.4 depicts a hypothetical relationship between an “input” whose quantity is designated by X on the horizontal axis and the resultant “output” of Y as measured on the vertical axis. In the example at hand, Y may be interpreted as the recovery from a lawsuit, measured in dollars, while X is the input of doses of “effort”. Obviously, X and Y may be suitably redefined to fit other similar situations; e.g., X may be units of care and Y reductions in the expected harm from a dangerous activity. Exhibit 2.4 is drawn according to the usual assumptions about the shape of such a relationship: the total output climbs steeply at first, but then the ascent “flattens out” as more and more units of X are employed. The slowing of the rate of increase in the curve indicates that successive marginal units of X have less and less efficacy in producing additional Y. This phenomenon of the declining efficacy of additional “inputs” is the same as that evidenced in the utility function example earlier in this chapter. (See Exhibit 2.2 above.)

If we examine any level of input X along the horizontal axis, the rate at which a small adjustment in X will affect Y is the *slope* of the output curve directly above. (Readers who have a little knowledge of calculus will recognize the marginal curve as the “first derivative” of the total curve. What is its algebraic form in the example?) For instance, for adjustments in the neighborhood of $X_A$, the slope is taken at point A directly above it. At the greater level of input, $X_B$, the corresponding slope is less. In words, these slopes are the marginal effects of X. If Y were a benefit, then the marginal effect in question would more
precisely be termed the marginal benefit at the associated level of \( X \), and similar reinterpretations would exist if \( Y \) were utility level and \( X \) income, if \( Y \) were cost and \( X \) some production level, etc.

**Exhibit 2.4 – Marginal Effect As Slope of Total Curve**

![Exhibit 2.4 – Marginal Effect As Slope of Total Curve](image)

Economic analysis usually employs marginal curves similar to Exhibit 2.5, rather than total curves such as Exhibit 2.4. The marginal curve in effect plots the slopes of its associated total curve for all possible values of \( X \). (Usually the vertical scale is appropriately adjusted since the total levels are much larger magnitudes than the marginal ones.) As a rough approximation, one can think of constructing a series of rectangular columns, the area of each corresponding to the increase in total output for some arbitrarily defined “dose” of input \( X \).

**Exhibit 2.5 - Area Under A Marginal Curve**

![Exhibit 2.5 - Area Under A Marginal Curve](image)
In Exhibit 2.5, the assumed sizes of doses are five-unit increments. Thus, the effect of the first dose is the first rectangular area. By a similar process, the effects of the second and successive doses are representable as the area of the second rectangle, the third rectangle, etc. Then, it is easy to see that the total output for any level of input is the sum of the incremental dose effects up to the current level of X. In Exhibit 2.5, the rectangles overlap the actual marginal curve drawn on the same diagram. This is only true, however, because we have been showing discrete, five-unit adjustments. Technically, a marginal curve presupposes a series of infinitesimal adjustments of X. Accordingly, imagine the width of the doses in Exhibit 2.5 getting smaller and smaller, skinnier and skinnier. If we redrew Exhibit 2.5 for smaller and smaller dose sizes, it would start to look like we were exactly shading in the area under total the curve with vertical lines. The smaller the width of the dose rectangles, the less “overlapping the curve” problem. On reflection, you should be able to see that the sum of the dose increments would, in the limit when the doses approach infinitesimal size, become equal to the area under the marginal curve.

The importance of this is that we can, in effect, work backwards from marginal curves and depict the total effects as the area under the marginal curve. Thus, if Y is a measure of benefits, then the gross benefits from $X_b$ units of input is the shaded area under the marginal cost curve between the zero output and the relevant output $X_b$. Frequently, the areas under two marginal curves will be of interest; e.g., if the area under a marginal benefits curve is gross benefits and the area under a marginal cost curve is total costs, then the difference between those two areas will be equivalent to net benefits. In applying economic analysis to legal problems, you may find it convenient to “shade in” such areas to depict in visible form the magnitude of some gain or loss.

We will now apply this simple diagrammatic technique to the situation of a lawyer working under a contingency fee. You should recognize that this situation closely parallels the real estate problem just discussed. Be alert, however, for the ways in which the diagrammatic analysis arguably improves on the tabular model. For instance, one obvious methodological improvement is that it does away with the arbitrarily-sized discrete “dose” used in our previous discussions, now allowing continuous variation in the level of the activities being analyzed. See how many other possible advantages you can identify.

7. ATTORNEY-CLIENT RELATIONSHIPS

Vic Timm, a widower suing for damages in the alleged wrongful death of his wife Clara, has retained a well known personal injury litigator, Kate Advo, as his attorney. Timm and Advo have agreed to a contingency fee arrangement under which Advo is to receive 33 1/3% of any recovery ultimately awarded to Timm.

Since Advo is a highly successful trial attorney, she is frequently compelled to turn down cases because of constraints on her available time. Defendant’s insurance company is engaged in vigorous attempts to negotiate a settlement of the case. In turn, Advo has done sufficient preparation of Timm’s case to have formed an estimate of the chances of success and expected recoveries. These probable outcomes vary depending on the “effort” expended by her in carrying forward the litigation. For simplicity, assume that there is a “dose” of such effort definable in some standard form such as a mix of attorney time, paralegal time, etc.
QUESTIONS

1. Construct a diagram in which “Expected Value of Recovery” is measured on the vertical axis and “Doses of Effort” on the horizontal axis. Unlike some of our previous discussions, effort is adjustable in other than discrete doses; indeed, its quantity is continuously adjustable in gradations as small as desired. Draw a curve that you think might be a plausible representation of the relationship between effort level and the expected recovery. Explain briefly the reasons why the curve has the particular shape given to it. How, if at all, does it differ from Exhibit 2.4 above?

2. The curve that you just drew is in terms of total recovery. Draw a new curve that shows the relationship between effort levels and marginal recovery from the defendant. On the same diagram, add a curve that shows Advo’s expected returns as a function of her effort inputs. (Be sure that the latter curve is drawn in approximately the correct scale to the former one.) Finally, assume that the marginal cost to Advo of additional doses of effort is constant and add a curve that plausibly represents these marginal costs. What analogy is there between the diagram just created and Exhibit 1.3 in the previous chapter?

3. Using the diagram just completed, explain (a) what level of effort will maximize Advo’s personal returns from the case; (b) what effort level would be preferred by Timm; and (c) what level seems to you to be optimal or cost-justified in some objective sense.

4. What are the “error costs” if the optimal result of (3) above is not achieved? What area in your diagram represents these costs? Who bears the error costs?

5. Advo might well be expected to want to “hold back” or shirk somewhat on her effort inputs to the case. In the real world, in what kinds of behavior might Advo’s desire be manifested? What mechanisms can be used to mitigate the problem? In your view, are they effective?

6. What “should” be the level of effort legally or ethically demanded of an attorney in the above situation? Why? Who gains or loses from your proposed rule?

7. Consider the arrangements wherein client and attorney each bore the same share of costs that they enjoy of the benefits. For instance, the attorney’s contingency fee might be reduced to 25% of the judgment but the client bears 75% of the costs of litigating the claim. First, explain why in theory this arrangement would produce an optimal amount of effort by the attorney. What practical considerations make this arrangement less attractive than it appears in theory?

8. SHOULD SUBROGEES KEEP “EXCESS” RECOVERIES?

Although obviously not a common occurrence, it sometimes happens that there is a “windfall gain” to be distributed between an insurance company and its insured. In the case of *The Livingstone*, 130 F. 746 (2d Cir. 1904), the owner of a ship which sank in a collision with *The Livingstone* was paid $25,000 by his insurer under the terms of a fixed-recovery “valued policy.” The insurance company, as subrogee, promptly recovered $37,500 from the tortfeasor. Both insurer and insured claimed the extra $12,500. Victory to the insured.

In another leading case, however, an insurance company paid its insured in full for loss due to a building deliberately destroyed by an arsonist. The company subsequently recovered not only compensatory but also substantial punitive damages from the parties responsible. The insurance company was permitted to keep the excess recovery. *Urban Industries Inc. v. Thevis*, No. c 75-0342 L(A) (W.D.Ky., 1978).

The basic issue here is whether a subrogee should be permitted to recover a net gain over what he pays out to the subrogor. That is, who “owns” the excess recovery?
QUESTIONS

1. Putting aside for a moment any distributional implications (fat, wealthy insurance companies versus impecunious orphans and crippled personal injury victims), is there any economic argument as to what the rule should be here?

2. Now let’s confront the distributional issue. Taking “insureds” as one class and “insurers” as another, who benefits or loses from a rule requiring insurance companies to turn back any “excess” recovery? Does anyone benefit?

3. In connection with the Victoria Nike scenario in Chapter I above, it was suggested that potential holders of legal claims may not receive the full expected value of that claim when the rights to it are transferred in the form of a contingency fee interest. What relationship is there between this conclusion and the present question about how to allocate the rights to “windfall” gains under subrogation?

C. LEAST-COST RISK BEARING

1. WHAT IS “LEAST-COST” RISK BEARING?

Following the seminal work of Guido Calabresi, scholars in the Law and an increasing number of judicial opinions now commonly invoke the concept of the “least-cost risk-bearer.” The meaning of this term can be clarified by considering the series of questions supplied below. In each case, there will be a least cost risk-bearer, but the reason for the existence of such a party will differ.

By design, it is possible to derive a precise numerical answer to many of the questions. Those who are modestly competent with high school algebra can “solve” some of the questions by expressing the problems in the form of equations. Even if you cannot remember enough algebra to solve the equations for actual numbers, it will be profitable to try and formulate some of the problems conceptually in this manner. In this age of the pocket calculator, the truly innumerate can even just “plug in” some numbers and search for an approximate numerical answer.

If you are able to see how one would go about getting actual numerical answers, your effort will be repaid by enhanced understanding of the underlying concepts. Fumbling with algebraic formulations is not absolutely necessary, however, if one seriously reflects on the underlying problem in each situation. Minimally, you should try to answer the questions before reading the brief explanation of the three conceptually distinct sources of optimal risk-bearing. See whether you can identify the forces involved in each scenario. Of course, in the real world, these conceptually distinct factors might all be operating at the same time, even in ways that are mutually counteractive.

QUESTIONS

1. Assume two individuals who have the same square root utility functions as indicated in the text earlier in this section: \( U = 100M^{1/2} \). (Remember that raising any number to the \( \frac{1}{2} \) power means taking its square root, to the \( \frac{1}{3} \) power its cube root, etc.) Sehler is about to sell a Whizbang to Beyer. The agreed price is $100. After payment of the $100, Beyer would have $1000 and Sehler $500. Whizbangs are purchased in the hope that they will whiz, but there is a 50% chance that they will bang instead. When

---

they do bang, they cause $250 worth of damage to the holder, in this case Beyer. Whizbangs cannot be made any safer by exercising more care; they are sold in their natural state, just as they are found by trained swine rooting underneath poplar trees by the light of a full moon. The existing law on whizbang damages is caveat emptor: tough luck, Beyer. Nothing, however, prevents Beyer and Sehler from voluntarily entering into an agreement whereby the risk of whizbang malfunction is reallocated from Beyer to Sehler.

a. Beyer suggests that Sehler guarantee the product, accepting the liability to pay all damages if the whizbang bangs instead of whizzes. In exchange for this contractual modification, Beyer will agree to negotiate some additional consideration above and beyond the $100. First, see whether you can confirm that Beyer will pay at most $129.48 to get rid of the risk he now bears and that Sehler would need a minimum of $132.82 in order to take on the risk. What can you say about the likely outcome of the suggestion that a guarantee be sold? Why is one person more willing than the other to bear the risk? In what sense do we mean “willing”?

b. Now change Sehler’s utility function by placing a 1000 in front of the M on the right-hand side instead of a 100:

\[ U = 1000M^{\frac{1}{2}}. \]

Thus, at every level of money, changes in income have ten times as much impact on Sehler’s welfare as before. Re-answer the previous question in the light of this new fact situation. What explains the change in result, or lack thereof?

c. Assume that, after the payment of the $100 original price, both parties would have $500 left. Would Beyer be willing to pay as much as $135.72 for a guarantee? What are the possibilities for a change in the original risk allocations now?

d. Finally, make Sehler’s utility function a cube root, rather than a square root, function. Would $135.72 be enough to compensate Sehler for taking on the risk? Who will predictably wind up bearing the risk of a bang and why?

e. What is the nature of the “error cost” if the law does not allow the risk to be transferred, if necessary, to the least-cost risk bearer? (Use specific numbers from the above questions as illustrations.)

f. Assuming that the law does not prevent risk reallocation between the parties, does it make any difference where the law initially assigns the liability for the risk?

2. Suppose that Beyer is the predictable bearer of the risk, either because he is least-cost as compared to his possible vendors or because the law prevents reassignment of this particular risk. Nobody ever buys more than one whizbang in a lifetime. The good news is that whizbangs have become a lot safer than before; the probability of a bang is only 10% now. Remember that, by any of the utility functions attributed to him, Beyer places negative value on the dispersion of his risks.

a. If Beyer bears his risk “as is,” what is the expected value of his loss? What is the expected range or average dispersion of his loss?

b. Someone shows Beyer the attached Exhibit 2.6, which is a portion of a binomial distribution table from a mathematical handbook. They explain to him that the columns show the probability of loss assumed for a single “trial” of any experiment. The row headings indicate the number of losses that might conceivably be observed. Finally, the entries in the cells are the probabilities of observing a particular number of losses, given the single-loss probability at the column head. Beyer runs out and finds 19 other whizbang purchasers and induces them to join with him in a mutual loss-sharing arrangement, each bearing 1/20th of any losses occurring within the group. Does this reduce the expected value of Beyer’s risk? Why does he find this “risk spreading” or “pooling” arrangement attractive?

c. Suppose that Beyer found 100 fellow whizbang consumers with whom to share the risk. What do you think the analogue of Exhibit 2.6 would say about the probability distribution of different losses? What
The general principle is at work here as the number of possible risk-sharers goes up to 1,000,000, then 2,000,000 etc.?

d. Beyer is a consumer of whizbangs. Do you think that it would be costly for him to organize this risk-sharing scheme? Who is likely to be a cheaper organizer, consumers or sellers? Why?

e. How is all of the above changed if whizbangs cause different damages for different consumers? Is there a “Gresham’s Law” of insurance, i.e., that bad risks tend to drive out good risks\(^{20}\)? Explain. What relationship does this have to what is sometimes termed “adverse selection”? In a world of full information, would there be a tendency for risk pooling arrangements to consist of relatively homogeneous risks?

### Exhibit 2.6 – Binomial Distribution

| Probabilities of \( x \) Occurrences in \( n \) Trials |
|---|---|---|---|---|---|---|---|---|
| \( n \) | \( x \) | 0.05 | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 | 0.35 | 0.4 | 0.45 | 0.5 |
| 19 | 0 | 0.3774 | 0.1351 | 0.0456 | 0.0144 | 0.0042 | 0.0011 | 0.0003 | 0.0001 | 0.0000 | 0.0000 |
| 1 | 0.3774 | 0.2852 | 0.1529 | 0.0685 | 0.0268 | 0.0093 | 0.0029 | 0.0008 | 0.0002 | 0.0000 |
| 2 | 0.1787 | 0.2852 | 0.2428 | 0.1540 | 0.0803 | 0.0358 | 0.0138 | 0.0046 | 0.0013 | 0.0003 |
| 3 | 0.0533 | 0.1796 | 0.2428 | 0.2182 | 0.1517 | 0.0869 | 0.0422 | 0.0175 | 0.0062 | 0.0018 |
| 4 | 0.0112 | 0.0798 | 0.1714 | 0.2182 | 0.2023 | 0.1491 | 0.0909 | 0.0467 | 0.0203 | 0.0074 |
| 5 | 0.0018 | 0.0266 | 0.0907 | 0.1636 | 0.2023 | 0.1468 | 0.0933 | 0.0497 | 0.0222 |
| 6 | 0.0002 | 0.0069 | 0.0374 | 0.0955 | 0.1574 | 0.1451 | 0.0949 | 0.0518 |
| 7 | 0.0000 | 0.0014 | 0.0122 | 0.0443 | 0.0974 | 0.1525 | 0.1443 | 0.0961 |
| 8 | 0.0000 | 0.0002 | 0.0032 | 0.0166 | 0.0487 | 0.0981 | 0.1489 | 0.1771 |
| 9 | 0.0000 | 0.0000 | 0.0007 | 0.0051 | 0.0198 | 0.0514 | 0.1464 | 0.1762 |
| 10 | 0.0000 | 0.0000 | 0.0000 | 0.0013 | 0.0066 | 0.0220 | 0.0528 | 0.1762 |
| 11 | 0.0000 | 0.0000 | 0.0000 | 0.0003 | 0.0018 | 0.0077 | 0.0233 | 0.0532 |
| 12 | 0.0000 | 0.0000 | 0.0000 | 0.0004 | 0.0022 | 0.0083 | 0.0237 | 0.0529 |
| 13 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0005 | 0.0024 | 0.0085 | 0.0233 |
| 14 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0006 | 0.0024 | 0.0082 |
| 15 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0001 | 0.0005 | 0.0074 |
| 16 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0018 |
| 17 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0003 |
| 18 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| 19 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |

---

\(^{20}\) Gresham’s Law was originally coined (pun intended) in the days when more than one type of coinage, e.g., both silver-based and gold-based, was put into circulation. As the values of the underlying precious metals fluctuated, one type of money would tend to become relatively overvalued in respect to its nominal rate of conversion to the other type of money. Hence, the original form of the adage, that “Bad money drives out good” because the “good” money would be hoarded and the “bad” money used for exchange.
3. Assume that both Sehler and Beyer are risk neutral; they care about the expected value of the loss but not the dispersion. It is discovered that specially trained swine can discriminate between bang-prone whizbangs and a safer variety, but only if special precautions are taken at the time the whizbangs are first found in the wild. Specifically, the specially trained swine can filter out half of the bang-prone whizbangs, thus reducing the probability of loss from the original 10%.

a. The extra cost of using the trained pigs instead of ordinary swine comes to $X per whizbang found. What is the highest value that X can have in order for you to predict that Sehler will assume the liability for “safer” whizbangs? Explain your answer.

b. Suppose that the cost of using the trained pigs is greater than the $X computed above, but the law places a non-reassignable strict liability on the seller. Will the trained pigs be used?

c. The cost of the pigs is still greater than $X—$X+1, to be exact—but the seller is legally liable only when he fails to mitigate expected damages by using trained pigs to preselect the whizbangs that he offers. What will happen? Who, if anyone, will be better or worse off?

2. SOURCES OF LEAST COST RISK-BEARING

The three sets of questions above deal with three distinct ways of reducing risk. The first scenario considers party-to-party transfers that can reallocate who bears the consequences of a risk. Such transfers reduce risk only in a purely subjective sense. They depend upon differences in the risk preferences of the two persons, i.e., the way in which they subjectively assess the certainty equivalent of what is objectively precisely the same risk. If there is a difference in the certainty equivalents, then one can predict who will wind up bearing the risk when the parties are permitted to trade money for risk assumption: the risk will tend to be borne by the person who places the lowest money cost on it. If the risk is not already allocated to the least-cost bearer, then a transfer to that party, together with at least a compensatory payment to the risk-bearer, will be mutually advantageous. One caveat, however, is that the transactions costs of actually effecting the transfer may diminish or even overcome the potential gain. Risk transferal thus reduces risk costs when an otherwise unchanged risk is reallocated to a person who, merely for psychological reasons, attaches a lower certainty equivalent to the risk.

If the underlying utility functions of the two individuals are different, then the possibility for a determinate least cost bearer seems greater because the individuals will evaluate the risks in variant ways. But the specific fact situations imbedded in the questions should demonstrate that different utility functions do not guarantee a difference in risk preference. For instance, when the difference in the utility functions is merely by some multiplicative factor, e.g., 100 instead of 100, the relative rankings of different outcomes remains unchanged; replacing the 100 by a 1000 makes the impact of both gains and losses ten times greater. Similarly, identical utility functions do not preclude diverse evaluations of a risk; if the two people are assessing movements from different starting positions on those identical functions, then their assessments of the same risk will not be the same. This latter phenomenon pops up in many different economic contexts as an “income effect” or “wealth effect,” i.e., a behavioral impact based on income level rather than underlying preferences. Finally, you should also see from the first set of questions that it is not true that the person who would feel the greatest utility loss from the risk will necessarily be the least-cost bearer. A person who feels losses very intensely may also feel gains very intensely! It is the relative difference in the intensity of feeling gains and losses that produces risk preference.

It is very important to perceive that reduction in the subjective cost of risk is nonetheless a very real reduction in costs. Many extremely significant economic costs are essentially subjective in origin. The
fact that people are willing to pay real things, money or other advantages, in order to avoid subjective costs is a dramatic proof of the potential “objectification” of subjective costs.

The second process of cost reduction is also mainly personal or subjective in its operation, although there is some objective change in the form in which the risk is borne. What is involved is an alteration in the distribution of an unchanged set of risks, a process sometimes called risk “spreading” or pooling. Each of the mutual risk-bearers now faces a different, more complex set of objective risks than before. The expected value of the loss is exactly the same, but the accompanying amount of dispersion falls. Note, however, that the aggregate amount of risk facing “the world” is not altered in any way, although the prospect facing each individual is improved. Because the individuals are assumed to be risk averse, the reduction in dispersion due to the pooling arrangement translates into a reduction in the subjective cost to each of the mutual riskbearers. As more and more sharers get together, the “law of large numbers” works, and in the limit each person is faced by an almost-certainty of the expected value of the loss. It is only an “almost-certainty” because the dispersion never falls to zero, but does grow narrower and narrower.

A diagrammatic illustration is sometimes helpful in seeing the effects of risk pooling. The accompanying Exhibits 2.7, 2.8, and 2.9 assume that the probability of loss associated with any single risk is 40% and the magnitude of a single loss equals $1. The histograms in each Exhibit illustrate the changing shapes of the probability distributions of the possible losses for any individual when 5, 10, and 20 people bear their $1 risks by “pooling.” The plotted shapes of these distributions can be related to the table of probabilities provided above in Exhibit 2.6.
Pooling involves summing the losses, if any, from each separate risk and then sharing these losses in pro rata fashion among the participants in the pool. Hence the 20 cents column of Exhibit 2.7 represents an outcome consisting of one $1 loss shared among 5 people ($1/5 = 20 cents) whereas the same column of Exhibit 2.8 represents the occurrence of two losses out of ten risks ($2/10 = 20 cents). Each X in a column of the histogram corresponds to approximately one percent of probability, although the sum of X’s in each histogram may not add to 100 because of rounding. The probabilities plotted in the Exhibits can be found in any standard table of probabilities for the so-called “binomial distribution” and correspond to the results indicated in Exhibit 2.6.

The third set of questions deals with a situation in which one party or the other can act in a way that alters the objective amount of risk created in the world. When such actions are cost-beneficial, there is a potential social gain from seeing to it that they are undertaken. In this third factor, we are dealing with purely objective considerations, perhaps best thought of as the “technology of risk variation.”

In the next few sections, we shall illustrate optimal risk-bearing principles in several different legal contexts.

**QUESTIONS**

1. Note that the questions were constructed so as to carefully separate these three conceptually distinct factors. Can you think of situations where more than one such factor is present? Where the directionality of the factors may be at cross purposes among the parties?

2. Can you supply an example of a situation wherein the placement of a legal liability is allegedly based upon consideration of who is the least-cost risk bearer?

3. An argument may be made that pharmaceutical companies, rather than their customers, are the least-cost bearers of risks from consuming the product. What is the argument based upon, and how persuasive do you find it? Is the pharmaceutical company an obviously superior bearer of the
kinds of “generic” risk involved when, despite all proper precautions in research and testing, a drug turns out years later to have extremely harmful side effects? What about the risk, that, notwithstanding all proper efforts at quality control in production, every now and then a “batch” of improperly produced drugs causes harm to users?

3. PORTFOLIO ANALYSIS AND THE PRUDENCE STANDARD

Professional investment managers are held by the law to standards of both prudence and competence in regard to the handling of their clients’ accounts. The manager is expected to abstain from speculation, to sell off unsuitable assets, to maintain productivity, and to “diversify” a client’s account. Unfortunately, such duties are stated in broad terms and their precise content is susceptible to considerable variation in interpretation.

Although neither the federal nor the common law restrictions on “speculation” provide very precise definitions of the term, the legal system’s understanding of speculation seems to place important weight on the existence of exceptional risk both of gain and loss. On this basis, some specific types of investment media have been condemned as unsuitable for purchase by trustees: junior mortgages on property, unsecured loans, new issues, etc. Whatever the risks of any single investment, the risk-pooling effect described above can be recognized as relevant to the duty to “diversify.” If the risks associated with different investments are random, then a large “portfolio” of such investments may have little risk, even though the individual components of that portfolio seem to involve heavy risk.

Why would one be willing to put very risky investments, i.e., those with a high variance in their possible results, into a portfolio at all? Some of our earlier discussion suggests the answer: an investor may be willing to “trade off” between rate of return and risk, that is, accept greater variability of results if the expected value of the return is higher. Conversely, the “sellers” of such risks would plausibly be willing to offer a higher return when the variability is great. If “risky” investments are barred from portfolios, opportunities for higher return may be lost.

The legal system might still have an understandable solicitude for the construction of portfolios with “too many” risky components when the actual bearer of the risks is not competent to understand or evaluate what his investment manager is doing. “After all,” it may be argued, “pooling only reduces risks. The overall portfolio would always have less risk if, in a choice between two investments, the one with the lower variability were added to the portfolio.” Right? No. Modern portfolio theory suggests that overall risk can actually be reduced by adding more risk!

The trick is that the added risk must be of the right kind. In our previous discussions of pooling, we implicitly assumed that the risks were random and uncorrelated. Hence, the pooling arrangement really just gave us the benefit of the “law of large numbers” or the ability to in effect count on something close to an “average” result. Suppose, however, that we could find two risks, A and B, that are correlated, that tend to move in a predictable relationship to one another. When risks vary in this way, investment theorists refer to them as having significant “covariance.” If the returns on A and B are closely related in the sense that, when the return on A is known, one can make a good prediction about the return on B, the risks are highly covariant.

Covariance can be either negative or positive. When it is positive, the returns move together: if A is a loser, so is B. Negative covariance, on the other hand, implies that when A is up, B will tend to be down, and vice versa. Positive covariance among the investments in a portfolio exacerbates the swings and increases the risks, but negative covariance permits a swing in one direction to counterbalance a swing in the other direction.

---

22 Restatement of Trusts, Sec. 227, comments, f-1 (1935).
In *Trustees of Hanover College v. Donaldson, Lufkin & Jenrette, Inc.*, plaintiffs sued an investment manager, alleging that certain specific securities purchases were unsuitable. In their motion to dismiss, defendants alleged, *inter alia*, that the commitments in question were not objectionable when judged in light of the performance of the portfolio as a whole. The court refused to adopt a full portfolio approach in judging the manager’s performance and the parties settled before final judgment.

**QUESTIONS**

4. What is the policy argument to be made against the court’s refusal to adopt the overall portfolio approach? How persuasive is this argument? Are there counterbalancing considerations?

5. Obviously, the judge in Hanover College was not impressed with the policy argument in the form it was actually made. How would you try to communicate the argument to the judge? Would you use an expert witness, your own oral argument, briefs, homely examples? Be as specific as possible.

6. If the portfolio approach were to be implemented, who should have the burden of proof as to the effect of a security? How could that burden be met? Should the traditional classifications of “risky” investments play any role?

7. Merger of businesses that have no apparent commercial or productive relationships are sometimes criticized as sterile financial exercises. Does the portfolio approach suggest any other explanation for such combinations?

**D. OPTIMAL INSURANCE AND RISK ADJUSTMENT**

1. **“INSURANCE” AGAINST CONTRACT BREACH**

[The material in this section is based upon the analysis in Section II of Goetz & Scott, “Liquidated Damages, Penalties, and the Just Compensation Principle: Some Notes on an Enforcement Model and a Theory of Efficient Breach,” 77 Colum.L.Rev. 554 (1977).]

Martin Wurrier, IV, a graduate of his state’s major law school, is a wealthy senior partner in a prosperous land development firm and, even more importantly, a truly fanatical fan of the State University’s basketball team. After years as an also-ran in its highly competitive conference, State University produced a conference championship basketball team that has advanced to the regional finals of the NCAA championship tournament. Never one to pinch a penny in following his beloved State U. hoopers, Wurrier acquired twenty-five tickets to the regional championship game in a metropolis 250 miles away. He then entered into contract negotiations with the Reliable Charter Service, Inc. to arrange for a bus to transport himself and twenty-four good buddies, also loyal State U. alumni, to the site of the Big Game. The standard price for a bus charter of this type is $500.

Wurrier considered his attendance at the game to be of supreme importance and began to suffer many anxieties over whether the bus might break down and deprive him of witnessing State U.’s moment of glory. He was eager to quiet his fears by securing adequate compensatory “insurance” to salve his feelings in case Reliable failed to perform. Ordinarily, damages for non-performance of a contract in effect provide this insurance by paying the disappointed breaches compensatory damages. Wurrier realizes, however, that the $11,000 sum required to compensate him for not seeing the game would be so large as to be deemed “fanciful” by a court and, therefore, not legally cognizable as breach of contract damages. Wurrier would expect only about $1,000 in damages to be legally recoverable.

Under such circumstances, what can a worried promisee do to insure against the $10,000 of unprotected potential losses from breach? One option is to attempt to insure with a third party (Lloyd’s of London, for example). For instance, adding the proceeds of a $10,000 policy to the award of $1,000 legally cognizable expectation damages would provide Wurrier with full insurance of losses resulting from breach by Reliable.
Alternatively, Wurrier could negotiate for direct insurance from Reliable. He might propose to pay Reliable $500 above the usual price for the charter service if, in return for the additional premium, Reliable would agree to a penal sanction of $10,000 upon failure of performance. The stipulated sum of $10,000 would represent that amount at which Wurrier would be indifferent between performance and breach. Unfortunately, insurance purchased directly from the promiser, Reliable, is not a real alternative; the mere labeling of the idiosyncratic damages as pursuant to an “insurance” contract is unlikely to prevent a perceptive court from recognizing that such payments are de facto equivalent to a “penalty” and are therefore legally unenforceable under the penalty doctrine. Hence, legally collectible insurance for losses not otherwise recoverable as breach damages is in practice obtainable only from third parties.

As between the third-party insurance company and the promiser, which would be able to bear the risk at least cost? Identifying the efficient insurer requires an analysis of the costs of providing insurance. Perhaps the overwhelming element in the cost of this insurance would be the expected value of the underwriting loss to the insurer. Algebraically, this expected value can be defined as the product \( pR \), where \( p \) is the probability of non-performance and \( R \) is the recovery payable to the insured. In addition, an insurer will also have other transactions costs, such as the costs of ascertaining the true probability \( p \) and the costs of negotiation and communication with the insured. These transaction costs will be subsumed in the portmanteau variable \( T \), so that the total cost \( C \) of a policy paying \( R \) on the occurrence of an event with probability \( p \) can be summarized as \( C = pR + T \). If the services in question are marketed competitively, the cost of breach insurance to Wurrier will be \( (1 + a)C \) where \( a \) is the competitive rate of return or profit for the insurer. The question, then, is whether \( C \) would differ between the bus company and the third-party insurer.

One focal point of interest is the transaction cost element \( T \). Here, it is tempting to argue that the advantage lies with the bus company. In the first place, the bus company is in a superior position to know the breakdown probability \( p \). Secondly, many of the other transactions costs normally incident to customer communication may be negligible when communication is already being undertaken relative to the carriage service itself. Hence, \( T \) may be lower for the bus company and thus so would \( C \) and the offering price of the insurance to Wurrier.

Actually, however, the transaction cost element is not the strongest argument in favor of the bus company as the most efficient insurer. The bus company’s main advantage derives from its power to exercise some control over the breakdown probability \( p \). This can be illustrated by examining how the rational enterprise will make the maintenance and repair decisions which affect the breakdown probability.

Exhibit 2.10 reflects the fact that there is an inverse relationship between the probability of breakdown and the level of maintenance. (“Maintenance” is used as a summary term for all bus company activities that potentially affect the breakdown probability.) Each upward adjustment of the level of maintenance reduces \( p \), presumably by smaller and smaller amounts as the level increases.

In Exhibit 2.11, the technological facts of Exhibit 2.10 are converted into their investment implications. If \( D \) is the damage liability upon non-performance (here $10,000), the expected damages decline as increasing maintenance decreases the probability of breakdown. Since these expected damages are \( pD \), the saving in expected damage costs attributable to a marginal unit of maintenance is \( rD \) where \( r \) is the marginal reduction in \( p \) brought about by the maintenance increase. Hence, the “marginal damage reduction” curve in Exhibit 2.11 is simply a schedule of the value \( rD \) for alternative maintenance levels. This is an inverse relationship, as it is closely related to the curve in Exhibit 2.10. As the maintenance level grows (from \( q_0 \) to \( q_1 \)), the marginal damage reduction decreases, eventually tending to zero, a point where increased maintenance does not reduce the probability of breakdown at all.
The “marginal cost” curve is, in turn, a reflection of the incremental costs incurred as the maintenance level is varied. In Exhibit 2.11, this marginal cost is a constant; i.e., each additional unit of maintenance has the same cost. If it seeks to maximize its own net benefits, the bus company will maintain up to the point where the last or marginal unit of maintenance is “just worth it,” the quantity $q_0$ at which the marginal damage reduction curve intersects the marginal cost curve. If the company spent another unit of maintenance beyond $q_0$, any such unit would cost $C_m$ but reduce the expected damage liability by a lesser amount. Note that this predicted decision of $q_0$ maintenance also determines the Exhibit 2.10 probability $p_0$, the probability of breakdown when the bus company has completed its maintenance and the bus actually sets out.

Absent any specially arranged damage provision, the bus company anticipates that $D$ will embody only the standard objective damage recovery of $1,000 for the Wurrier bus trip. This anticipation underlies the marginal damage reduction curve in Exhibit 2.11 and is the ultimate basis for the $p_0$ breakdown probability. In computing the expected underwriting loss, a third party insurance company will therefore arrive at a value $(p_0 \times 10,000)$.

Suppose, however, that the bus company can offer the same insurance. The expected value of damages is now based, not on a $D$ of $1,000, but on a $D$ composed of $1,000 actual provable damages plus the $10,000 insurance recovery. Therefore, the new marginal damage reduction curve ($M$) for maintenance will be exactly eleven times higher than the original one. The company will expend additional
maintenance costs $A$, equivalent to the shaded area in Exhibit 2.11 or $C_m (q_1 - q_0)$, and the breakdown probability will consequently decline to $p_1$.

What are the implications of these adjustments on cost? For the bus company, the insurance cost must now be modified to reflect the net benefits of possible risk-avoidance efforts. Hence, the appropriate cost function for the provision by the bus company of $10,000 coverage is

$$C = (p_0 \times 10,000) - [11,000 (p_0 - p_1) - A] + T$$

where the terms in the square bracket are net gains from adjusting maintenance levels; $(p_0 - p_1)$ is the change in breakdown probability; its product with $11,000$ is the expected damage reduction; and $A$ is the added cost of maintenance. Furthermore, we know: (1) that these net gains are positive; and (2) that they would not be achieved when the third-party insurance is purchased. Hence, even where the insurance cost component $T$ is identical for the alternative insurers, the bus company has a cost advantage equal to the square-bracketed term in the equation above. (Graphically, this quantity is represented in Exhibit 2.11 by the crosshatched triangular area of net gains above the shaded area of incremental costs $A$ and below the new marginal benefits curve $M$.)

The preceding argument has been that non-enforcement of liquidated damages provisions has the result of inducing individuals to protect against otherwise non-recoverable losses through special third-party insurance. This is likely to be an inferior alternative since strong economic arguments suggest that the vendor himself would be the lowest-cost insurer against his own non-performance. Although this argument has been illustrated in terms of a bus company, a similar conclusion may be generalized to all cases in which the vendor has some control over the probability of externally caused non-performance.

In sum, many people may not want to make deals unless they can shift to others the risk that they will suffer idiosyncratic harm or otherwise uncompensated damages. To the extent that the law altogether prevents such shifts from being made or reduces their number by unnecessarily high costs, it creates efficiency losses; that is, it prevents some welfare-increasing deals from being achieved.

**QUESTIONS**

1. Notwithstanding the argument in the text, is it possible that a third-party insurer could bear the risk at lesser cost? Why?

2. Suppose that the conduct of the insured also affected in some way the expected value of the risk. In Wurrier’s case, you may hypothesize that a boozy, raucous party on the way to the game might distract the driver. Better yet, you may change the facts to generate a more important or plausible real-world example. A care-reduction response by the insured is commonly known by the term “moral hazard.” How can the same diagrammatic models used for the insurer be adapted to analyze moral hazard?


4. One important purpose of legal liability presumably is to channel the behavior of people. Yet, insurance against many forms of legal liability (e.g., for automobile accidents) seems to take away the very incentives that the imposition of liability supposedly creates. Then why is such insurance allowed?

**2. AFTER-THE-FACT INDEMNIFICATION: REALLY INSURANCE?**

We are accustomed to thinking of insurance as provision against the consequences of events that have not already happened, and perhaps never will. Although major examples are rather unusual, it is not impossible to find instances wherein the “risk” against which protection is sought seems to be one that is
no longer a mere prospect, but already has materialized. One such case stirred at least a minor controversy, as is described in the following excerpt, representative of the press coverage at the time.

**PLAN FOR MGM FIRE INSURANCE BLASTED**

A consortium of insurance companies plans to “insure” retroactively liabilities stemming from the November 1980 fire at the MOM Grand Hotel in Las Vegas in which 84 persons were killed.

Retroactive insurance of $170 million above the existing coverage of $30 million is being purchased to cover claims from the fire. The premium for the coverage is an estimated 37.5 million.

Frank B. Hall, Inc., the brokerage firm handling the insurance package, is spreading the coverage among insurers here and abroad. Those firms, in turn, will reinsure portions of the coverage to spread the risk broadly.

The National Consumer Insurance Organization has criticized these plans, contending that the added cost of providing such retroactive “insurance” may fall on small policyholders and the tax-paying public.

J. Robert Hunter, president of NCIO, argues that “all the problems associated with this unusual deal stem from the fact that this is simply not insurance. * * * Insurers have always said that you can’t insure a burning house, but when the client is big enough, apparently you can insure one that’s already burnt.”

Hunter claims that one result of this “non-insurance” could be higher rates for “true insurance policyholders”.

John F. McCaffrey, executive vice-president of the Hall firm, feels that uncertainty about the ultimate amounts of the claims settlements makes it a valid form of insurance.

“It’s not like property insurance where a building is insured for x-dollars and it burns down—then you’re covered for the amount of insurance. In liability, we don’t know what the costs will be until all the claims are settled. Therefore, you have risk, so you also have something that’s insurable.”

Don Heath, Nevada Insurance Commissioner, sided with Hunter on the issue of whether the coverage is truly “insurance”.

“My chief concern throughout the whole thing is determining what it is,” Hunter stated. “If it’s not insurance, maybe it’s simply a reserving program. But I’m not going to speak to that until I have all the facts before me.”

**QUESTIONS**

1. Why do some people think that this is not legitimate insurance? What argument is there that the arrangement really is insurance? What is your own opinion?

2. Assuming that this is insurance, why would the insurance consortium be a better bearer of the risk than MOM? (If insurance companies enjoy no superiority over MOM, then why are the parties agreeing to “transfer” the risk for valuable consideration?)

3. Arrangements of this type are unusual. Why do you think they are not more frequently seen?

4. Commentators have attempted to identify the characteristics of a risk that make it “insurable” in practice. For instance, it is said I that an insurable risk must possess the following characteristics: (1) it must entail a real loss; (2) the loss must be of significant magnitude; (3) the amount and circumstances (cause, time, place, etc.) must be accurately ascertainable; (4) the cost must not be excessive; and (5) the magnitude and probability of the loss must be susceptible of accurate calculation by the insurer. Why would each factor on the list arguably be important? Are there any that you think might be of lesser importance, or even not necessary at all? How do the listed factors apply to the MOM situation?
McCORMICK, Justice.

The United States Court of Appeals for the Eighth Circuit has certified the following question to us: “Does the theory of strict liability for abnormally dangerous activities apply to a common carrier under the circumstances of this case?” We answer the question affirmatively.

The court of appeals recited the facts it deemed relevant to the certified question:

This is a civil action brought by National Steel against William Gibbons, the bankruptcy trustee of the Chicago, Rock Island and Pacific Railroad Company (Rock Island) for damages resulting from a train accident on September 1, 1975. On that date, the Rock Island operated a train along its right-of-way which consisted in part of eleven tank cars loaded with propane gas. The train derailed and four tank cars exploded, resulting in extensive damage to a warehouse owned by National Steel. National Steel sought recovery under theories of res ipsa loquitur, specific negligence, and strict liability.

At trial, the jury ruled in favor of the defendant on the res ipsa loquitur claim. The district court directed a verdict for the defendant on the specific negligence theory. The court entered a directed verdict for the plaintiff on the strict liability claim. In a special interrogatory, the jury found that National Steel suffered $443,623 in damages as a result of the explosion, and judgment was entered accordingly.

The certification was in response to a motion by the Rock Island asking the court of appeals to certify to this court the question of whether the theory of strict liability was properly applied in the federal action.

The doctrine is incorporated in Restatement (Second) of Torts Sec. 519 (1977):

(1) One who carries on an abnormally dangerous activity is subject to liability for harm to the person, land or chattels of another resulting from the activity, although he has exercised the utmost care to prevent the harm.

(2) This strict liability is limited to the kind of harm, the possibility of which makes the activity abnormally dangerous.

The reason for the rule is explained in Comment d:

The liability arises out of the abnormal danger of the activity itself, and the risk that it creates, of harm to those in the vicinity. It is founded upon a policy of the law that imposes upon anyone who for his own purposes creates an abnormal risk of harm to his neighbors, the responsibility of relieving against that harm when it does in fact occur. The defendant’s enterprise, in other words, is required to pay its way by compensating for the harm it causes, because of its special, abnormal and dangerous character.

Rock Island appears to concede that the strict liability doctrine would apply in this case if it were not a common carrier. * * *

The determinative issue presented by the certified question is whether this court will adopt the common carrier exception to the strict liability rule. In making that choice we are confronted with two lines of authority.

One line holds that a common carrier that is required to carry abnormally dangerous cargo offered to it for carriage should not be held strictly liable. The leading case for this view is Actiesselskabet Ingrid v. Central Railroad Co., 216 F. 72 (2nd Cir.1914):

We think there can be no doubt, so far as a common carrier is concerned, that such danger as necessarily results to others from the performance of its duty, without negligence, must be borne by them as an unavoidable incident of the lawful performance of legitimate business. * * * It certainly would be an extraordinary doctrine for courts of justice to promulgate to say that a common carrier is under legal obligation to transport dynamite and is an insurer against any damage which may result in the course of
transportation, even though it has been guilty of no negligence which occasioned the explosion which caused the injury. It is impossible to find any adequate reason for such a principle.

Id. at 78; [other citations omitted].

The second line of authority, shorter but more recent, does not recognize the common carrier exception. The leading case for this view is *Chavez v. Southern Pacific Transportation Co.*, 413 F.Supp. 1203 (E.D.Cal.1976), where the court found that the common carrier exception would not be recognized in California. The court based its decision on the California Supreme Court’s use of risk distribution analysis in imposing strict liability in tort in such cases as *Greenman v. Yuba Power Product, Inc.*, 59 Cal.2d 57, 27 Cal.Rptr. 697, 377 P.2d 897 (1963). The *Chavez* court reasoned that the common carrier or “public duty” exception is based only on a consideration that it is unfair to impose liability on a carrier for engaging in an activity required by the public. It added:

But, there is no logical reason for creating a “public duty” exception when the rationale for subjecting the carrier to absolute liability is the Carrier’s ability to distribute the loss to the public. Whether the carrier is free to reject or bound to take the explosive cargo, the plaintiffs are equally defenseless. Bound or not, Southern Pacific is in a position to pass along the loss to the public. Bound or not, the social and economic benefits which are ordinarily derived from imposing strict liability are achieved. Those which benefit from the dangerous activity bear the inherent costs. The harsh impact of inevitable disasters is softened by spreading the cost among a greater population and over a larger time period. A more efficient allocation of resources results. Thus, the reasonable inference to be drawn from the adoption of the risk distribution rationale in *Smith v. Lockheed Propulsion Co.* is that California would follow the path of the Supreme Court of Washington in *Sieglr v. Kuhlman*, supra, and find that carriers engaged in ultra hazardous activity are subject to strict liability.

413 F.Supp. at 1214. Strict liability was also imposed in *Sieglr v. Kuhlman*, and the *Chavez* rationale was endorsed and applied in *Indiana Harbor Belt Railroad Co. v. American Cyanamid Co.*, 517 F.Supp. 314 (N.D.Ind.1981), a suit against a manufacturer.

Rock Island contends this court has already recognized the common carrier exception, and since it is the better view, should adhere to that position. We do not believe, however, that the issue has been answered in prior cases, and we believe, in any event, that the *Chavez* line of authority represents the better view.

* * *

In arguing that the common carrier exception represents the better view, Rock Island attacks risk distribution analysis as applied to common carriers generally and as applied in the circumstances of this case. It also points out that the exception has been adopted by the American Law Institute in Restatement (Second) of Torts section 521.

In attacking risk distribution analysis as applied to carriers generally, Rock Island asserts carriers cannot raise their tariffs to cover past losses. The fallacy in this argument, however, is that risk of liability is a factor to be considered in determining tariffs. See *Akron, Canton & Youngstown Railroad Company v. Interstate Commerce Commission*, 611 F.2d 1162, 1170 (6th Cir. 1979) (“A question of possible liability for damage resulting from carriage of a commodity is therefore within the Commission’s jurisdiction as the regulator of the economics of interstate rail transport.”) Thus, assuming tariffs cannot be increased to cover past losses, they can be adjusted to cover potential liability. * * *

In challenging application of risk distribution analysis in the circumstances of this case, Rock Island points out that it is bankrupt and that, in contrast, the corporate plaintiff was insured for all but $5,000 of its loss. We decline, however, to decide the issue of a carrier’s liability for abnormally dangerous activities on the basis of the parties’ relative abilities to spread the risk of loss in a particular case. To achieve a measure of uniformity and predictability, we prefer to adopt a rule for general application. We believe it is more likely in the generality of cases that a carrier will be better able to bear the loss than the party whose property is damaged. Moreover, the carrier is in a position to spread the risk of liability among the beneficiaries of its services. As in *Lubin v. Iowa City*, 257 at 391, 131 N.W.2d at 770, we think
“they should bear the loss and not the unfortunate individual whose property is damaged without fault of his own.” Although tort liability should not always be determined on the basis of which party can ordinarily best stand or distribute the loss, that basis is appropriate in this kind of case. See id. at 392, 131 N.W.2d at 771.

Here we have two parties without fault. One of them, the carrier, engaged in an abnormally dangerous activity under compulsion of public duty. The other, who was injured, was wholly innocent. The carrier was part of the dangerous enterprise, and the victim was not. The carrier was in a better position to investigate and identify the cause of the accident. When an accident destroys the evidence of causation, it is fairer for the carrier to bear the cost of that fortuity. Apart from the risk distribution concept, the carrier is also in a better position than the ordinary victim to evaluate and guard against the risk financially.

Furthermore, the carrier is in a superior position to develop safety technology to prevent such accidents, and assessment of accident costs is one means of inducing such developments:

In some cases it may be reasonably clear that only injurers, or only victims, can be looked to for advances in safety technology or other adjustments that might minimize accident costs ***.

This analysis might explain the major pockets of strict liability in the law. These include liability for damage caused by “ultra hazardous” activities ***. All are cases where the potential victims of the injury are not in a good position to make adjustments that might in the long run reduce or eliminate the risk of injury. It is difficult to imagine the development of cost-justifiable technologies that would *** enable a traveler at a railroad crossing to supervise the selection and monitoring of the railroad’s crew.


We conclude that the doctrine of strict liability for abnormally dangerous activities should be applied in the circumstances of this case. In so holding we do not overlook the Restatement position. We note, however, that we are committed to a broader application of the strict liability doctrine of *Rylands v. Fletcher* than is reflected in the Restatement. We do not limit it to “ultra hazardous activity.” See, e.g., *Healey v. Citizens’ Gas & Electric Co.*, 199 Iowa 82, 201 N.W. 118 (1924); cf. W. Prosser, The Law of Torts, supra, at 512 (“The Restatement of Torts has accepted the principle of *Rylands v. Fletcher*, but has limited it to an ‘ultra hazardous activity’ of the defendant. ** **”). Nor have we limited risk distribution analysis to the products liability field. See, e.g., *Lubin*, 257 Iowa at 391, 131 N.W.2d at 770. For the reasons expressed in this opinion, we decline to adopt the common carrier exception in Restatement (Second) of Torts, supra, section 521. Other Restatement provisions affecting the strict liability doctrine are not at issue here, and we do not pass on them.

We answer the certified question in the affirmative.

All Justices concur except ALLBEE and SCHULTZ, JJ., who dissent, and UHLENHOPP, J., who takes no part.

ALBEE, Justice (dissenting).

I dissent because I believe that the court should adopt as our own rule the common carrier exception to the strict liability rules for abnormally dangerous activities embraced by the American Law Institute in the Restatement (Second) of Torts (1977), to wit:

§ 521. Abnormally Dangerous Activity Carried on in Pursuance of a Public Duty

The rules as to strict liability for abnormally dangerous activities do not apply if the activity is carried on in pursuance of a public officer or employee or as a common carrier.

This exception is explained by Comment a to section 521, which states in pertinent part:

[A] common carrier, in so far as it is required to carry such explosives as are offered to it for carriage, is not liable for harm done by their explosion, unless it has failed to take that care in their carriage which their dangerous character requires.

The considered view of the American Law Institute seems to me to be more authoritative and to carry much greater weight than the reasoning of the single district judge in *Chavez v. Southern Pacific Transportation Co.*, the decision this court builds upon to reach its ultimate rationale.
QUESTIONS

1. Who are the potential bearers of the risk under alternative assumptions as to the liability rule, regulatory reaction, etc.?

2. What do you understand to be the rationale of the traditional “common carrier exception”? In what sense, if at all, is it a “fairness” approach? [You may wish to re-answer this question after considering the material on Justice and Fairness below.]

3. The Court notes that “assuming tariffs cannot be increased to cover past losses, they can be adjusted to cover potential liability.” If past liabilities are not provided for, to what extent will this affect the behavior of a business enterprise? Aren’t bygones bygones? Under the opportunity cost doctrine, how would past liability affect the cost of remaining in the business in the future? What difference does it make whether the liabilities imposed on railroads can be “passed on” to users of the railroad?

4. What sources of least-cost risk-bearing seem to play a part in the Court’s reasoning? Cite the passages in the opinion that support, either implicitly or explicitly, your view. Is there any doubt in your mind as to whether the court’s assumption about who constitutes the least-cost risk-bearer is empirically sound?

5. Does the majority opinion address the rationale underlying the common carrier exception that it rejects? Does the dissenting opinion address the majority’s rationale for the rejection?

6. Who would you personally prefer to see saddled with the cost of the explosion risk? If the court did not adopt that ideal solution, why do you think it did not do so?

E. JUSTICE AND FAIRNESS

“Justice” is a slippery concept which, like beauty, often shifts with the eye of the beholder. What contribution, if any, can formal analysis contribute to our understanding of justice and fairness?

1. A CONTRACTARIAN THEORY OF JUSTICE AS FAIRNESS

One persistent strain in the historical efforts to objectify norms for justice is that of ex ante fairness, the opportunity to enjoy fair and attractive prospects which, considered ex post, may or may not yield ultimately happy results. A contractarian analysis of social institutions falls within this genre because the social contract is, just as any garden-variety commercial contract, an agreement that is entered into precisely because it offers attractive prospects but which in any particular case may nevertheless retrospectively generate considerable regret and disillusionment for some parties. Contracts allocate risks of losses, presumably in exchange for prospectively more attractive chances for gains. Yet, things sometimes turn out badly for one party to a contract. As the ancient cliché puts it, one pursuing some potential benefit may have to “take the bitter with the better” and bear some downside risk. In legal terms, ex ante approaches correspond most closely to process fairness rather than to a requirement of substantive fairness in the ultimate results.

Along these lines, recent work by philosopher John Rawls has been the subject of vigorous discussion by legal scholars and social scientists alike. [See his A Theory of Justice, 1971. Many of the same basic ideas are contained in Rawls’ earlier article “Justice As Fairness,” 57 Philosophical Rev. 164-94 (1958).] One major theme in Rawls’ writings is that basic agreements are made “too late,” by which he means a time at which individuals already know their social positions, abilities, and those other characteristics that enable a person to predict the impact of social institutions on his own narrow self interests. This knowledge distorts the process of rendering justice because it induces self-serving choices among institutional alternatives.

In an effort to remedy the problem of self-interested choice, the notion of an “original position” is introduced. In this original position, individuals know generally the consequences of alternatives, but they do not know their own personal positions in the different outcomes. For instance, a similar situation
would exist if a person were looking down at Earth, understanding a great deal about how any particular body observed below will fare under each of a set of policy alternatives, realizing also that his own being is about to be infused into a body, but not knowing anything about which body will be his own. Subject to certain other conditions, it is contended that choices made in this situation are likely to be “fair” and confirm to a Rawlsian conception of social justice. This lack of “morally irrelevant” information about personal consequences is sometimes called the “behind the veil [of ignorance]” assumption.

Decisionmaking “behind the veil” may be illustrated by a simple matrix model. The rows labeled $S_1$, $S_2$, and $S_3$ represent three different sets of social institutions, differing in their constitutional guarantees, economic regulations, systems of property rights, etc. Following the usual conventions for these matrix models, the row states are what the decision-maker gets to choose, whereas the columns are the circumstances that external forces impose. In this case, the uncontrolled conditions represented by the columns will be the status of the decision-maker in society, designated by the letter A for an Alpha class individual and B for a Beta type. In our example, there will be equal numbers of Alphas and Betas in society.

<table>
<thead>
<tr>
<th>SOCIAL STATE</th>
<th>INDIVIDUAL’S STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>Rule Set $S_1$</td>
<td>100</td>
</tr>
<tr>
<td>Rule Set $S_2$</td>
<td>60</td>
</tr>
<tr>
<td>Rule Set $S_3$</td>
<td>40</td>
</tr>
</tbody>
</table>

Exhibit 2.12 – Social Choice Matrix

By the “behind the veil” assumption, the decision-maker does not know who he will be, whether an Alpha or a Beta. Suppose that the numbers in the cells of the Exhibit 2.12 matrix are the utility or satisfaction payoffs of the various alternatives. Which set of social institutions would the person choose?

The certainty equivalent techniques discussed above may be applied to this situation. Since the probabilities of A and B are each .5, the expected values of states $S_1$, $S_2$ and $S_3$ are 50, 45, and 40, respectively. Hence, a risk-neutral chooser would rank the alternative social states in the same order as they are listed in the matrix. Mild risk aversity might result in $S_2$ being ranked first because it achieves considerably less dispersion of the results at relatively small cost in terms of decreased expected value. Strong risk aversity might cause $S_3$ to be highest-ranked, however, if the better average results under the other two states were deemed not “worth it” because of their higher risk cost.

The behind the veil methodology allegedly leads to disinterested choice and, in effect, results closely allied to those predictable under other well known ethical principles, e.g., the Biblical “Golden Rule” of doing unto others as you would have done unto yourself and the Kantian categorical imperative. Since the chooser does not know how to serve his own selfish interests, he will presumably be led to give equal weight to the interests of each person affected by the policies chosen.

One fact deserves special emphasis here: the insight embodied in the “behind the veil” approach is not unique to Rawls. Indeed, this general approach of impersonal ex ante decision making is quite consistent with the prescriptions of Utilitarians, who would maximize the aggregate welfare of society, thus...
producing the highest possible expected value or “average” result. However, the extreme Utilitarian approach seems to imply risk neutrality, since it attaches no importance to the dispersion of individual results around the average. Thus, extreme Utilitarianism is analogous to a policy of maximizing per capita income without giving any thought to possible extreme disparities in the way that income is distributed among individuals.

While extreme Utilitarianism stands at the one polar position imposing no distributional norms whatever, Rawls is representative of approaches at the other polar extreme where rather inflexible distributional norms are superimposed upon the behind the veil decision process. Specifically, one of Rawls’ additional normative requirements is the so-called “difference principle” which requires that choices always be made so as to maximize the prospects for the “least favored” person or group in society. This principle dictates what in game theory is known as a “maximin strategy” wherein the decision maker assumes that the worst will happen and therefore maximizes the minimum possible payoff. For instance, in Exhibit 2.12, the worst result under each rule set is column B and S₃ would therefore be chosen because it leads to the largest minimum (maximin) expectation. Such a strategy is, of course, an example of extreme risk aversion.

Exhibit 2.13 has been suggested as a device to demonstrate the difficulties with Rawls’ difference principle. [See D. Mueller, R. Tomlinson and T. Willett, “The Utilitarian Contract: A Generalization of Rawls’ Theory of Justice,” Theory and Decision 345-347, at 348.] Assuming that n is some number greater than or equal to 1, rule S₅ is mandated by the difference principle because the “underdog” Bs are always better off under that system (1/n is better than 0 for the “least advantaged” group). On the other hand, the expected value or average prospect under S₄ is always higher. Specifically, the expected value of S₄ is .5n + .5(0) = .5n whereas the expected value of S₅ is .5(1) + .5(1/n) = .5(1 + 1/n). By letting the value of n start at one and then take on increasingly bigger values, one can make the prospective advantage of S₄ grow larger and larger because the expected value of S₄ approaches infinity at half the rate of n while the prospect of S₅ quickly shrinks toward the limit of .5 as 1/n approaches a value of zero.

The purpose of the Exhibit 2.13 construction is therefore to dramatize the question of whether, at some sufficiently high value for n, even a fairly strong proponent of egalitarian distributional principles would not willingly trade off the modest distributional superiority of S₅ for the far higher average results under S₄. In sum, the Exhibit is an attempted reductio ad absurdum suggesting that, while distributional goals are one plausible choice criterion, few people regard them as absolutely inflexible when they can be met only at an exorbitant cost in terms of other desiderata.
Chapter 2  BEHAVIOR INVOLVING UNCERTAINTY  91

<table>
<thead>
<tr>
<th>SOCIAL STATE</th>
<th>INDIVIDUAL'S STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>Rule Set (S_4)</td>
<td>(n)</td>
</tr>
<tr>
<td>Rule Set (S_5)</td>
<td>1</td>
</tr>
</tbody>
</table>

**Exhibit 2.13 – Maxi-Min Social Choice Matrix**

As a practical matter, the Utilitarians and strong distributional advocates such as Rawls would opt for similar policies under a great many circumstances, especially where distributional consequences are reasonably neutral. It is, for instance, difficult to maintain that the question of whether to observe the rule of Hadley v. Baxendale has any very powerful distributional implications. In any particular application, such a rule will of course have distributional ramifications for the parties involved, but the rule's application over the generality of cases does not seem likely to generate any systematic redistributive trends. Indeed, rules that have apparently clear distributional consequences are frequently seen, upon careful economic analysis, to have more diffuse or unexpected results than first impressions suggested. (Consider, for example, the rule determining the rights of subrogors versus those of subrogees in the distribution of “excess” recoveries, a question discussed above.) Where the distributional implications are, for one reason or another, unimportant, a Rawlsian and an extreme Utilitarian stand revealed as very close methodological kinfolk, both depending crucially on an ex ante analysis of prospects.

Whatever its limitations, the criterion of ex ante fairness as a test of justice is an intellectually intriguing notion. The following question sections are designed to reflect how ex ante fairness may be applied in different contexts and, incidentally, how tricky the impression of fairness may be: in each of the scenarios presented, the application of a rigorous formal analysis raises hidden problems, demonstrating that the notion of justice and fairness is not as simple or intuitively clear as it may first seem.

2. **FAIR DIVISION SCHEMES: DIVIDE AND CHOOSE**

Mrs. Bufförpington is an extremely wealthy widow who knows that she is suffering from a terminal illness. She wishes to divide the large and extremely heterogeneous stock of assets in her estate between her two daughters, Doris and Cloris. The daughters are known to be extremely antipathetic to each other; past treatment at their mother’s hands, however equitable in both intent and practice, has produced frequent bouts of jealousy and animosity. Nonetheless, Mrs. Bufförpington loves both daughters dearly and wishes to distribute the estate in a manner that will minimize eventual recriminations about their respective legacies.
Dividing the cash and other liquid assets of the estate presents no problem. As both sisters are aware, however, the division of certain real property, heirlooms, family jewelry, artwork, etc. involves subjective valuations which are highly variant from person to person. The sisters, not unexpectedly, each have considerable knowledge about each other’s tastes and preferences for various items that are included in the estate. Although one possible solution would be to liquidate all assets and divide the proceeds equally, Mrs. Bufforpinion regards the diffusion of many family assets to “outsiders” to be highly objectionable and wants the division of the estate to be physical or in kind — i.e., each daughter is to receive a fair share of the land and other personal property.

The troubled woman has previously sought the advice of a “fair division” expert who suggests that Doris be empowered to divide the physical assets into two shares from which her sister Cloris will then be entitled to choose whichever one she wants. Doris, of course, would receive the share not chosen by Cloris.

You are the trusted old family attorney for the Bufforpinion family. The mother contacts you and asks for a will to be drafted that incorporates the expert’s proposed fair division scheme. Assume that there are no strictly legal objections to the process proposed. However, she also solicits your opinion as to whether the scheme will achieve the desired result, that is, a division that both sisters will acknowledge as fair.

**QUESTIONS**

1. In a dazzling exhibition of the cunning of the legally trained mind, you impress the be jabbers out of the old woman by coming up with an explanation of how the process is arguably quite unfair to one of the sisters. Who is placed at a disadvantage, and why?

2. Assume that Mrs. B. nevertheless has her will written according to the original divide-and-choose scheme. When the mother dies, daughter Doris is notified by the executor that she is to be the Divider. The process of dividing is carried out and the moment comes when Cloris is entitled to choose between the divisions of the estate that Doris has made. Just as Cloris is about to choose, the executor rushes in and informs the sisters that a terrible error has been made: due to a misreading of the will, the roles of the sisters have been reversed. Why does Cloris now rejoice, expecting to improve her position as she takes over the role of Divider? (Note that this scenario lays the basis for a generalization and “proof” of the answer to # 1 above.)

3. Parents commonly use a similar scheme to divide a piece of cake between two children. How does this situation differ from the Doris-Cloris scenario?

4. Dividing cake among the children in a larger family requires a more complex process. Suppose there is a rectangular layer cake to be divided among more than two children. A parent proposes to pass a knife over the cake from left to right. At any moment, any child can say “I’ll take it.” A cut is made at the current position of the knife and the child who spoke out gets the piece to the left of the knife. The last child, of course, gets the residual piece. Do you think that this is a fair scheme?

5. Some asset or collection of assets is required to be divided up among \( n \) people. It is proposed to create exactly \( n \) shares with random — and possibly quite disparate — composition. Slips of paper with \( n \) numbers corresponding to each of the shares are placed in an urn and each beneficiary assigned a share by random selection of a slip. Is this a “fair” scheme? In what sense might it be argued that it is not fair?

**3. EQUITABLE DISTRIBUTION OF VOTING POWER**

*a. Pivotal-Voter Measure of Political Power*

In a majority-rule voting system, it might plausibly be contended that a voter exercises “power” only when the casting of his vote turns a non-winning coalition into a winning one. In other words, a voter has power only when the casting of the vote tips the result in a direction it would not otherwise have taken. In
several articles, John Banzhaf has proposed that this concept of voting power be used as a benchmark in assessing the equity and constitutionality of various representation schemes, such as multi-member districts and weighted voting. [See J. Banzhaf, “Weighted Voting Doesn’t Work: A Mathematical Analysis,” 19 Rutgers L.Rev. 317 (1965); “Multi-member Electoral Districts—Do They Violate the ‘One Man, One Vote’ Principle;” 75 Yale L.J. 1309 (1966).] Banahaf suggests that, since the ways in which coalitions of voters will form cannot be predicted a priori, all possible coalitions should be regarded as equally likely. Then, a voter’s power will be measured by the number of winning coalitions in which he is “necessary” divided by the number of all possible coalitions. Basically, this “power index” can be interpreted as the probability that a vote will “matter” as far as the ultimate result is concerned. In turn, a single vote will make a difference to the result when the other voters have divided themselves into two equal-sized coalitions such that the casting of the last vote will break a tie and determine which alternative will win.

QUESTIONS

1. In a system of representative delegations, the power of different delegations ought presumably to be proportional to the populations of their respective constituencies. I.e., the power of a delegation from a district of 300,000 ought to be three times that of one having a population of 100,000. In practice, political “weights” may be attached to districts of different sizes either by actual weighted votes cast by a single member or by multimember delegations. Consider a four-district legislature with district populations of 300,000, 500,000, 700,000, and 900,000. Each district has a representative for every 100,000 constituents. For simplicity, assume that the representatives of each particular district always bloc vote. Is the power of each delegation proportional to its population? [Hint: There are 16 different possible ways the delegations can vote. In how many of these does a particular delegation “make the difference” in forming a winning coalition?] See Morris v. Board of Supervisors of Herkimer County, 273 N.Y.S.2d 929 (1966) in which the Banzhaf analysis is cited as a basis for striking down a weighted representation scheme.

2. Assume that the relative power of each delegation is properly proportional to its constituency size. The next question is the amount of power that any individual voter has over deciding the nature of the delegation. Again, for simplicity assume that each delegation is elected as a “slate.” There are two competing slates running in each district. In a district that has n+1 voters, it can be argued that a voter has power in those cases in which all other voters divide up into two coalitions of identical n/2 size, thus permitting a single voter to break the tie. Banzhaf shows that the proportion of such “tie” combinations is approximately equal to one divided by the square root of two times Pi times the number of voters, i.e., 1/(2Πn)^1/2. [Banzhaf, “Multimember Districts * * *”, n. 28] Thus, the probability of being a tie-breaker is not inversely proportional to n, but rather to the square root of n; i.e., an individual’s power over his delegation declines less than proportionally as the size of the constituency increases. Does the Exhibit 2.14 accurately summarize these conclusions?
3. What is the implication of this result for the fairness of the allocation of power among voters in different-sized districts? What remedy does it suggest about the size of district delegations?

4. This measure of voting power can be applied to the Electoral College used to determine the U.S. presidency.\(^{23}\) As is well known, less populous states are somewhat over-represented (and, therefore, larger states under-represented) in terms of electoral votes. Why? As the dots in the Exhibit 2.16 scattergram indicate, however, application of the power index yields the result that a high-population state voter (e.g., a Californian) has more power that a low-population state voter (e.g., an Alaskan). Do you understand why this conclusion emerge?

5. Note that the definition of power being used above implies that, in the real world, no voter exercises power in any except a very small fraction of elections: those decided by a one-vote margin. Do you think that there is anything wrong with this concept of power? Why or why not? Can you suggest a better alternative? Reconsider this question after reading the next section.

\(b.\) An Equilibrium-Displacement Measure of Voter Power

In a seminal model of spatial competition,\(^{24}\) economist Harold Hotelling provided some insights into how competing businesses would locate in space. Adapting some of Hotelling’s results, we will first consider how two competing hot dog stands would locate on a beach and then analogize this spatial competition to political competition between two opposing parties. Then, by applying probability principles, an alternative to Banzhaf’s pivotal power index can be developed. These models will be constructed under what may seem to be extreme simplifying conditions. Nonetheless, the analysis is designed to reveal the commonsense rationale of an argument that points in quite the opposite direction from Banzhaf’s. The question will be: Which is (more) correct?

Consider a beach that is, say, a mile long from West to East. Bathers arrive and distribute themselves along the beach. Each bather will buy one hot dog and will always buy from the hot dog stand that is closest. Under these circumstances, Hotelling’s model predicts that the hot dog stands will wind up locating right next to each other. Further, the stands will attempt to locate at the West-East median of the mass of the bathers, i.e., with half of them on each side. Do you see why these predictions about the “equilibrium” behavior of the hot dog stands would be made?

---

\(^{23}\) See Banzhaf, One Man, 3.312 Votes: A Mathematical Analysis of the Electoral College, 13 Villanova L. Rev. 304-346 (1968).

Anthony Downs\textsuperscript{25} recognized that the concept of a linear market (the beach) in Hotelling could be analogized to a linear issue-space in a political competition where the “hot dog stands” are rival political parties and his insight has been elaborated and applied by other scholars.\textsuperscript{26} Accepting \textit{arguendo}, the simplifications in the most basic forms of these models, some interesting issues surface. Try your hand at the questions below.

\textbf{QUESTIONS}

6. Suppose that voters’ preferred issue positions are scattered along an issue space in the manner suggested by the simple seven-voter example in Exhibit 2.15 below. Indicate the basis for the following prediction: the winning candidate’s issue-positioning is likely to be very close to the position of the median voter on the left-right spectrum (i.e., the desired position of Voter 4 in the Exhibit 2.15 example).

7. How, then, can the “power” of a single voter over the political equilibrium result be estimated? Imagine that Voter 6 in Exhibit 2.15 got up and figuratively “moved his blanket down the beach” leftward to the new position indicated by the grey marker at the end of the arrow. If and when this movement is detected, what happens to the equilibrium “winning” position? That is, by how much is a single voter likely to be able to “displace” or affect the likely outcome? How, if at all, does the analysis differ for the other Voters in the Exhibit 2.15 model?

8. Besides changing position on the issue spectrum, a voter might also displace the equilibrium by “dropping out” and choosing not to vote. By how much would this non-participation affect the equilibrium result?

9. The expected reaction of the rival parties, and therefore the proffered and winning choices, depends on parties’ knowing where the voter are “located” and who is going to “come to the beach” and vote. How does this affect anything that has been said so far?

\textsuperscript{25} An Economic Theory of Democracy (1957)
10. States with different numbers of voters are analogous to beaches with different degrees of crowd density. Suppose there were two beaches of equal size and one of them has 10 times as many people on it. If that’s all you knew, what would you predict about the likely distance between adjacent bathers on the two beaches? How does that insight carry over into the equilibrium displacement voter power context?

11. We have developed a very simple model which suggests that the equilibrium displacement power of a voter is, in fact, likely to be proportional to the number of voters, i.e., it declines as the fraction $1/n$. (Remember, in the Banzhaf pivotal power model, individual voter power varies only as the square root of $n$.) Importing these alternative conceptions of power into the Electoral College context produces the starkly contrasting results in Exhibit 2.16 below. So, bearing in line the respective limitations of these alternatives, which concept is better?

12. Finally, the discourse thus far has been conducted in terms of assessing the power of a single voter. What, if anything, do these two theories suggest about the measurement of the power of a group of voters?

![Exhibit 2.16: Electoral College Power Under Alternative Theories](image)

**4. EFFICIENCY ASPECTS OF FAIRNESS**

**STATE OF IOWA v. HENRY PARRISH**

Supreme Court of Iowa, 1975. 232 N.W.2d 511

HARRIS, Justice.
Henry Parrish (defendant) was involved in a fight on August 17, 1974 with Leslie Ford, during which he struck Ford over the head with a beer bottle. Ford has since married defendant’s sister. At the time of sentence defendant’s counsel said the fight occurred as the result of defendant’s feeling Ford was attempting to make defendant’s sister “into what’s called a dealer’s woman.” Defendant also believed Ford had gotten his sister pregnant.

On August 30, 1974 defendant appeared with counsel in district court before Judge Lewis C. Schultz. Following arraignment defendant entered a plea of guilty to the charge. The proceedings included the following:

THE COURT: Are you ready at this time to enter a plea? DEFENDANT: Yes.

THE COURT: And what is that plea? DEFENDANT: I guess I’m going to plead guilty to the charge.

THE COURT: Do you wish to plead guilty? DEFENDANT: Yes.

THE COURT: Now is this the result of a plea bargain, Mr. County Attorney?

MR. HORAN: Yes, Your Honor. We have agreed that on a plea of guilty we would recommend that the defendant be sentenced to one year in the county jail and be given credit for time served.

THE COURT: How long has he served? MR. MONROE: Since August 13th, I believe, Your Honor.

THE COURT: Did you understand that, that this is the recommendation of the County Attorney and the Court would be free to pass whatever sentence the Court would determine would be necessary? DEFENDANT: I believe so.

THE COURT: I am going to have to ask you some questions, Mr. Parrish, which I do on any plea of guilty, to determine the voluntariness of the plea and determine whether or not you understand it, and also whether or not you understand the rights you are about to give up. The sentence under 694.6 is imprisonment in the county jail not to exceed one year or by a fine not exceeding $500, or by imprisonment in the penitentiary not exceeding one year. Any one of those things. Do you understand that? DEFENDANT: I believe so.

* * *

THE COURT: By your plea of guilty you in fact admit that the State can prove the necessary elements of the crime. You are entitled to have a jury trial, and the jury would have to find beyond a reasonable doubt that you were guilty of the following elements: That on or about August 17, 1974, in Linn County, you did willfully and unlawfully assault one Leslie Gene Ford with intent to inflict upon him great bodily injury. ‘Willfully’ means intentionally. ‘Assault’ means to put another person in fear or threaten their body, and this assault must be with intent to inflict great bodily injury. By pleading guilty you are admitting the State can prove these elements beyond a reasonable doubt. You understand that? DEFENDANT: Yes.

* * *

“THE COURT: By your plea of guilty you also give up your right to a jury trial, to confront the witnesses against you, they must testify in open court before you, and you have the right to cross examine them and to have an attorney assist you. If you don’t have funds to hire an attorney one will be appointed at county expense. You have the right to call witnesses on your behalf, you have the right to testify yourself, or you have the privilege against self-incrimination, which means you don’t have to testify or admit anything, and if you would request it, the jury would be told that your failure to take the stand was done as a matter of right. By a plea of guilty you give these rights up. Do you understand that? DEFENDANT: Yes.

* * *

THE COURT: And you understand that the Court is the one that passes punishment and that I can disregard the County Attorney’s recommendation of one year in the county jail, I can pass a lesser sentence or fine you, or I could pass a sentence to the Men’s Reformatory. DEFENDANT: Yes.
BEHAVIOR INVOLVING UNCERTAINTY

THE COURT: You have never been convicted of a felony previously, have you? DEFENDANT: I have been convicted of a felony and received a bench parole.

THE COURT: But you have never been to the reformatory? DEFENDANT: No.

THE COURT: You understand the Court is the one that passes judgment and that I am not bound by your agreement with the County Attorney? DEFENDANT: What you are saying is that the Court here, the judge, is the one that says what happens, that the people I have talked to out here can’t tell me what will happen, they can only make recommendations.

THE COURT: That is exactly right. You said it better than I can. DEFENDANT: Yes, I understand.

THE COURT: And what is your plea? DEFENDANT: A plea of guilty, Your Honor.

On the same date, August 30, 1974, Judge Schultz also ruled “an actual basis” existed for the plea and that the plea had been voluntarily made with knowledge and understanding of the charge, plea, and consequences of such plea. A presentence investigation was ordered.

Defendant appeared with counsel in district court for sentencing before Judge James H. Carter on September 25, 1974. At this time the State again noted it had entered into a plea bargain with defendant and stated the terms thereof. At this time the State recommended “* * * the defendant be sentenced to a six month term in the Linn County jail and that he be credited for time served.” Judge Carter never advised defendant he had a right to withdraw his guilty plea or that the county attorney’s proposal was not binding on the trial court. Judge Carter ignored both recommendations of the county attorney. Defendant was sentenced to one year in the state reformatory at Anamosa with the recommendation he be transferred to the security medical facility for psychiatric evaluation and treatment.

* * *

In the two recent opinions we have considered the obligation of a trial court which rejects the terms of a plea bargain after accepting a guilty plea. State v. Fisher, 223 N.W.2d 243 (Iowa 1974); State v. Runge, 228 N.W.2d 35 (Iowa 1975). Defendant relies on language in our opinion in Fisher:

“* * * [P]rior to entry of judgment, trial court, having elected not to honor the aforesaid plea bargain, neither so advised defendant nor accorded him related opportunity to stand on his guilty plea or move to withdraw same. Upon the record here made we cannot say defendant’s guilty plea was voluntarily entered.” 223 N.W.2d at 246.

In Runge the trial court had actual knowledge of a plea bargain and made no record of the fact. In reversing for failing to allow the plea to be withdrawn we explained:

“Briefly stated, defendant’s guilty plea was not shown of record to have been free from inducement flowing from the county attorney’s pre-plea concessions. Thus the controverted plea was not knowingly and voluntarily entered. See United States v. Mancusi, [275 F.Supp. 508, 515-519 (E.D.N.Y.1967)]; State v. Fisher, 223 N.W. Ed 243, 244-246 (Iowa 1974), and citations.” 228 N.W.2d at 37.

We believe the facts in the instant case fall outside the situations outlined in Fisher and Runge. The record now before us clearly shows there was no taint of false inducement of the plea by the judge accepting it. The trial court clearly renounced any willingness to grant a sentence concession on the basis of the bargain. Defendant noted his understanding the judge would not be bound by any recommendations in his own words: “What you are saying is the Court here, the judge, is the one that says what happens, that the people I have talked to out here can’t tell me what will happen, they can only make recommendations.”

Our holdings in Fisher and Runge do not render all guilty pleas tentative and subject to withdrawal as a matter of defendant’s right where there has been a plea bargain. The court is not bound to allow withdrawal of the plea where each of the following conditions is met: (1) the court renounces participation in the bargain, (2) denies any inclination to be bound thereby, (3) such renunciation and denial are made known to the defendant, and (4) the defendant thereafter enters or (as in this case) affirms his plea of guilty. All four conditions are established in the record before us.
Accordingly the question is governed by our cases which grant a trial court discretion in deciding whether to allow withdrawal of a guilty plea under Section 777.15, The Code. As stated in *State v. Vantrump*, 170 N.W.2d 453, 454 (Iowa 1969): “A defendant should not be permitted to enter a guilty plea, gamble on the sentence and then move to withdraw the plea if he is disappointed with the severity of the sentence imposed. [Authority].”

Affirmed.

Mc Cormick, Justice (concurring specially).

I agree with the dissent by Rawlings, J., that the procedure in this case did not comport with ABA Standards* * *approved unanimously by this court in *State v. Fisher*, 223 N.W.2d 243 (Iowa 1974), and again in *State v. Runge*, 228 N.W.2d 35 (Iowa 1975). The standard does not authorize denial of a motion to withdraw a guilty plea when the judge has forewarned the defendant he will not be bound by the prosecutor’s recommendation. It requires the judge in any case in which he decides in fact not to follow the prosecutor’s recommendation, whether he has forewarned the defendant or not, to notify the defendant he will not do so and accord the defendant the opportunity at that time to withdraw his plea. *United States ex. ref. Culbreath v. Rundle*, 466 F.2d 730 (3 Cir. 1972); *Commonwealth v. Wilson*, 335 A.2d 777 (Pa.Super. 1975); *Quintana v. Robinson*, 31 Conn.Sup. 22, 319 A.2d 515 (1973).

The ABA standard approved by the court in *Fisher* does not depend for its applicability upon proof the defendant entered his plea from a belief the court would follow the prosecutor’s recommendation. As the commentary to the rule makes clear, the standard is preventive. It seeks to avoid false inducement in some cases by removing the occasion for false inducement in all cases. ABA Standards, The Function of the Trial Judge, supra, commentary at 57 (“Even though the judge has said nothing to the defendant * * * except that he need not follow the prosecutor’s recommendations, there nevertheless remains at least the taint of false inducement.”).

I concur in the result because this guilty plea and sentencing proceeding antedated the decision of this court in *Fisher*, and defendant’s guilty plea has not been shown to have been the result of unlawful inducement under standards then applicable. The sentencing occurred here September 25, 1974, and Fisher was decided November 13, 1974.

**QUESTIONS**

1. In both the majority opinion and concurring opinion, concern is raised with “false inducement.” What is meant by this? Is it a “fairness” issue? From the defendant’s perspective, does it matter whether some probability $p$ of non-performance by the State is due to false inducement or to bona fide disapproval by a judge who regards the plea bargain as not being in the public interest?

2. Assume that the possibility of false inducement can be ignored because of other safeguards. Does the inability of the defendant to withdraw his plea then seem unfair to you?

3. Consider whether inability to withdraw the plea is a poor rule on other grounds. Specifically, who gains or loses from the rule in future cases? How will defendants and prosecutors alter their bargaining behavior? (In answering these questions, you may wish to give specific attention to the treatment of risk by defendants and the prosecutor’s office, respectively.)

4. What possible relevance does this case have to private law situations? Here is a hint. Suppose that the legislature of your state passed a statute pertaining to sales of new automobiles. Under the statute, the seller-dealer would have the right to spin a roulette wheel with 100 numbers on it. If the number 7 comes up, the buyer is bound to pay the purchase price, but the dealer is discharged from the responsibility of
delivering the vehicle. What would be the harmful or beneficial consequences of this arrangement? To whom?

5. “PATERNALISM” IN THE LAW

Paternalism—or maternalism, if you prefer—is not necessarily a factor to be excluded from a system of justice. One problem, of course, is that rules which protect may also hinder valuable personal freedoms.

The decision in Williams v. Walker-Thomas Furniture Co., 350 F.2d 445 (1965) is based on a rather narrow procedural ruling as to the possibility under D.C. law of offering an unconscionability defense for contract breach; the appellate court does not actually decide on the factual merits of the allegation of unconscionable acts by the seller. It is not hard to find in the full text of the opinion, however, a certain interstitial sense of uneasiness about any vendor who would sell, under seemingly very onerous terms, a very expensive stereo set to an impecunious person subsisting on welfare payments. (For comparative purposes, it may be instructive to note that the 1962 price of the stereo set would be the equivalent in real purchasing power terms to over $1600 in 1983 dollars.)

Unconscionability is ordinarily defined in terms of the conditions of the bargaining process, either absence of “meaningful choice” by one of the parties or unreasonably imbalanced contractual terms. But suppose that the Williams court had gone the whole distance and ruled that, notwithstanding otherwise acceptable terms and bargaining conditions, it is unconscionable to sell to a person who is recognizable to any prudent seller as one who later will almost certainly regret making the contract. One prediction as to the effect of such a rule is, of course, that a welfare mother, such as Mrs. Williams, would thereafter experience considerable difficulty in buying an expensive stereo on credit.

What is being suggested here is the possibility of examining not only the conditions of the bargaining process but also the capacities of the participants. Arguably, such an approach is already codified in the Wisconsin Consumer Act, Sec. 425.107(3), which allows courts to consider the following, inter alia, as possible indices of unconscionable practice:

(a) That the practice unfairly takes advantage of the lack of knowledge, ability, experience, or capacity of customers;
(b) That those engaging in the practice know of the inability of customers to receive benefits properly anticipated from the goods or services involved.

Subsection (b) above might be interpreted as dealing with the case of easily foreseeable mistake.

QUESTIONS

1. Is a preference-inference role of the type suggested in the Williams dictum or the Wisconsin statute justifiable within the context of an ex ante theory of “behind the veil” justice (Rawls, the Utilitarians, etc.)? How can it be defended, if at all?

2. What are likely to be the arguments offered against a rule of the type in question?

3. How do you personally assess the balance of the arguments? Would it make any difference if you were told that an empirical study showed that, in a random sample of 100 welfare mothers who bought similar stereo sets for over $1500, 92% responded that they “regretted having spent my money on this item” when surveyed six months after the purchase?

4. Can the Wisconsin statute reasonably be stretched to cover a situation of the Williams type? Is this an appropriate area for “judge-made law” or is it only a legislative concern?

5. What are some actual statutes or legal doctrines that you believe to be plainly motivated by paternalistic objectives? Why is this whole line of inquiry indulged in a section on “uncertainty”?

6. It is easy to muddle a relevant distinction between paternalistic and self-protective rules. You may even have done so in thinking about some of the preceding questions. Consider, for instance, the
“cooling off period” for door-to-door sales contracts or the waiting period traditionally required for a marriage license. How do these rules differ from prescription drug restrictions, the special provisions affecting minors, or an arguably “foolish” buyer such as Mrs. Williams?
A. INTEREST RATES AND “PRESENT VALUES”

In deriving an expected value, one converts a possibly very complex set of possibilities into a single value, the certainty equivalent of the prospect being considered. An analogous computation in temporal problems is the conversion of one or more future payments, perhaps differing greatly in timing and magnitude, into a single-valued sum that is the equivalent of the future payments. This time-adjusted equivalent is typically expressed in terms of a lump sum payment transferred immediately, i.e., a “present value.” The application of this concept would be exemplified by a settlement negotiation where-in defendant offers either $800,000 paid in two annual installments beginning one year from now or 24 monthly payments of $40,000 beginning one month from now. Which is a better offer? In order to make the two rather different flows comparable, the most common procedure is to convert each individual future payment to its present lump sum-equivalent, i.e., “discount” the future payments to their “present values.” By summing the present values of the future payments called for in alternative streams, a decisionmaker reduces the information to intelligible and comparable terms. Although less frequently encountered, the reverse procedure may also be performed, i.e., computation of the flow of payments that can be achieved over a certain period by converting a lump sum to an equivalent stream of future receipts.

This section explains very briefly the logic behind the use of interest rates to determine the comparability of different time streams of payments. Much more elaborate discussions may be found in elementary economics and business texts, but the treatment herein should suffice to deal with the subsequent legal application problems. Notwithstanding any attempt to render the principles involved in as simple a “commonsense” fashion as possible and to provoke a variety of examples, the material of this chapter is inevitably somewhat quantitative and may be offputting to some. Nevertheless, one should be prepared to face the reality that there is no short-cut substitute for performing calculations when a quantity must be expressed as its equivalent at a different time period; the answer will not leap out of the page at you, even if the principles involved are thoroughly understood. Fortunately the growing accessibility of calculators and computers to lawyers will obviate any real computational drudgery. Still, one must take the time and have the diligence to ask even a computer the correct questions.

Most people have very little difficulty in understanding how a present sum of money grows into a larger sum in the future. The holder of money can invest his original principal and earn interest during each period. The interest from the first period then becomes available as an addition to invertible capital and, in turn, itself becomes a source of interest accumulation. Assuming a $1.00 initial sum and a 10% interest rate per period, the following numerical example could be developed. After one period, the original $1.00 would have grown to:

$1.00 (110%) = $1.10

After the second period, the growth would have attained the level:

[$1.00 (110%) ] (110%) = $1.21

where the square-bracketed term represents the new principal available after the first period’s growth. At the end of three periods, the original $1.00 would be:

{ [$1.00 (110%) ] (110%) } (110%) = $1.33

where the numbers in the curley brackets indicate the reinvestible sum available from the previous two periods of growth.

Just a glance at the numerical example suggests a way of creating a general formula. For each period, we are just multiplying the original principal by 110% one more time. Hence, we could just as easily say that the three-period result is equal to $1.00 (110%)3. Even better, the same reasoning supports the
assertion that, after any arbitrary number of periods symbolized by the variable n, the value of a $1.00 invested at 10% compound interest will be $1.00 \times (1.10)^n$. A full generalization occurs when we substitute symbols for all of the relevant parameters, not just the number of periods. This substitution yields the standard compound interest formula:

\[ F = V(1+r)^n \]

where \( F \) is the future value at the end of \( n \) periods, \( V \) is the present or starting value and \( r \) is the periodic interest rate expressed as a decimal.

Note that the standard compound interest formula really tells us how a present sum \( V \) transforms into a future sum \( F \). The formula answers the question: “If I have \( V \) dollars now and invest it at a compound interest of \( r \) per period, how much will I have after \( n \) periods?” In a relevant sense, \( F \) and \( V \) are “equivalents,” assuming a passage of \( n \) periods and an earning rate of \( r \). This “compounding” of interest and growth of the original capital is familiar to anyone who has a savings account.

A less familiar, but conceptually linked, question is: “If I need \( F \) dollars in the future, how much would I need to invest now in order to reach my goal?” A slight manipulation of the compound interest formula given above produces an equation that “solves” for the present value \( V \) rather than the future value \( F \):

\[ V = F/(1+r)^n \]

If I were originally entitled to receive \( F \) dollars at a time \( n \) periods from today, the formula suggests that I would presumably be indifferent to receiving the lesser sum \( V \) now, since it would grow to exactly the same sum by the appointed date. For similar reasons, I would presumably be indifferent as to whether I had to pay \( F \) in the future or \( V \) right now. “Discounting” a future sum to its “present value” equivalent involves dividing it by \((1 + r)^n\) or, what amounts to the same thing, multiplying it by a “discount factor” equal to \(1/(1 + r)^n\).

The two tables provided below are very closely related. Exhibit 3.1 is a table of values of \((1+r)^n\) for various values of \( n \) and \( r \). It may be used either for the compound interest formula (as a multiplicative “growth factor” applied to \( V \)) or for the present value formula (as the divisor applied to \( F \)). In essence, the entries in this table can be interpreted as the number of dollars to which one dollar will have grown when invested at the interest rate indicated on the column heading and for the number of periods indicated in the row heading. Exhibit 3.2 is a table of the reciprocals of the values in Exhibit 3.1, i.e., \(1/(1+r)^n\). These are the multiplicative form of the “discount factors” that, when multiplied by the relevant future value \( F \) will yield the present value equivalent \( V \). The multiplicative factors table is perhaps more suggestive and easier to understand for discounting purposes because the numbers have an appealing commonsense interpretation as the cents-per-dollar that a future payment is worth in present value terms. The entries in both of the tables have been rounded and should not be regarded as exact.
### Exhibit 3.1: (1+r)n Compound Interest Growth Factors

<table>
<thead>
<tr>
<th>Periods</th>
<th>1.00%</th>
<th>1.50%</th>
<th>2.00%</th>
<th>2.50%</th>
<th>5.00%</th>
<th>10.00%</th>
<th>12.00%</th>
<th>15.00%</th>
<th>20.00%</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=1</td>
<td>1.01</td>
<td>1.02</td>
<td>1.03</td>
<td>1.05</td>
<td>1.10</td>
<td>1.12</td>
<td>1.15</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.02</td>
<td>1.05</td>
<td>1.08</td>
<td>1.16</td>
<td>1.30</td>
<td>1.45</td>
<td>1.69</td>
<td>1.91</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.04</td>
<td>1.08</td>
<td>1.16</td>
<td>1.22</td>
<td>1.38</td>
<td>1.57</td>
<td>1.80</td>
<td>2.10</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.06</td>
<td>1.11</td>
<td>1.23</td>
<td>1.29</td>
<td>1.46</td>
<td>1.72</td>
<td>2.08</td>
<td>2.50</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.08</td>
<td>1.13</td>
<td>1.27</td>
<td>1.36</td>
<td>1.54</td>
<td>1.82</td>
<td>2.23</td>
<td>2.80</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.10</td>
<td>1.18</td>
<td>1.33</td>
<td>1.44</td>
<td>1.63</td>
<td>2.03</td>
<td>2.55</td>
<td>3.26</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.12</td>
<td>1.23</td>
<td>1.39</td>
<td>1.52</td>
<td>1.73</td>
<td>2.22</td>
<td>2.88</td>
<td>3.72</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.14</td>
<td>1.28</td>
<td>1.46</td>
<td>1.62</td>
<td>1.86</td>
<td>2.41</td>
<td>3.13</td>
<td>4.13</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1.16</td>
<td>1.33</td>
<td>1.53</td>
<td>1.70</td>
<td>1.99</td>
<td>2.64</td>
<td>3.41</td>
<td>4.46</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.18</td>
<td>1.38</td>
<td>1.60</td>
<td>1.80</td>
<td>2.16</td>
<td>3.00</td>
<td>3.89</td>
<td>5.10</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1.20</td>
<td>1.43</td>
<td>1.68</td>
<td>1.91</td>
<td>2.30</td>
<td>3.40</td>
<td>4.37</td>
<td>5.79</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1.22</td>
<td>1.48</td>
<td>1.76</td>
<td>2.04</td>
<td>2.49</td>
<td>3.84</td>
<td>5.01</td>
<td>6.64</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1.24</td>
<td>1.53</td>
<td>1.84</td>
<td>2.18</td>
<td>2.69</td>
<td>4.33</td>
<td>5.80</td>
<td>7.83</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1.26</td>
<td>1.58</td>
<td>1.93</td>
<td>2.33</td>
<td>2.92</td>
<td>4.88</td>
<td>6.77</td>
<td>9.31</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1.28</td>
<td>1.63</td>
<td>2.02</td>
<td>2.49</td>
<td>3.18</td>
<td>5.51</td>
<td>8.00</td>
<td>10.99</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1.30</td>
<td>1.68</td>
<td>2.11</td>
<td>2.66</td>
<td>3.46</td>
<td>6.22</td>
<td>9.43</td>
<td>12.87</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1.32</td>
<td>1.73</td>
<td>2.20</td>
<td>2.84</td>
<td>3.76</td>
<td>7.02</td>
<td>11.10</td>
<td>15.02</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.34</td>
<td>1.78</td>
<td>2.29</td>
<td>3.03</td>
<td>4.08</td>
<td>7.92</td>
<td>12.98</td>
<td>17.54</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1.36</td>
<td>1.83</td>
<td>2.38</td>
<td>3.23</td>
<td>4.42</td>
<td>9.02</td>
<td>15.17</td>
<td>20.37</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1.38</td>
<td>1.88</td>
<td>2.48</td>
<td>3.44</td>
<td>4.78</td>
<td>10.24</td>
<td>17.63</td>
<td>23.51</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>1.40</td>
<td>1.94</td>
<td>2.58</td>
<td>3.66</td>
<td>5.16</td>
<td>11.61</td>
<td>20.39</td>
<td>27.07</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>1.42</td>
<td>1.99</td>
<td>2.69</td>
<td>3.89</td>
<td>5.56</td>
<td>13.12</td>
<td>23.52</td>
<td>31.04</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>1.44</td>
<td>2.05</td>
<td>2.80</td>
<td>4.13</td>
<td>5.98</td>
<td>14.78</td>
<td>27.10</td>
<td>35.39</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>1.46</td>
<td>2.11</td>
<td>2.92</td>
<td>4.38</td>
<td>6.42</td>
<td>16.61</td>
<td>31.03</td>
<td>40.15</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>1.48</td>
<td>2.17</td>
<td>3.05</td>
<td>4.64</td>
<td>6.89</td>
<td>18.61</td>
<td>35.33</td>
<td>45.31</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>1.50</td>
<td>2.23</td>
<td>3.19</td>
<td>4.91</td>
<td>7.38</td>
<td>20.81</td>
<td>40.02</td>
<td>51.00</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>1.52</td>
<td>2.29</td>
<td>3.33</td>
<td>5.19</td>
<td>7.90</td>
<td>23.20</td>
<td>45.10</td>
<td>57.20</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>1.54</td>
<td>2.36</td>
<td>3.48</td>
<td>5.48</td>
<td>8.44</td>
<td>25.79</td>
<td>50.61</td>
<td>63.92</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>1.56</td>
<td>2.43</td>
<td>3.63</td>
<td>5.79</td>
<td>9.01</td>
<td>28.64</td>
<td>56.62</td>
<td>71.16</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1.58</td>
<td>2.50</td>
<td>3.79</td>
<td>6.11</td>
<td>9.60</td>
<td>31.71</td>
<td>63.05</td>
<td>78.93</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>1.60</td>
<td>2.58</td>
<td>3.95</td>
<td>6.44</td>
<td>10.21</td>
<td>35.00</td>
<td>70.00</td>
<td>87.23</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>1.62</td>
<td>2.66</td>
<td>4.12</td>
<td>6.79</td>
<td>10.84</td>
<td>38.51</td>
<td>77.41</td>
<td>96.05</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>1.64</td>
<td>2.74</td>
<td>4.29</td>
<td>7.15</td>
<td>11.50</td>
<td>42.24</td>
<td>85.27</td>
<td>105.39</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>1.66</td>
<td>2.83</td>
<td>4.47</td>
<td>7.52</td>
<td>12.18</td>
<td>46.20</td>
<td>93.61</td>
<td>115.25</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>1.68</td>
<td>2.91</td>
<td>4.65</td>
<td>7.90</td>
<td>12.89</td>
<td>50.41</td>
<td>102.50</td>
<td>125.63</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>1.70</td>
<td>3.00</td>
<td>4.84</td>
<td>8.29</td>
<td>13.62</td>
<td>54.87</td>
<td>112.00</td>
<td>136.53</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>1.72</td>
<td>3.09</td>
<td>5.03</td>
<td>8.69</td>
<td>14.37</td>
<td>59.61</td>
<td>122.01</td>
<td>148.04</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>1.74</td>
<td>3.19</td>
<td>5.23</td>
<td>9.10</td>
<td>15.14</td>
<td>64.61</td>
<td>132.53</td>
<td>160.10</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>1.76</td>
<td>3.29</td>
<td>5.43</td>
<td>9.52</td>
<td>15.93</td>
<td>70.00</td>
<td>143.61</td>
<td>172.71</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>1.78</td>
<td>3.40</td>
<td>5.64</td>
<td>9.95</td>
<td>16.74</td>
<td>75.71</td>
<td>155.24</td>
<td>185.87</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>1.84</td>
<td>3.92</td>
<td>6.79</td>
<td>12.48</td>
<td>21.19</td>
<td>117.39</td>
<td>4,913.06</td>
<td>868,147.37</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>2.11</td>
<td>4.42</td>
<td>7.83</td>
<td>17.10</td>
<td>38.83</td>
<td>1,271.90</td>
<td>7,652.87</td>
<td>69,014,979.77</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>2.68</td>
<td>4.37</td>
<td>7.10</td>
<td>11.53</td>
<td>125.24</td>
<td>12,527.83</td>
<td>74,573.45</td>
<td>1,021,142.13</td>
<td></td>
</tr>
</tbody>
</table>
Exhibit 3.2: $1/(1+r)^n$ Discount Factors

<table>
<thead>
<tr>
<th>Periods</th>
<th>1.00%</th>
<th>1.50%</th>
<th>2.00%</th>
<th>2.50%</th>
<th>5.00%</th>
<th>10.00%</th>
<th>12.00%</th>
<th>15.00%</th>
<th>20.00%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n=1$</td>
<td>0.99</td>
<td>0.99</td>
<td>0.98</td>
<td>0.95</td>
<td>0.91</td>
<td>0.89</td>
<td>0.87</td>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>$n=2$</td>
<td>0.98</td>
<td>0.97</td>
<td>0.96</td>
<td>0.95</td>
<td>0.91</td>
<td>0.83</td>
<td>0.80</td>
<td>0.76</td>
<td>0.69</td>
</tr>
<tr>
<td>$n=3$</td>
<td>0.97</td>
<td>0.96</td>
<td>0.94</td>
<td>0.93</td>
<td>0.90</td>
<td>0.75</td>
<td>0.71</td>
<td>0.66</td>
<td>0.58</td>
</tr>
<tr>
<td>$n=4$</td>
<td>0.96</td>
<td>0.94</td>
<td>0.92</td>
<td>0.91</td>
<td>0.88</td>
<td>0.68</td>
<td>0.64</td>
<td>0.57</td>
<td>0.48</td>
</tr>
<tr>
<td>$n=5$</td>
<td>0.95</td>
<td>0.93</td>
<td>0.91</td>
<td>0.89</td>
<td>0.86</td>
<td>0.75</td>
<td>0.66</td>
<td>0.51</td>
<td>0.43</td>
</tr>
<tr>
<td>$n=6$</td>
<td>0.94</td>
<td>0.91</td>
<td>0.90</td>
<td>0.88</td>
<td>0.86</td>
<td>0.75</td>
<td>0.66</td>
<td>0.51</td>
<td>0.43</td>
</tr>
<tr>
<td>$n=7$</td>
<td>0.93</td>
<td>0.90</td>
<td>0.87</td>
<td>0.84</td>
<td>0.81</td>
<td>0.74</td>
<td>0.67</td>
<td>0.57</td>
<td>0.48</td>
</tr>
<tr>
<td>$n=8$</td>
<td>0.92</td>
<td>0.89</td>
<td>0.85</td>
<td>0.82</td>
<td>0.79</td>
<td>0.71</td>
<td>0.64</td>
<td>0.55</td>
<td>0.46</td>
</tr>
<tr>
<td>$n=9$</td>
<td>0.91</td>
<td>0.87</td>
<td>0.84</td>
<td>0.81</td>
<td>0.78</td>
<td>0.71</td>
<td>0.64</td>
<td>0.55</td>
<td>0.46</td>
</tr>
<tr>
<td>$n=10$</td>
<td>0.90</td>
<td>0.85</td>
<td>0.82</td>
<td>0.79</td>
<td>0.76</td>
<td>0.70</td>
<td>0.63</td>
<td>0.54</td>
<td>0.45</td>
</tr>
<tr>
<td>$n=11$</td>
<td>0.89</td>
<td>0.84</td>
<td>0.81</td>
<td>0.78</td>
<td>0.76</td>
<td>0.70</td>
<td>0.64</td>
<td>0.55</td>
<td>0.46</td>
</tr>
<tr>
<td>$n=12$</td>
<td>0.87</td>
<td>0.77</td>
<td>0.73</td>
<td>0.68</td>
<td>0.65</td>
<td>0.63</td>
<td>0.55</td>
<td>0.47</td>
<td>0.39</td>
</tr>
<tr>
<td>$n=13$</td>
<td>0.86</td>
<td>0.77</td>
<td>0.72</td>
<td>0.68</td>
<td>0.65</td>
<td>0.63</td>
<td>0.55</td>
<td>0.47</td>
<td>0.39</td>
</tr>
<tr>
<td>$n=14$</td>
<td>0.85</td>
<td>0.74</td>
<td>0.71</td>
<td>0.66</td>
<td>0.65</td>
<td>0.62</td>
<td>0.54</td>
<td>0.45</td>
<td>0.37</td>
</tr>
<tr>
<td>$n=15$</td>
<td>0.84</td>
<td>0.71</td>
<td>0.69</td>
<td>0.64</td>
<td>0.61</td>
<td>0.55</td>
<td>0.47</td>
<td>0.39</td>
<td>0.31</td>
</tr>
<tr>
<td>$n=16$</td>
<td>0.83</td>
<td>0.69</td>
<td>0.66</td>
<td>0.61</td>
<td>0.58</td>
<td>0.51</td>
<td>0.43</td>
<td>0.35</td>
<td>0.29</td>
</tr>
<tr>
<td>$n=17$</td>
<td>0.82</td>
<td>0.65</td>
<td>0.62</td>
<td>0.58</td>
<td>0.53</td>
<td>0.45</td>
<td>0.37</td>
<td>0.30</td>
<td>0.25</td>
</tr>
<tr>
<td>$n=18$</td>
<td>0.81</td>
<td>0.61</td>
<td>0.59</td>
<td>0.55</td>
<td>0.48</td>
<td>0.39</td>
<td>0.30</td>
<td>0.24</td>
<td>0.19</td>
</tr>
<tr>
<td>$n=19$</td>
<td>0.80</td>
<td>0.57</td>
<td>0.54</td>
<td>0.50</td>
<td>0.44</td>
<td>0.32</td>
<td>0.25</td>
<td>0.19</td>
<td>0.14</td>
</tr>
<tr>
<td>$n=20$</td>
<td>0.79</td>
<td>0.54</td>
<td>0.50</td>
<td>0.45</td>
<td>0.38</td>
<td>0.26</td>
<td>0.19</td>
<td>0.13</td>
<td>0.09</td>
</tr>
<tr>
<td>$n=21$</td>
<td>0.78</td>
<td>0.50</td>
<td>0.45</td>
<td>0.39</td>
<td>0.33</td>
<td>0.19</td>
<td>0.14</td>
<td>0.09</td>
<td>0.07</td>
</tr>
<tr>
<td>$n=22$</td>
<td>0.77</td>
<td>0.45</td>
<td>0.41</td>
<td>0.34</td>
<td>0.28</td>
<td>0.12</td>
<td>0.08</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>$n=23$</td>
<td>0.76</td>
<td>0.40</td>
<td>0.36</td>
<td>0.29</td>
<td>0.23</td>
<td>0.06</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>$n=24$</td>
<td>0.75</td>
<td>0.35</td>
<td>0.31</td>
<td>0.24</td>
<td>0.19</td>
<td>0.03</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>$n=25$</td>
<td>0.74</td>
<td>0.30</td>
<td>0.26</td>
<td>0.18</td>
<td>0.13</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>$n=26$</td>
<td>0.73</td>
<td>0.25</td>
<td>0.21</td>
<td>0.13</td>
<td>0.08</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$n=27$</td>
<td>0.72</td>
<td>0.19</td>
<td>0.16</td>
<td>0.10</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$n=28$</td>
<td>0.71</td>
<td>0.14</td>
<td>0.11</td>
<td>0.07</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$n=29$</td>
<td>0.70</td>
<td>0.09</td>
<td>0.06</td>
<td>0.04</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$n=30$</td>
<td>0.69</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

At times, it is necessary to convert a stream of many future sums, each received at different times, into a single present value equivalent. The method for this is straightforward enough. Each of the individual future payments is separately discounted to its present value and then the present values are summed. The method is the same whether one is discounting future costs or future benefits.

There is a special formula that simplifies the computation of the value of a stream of payments when the payments are the same in each future period. Such a stream is called an “annuity” and its present value is:
\[ V = A \left[ 1 - (1 + r)^{-n} \right]/r \]

where \( A \) is the amount of each periodic payment. Although the general annuity formula is complex to derive and difficult to remember accurately, it is worth committing to memory the special case of an annuity that goes on forever, i.e., a so-called “perpetuity.” The present discounted value of \( SA \) per period received forever is simply:

\[ V = A/r \]

Although the perpetuity may seem to have little real-world applicability, it is actually quite useful as a quick “rule of thumb” way of estimating the value of a stream. Remember that the value of a perpetuity of \( SA \) always constitutes an upper bound on how much a shorter stream could be worth. In fact, for annuities of any considerable length, the degree of overestimate is surprisingly small, especially for higher interest rates: for instance, at a 10% rate a 30-year annuity of $1 is worth $9.48 as compared to a perpetuity’s $10; at a 20% rate it is worth $4.98 compared to $5.00 for a perpetuity.

The interest rate itself is, of course, the transformation rate between present and future money or, quite simply, the price of having money earlier in time. Lenders will require a higher-than-ordinary interest rate to the extent that they must bear certain risks regarding repayment. One of these is the risk of outright non-payment or default by the borrower. Thus, poor risks must generally pay higher interest rates because the “cost” of lending to them is in effect higher. Another risk borne by lenders is that the money returned on their investment will have depreciated in terms of purchasing power because of the impact of inflation. For this reason, expectations about inflation tend to cause the market interest rate to fluctuate in order to maintain an adequate “real” rate of interest.

**QUESTIONS**

1. Construct a rough graph with dollar values on the vertical axis and time on the horizontal axis. Draw a curve that you think reflects the “growth path” of $1 invested at 5% compound interest and also at 10% interest. Check your results against Exhibit 3.1. Make sure that you understand why the curves have the shape that they do.

2. Construct another similar graph and plot a curve that shows the present discounted value of $1.00 received at varying times in the future. Do this for 10% and 20% interest rates. How do your rough plots compare with the values in Exhibit 3.2? Is there any relationship between the shapes of the present curves and the curves in the previous question?

3. Suppose that you had an 18-year old client who has been seriously injured in an accident. The tortfeasor’s insurance company offers $100,000 per year for life to the victim in lieu of the $1,000,000 lump-sum judgment (net of litigation costs) that would be certainly achievable by going through to final judgment. Your client finds the offer attractive. What is your advice? (If the answer depends on facts not given, say what they are.)

4. Look at the general formula for an annuity in the text above. Can you see why the formula for a perpetuity is so much simpler? What is the common sense of why the perpetuity formula is a close approximation of the true annuity value at reasonably long terms? And why is it an even closer approximation when the interest rates get higher?
B. VALUING THE BENEFITS OF INTERESTS IN LAND

1. DETERMINING THE VALUE OF A LEASEHOLD

In eminent domain condemnations, the taking of land under long-term lease frequently causes confusion as to the allocation of damages between the owner of the fee simple and the lessee. In the following segment of *Kentucky Department of Highways v. Sherrod*, 367 S.W.2d 844, at 850 (Ky.1963), the Commissioner quite properly is troubled by the damage calculation procedure adopted in another case:

As related to the leased parcel, and the determination of the respective damages of the landowners and the lessees, the evidence in this case was directed in a loose way towards application of the method outlined in *City of Ashland v. Price*, Ky., 318 S.W.2d 861. The key to this method is “to ascertain the present fair rental value, compare it with the rent stipulated in the contract and allow the aggregate difference for the period of the unexpired term of the lease” (318 S.W.2d 863).

Aside from the fact that the foregoing method furnishes no criterion or basis for determining the lessee’s damages where only part and not all of the leased property is taken, we have come to the conclusion that the method is completely unsound, unfair, and unworkable.

Let us use an illustration that is not far from the actual facts of this case: The property as whole, if sold free and clear of the lease, would have a value of $100,000. The fair rental value would be $12,000 per year (there is testimony in this case that 12 rent is fair for commercial property). The lessees have contracted to pay only $6,000 per year. Their lease has 20 years to run. The state condemns the entire tract. By applying the method outlined in the Ashland case, the lessees would be damaged $120,000 ($6,000 per year for 20 years) or $20,000 more than the whole property was worth. This is so patently absurd as to establish beyond any question the fallaciousness of the method. The foregoing illustration is not at all inappropriate because in the instant case the lessees were paying $7,200 a year rent; they said the fair rental was $14,400 a year; their lease had 15 years to run; therefore their damage was $108,000. Yet they testified that the leased property as a whole was worth only $120,000 (one said it was worth only $105,000) and the state took less than one-fourth of the area and none of the structures.

QUESTIONS

1. Is the approach in *City of Ashland v. Price* “completely unsound and unworkable”? What is really wrong with the procedure?

2. Assume that the state took the entire property (and leasehold). On the other facts of the above case, about how much should have been awarded to the lessees as compensatory damages?

2. VALUING A PROPERTY ENCUMBERED BY A LEASE

Raines Land Co. is the owner of 70 acres of commercially zoned land on the outskirts of a small Virginia city. Fashion Shopping Centers, Inc. has entered into negotiations with Raines for a 99-year lease on the 70 acres. Under the proposed lease, Fashion will construct $30,000,000 worth of buildings and other permanent improvements on the property. Title to all such improvements vests in Raines immediately, but Fashion has full use and certain other specified rights, including that of subletting, during the term of the lease. Financial arrangements are that Fashion makes an $800,000 lump sum payment immediately, followed by monthly payments of $16,667 for the period of the lease. Fashion bears all normal expenses of ownership (e.g., maintenance of buildings, etc.) except for payment of any applicable local real property taxes, the latter being the responsibility of the landowner. The reversion value of the land and improvements 99 years hence is estimated at $3,000,000.
Local real estate taxes are levied at the rate of 2% of the assessed value of land and improvements. In turn, state statutes require assessed value to be based on “fair market value.” State courts have construed fair market value to be the price that would be paid in an arms-length transaction by “a willing buyer and a willing seller.”

Raines has other current offers to sell the land for $40,000 per acre. It believes that it could secure a long-term return of 10% on any investments made currently. If you believe that other relevant facts have been omitted, please make plausible assumptions.

**QUESTIONS**

1. Briefly relate the common sense of why the market value of any asset, such as the title to a land parcel, can be expected to be approximately equal to the present discounted value of the future flow of net receipts derivable from owning the asset. Net receipts should be understood as gross receipts (rental payments, etc.) minus any costs attributable to the owner (taxes, operating costs, etc.)

2. If Raines executes the lease, collects the $800,000, and then attempts to sell the property, what is your estimate of how much it could be sold for? Why?

3. The county tax assessor expects to bill Raines for $650,000 on an assessed valuation of $32,500,000 for the land and improvements. As Raines’ attorney, would you advise him that he is liable for this full amount? Why?

4. Is Fashion liable for any part of the real estate taxes? Assume that the relevant state statute reads as follows:
   
   * * * For purposes of this chapter, and other provisions of law relating to the assessment of real estate for taxation, the term “taxable real estate” shall, include a leasehold interest in every ease in which the land or improvement, or both, as the case may be, are exempt from assessment or taxation to the owner. [Va. Code § 58-758.]

   Except for the above section, the state code appears to make no provision for the taxation of a leasehold interest.

5. The “Bottom Line.” What do you think of the proposed lease overall? Is it a good deal for Raines? For Fashion?

6. Should the value of a leasehold be taxable as real estate? What are the consequences of taxing or not taxing it?

**C. INTEREST ON DAMAGES AND PENALTIES**

1. **PREJUDGMENT INTEREST ON DAMAGES**

   Since damages that would in principle accrue only in the future are subjected to “discounting” to present values, the awarding of interest on damages that ought in principle to have been paid at some past date may seem logically symmetrical. Nonetheless, the law has traditionally been hostile to the award of prejudgment interest on damages, especially when the claim is “unliquidated” or the cause of action sounds in tort rather than contract. The case of Busik v. Levine below represents a trend by state courts to award prejudgment interest on damages more liberally than heretofore.

**BUSIK V. LEVINE**

307 A.2d 571 (Supreme Court of New Jersey, 1973)

WEINTRAUB, C.J.
These cases involve the validity of R. 4:42-11(b), adopted on December 21, 1972, which authorizes prejudgment interest in tort actions.

* * *

Although it is said with respect to interest on claims that “in this country interest is generally of statutory origin,” Consolidated Police and Firemen’s Pension Fund Commission v. City of Passaic, 130 A.2d 377, 382 (N.J.,1957), the contrary is true in our State. The Legislature has dealt with usury; that is, it has fixed the upper limit of the interest for which an ordinary loan may be made, but there is no statute dealing with interest upon other obligations or claims or with interest upon judgments. The controlling rules have always been and remain judge-made.

And, briefly, these are the rules which the courts of this State developed: As to interest upon judgments, the practice permitted collection at the “legal” rate, that is, at the rate permitted to be contracted under the usury statute to which we have referred. With respect to prejudgment interest, our courts of law assessed interest, if the demand was liquidated, at the same “legal” rate on the assumption that the creditor could have earned such interest if his obliger had paid him what was due, see Jersey City v. O’Callaghan, supra, 41 N.J.L. at 354, but declined to allow interest on claims that were unliquidated. The justice of that limitation has been questioned, as we will develop in a moment. Our courts of equity allowed or withheld interest or fixed the rate as justice dictated. [Citations omitted.]

In this setting we adopted R.4:42-11 which reads:

**Interest: Rate on Judgments; in Tort Actions.**

(a) Rate. Judgments, awards and orders for the payment of money and taxed costs shall bear interest at 6% per annum from the date of entry, except as otherwise ordered by the court.

(b) Tort Actions. In tort actions, including products liability actions, the court shall include in the judgment interest at 6% per annum on the amount of the award from the date of the institution of the action or from a date 6 months after the date of the tort, whichever is later. The contingent fee of an attorney shall not be computed on the interest so included in the judgment. * * *

We turn then to the merits. Interest is not punitive, Wilentz v. Hendrickson, 135 N.J.Eq. 244, 255-257 (1944); here it is compensatory, to indemnify the claimant for the loss of what the moneys due him would presumably have earned if payment had not been delayed. We mentioned earlier the judge-made limitation that interest should not be allowed if the claim was unliquidated. That limitation apparently rested upon the view that a defendant should not be deemed in default when the amount of his liability has not been adjudged. But interest is payable on a liquidated claim when liability itself is denied, even in good faith, Kamens v. Fortugno, 108 N.J.Super. 544, 552-553 (Ch.Div.1970). The fact remains that in both situations the defendant has had the use, and the plaintiff has not, of moneys which the judgment finds was the damage plaintiff suffered. This is true whether the contested liability is for a liquidated or for an unliquidated sum. For that reason, the concept of a “liquidated” sum has often been strained to find a basis for an award of interest.

It is said there is now, in general, a willingness to allow interest on unliquidated claims as justice may dictate. 22 Am.Jur.2d, Damages, Section 181, pp. 259-260. In upholding the retrospective application of a New York statute providing for interest in contract actions upon unliquidated damages, the United States Supreme Court observed that “The statutory allowance is for purpose of securing a more adequate compensation by adding an amount commonly viewed as a reasonable measure of the loss sustained through delay in payment,” and that “It has been recognized that a distinction, in this respect, simply as between cases of liquidated and unliquidated damages, is not a sound one.” Funkhouser v. J.B. Preston Co., 290 U.S. 163, 168 (1933).

So also, a refusal to allow interest in tort matters has been criticized. See Moore-McCormack Lines v. Amirault, 202 F.2d 893 (1 Cir. 1953). It is questioned whether justice is thereby done as between parties
to the suit. But beyond their interest, there is also a public stake in the controversy, for tort litigation is a major demand upon the judicial system. Delay in the disposition of those cases has an impact upon other litigants who wait for their turn, and upon the tax-payers who support the system. And here there is a special inducement for delay, since generally the claims are covered by liability insurance, and when payment is delayed, the carrier receives income from a portion of the premiums on hand set aside as a reserve for pending claims. See *In re Insurance Rating Board*, 55 N.J. 19, 258 A.2d 892 (1969). Hence prejudgment interest will hopefully induce prompt defense consideration of settlement possibilities. In that meaningful way, prejudgment interest bears directly upon the judicial machinery and the problems of judicial management. It is this facet, added to the consideration of justice between the litigants, which warrants our holding that prejudgment interest be payable in these matters.

The proposition we thus accept is not uniquely ours. There are a number of States which so provide by statute. [Citations omitted.] We add, parenthetically, that it is of no moment that there the principle was established by Legislature. As we have already noted, the subject of interest on claims rests wholly in case law in our State, and, assuming as we do for the moment that the issue is one of substantive law, the power and responsibility of the judiciary to deal with the subject in the absence of a statute cannot be questioned.

We see no strength in the assertion that the allowance of interest duplicates some element of damage or constitutes a payment with respect to damages not yet experienced. The jury is not instructed to add interest to its verdict in tort cases. In any event an instruction to the jury can obviate the risk. And with respect to the criticism that a verdict may embrace losses not yet suffered, the answer is that a verdict necessarily anticipates future experience, and the interest factor simply covers the value of the award for the period during which the defendants had the use of the moneys to which plaintiffs are found to be entitled. We think the equities are met when the date for the commencement of liability for interest is fixed as set forth in paragraph (b) of R.4:42-11 (the rule here under attack).

* * *

Finally we add that we see no merit in the proposition advanced by the intervener, that the usual insurance policy, which provides for payment of interest upon a judgment, does not cover prejudgment interest and therefore we should not impose such liability. It is enough to say the carrier’s obligation to pay the judgment plainly includes the obligation to pay the constituent elements of damage incorporated in that judgment, among which, of course, is the item of prejudgment interest.

* * *

The final question is whether the rule denies equal protection because it is limited to tort actions.

We see no serious issue. To begin with, tort actions are a distinctive class of litigation having a special impact upon the administration of the judicial system in terms of volume. They are distinctive too insofar as overall there is insurance coverage permitting to invest for gain portions of the premiums allocated to the reserve for pending claims. Further, problems of this kind may be met in stages or on the basis of intensity. Hence it is no solid objection that a rule, whether announced in a judicial opinion or in a rule of court, does not go as far as it might. See *New Jersey Chapter, American Institute of Planners v. New Jersey State Board of Professional Planners*, 48 N.J. 581 (1967). Finally the adoption of the rule does not foreclose holders of other liquidated claims from contending the circumstances which attend their scene are so like the circumstances here involved that interest should also be allowed to them.

* * *

The judgments are affirmed.

CONFORD, P.J.A.D., Temporarily Assigned, dissenting.
Were it necessary to decide whether the substance of the rule should be embraced judicially by the Court as a matter of substantive law, I would conclude in the negative, at least as to the rule as drawn. The primary objection to it is its mandatory nature. There would be no serious objection to the principle of prejudgment interest if its application were subject to the discretion of the trial court in the particular case, as it continues to be the rule on interest in equity and in relation to unliquidated damages in contract cases. [Citations.] The justice of making the defendant pay the claimant for the use of the money declared due during the pendency of the action is as pertinent to the contract or equitable obliger as to the tort obliger. Yet the sanction of interest remains discretionary in the former instances. And for good reason! Experience teaches that in many situations the winner of a money award is nevertheless denied costs by the court because of equities the other way. Why disarm the trial judge of discretion to deny, or to allow only partial prejudgment interest in tort cases when a comparable balance of equities points in that direction?

As one of numerous illustrations which might be afforded, there might be intervening appeals between institution of action and final judgment, inordinately protracting the interest-payment period under the rule, yet where the occasion for the appeal was not attributable to acts of the defendant but to those of the trial court or the plaintiff. Defendant may even have prevailed on such appeals. Must interest be nevertheless inexorably awarded the plaintiff for the entirety of the intervening period before final judgment without the tempering discretion of the judge? Or suppose the defendant tenders at the outset a settlement offer which the plaintiff unreasonably refuses. Under the rule as framed this is irrelevant to the accrual of the interest.

The main rationale advanced for the rule is that it merely transfers to the plaintiff interest which defendant or his insurer was drawing on money which theoretically was due the plaintiff the moment the tort took place or the action instituted. But not all tort defendants are insured; many are insured for substantially less than the trial award; and many, who are not insured at all or are insufficiently insured, may be relatively impecunious. Torts encompass not only automobile and product liability cases, but defamation, malicious prosecution, interference with economic opportunity, household negligence and miscellaneous other causes of action where insurance may well be non-existent, spotty or inadequate, and where defenses may be advanced in good faith. In such instances the defendant may not have actually earned interest on the award during the intervening period, or on the total amount of the award, and allowance of prejudgment interest on the whole award could well operate oppressively or unfairly. Surely discretion should be reposed in the trial court in all cases to control the matter.

It is difficult to resist the impression that the impelling motivation for adoption of the rule was clearance of trial dockets in automobile and kindred insured-tort situations by imposing coercive pressure on insurance companies to settle cases early. But beyond the observation above that others than insurers may be penalized by the rule, it does not seem fair to assume, as does the rule, that only defendants or their insurers are responsible for unreasonable failures of litigants to arrive at pretrial settlements. The rule may well operate to encourage unreasonable recalcitrance on the part of some plaintiffs. It is a blunderbuss which strikes its objects indiscriminately and without necessary regard to the justice of its effect in particular cases. In my view it does not, as drawn, calendar control in a justifiable manner.

QUESTIONS

1. If a claim is unliquidated, it is alleged, the defendant does not know how much is owed. Compare the situation of a liquidated claim and uncertain liability to that of certain liability but uncertain magnitude of damages. Specifically, Defendant A knows that, if he is found liable, the damages payable are a liquidated claim of exactly $100,000, but it is a “coin flip” as to whether liability will
be found. In another case, Defendant B has no doubt that he is liable for some sum, but there are equal probabilities that the judgment against him will be $50,000 or $150,000. Is there any relevant distinction with respect to the equities of charging prejudgment interest?

2. What effect, if any, would the giving of prejudgment interest have on the incentive to settle cases out of court?

3. The dissent makes the point that the defendant may not have been earning money on the funds. First of all, is this correct? If it is correct, does it matter?

2. “BACK DOOR” PREJUDGMENT INTEREST?

In many jurisdictions, an attorney may confront a statute that appears to bar interest on past damages. All is not necessarily lost, however, as the following case demonstrates. As you read this case, reflect on what legal advice might be useful to a client as a similar situation enters its very early, pre-litigation stages.

UNITED TELECOMMUNICATIONS, INC. v. AMERICAN TELEVISION AND COMMUNICATIONS CORP.

536 F.2d 1310 (10th C., 1976)

[Telephone company which exchanged its interest in a cable television company for shares of defendant corporation brought action for breach of contract and breach of warranty in that defendant failed to use its best efforts to register its shares. Plaintiff was therefore unable to sell the shares and was deprived of the use of the business capital that would have been secured from the anticipated sale. The U.S. District Court for the District of Colorado entered judgment for plaintiff and defendant appealed. Among other allegations of error, defendant alleged that it was error for the court to charge the jury that United was entitled to recover interest costs or borrowing costs paid as a result of ATC’s breach of the agreement.]

WILLIAM E. DOYLE, Circuit Judge.

This is a diversity action in which the appellant was defendant in the trial court and suffered the award of a judgment based on a jury verdict in the amount of $2,021,500. The action brought by plaintiff-appellee was for a breach of contract and breach of warranty. The case was filed December 5, 1973 and was tried in March 1975.

* * *

At the trial United presented testimony which showed the amount of its borrowing and the nexus between these costs and the breach of ATC. United also presented testimony showing that ATC was aware before entering into the agreement that United had planned to convert the ATC stock to cash and would use that cash to reduce its short-term debts. Evidence was also offered to show that parties were aware that registration was necessary in order to have a successful public sale of the stock and that they were also aware that a private sale could be obtained only at a significant discount. For this element of damage United claimed the sum of $1.6 million.

The trial court explained to the jury that the plaintiff was claiming that it suffered an additional direct loss “by being unable to reduce its short term debt by the amount of money which it would have received on the sale to the public of the registered shares of ATC stock.” The trial court also explained that the claim was that the cost of continued borrowing was a part of actual damages from the breach. The court then continued:
“As to this claim, you must find from a preponderance of the evidence that this was a natural and probable consequence of the claimed breach of contract by the defendant and also that at the time the parties entered into the contract, the defendant reasonably could have anticipated from the facts and circumstances which it then knew or should have known, that these damages would probably be incurred by the plaintiff if the defendant breached the contract.”

ATC claims that this instruction was erroneous. Its argument is that the instruction allowed United to recover prejudgment interest, and that applicable Colorado law bars such interest. United responds that this is not prejudgment interest in that it is not the kind of interest contemplated by the Colorado statutes and, instead, it is, according to United, borrowing costs that are recoverable like any other element of contract damage.

In general, we agree with the United argument and with the position taken by the trial court that interest as an item of damages or loss related to the substantive claim differs from interest measured by a percentage of the judgment. In our view the kind of interest that was here awarded is not barred or limited by the Colorado interest Statute and the Colorado cases. Here our concern is with United’s borrowing costs and expenses. The Colorado statute does not deal with interest charges which arise from an independent debt owed by the plaintiff to a third party. So even on the assumption that the law of Colorado governs, it would not be the Hays v. Arbuckle [72 Colo. 328, 211 P. 101 (1922)] line of cases together with the interest statute that apply.

Had United claimed damages of $8 million for lost market value stemming from the stock not being registered together with interest measured by a percentage of that loss, the claim would have been one for prejudgment interest subject to the limitations of the Colorado law. When United claims, however, that it is entitled to recover from ATC expenses which resulted from interest which it was obligated to pay on money which it was required to borrow as a result of ATC’s unwarranted delays and manipulations, a wholly different theory is presented. This latter type of interest is not regulated by the Colorado interest statute.

So, having determined that the substantive rights here asserted by United are governed by general damage principles, we need not pursue the Hays v. Arbuckle doctrine as to the choice of law in the area of prejudgment interest. Contractual rights such as we are here considering are under Colorado law governed by the place where the contract was made and the place designated by the parties as governing. See Western Enterprises, Inc. v. Robo Sales, Inc., 470 P.2d 931 (Colo. App., 1970). In this case, that is North Carolina.

Finally, we consider whether the North Carolina court would allow recovery of borrowing costs and expenses as an element of damages. We conclude that it would and does. North Carolina recognizes that a party injured by a breach of contract has a right to be compensated “insofar as this can be done by money.” Also, “he is entitled to be placed in the same position he would have occupied if the contract had been performed.” Perfecting Service Co. v. Produce Dev. & Sales Co., 131 S.E.2d 9, 21 (1963). Accord, Tillis v. Calvin Cotton Mills, Inc., 251 N.C. 359, 111 S.E.2d 606 (1959). Indeed, North Carolina has specifically endorsed the doctrine of Hadley v. Baxendale, 9 Ex. 341 (1854). See Perkins v. Langdon, 237 N.C. 159, 74 S.E.2d 634, 643 (1953), holding that a defendant is liable for damages based upon proximate cause and foreseeability, including damages flowing from “special circumstances known and communicated.” [Citations omitted.] The instruction given by the court in which foreseeability is the guide is in accord with North Carolina law on the subject. ***

Cases from other jurisdictions do give credence to the proposition here approved. See New Amsterdam Cas. Co. v. Mitchell, 325 F.2d 474 (5th Cir. 1963) (additional interest on construction loan because of builder’s delay); Thompson v. Hanson, 491 P.2d 1065 (Wash., 1971) (semble); Herbert & Brooner Construction Co. v. Golden, 499 S.W.2d 541 (Mo.App.1973) (semble); Repinski v. Clintonville Federal S
We disagree with ATC’s contention that in its instruction on damages the court gave the jury a license to speculate. It is well settled that the difficulties in assessing damages do not preclude an injured party from recovering compensation. See Bigelow v. RKO Radio Pictures, Inc., 327 U.S. 251, 265, 66 S.Ct. 574, 90 L.Ed. 652 (1945).

The judgment of the district court is affirmed.

QUESTIONS

1. Suppose that United admitted that, had it received the funds on time, it still would have borrowed the same amount of money. E.g., suppose that they would have declared it as dividends to stockholders. Logically, should that make any difference to their damage claim?

2. As a practical matter, what does the court’s view suggest as to how a potential plaintiff should act when a cause of action for some loss arises? Are courts equally likely to recognize outlays and “opportunity costs” (foregone returns) as damages?

3. PREJUDGMENT INTEREST ON PENALTIES

As the above cases suggest, the old rules against prejudgment interest on compensatory damages in civil cases are gradually breaking down in most jurisdictions. By contrast, the right to moratory interest on penalties or other sums whose purpose is clearly punitive — receives surprisingly little attention. The general rule in federal cases derives from Rodgers v. United States, 332 U.S. 371, at 373-74 (1947):

As our prior cases show, a persuasive consideration in determining whether such obligations shall bear interest is the relative equities between the beneficiaries of the obligation and those upon whom it has been imposed. And this Court has generally weighed these relative equities in accordance with the historic judicial principle that one for whose financial advantage an obligation was assumed or imposed and who has suffered actual money damages by another’s breach of that obligation, should be fairly compensated for the loss thereby sustained. [Citations.]

The contention is hardly supportable that the Federal Government suffers money damages or loss, in the common law sense, to be compensated by interest, when one convicted of a crime fails promptly to pay a money fine assessed against him. The underlying theory of that penalty is that it is a punishment or deterrent and not a revenue-raising device; unlike a tax, it does not rest on the basic necessity of the Government to collect a carefully estimated sum of money by a particular date in order to meet its anticipated expenditures. For the foregoing reasons, this Court’s holding that a criminal penalty does not bear interest, Pierce v. United States, 255 U.S. 398, 405, 406 is consistent with its holding that the Government does suffer recoverable damages if a taxpayer fails to pay taxes when due and is therefore equitably entitled to interest. Furthermore, denial of interest on criminal penalties might well be rested on judicial unwillingness to expand punishment fixed for a criminal act beyond that which the plain language of the statute authorizes.

QUESTIONS

1. The language of the Court suggests that legislatures may have set penalties as deterrents. If interest is not added to penalties, does the amount of deterrent (or retributive punishment) associated with the same nominal fine change with circumstances?
2. Treble damages under the antitrust statutes provide an interesting mixed case of compensatory and “other” damages. Clearly, the single damages are compensatory. Various hypotheses exist as to the motivation of Congress with respect to the other two-thirds of the damage assessment. It has been suggested, for instance, that the additional damages are deterrent in nature and/or that they operate as an incentive “bounty” to plaintiffs who provide what are in effect private enforcement services. Federal courts have been generally unsympathetic to the award of prejudgment interest in antitrust cases. In addition to the Rodgers rule, many courts seem to feel that interest would result in excessive recovery:

Interest should not be allowed in antitrust actions where the statute provides for punitive damages. Treble damages compensate a plaintiff handsomely for all his losses, including the loss of the use of money rightfully his. [Trans World Airlines, Inc. v. Hughes, 308 F.Supp. 679, 696 (S.D.N.Y.1969).]

What is your own analysis of whether prejudgment interest should be granted in antitrust actions?

3. Plaintiff alleges that it was deprived of profits amounting to $2,000,000 during fiscal 1981. If prejudgment interest were permitted on any part of the damages, would 17% presumptively be the proper rate if plaintiff had contracted $2,000,000 or more of medium term indebtedness at that rate during 1981?

4. If plaintiff files for and recovers the full $2,000,000 trebled, what is the latest year in which it can receive the damages (after litigation, possible appeal, etc.) and still be fully compensated for its loss if no prejudgment interest is awarded?

5. Assume that defendant has the same rate of return on capital as plaintiff. In contemplating the original offense, defendant had reason to believe that the odds against its being successfully prosecuted and fined were three to two (three out of five). Would a change in the rule on prejudgment interest be likely to affect its conduct?

6. Under what circumstances might the interest rates of plaintiff and defendant differ materially? In such cases, which rate should be used if prejudgment interest were given?

7. Describe the possible rationale for a “bounty” to antitrust plaintiffs in terms of an externalities or “public goods” effect of bringing such suits.

D. EVALUATING LOST FUTURE INCOME AS DAMAGES

Exhibits 3.3 and 3.4 represent the types of damage calculations sometimes presented by forensic economists in wrongful death or personal injury cases wherein an important component of damages is the lost future income of the victim. It is instructive to consider the types of predictive input that underly such calculations and the legitimate bases for argument about the conclusions reached.

Columns (1) and (2) are self-explanatory. Column (3) reflects a projection of future income in terms of dollars of current purchasing power. Such constant purchasing power measures of income are frequently called “real” income, as opposed to “nominal” income measured in dollars that may be depreciated in value due to future inflation. Both Exhibits assume a base year income of $20,000 and project future gains in real income at a rate of 1% compounded annually. This rate of increase is presumably an estimate of gains in the individual’s productivity, the scarcity of his skills, etc.

Column (4) is a key difference between the two sets of calculations. The actual dollars earned in any future year will to some extent be determined by the rate of overall inflation in prices in the economy. Hence, column (4) adjusts the constant purchasing power income upward by an additional growth factor to account for the influence of inflation. Exhibit 3.3 assumes a 7% rate of inflation while Exhibit 3.4 assumes a higher rate of 10%. In effect, nominal income grows by 8% annually (1% real and 7%
inflation) in Exhibit 3.3 and 11% (1% real and 10% inflation) in Exhibit 3.4. Note that the alternative assumptions about inflation rates produce substantially different damage calculations.

Columns (5), (6), and (7) reflect the process of “discounting” the lost future income (in nominal dollars) to its present value equivalent. Thus, if \( r \) stands for the decimal equivalent of the column (5) term interest rate available in a year \( n \) years in the future, as indicated by column (1), then column (6) is the appropriate discount factor \( 1/(1+r)^n \) for the \( n \)th period. Next, column (7) is calculated as the product of the column (6) discount factor times the lost income from column (4).

Finally, column (8) aggregates the present values of the damages for each future year. If it were known, for instance, that the victim had been deprived of exactly 13 years worth of future income, then the relevant damage total as computed in Exhibit 3.3 would be $184,345.70. By contrast, the higher rate of inflation assumed in Exhibit 3.4 yields damages of $219,579.13.

### Exhibit 3.3: Lost Income Damages, Assumption Set #1

<table>
<thead>
<tr>
<th>Future Year</th>
<th>Constant Income</th>
<th>Inflated Income</th>
<th>Investment Yield %</th>
<th>Discount Factor ( 1/(1+r)^n )</th>
<th>Present Value</th>
<th>PV Sum</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60</td>
<td>20,200.00</td>
<td>21,600.00</td>
<td>10.50</td>
<td>0.90</td>
<td>19,547.51</td>
<td>19,547.51</td>
</tr>
<tr>
<td>2</td>
<td>61</td>
<td>20,402.00</td>
<td>23,328.00</td>
<td>11.10</td>
<td>0.81</td>
<td>18,899.46</td>
<td>38,446.97</td>
</tr>
<tr>
<td>3</td>
<td>62</td>
<td>20,606.02</td>
<td>25,194.24</td>
<td>12.37</td>
<td>0.70</td>
<td>17,756.20</td>
<td>56,203.17</td>
</tr>
<tr>
<td>4</td>
<td>63</td>
<td>20,812.08</td>
<td>27,209.78</td>
<td>13.47</td>
<td>0.60</td>
<td>16,413.48</td>
<td>72,616.66</td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>21,020.20</td>
<td>29,386.56</td>
<td>14.00</td>
<td>0.52</td>
<td>15,262.46</td>
<td>87,879.12</td>
</tr>
<tr>
<td>6</td>
<td>65</td>
<td>21,230.40</td>
<td>31,737.49</td>
<td>14.00</td>
<td>0.46</td>
<td>14,459.17</td>
<td>102,338.29</td>
</tr>
<tr>
<td>7</td>
<td>66</td>
<td>21,442.71</td>
<td>34,276.49</td>
<td>14.00</td>
<td>0.40</td>
<td>13,698.16</td>
<td>116,036.45</td>
</tr>
<tr>
<td>8</td>
<td>67</td>
<td>21,657.13</td>
<td>37,018.60</td>
<td>14.00</td>
<td>0.35</td>
<td>12,977.21</td>
<td>129,013.66</td>
</tr>
<tr>
<td>9</td>
<td>68</td>
<td>21,873.71</td>
<td>39,980.09</td>
<td>14.00</td>
<td>0.31</td>
<td>12,294.20</td>
<td>141,307.86</td>
</tr>
<tr>
<td>10</td>
<td>69</td>
<td>22,092.44</td>
<td>43,178.50</td>
<td>14.00</td>
<td>0.27</td>
<td>11,647.13</td>
<td>152,954.99</td>
</tr>
<tr>
<td>11</td>
<td>70</td>
<td>22,313.37</td>
<td>46,632.78</td>
<td>14.00</td>
<td>0.24</td>
<td>11,034.13</td>
<td>163,989.12</td>
</tr>
<tr>
<td>12</td>
<td>71</td>
<td>22,536.50</td>
<td>50,363.40</td>
<td>14.00</td>
<td>0.21</td>
<td>10,453.38</td>
<td>174,442.50</td>
</tr>
<tr>
<td>13</td>
<td>72</td>
<td>22,761.87</td>
<td>54,392.47</td>
<td>14.00</td>
<td>0.18</td>
<td>9,903.20</td>
<td>184,345.70</td>
</tr>
<tr>
<td>14</td>
<td>73</td>
<td>22,998.48</td>
<td>58,743.87</td>
<td>14.00</td>
<td>0.16</td>
<td>9,381.98</td>
<td>193,727.69</td>
</tr>
<tr>
<td>15</td>
<td>74</td>
<td>23,219.38</td>
<td>63,443.38</td>
<td>14.00</td>
<td>0.14</td>
<td>8,888.19</td>
<td>202,615.88</td>
</tr>
<tr>
<td>16</td>
<td>75</td>
<td>23,451.57</td>
<td>68,518.85</td>
<td>14.00</td>
<td>0.12</td>
<td>8,420.39</td>
<td>211,036.28</td>
</tr>
<tr>
<td>17</td>
<td>76</td>
<td>23,686.09</td>
<td>74,000.36</td>
<td>14.00</td>
<td>0.11</td>
<td>7,977.22</td>
<td>219,013.49</td>
</tr>
<tr>
<td>18</td>
<td>77</td>
<td>23,922.95</td>
<td>79,920.39</td>
<td>14.00</td>
<td>0.09</td>
<td>7,557.36</td>
<td>226,570.85</td>
</tr>
<tr>
<td>19</td>
<td>78</td>
<td>24,162.18</td>
<td>86,314.02</td>
<td>14.00</td>
<td>0.08</td>
<td>7,159.61</td>
<td>233,730.46</td>
</tr>
<tr>
<td>20</td>
<td>79</td>
<td>24,403.80</td>
<td>93,219.14</td>
<td>14.00</td>
<td>0.07</td>
<td>6,782.79</td>
<td>240,513.25</td>
</tr>
</tbody>
</table>

As might be imagined, the presentation of such damage calculations is frequently disputed quite vigorously at trial. Before reading further, you should reflect on how differences in the combinations of assumptions underlying the calculations will affect the ultimate damage figure. Also, it is worth attempting to determine exactly how exhibits such as Exhibits 3.3, above, and 3.4, below, are computed.
If you have access to a calculator, check a few of the entries in each Exhibit in order to test your understanding of the underlying calculations.

**Exhibit 3.4: Lost Income Damages, Assumption Set #2**

<table>
<thead>
<tr>
<th>Year</th>
<th>Age</th>
<th>Constant $ Income</th>
<th>Inflated Income</th>
<th>Yield %</th>
<th>Discount Factor</th>
<th>Present Value</th>
<th>Cumulative Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60</td>
<td>20,200.00</td>
<td>22,200.00</td>
<td>10.50</td>
<td>0.90498</td>
<td>20,090.50</td>
<td>20,090.50</td>
</tr>
<tr>
<td>2</td>
<td>61</td>
<td>20,402.00</td>
<td>24,642.00</td>
<td>11.10</td>
<td>0.81016</td>
<td>19,964.01</td>
<td>40,054.51</td>
</tr>
<tr>
<td>3</td>
<td>62</td>
<td>20,606.02</td>
<td>27,352.62</td>
<td>12.37</td>
<td>0.70477</td>
<td>19,277.37</td>
<td>59,331.88</td>
</tr>
<tr>
<td>4</td>
<td>63</td>
<td>20,812.08</td>
<td>30,361.41</td>
<td>13.47</td>
<td>0.60322</td>
<td>18,314.61</td>
<td>77,646.49</td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>21,020.20</td>
<td>33,701.16</td>
<td>14.00</td>
<td>0.51937</td>
<td>17,503.33</td>
<td>95,149.82</td>
</tr>
<tr>
<td>6</td>
<td>65</td>
<td>21,230.40</td>
<td>37,408.29</td>
<td>14.00</td>
<td>0.45559</td>
<td>17,042.71</td>
<td>112,192.53</td>
</tr>
<tr>
<td>7</td>
<td>66</td>
<td>21,442.71</td>
<td>41,523.20</td>
<td>14.00</td>
<td>0.39964</td>
<td>16,594.22</td>
<td>128,786.76</td>
</tr>
<tr>
<td>8</td>
<td>67</td>
<td>21,657.13</td>
<td>46,090.76</td>
<td>14.00</td>
<td>0.35056</td>
<td>16,157.53</td>
<td>144,944.29</td>
</tr>
<tr>
<td>9</td>
<td>68</td>
<td>21,873.71</td>
<td>51,160.74</td>
<td>14.00</td>
<td>0.30751</td>
<td>15,732.33</td>
<td>160,676.62</td>
</tr>
<tr>
<td>10</td>
<td>69</td>
<td>22,092.44</td>
<td>56,788.42</td>
<td>14.00</td>
<td>0.26974</td>
<td>15,318.32</td>
<td>175,994.95</td>
</tr>
<tr>
<td>11</td>
<td>70</td>
<td>22,313.37</td>
<td>63,035.15</td>
<td>14.00</td>
<td>0.23662</td>
<td>14,915.21</td>
<td>190,910.16</td>
</tr>
<tr>
<td>12</td>
<td>71</td>
<td>22,536.50</td>
<td>69,969.01</td>
<td>14.00</td>
<td>0.20756</td>
<td>14,522.71</td>
<td>205,432.86</td>
</tr>
<tr>
<td>13</td>
<td>72</td>
<td>22,761.87</td>
<td>77,665.60</td>
<td>14.00</td>
<td>0.18207</td>
<td>14,140.53</td>
<td>219,573.39</td>
</tr>
<tr>
<td>14</td>
<td>73</td>
<td>22,989.48</td>
<td>86,208.82</td>
<td>14.00</td>
<td>0.15971</td>
<td>13,768.41</td>
<td>233,341.80</td>
</tr>
<tr>
<td>15</td>
<td>74</td>
<td>23,219.38</td>
<td>95,691.79</td>
<td>14.00</td>
<td>0.14010</td>
<td>13,406.08</td>
<td>246,747.88</td>
</tr>
<tr>
<td>16</td>
<td>75</td>
<td>23,451.57</td>
<td>106,217.89</td>
<td>14.00</td>
<td>0.12289</td>
<td>13,053.29</td>
<td>259,801.17</td>
</tr>
<tr>
<td>17</td>
<td>76</td>
<td>23,686.09</td>
<td>117,901.85</td>
<td>14.00</td>
<td>0.10780</td>
<td>12,709.78</td>
<td>272,510.96</td>
</tr>
<tr>
<td>18</td>
<td>77</td>
<td>23,922.95</td>
<td>130,871.06</td>
<td>14.00</td>
<td>0.09456</td>
<td>12,375.32</td>
<td>284,886.27</td>
</tr>
<tr>
<td>19</td>
<td>78</td>
<td>24,162.18</td>
<td>145,266.87</td>
<td>14.00</td>
<td>0.08295</td>
<td>12,049.65</td>
<td>296,935.92</td>
</tr>
<tr>
<td>20</td>
<td>79</td>
<td>24,403.80</td>
<td>161,246.23</td>
<td>14.00</td>
<td>0.07276</td>
<td>11,732.55</td>
<td>308,668.48</td>
</tr>
</tbody>
</table>

**QUESTIONS**

1. What evidence is available to suggest how many years of future income have been lost? How controversial is the use of such evidence likely to be?

2. On what basis can any assumption about real income growth be supported? What kind of expert would be best qualified to do this projection, and what facts would he be relying upon?

3. How does one know what interest rates to use in the discounting process? The exhibits provided above use the “term” rates for each relevant maturity interval. Is that correct? Should a “risk free” rate, such as that on government securities, be used even if the plaintiff would predictably choose to invest in securities that impose some additional risk in exchange for a higher rate of return?

4. Would an economic expert be able to predict the future rate of inflation? With what accuracy? Would different experts be likely to disagree greatly on the future rate of inflation?

5. Compare the relative reliability of the different assumptions underlying tables such as Exhibits 3.3 and
6. Under all of the other assumptions of Exhibit 3.3, how would the damages change if the rate of growth in real income were assumed to be 4% instead of 1%? [No calculations will be necessary here if you carefully compare Exhibits 3.3 and 3.4.]

BEAULIEU v. ELLIOTT
434 P.2d 665 (Supreme Court of Alaska, 1967)

DIMOND, Justice.

As a result of an automobile accident on April 13, 1963, James Elliott suffered a fracture dislocation of his right ankle. He brought this action for damages against Richard Beaulieu, and the issue of damages was termed by the court without a jury. The trial court filed findings of fact and conclusions of law and entered judgment awarding Elliott $169,937.25 compensatory damages, costs of $82.40, and attorney’s fees in the amount of $13,870.29.

* * *

The trial court did not reduce the amount it found as damages for future impairment of earning capacity to present value. Instead, the court stated that “The interest rate reduction and decline in purchasing power of the dollar is off-set by pay increases plaintiff could have expected in the future from his military service.” Beaulieu contends that the failure to reduce the damages to present value was prejudicial error.

The general principle underlying the assessment of damages in tort cases is that an injured person is entitled to be replaced as nearly as possible in the position he would have occupied had it not been for the defendant’s tort. In the case of impairment of future earning capacity, it is reasoned that a failure to reduce damages to present value would be to place the injured person in a better position than he would have occupied except for the defendant’s tort, because the injured person would get all of his future wages long in advance and would be able to invest the lump sum and realize earnings on such investment during the intervening period.36 For this reason—that money has the power to earn money—it has become the generally accepted rule that damages awarded for future loss of earnings should be reduced to present worth.37

In applying the general rule, the Supreme Court of Washington has stated a formula for reducing awards of future earnings to present value which involves the “rate of interest (which) could fairly be expected from safe investments which a person of ordinary prudence, but without particular financial experience or skill, could make in that locality.” Wentz v. T.E. Connolly, Inc., 273 P.2d at 492. This formula, although empirical at best, is probably as definite as any that has been devised. But we believe that the rule for reducing awards, including the formula applied by the Washington court, ignores facts which should not be ignored. Annual inflation at a varying rate is and has been with us for many years. There is no reason to expect that it will not be with us in the future. This rate of depreciation offsets the interest that could be earned on government bonds and many other “safe” investments. As a result the plaintiff, who through no fault of his own is given his future earnings reduced to present value must, in order to realize his full earnings and not be penalized by reduction of future earnings to present value, invest his money in enterprises, other than those which are considered “safe” investments, which promise

---

36 McCormick, Damages, Section 86, at 304 (1935).
37 Wentz v. T.E. Connolly, Inc., 45 Wash.2d 127, 273 P.2d 485, 491 (1954); Borcherding v. Eklund, 55 N.W.2d 643, 650 (Neb., 1952); Daughtry v. Cline, 30 S.E.2d 322, 324 (N.C., 1944); Rigley v. Prior, 290 Mo. 10, 233 S.W. 828, 832 (Mo., 1921); Restatement, Torts, Section 924 comment d, at 634-35 (1939); McCormick, Damages, Section 86, at 304 (1935); Annots., 77 A.L.R. 1439, 1446 (1932) 154 A.L.R. 796, 797 (1945).
a return in interest or dividends greater than the offsetting rate of annual inflation. But ours is a competitive economy. By their very nature some enterprises backed by investors’ money are going to fail with resulting loss to individuals. Thus, instead of being assured of earnings at rates greater than the annual rate of inflation, the injured plaintiff stands a chance of entirely losing his future earnings by unlucky or unwise investments. Since the plaintiff, through the defendant’s fault and not his own, has been placed in the position of having no assurance that his award of future earnings, reduced to present value, can be utilized so that he will ultimately realize his full earnings, we believe that justice will best be served by permitting the trier of fact to compute loss of future earnings without reduction to present value. The plaintiff is more likely to be restored to his original condition under the rule we adopt than under the prevailing rule which calls for a discounting of the award for future earnings.

Our conclusion is fortified by another factor which also may not be ignored. This is the factor, relied upon by the trial judge, which involves wage increases that the injured plaintiff might have expected to receive in the future had he not been injured. It is a matter of common experience that as one progresses in his chosen occupation or profession he is likely to increase his earnings as the years pass by. In nearly any occupation a wage earner can reasonably expect to receive wage increases from time to time. This factor is generally not taken into account when loss of future wages is determined, because there is no definite way of determining at the time of trial what wage increases the plaintiff may expect to receive in the years to come. However, this factor may be taken into account to some extent when considered to be an offsetting factor to the result reached when future earnings are not reduced to present value. Thus, if there is any fear that failure to reduce the present value will give the plaintiff more than he is entitled to because of the possibility of his making successful investments of the sum awarded at returns greater than the annual rate of inflation, such fear is obviated by the fact that the award may well be deficient in that it does not take into account probable wage increases that the plaintiff would ordinarily be expected to receive in the future.

Elliott testified that he would receive disability retirement pay from the Air Force in the amount of $191.00 a month for the remainder of his life. Beaulieu contends that the trial judge committed prejudicial error in refusing to deduct the net present value of future retirement pay from the award for future loss of earnings. Beaulieu’s argument is that to allow Elliott damages for future wage loss, in addition to his retirement pay, is to unjustly enrich Elliott by allowing him double compensation for his injuries.

The general principle underlying the assessment of damages in tort cases is that an injured person is entitled to be replaced as nearly possible in the position he would have occupied had it not been for the defendant’s tort. Elliott has been in the Air Force for about 18 years at the time of his discharge and he testified that he had intended to remain the service for at least 20 years. If he had not been injured, Elliott could have continued to earn to his full capacity and, in addition, after 20 years’ service, would have been entitled to retire and draw retirement pay. By reason of his injuries, Elliott was entitled under law to be retired early for disability and draw retirement pay in lieu of retirement on a regular basis after completion of 20 years’ service. The award of damages for impaired earning capacity has the effect of putting Elliott in the same position he would have occupied had it not been for the injury, because the damages represent what Elliott could have earned had he not been injured and the disability retirement pay represents that which Elliott had earned and become entitled to under law by reason of his years of service in the Air Force. In other words, Elliott now receives an amount representing wages he could have earned were it not for the injury, plus retirement pay; had he not been injured, he would have received the full wages he could have earned during his remaining work life, in addition to receiving the retirement pay to which he would become entitled by reason of his years of service in the Air Force. Thus, Elliott, under the court’s award, is getting no more than he would have gotten had he not been injured. The disability retirement pay Elliott is receiving should not be used to mitigate damages and reduce the award for loss of future earnings.
Beaulieu argues that the trial judge erred in failing to deduct from the damages awarded for impairment of future earning capacity an amount representing income taxes that Elliott would have had to pay on future income.

The courts are divided on this question. It is the more general view, supported by a majority of American decisions, that an amount representing future income taxes should not be deducted from the award. Annot., 63 A.L.R.2d 1393, 1396 (1959). As was stated by the Supreme Court of Rhode Island:

This view has been adopted by the various courts on diverse grounds but primarily on the ground that the quantum of such taxation is of necessity in the realm of conjecture.


We adopt the majority rule. Income tax rates, provisions relating to deductions and exemptions, and other aspects of income tax laws and regulations are so subject to change in the future that we believe that a court cannot predict with sufficient certainty just what amounts of money a plaintiff would be obliged to pay in federal and state income taxes on income that he would have earned in the future had it not been for a defendant’s tortious conduct. We hold that a damage award for impairment of earning should not be reduced by an estimated amount representing income taxes that the injured party may be required to pay on future income. In awarding damages to Elliott for impaired earning capacity, the court did not err in failing to take income tax consequences into consideration.

The rule we adopt has no application, however, as to the court’s award of past wages in the amount of $10,000. The reason for the rule-inability to predict with sufficient certainty what taxes would have to be paid—does not exist here, because taxes on income earned prior to trial can be easily calculated based on income tax laws and regulations as they existed at the time the wages would have been earned. The court erred in failing to deduct from the award for past loss of wages the income taxes Elliott would have had to pay had he earned the amount awarded prior to the trial.

Elliott testified that he had not lost any military pay or allowances between the date of the accident in April 1963 and the date of his military discharge in January 1966. During that period of time Elliott was either hospitalized or on leave, except for the period January to August, 1964, when he was on duty status. The trial court awarded $10,752.85 for a partial past wage loss covering the period from the date of the accident to the day of Elliott’s discharge from the Air Force, but excepting the period between January and August, 1964, when Elliott was on duty status.

Beaulieu contends that this award for past wages was error. His argument in essence is that the general principle underlying the assessment of damages in tort cases is that the injured person is entitled to be replaced as nearly as possible in the position he would have occupied had it not been for the defendant’s tort, and that since Elliott suffered no loss of wages during the period involved he should be awarded none.

In arguing that the award should be sustained, Elliott urges the adoption of the collateral source rule, which provides that damages may not be diminished or mitigated on account of payments received by plaintiff from a source other than the defendant. Bell v. Primeau, 104 N.H. 227, 183 A.2d 729, 730 (1962). The collateral source rule is followed in most jurisdictions. Annot., 7 A.L.R.3d 516, 520 (1966). We applied this rule as to workmen’s compensation benefits in Ridgeway v. North Star Terminal & Stevedoring Co., 378 P.2d 647, 650 (Alaska 1963). We apply the rule in this instance. By entering the military service, Elliott in effect agreed to perform certain duties and functions in exchange for certain benefits to be given him by the government. One of those benefits was that he was to receive military pay and allowances during periods of physical incapacity from performing his duties. This was in the nature of a contractual arrangement between Elliott and the government when he became a member of the armed forces.
forces, and which he may have paid for by accepting wages lower than those he might have obtained from
the performance of like duties in civilian life. The income that Elliott received from the government is not
the result of earnings, but of such previous contractual arrangement.

Such a contractual arrangement was made for Elliott’s own benefit, and not for the benefit of a tort
feasar, such as Beaulieu. The latter has no right to claim the benefit of such an arrangement by having the
damages awarded against him reduced by the amount that Elliott was paid by the government during the
period of his disability. The trial court did not err in awarding damages for loss of wages during the
period of Elliott’s disability while he was still in the military service.

The court awarded Elliott $71,244 for pain and sufferi ng that he would experience for the remainder of
his life. Beaulieu contends that the evidence does not justify such an award. * * * We agree with the
foregoing. As we stated in Patrick v. Sedwick, 413 P.2d 169, 176 n. 21 (Alaska 1966), there is no fixed
measure of compensation in awarding damages for pain and suffering, and such an award necessarily
rests in the good sense and deliberate judgment of the tribunal assigned by law to ascertain what is just
compensation. We can see nothing manifestly unfair or unjust about the method used by the trial court in
assessing damages for future pain and suffering. In fact, as we suggested in the dissenting opinion in the
Kansas case of Caylor v. Atchison, Topeka & Santa Fe R. Co., 374 P.2d 53, 64 (1962), it appears to be a
fair argument and a rational approach to treat damages for pain the way it is endured-day by day, month
by month, year by year. Ultimately, however, the question for decision is whether the total sum is
reasonable or not, regardless of how it was arrived at. We find no error in the method used by the trial
court in awarding damages for future pain and suffering.

* * *

Beaulieu contends that the total sum awarded is unreasonable and is grossly excessive. We shall not set
aside an award on a claim of excessiveness unless it is so large as to strike us that it is manifestly unjust,
such as being the result of passion or prejudice or a disregard of the evidence or rules of law. Considering
the evidence of permanent damage to Elliott’s ankle, the osteomyelitis, and the pain and suffering that he
is likely to endure for the remainder of his life, it is our opinion that the award for future pain and
suffering was not manifestly unjust. And Beaulieu did not contend in his brief or in oral argument that the
trial court acted through passion or prejudice.

Beaulieu contends that the court erred in not reducing the future pain and suffering award to present
value. He relies principally on the case of Affett v. Milwaukee & Suburban Transp. Corp., 11 Wis.2d 604,
106 N.W.2d 274, 279 (1960), where the court, after disapproving of the use of a mathematical formula for
computing damages for pain and suffering, said: “Logically, if this method were followed, the gross
amount arrived at should be discounted to its present worth.”

If an award for future pain and suffering must be reduced to present value when a mathematical
formula is used, it must be for the same reason that an award for future earnings is discounted under the
prevailing rule-- i.e., because the plaintiff receives his damages for the future long in advance and is able
to invest the sum awarded and realize earnings during the intervening period. But we have held that as to
impairment of future earning capacity, the award should not be reduced to present value. The same
reasoning applies here as to an award for future pain and suffering. Because of the annual rate of inflation
offsetting dividends or interest that may be expected on “safe” investments, and of the risk of loss
involved in making other investments, a plaintiff is more likely to be restored to his original condition had
defendant not committed his tort by allowing the plaintiff his award for future pain and suffering without
reduction to present worth.

[At this point, the court commenced a long discussion of factual findings which, it determined, were not
adequate to permit a clear understanding by the reviewing court as to the basis for some aspects of the
disability calculus. Hence, the case was remanded to superior court for purposes of clarifying certain
specific factual issues.]
The final and most significant question raised by this appeal is whether and to what extent an award for lost future wages should be adjusted because of inflation. If an adjustment for inflation is to be made, two approaches are available. Under the first method the projection of current wages is increased to reflect the fact that the employee would have received cost of living increases to keep pace with inflation. The sum of those increased annual wage figures is then discounted to present value by the prevailing interest rate. See United States v. English, 521 F.2d 63, 75 (9th Cir. 1975). Under the second method the discount rate is decreased below the prevailing interest rate to reflect the estimated rate of inflation. This reduced discount rate is then applied to the sum of the current wages projected to continue for each of the years for which wages will be lost. See Feldman v. Allegheny Airlines, Inc., 382 F.Supp. 1271 (D.Conn.1974), aff’d, 524 F.2d 384, 387-88 (2d Cir. 1975). If the same estimated inflation rate is used, the “outcome of the calculation under either approach would be very nearly identical,” Feldman, supra, 524 F.2d at 391 (Friendly, J., concurring).

In this case the District Court took inflation into account, to an unspecified extent, apparently by using the second approach and adjusting for future inflation by decreasing the present value discount. The Court did not provide a computation for its award for loss of future earnings; however, Judge Duffy stated that he had reached his result by “considering both probability of continued inflation and a discount to present value.” Doca infers that Judge Duffy used a discount factor of 1%.\(^4\) Defendants object to this computation for lack of evidence to support an estimate of future inflation and, more fundamentally, on the ground that any adjustment for inflation is impermissible. Since the significance of inflation on damage awards for lost future wages seems to be as persistent a problem as inflation itself, the matter merits some examination.

In this Circuit, the issue has been considered but not resolved. In McWeeney v. New York, New Haven, & Hartford R.R. Co., 282 F.2d 34 (2d Cir. 1960), the issue was whether the calculation of lost future wages should be made on the basis of gross wages or net wages after deduction for income taxes. In rejecting a deduction for taxes, but see Norfolk and Western Railway v. Liepelt, 444 U.S. 490 (1980), this Court noted that whatever benefit this gives to plaintiffs is partially offset by the absence of an upward adjustment for inflation. A decade later in Yodice v. Koninklijke Nederlandsche Stoomboot Maatschappij, 443 F.2d 76 (2d Cir. 1971), a damage award for lost wages was reversed for reasons that included the failure to apply any discount for present value. Judge Friendly presciently noted that “if inflation should continue at its present pace, courts may have to reconsider the propriety of the long recognized charge with respect to discount.”

\(^4\) Doca derives this 1% discount rate in the following manner. The difference between the award for lost wages, $325,560, and Doca’s projection of future lost wages, $394,861, is $42,301. This difference represents approximately 1% of $352,560, multiplied by the 12 years for which future wages will be lost. Expressed as a formula, the computation would have been \(x + (12i)x = w\), where \(x\) is the discounted value, \(i\) is the interest rate, and \(w\) is the projection of total future wages. If the District Court applied a 1% rate in this fashion (and we cannot be certain if the award for lost future wages was determined on this basis), the discount factor was incorrectly applied. The correct method is to apply to the future wages for each year a discount factor of \(1/(1 + i)^n\), where \(i\) is the interest rate and \(n\) is the year for which present value of income is sought. The discounted amounts for each year are then added together to determine the total award.
More recently this Court, in a diversity case, approved District Judge Blumenfeld’s use of an “inflation-adjusted discount rate” of 1.5%. *Feldman v. Allegheny Airlines, Inc.,* supra, 524 F.2d at 387. This rate was determined by subtracting the average annual change in the consumer price index over each of several periods of years from the effective annual interest on government bonds held during each of the same periods, and then selecting the average of the rates resulting from these subtractions. The purpose of this method was to determine the “real yield” of money, the portion of interest charged on virtually risk-free investments that represents only the real cost of the money and not the additional cost the lender exacts as a hedge against future inflation. Judge Friendly concurred *dubitante,* noting that the decision represented a prediction as to unsettled state law and not “a precedent on the inflation problem in a case arising under federal law.” In another diversity case, *Dullard v. Berkeley Associates Co.,* 606 F.2d 890, 896 (2d Cir. 1979), this Court was willing to assume the propriety of an inflation-adjusted discount rate only for purposes of ruling that even with such an adjustment applicable to lost future wages, the total award in question was excessive. Most recently, the Court, in a Federal Tort Claims Act case, approved use of a present value discount rate that had been reduced to 5% from the prevailing interest rate to reflect future inflation. *Espana v. United States,* 616 F.2d 41 (2d Cir. 1980).

Among other courts, there has been a growing recognition that inflation should be taken into account in damage calculations. At least four federal Courts of Appeals have explicitly so ruled, see *Steckler v. United States,* 549 F.2d 1375-78 (10th Cir. 1977); *United States v. English,* supra, 521 F.2d at 72-76; *Riha v. Jasper Blackburn Corp.,* 516 F.2d 840, 843-45 (8th Cir. 1975); *Bach v. Penn Central Transportation Co.,* 502 F.2d 1112, 1122 (6th Cir. 1974), as have a number of state courts, see e.g., *Beaulieu v. Elliott,* 434 P.2d 665 (Alaska 1967); *Schnebly v. Baker,* 217 N.W.2d 708 (Iowa 1974); *Halliburton Co. v. Olivas,* 517 S.W.2d 349 (Tex.Civ.App.1974). In addition, the commentators generally favor this position. [Citations.]

Though the Supreme Court has yet to rule definitively on the issue, it clearly indicated its approval of including inflation estimates in future wage loss awards in *Norfolk and Western Railway v. Liepelt,* supra. The issue in that case was whether wage loss should be estimated without deduction for estimated income taxes. The Court ruled that such a reduction should be made. The opinion is instructive on the issue of inflation for two reasons. First, the Court noted that the plaintiff’s economic expert had estimated that the lost future earnings would have increased by five percent a year until the wage-earner’s retirement. More significantly, the Court rejected the argument that reduction for future income taxes would be too speculative by noting that other aspects of a damage award computation, specifically including “future inflation,” are also matters of “estimate and prediction.” 100 S.Ct. 758. Thus, the Court not only noted that the award included an estimate for inflation, but also used the reasonableness of making such an estimate as support for similar estimates of future taxes. * * *

Although the view that inflation should be considered is clearly on the increase, several federal courts including three Courts of Appeals, have prohibited the adjustment of awards for future wages to account for inflation. See *Johnson v. Penrod Drilling Co.,* 510 F.2d 234 (5th Cir. 1975) (en bane); *Magill v. Westinghouse Electric Corp.,* 464 F.2d 294, 299-301 (3d Cir. 1972); *Williams v. United States,* 435 F.2d 804 (1st Cir. 1970). The principal rationale for these decisions is that estimates of future inflation are too speculative to serve as basis for calculating damages.

We believe that the developing majority view is sound and that inflation should be considered in estimating the present value of lost future wages. According to the figures computed by the United States Department of Labor, inflation has become a constant feature of our modern economy. The Consumer Price Index, perhaps the leading indicator of inflation has increased in all but two years since the beginning of World War II, and in every year since 1955. This increase was particularly pronounced during the past decade, when the Index increased by five times as much as it had during the previous
decade. In 1979, the year in which the District Court decided this case, the average annual Consumer Price Index was 217.4, compared to 100 in 1967, the base year, and 72.1 in 1950.\footnote{Thus, if Doca had received a damage award for twelve years’ future wages in 1967, fully discounted to present value at then prevailing interest rates, and had invested this award at those prevailing interest rates with the intention of drawing out one-twelfth each year, his yearly amount, despite the interest, would have declined steadily in purchasing power. For the final year, 1979, the money would have been worth only 42\% of what it was worth when the award was made.}
In short, courts cannot fail to recognize that inflation is a dominant factor on the current economic scene and, despite episodic recessions, is likely to be so for the foreseeable future.

To be sure, the idea that a person’s future income will increase in dollar amount is a prediction, one based largely on our inductive assumption that the future will resemble the past. But using inductive inferences about economic conditions to adjust the value of an award for lost future income is hardly novel. In fact, it underlies the practice of discounting future lost wages to the present value of the right to receive such wages. This practice, first required by the Supreme Court in 1916, *Chesapeake & Ohio Ry. v. Kelly*, 241 U.S. 485 (1916), and routinely followed by courts throughout the nation ever since, rests on the inductive inference that interest rates will continue at their present value. If interest rates decline after the lump sum award is received, the plaintiff will normally fail to realize the full dollar amount of the lost future wages. If it is permissible, indeed required by *Kelly*, to make a prediction about interest rates, it seems as reasonable to make a prediction about inflation as well. See *Steckler*, supra, 549 F.2d at 1377.

In fact, as Judge Blumenfeld demonstrates by the data collected in *Feldman*, there is a fairly constant relationship between interest and inflation rates, so that it is more reasonable to make a prediction about the relationship of both rates than about the level of interest rates alone. Discounting without regard to inflation charges the plaintiff for that portion of the prevailing cost of money that represents the lenders’ anticipation of inflation without allowing the plaintiff an offsetting addition for inflation, either by increasing the sum to be discounted or reducing the discount rate. Or, to put in another way, discounting necessarily includes a prediction about inflation (the prediction made by those who determine the interest rate), and it is neither reasonable nor fair to use that prediction only to reduce an award instead of endeavoring to determine an award that approximates the present value of what the future wages would have been. Economics should not be an instrument for the undercompensation of plaintiffs. It should be used to achieve a realistic estimate of future conditions, so that an injured plaintiff can be compensated for his loss as fairly and completely as possible.

The principal argument advanced against taking inflation into account is that it is too speculative. There can be no doubt that predicting next year’s inflation rate is at least as hazardous a task as forecasting next year’s weather. This concern about speculation persuades us to be rather cautious in determining the way in which inflation should be taken into account, but, for three reasons, is not sufficient to preclude any adjustment for inflation. First, the law of damages frequently recognizes the need to make some predictions about the future, despite the uncertainties of doing so. Estimates of lost future profits are an obvious example. See *Autowest, Inc. v. Peugeot, Inc.*, 434 F.2d 556, 564-67 (2d Cir. 1970). Second, since discounting future lost wages to present value uses an interest rate that reflects inflation, the issue is not

---

38 At a minimum the plaintiff will earn compound interest only at the reduced prevailing interest rate. If the plaintiff invests his lump-sum award (plus earned interest) annually, he will also suffer a decrease of simple interest because of the reduced rates. It is possible for the plaintiff to avoid this latter reduction by investing his lump-sum award in long-term bonds, corresponding to the length of time for which the wages are lost. Yet if the plaintiff makes that type of investment, he takes the risk that if interest rates increase, the market value of his bonds will decrease, and he will suffer a loss each time he sells a bond to meet current needs. One way to avoid that risk is to purchase a series of bonds ranging in maturities from one year up to the number of years for which the wages were lost. That assures that each year’s needs will be met from the proceeds of a maturing bond, without the risk of selling a bond at a discount. However, the average rate of such a series of short, medium, and long-term bonds will most likely be less than the rate on long-term bonds, which was used to discount the future wages to present value.

39 It is worth noting that when the Supreme Court declared its present value discount rule in 1916, it was able to look back on an era when prices, though fluctuating, were not generally increasing. The estimated Consumer Price index for 1916 stood at 32.7, having varied between 25 and 51 during the course of the previous century. Looking backward over the past four decades from 1980, rather than 1916, reveals an entirely different economic landscape, in which prices have increased almost every year, often at an accelerating rate.
whether inflation is too speculative to be considered at all; it is whether inflation, as a component of interest rates, should be considered for the defendant (by discounting at the prevailing interest rate) but ignored for the plaintiff (by rejecting any compensating adjustment for inflation). Finally, and most important, as Feldman illustrates, it is entirely feasible to take inflation into account without making any prediction as to the specific level of future inflation rates. All that is needed is a prediction that in the future inflation rates will bear approximately the same relationship to long-term interest rates that they have in the past. Essentially the only premise for that prediction is that over the course of years investors in long-term bonds will demand returns that reflect the prospects of future inflation. They may guess wrong in some years and accept rates that fail to keep pace with inflation. “Yet common sense suggests,” as Judge Friendly has pointed out, “that investors will not tolerate such a situation indefinitely.” Feldman, supra, 524 F.2d at 392. These considerations persuade us that inflation is not too speculative to be taken into account in determining awards for future lost wages. See Norfolk and Western Railway v. Liepelt, supra.

The next question is how inflation should be considered. Courts have favored various methods, which can generally be divided into two basic categories. In the first category are those methods that submit the question of the inflation rate to the factfinder, so that a separate determination is made in each case. In this category, one method treats the rate of future inflation as a fact to be proved by evidence, presumably expert opinion. See, e.g., United States v. English, supra, 521 F.2d at 75-76. Another method forbids evidence, but permits the fact-finder to apply his own knowledge and prediction in estimating a future inflation rate. See, e.g., Bach v. Penn Central Transportation Co., supra, 502 F.2d at 1122-23. A third method permits the fact-finder to apply his own knowledge but also permits expert testimony. See, e.g., Riha v. Jasper Blackburn Corp., supra, 516 F.2d at 845. In the other basic category are methods that remove the issue from the fact-finder and instead adopt a rule of law that focuses on a constant relationship between inflation and interest rates. Using this approach, the Alaska Supreme Court held that the inflation rate will be assumed to equal the interest rate, thereby eliminating any discount for present value. See Beaulieu v. Elliott, supra. A somewhat different technique was used in Feldman, where it was observed that over a period of years the inflation rate bore a relatively constant relationship to the interest rate, averaging 1.5% less; the Court used this 1.5% rate as a discount rate, reflecting the true cost of money with the effect of inflation removed.

Having considered these alternatives, we are not prepared to require any one particular method by which inflation should be taken into account in estimating lost future wages. Predictive techniques are constantly being refined, as are methods of correctly assessing the immediate past. If litigants prefer to offer evidence as to future rates of both inflation and interest, they are entitled to do so. We note, however, that one virtue of the Feldman approach is that it lessens trial disputes concerning the future impact of inflation. The average accident trial should not be converted into a graduate seminar on economic forecasting. The adjusted discount rate approach of Feldman avoids all predictions about the level of future inflation, and focuses instead only on the relationship between the inflation rate and the interest rate. Since that relationship is relatively constant in periods of stable inflation, as the historical data set forth in Feldman reveal, it may frequently be possible for litigants to agree on an adjusted discount rate, derived from that type of historical data. Without requiring a particular adjusted rate, we suggest that a 2% discount rate would normally be fair to both sides.

A 2% discount rate appears to be the estimate of the true cost of money appropriate for use in a computation whose purpose is to determine the present value of lost future wages. As Judge Blumenfeld recognized in Feldman, 2% is the real yield normally attainable during periods of low, stable rates of inflation. Feldman, supra, 383 F.Supp. at 1239, 1310 (Table A-II). It is also a figure that lies within the
narrow range bracketed by many economists as representing the true yield of money.\footnote{Although economists disagree over the validity of the assumption that the real rate of interest is constant and consequently independent of inflation, cf. Fama, Interest Rates and Inflation: The Message in the Entrails, 67 Amer.Econ.Rev. 487 (1977) (real estate constant) with Carlson, Short-Term Interest Rates as Predictors of Inflation: A Comment, 67 Amer.Econ.Rev. 469 (1977) (real estate varies), there is substantial opinion that during periods of stable rates of inflation, the real yield of money, whether constant or slightly fluctuating, is approximately 2%, as estimated in Feldman. As Sharpe has suggested, an appropriate period for determining investor expectations of inflation and real rates of return is 1959-1972, when actual real returns may have been fairly close to expected real returns. W. Sharpe, Investments 170 (1978). For this period the Feldman calculation approximates 2%. Moreover, studies by economists, though taking different positions concerning the relation between the real rate and expectations about inflation, have estimated that the real rate of interest was between 1.5% and 35 for a similar time frame. See, e.g., Carlson, supra at 471 (expected real rate fluctuates around 2.5% during 1953-1975); Elliott, Measuring the Expected Real Rate of Interest: An Exploration of Macroeconomic Alternatives, 67 Amer.Econ.Rev. 429, 440 (1977) (mean expected real rate 1.44% for 1960-1974); Gibson, Interest Rates and Inflationary Expectations: New Evidence, 62 Amer.Econ.Rev. 854, 856 (1972) (real rate between 2% and 3% during 1952-1970); But cf. W. Sharpe, supra, at 170 (extrapolating from 1976 study by Ibbotson & Sinquefield, average actual real rate 1.2% for 1959-1972).} Feldman used the slightly lower rate of 1.5% to allow “a margin of error to compensate for unexpected increases in the rate of inflation which may depress long term real yields.” Id. at 1294 (footnote omitted). It may well be that during periods of unusually high inflation rates, interest rates will lag somewhat behind the inflation rate, at least temporarily, with the result that the interest return on a discounted lump sum award will not fully keep pace with inflation. But in such periods of high inflation, wages too will not keep pace with inflation.

In the only three years since 1947 when the rate of the annual increase in the Consumer Price Index (CPI) exceeded 8%, the annual rate of wage increases (non-farm weekly earnings) has been less than the annual rate of CPI increases by at least 3 percentage points; in all other years since 1947, wages have increased at a rate greater than the rate of increase within 1 percentage point of the CPI percentage increase. Thus, since there is little reason to expect that injured plaintiffs would have received wage increases sufficient to keep pace with unusually high rates of inflation, a present value discount rate normally need not be reduced below 2% just to compensate for unusually high inflation. To do so would ignore the basic objective of selecting a present sum of money that will replace what the future wages would have been.

We emphasize that we are not requiring the use of an adjusted discount rate, nor specifying that when such a rate is used, it must be set at 2%. Litigants are free to account for inflation in other ways, or, if they use the adjusted discount rate approach, to offer evidence of a rate more appropriate than 2%. But in the hope that disputes about the appropriate rate may be minimized, we simply suggest the 2% rate as one that would normally be fair for the parties to agree upon, and we authorize district judges to use such a rate if the parties elect not to offer any evidence on the subject of either inflation or present value discount.

Obviously, in this case the District Judge could not have been expected to anticipate this opinion. Nevertheless, his calculation of an award for lost future wages must be reconsidered for two reasons. First, if we assume, as the plaintiff urges, that a 1% discount rate was used, there is nothing in the evidence to support a reduction of the discount rate from prevailing interest rates all the way down to 1%. As indicated, the lowest discount rate we are willing to approve, in the absence of historical data justifying a different rate, is 2%. Second, the District Court to some extent gave duplicate consideration to the impact of inflation. In estimating what Doca’s future wages would have been, the Court took into...
account a post-trial wage increase of 80¢ an hour, specified in Doca’s union contract to take effect on October 1, 1970. Having based future lost wages on this cost of living increase, the Court then further reflected the impact of inflation by reducing the discount rate to 1%. When a discount rate is adjusted to reflect inflation, it must be applied to a stream of earnings calculated without regard to inflation, even to a cost of living increase already scheduled to occur. The certainty of such an increase does not lessen the fact that it reflects inflation, which is already accounted for by the reduced discount rate.

We therefore vacate the damage award and remand to the District Court for recomputation of the award for lost future wages in accordance with this opinion.

QUESTIONS

1. What are the courts in Beaulieu and Doca trying to do? Compare the approaches in the two cases. Which is more likely to be “right” in the sense of giving the most nearly correct quantum of damages? Is there any possible counter argument that the less exact approach is, all things considered, preferable? [For a review of recent cases on this general subject, see Note, “Inflation, Productivity, and the Total Offset Method of Calculating Damages for Lost Future Earnings,” 49 U.Chi.L.Rev. 1003 (1982).]

2. For purposes of discounting damages, what is the appropriate interest rate? At any moment of time, there are a host of different interest rates varying with the degree of risk of the creditor and the term of the loan or investment. Why is a “riskless” rate, such as the government bond rate usually used, rather than some higher rate?

3. Can you generate any argument that, regardless of the term of years by which the lump sum award anticipates future damages, the discount rate should be based upon short or intermediate-term interest rates rather than the (usually higher) long-term interest rates? What relevance does this have to the 2%/to discount rule suggested in Doca?

4. In Norfolk and Western Railway Co. v. Liepelt, 444 U.S. 490 (1980), the Supreme Court gave its approval to the consideration of the effect of taxes on the lost future income for which compensation is being given. Heretofore, damages have traditionally been calculated by discounting the stream of before-tax income to its present value. Suppose that the full effect of taxation were reflected in damage awards. Is it possible to say whether awards would be higher or lower than under the method where taxes are not considered? How difficult would it be to make the requisite adjustments? We return to this topic later in the present chapter, but it is worth forming a tentative opinion at this time.

E. COMBINED UNCERTAINTY AND DISCOUNTING

In most of the previous applications, the facts of the situations allowed uncertainty and present value adjustments to be considered separately. In this section, we consider what happens when, as is usually the case, risk and time-values are both inherent aspects of the phenomenon being analyzed.

1. INTRODUCTORY CONSIDERATIONS

The following set of exercises is designed to demonstrate a few important considerations when uncertainty and present value factors are presented in virtually inseparable fashion. One issue that is raised is the extent to which it is proper, either analytically or in equity, to rely on “average” values. These questions are designed to lead into a difficult set of real world policy issues implicating the “fairness”—or lack thereof—of recognizing certain alleged empirical differences between the sexes, races, or other distinctive sub-sets of persons.
a. Tintd Hypothetical

On the planet of Tintd, all citizens may be distinguished by the color of their fur. Only three pelt shades exist: mauve, fuschia, and calico. All workers are entitled to an annual retirement pension of 10,000 dize. Each worker must retire on the birthday marking its 42nd Tintd year and may then collect a 10,000 dize pension payment on each succeeding birthday. The life spans of Tintdians are much more predictable than that of humans, but vary systematically with fur color. All Tintdians die at precisely 27:00 hours on one of their birthdates. Exhibit 3.5 indicates the probability that a Tintdian of a particular color will live until various birthdays.

Exhibit 3.5: Tintds’ Probability of Survival to Various Birthdays

<table>
<thead>
<tr>
<th>Year</th>
<th>Mauve</th>
<th>Fuschia</th>
<th>Calico</th>
<th>PV @ 10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>9090.91</td>
</tr>
<tr>
<td>44</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>8264.46</td>
</tr>
<tr>
<td>45</td>
<td>0.500</td>
<td>1.000</td>
<td>1.000</td>
<td>7513.15</td>
</tr>
<tr>
<td>46</td>
<td>0.500</td>
<td>0.667</td>
<td>1.000</td>
<td>6830.13</td>
</tr>
<tr>
<td>47</td>
<td>0.000</td>
<td>0.333</td>
<td>0.000</td>
<td>6209.21</td>
</tr>
<tr>
<td>48</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>5644.74</td>
</tr>
</tbody>
</table>

The last column of the table indicates the discounted present value of the pension payment in the corresponding year, assuming that the valuation is being done at the retirement date and that the ruling interest rate is 10% compounded annually.

QUESTIONS

1. What are the average lengths of time (in Tintdian years) that the Mauves, Fuschias and Calicos, respectively, will collect their pensions? [Note: Half of the Mauves must die on their 44th birthday because the table indicates that, whereas all live to 44, only half are left at the 45th birthday; the rest die on their 46th birthday because the tables says that no Mauves will make it to 47. Apply a similar logic to the other types of Tintdians.]

2. Tintdian law allows pensioners to assign their life pension proceeds to third parties. Duration of the pension is, of course, determined by the lifespan of the assignor, not the assignee. If you were a Calico and wished to sell your pension rights on the day of retirement, about how much do you think you could get if the interest rate were 10%. How does this compare to the prices that Mauves and Fuschias could expect?

3. What explains the difference in the values of the Fuschia and Calico pensions?

4. How would the ratios of the three different pension values change if the interest rate were 20% instead of 10%?

5. Each worker also receives a free 20,000-dize life insurance policy. If the rights to this policy can likewise be assigned, how much are they worth in the three relevant cases?

b. Annuity Hypothetical

Assume that a $10,000 annual annuity is available to be paid in equal daily payments (of about $27.40) during each year and that any time length is possible. The interest rate is 15%.
QUESTIONS

1. Draw a diagram that has time in years on its horizontal axis and the present value of the $10,000 annual-rate annuity for various time terms on the vertical axis.

2. Next, draw a curve that you think approximately reflects the shape of a function that relates the annuity’s length to its discounted present value. (The diagram does not have to be precise, but think about what the shape of the curve should be, at least roughly.)

3. Select a point on the curve corresponding to an annuity of 10-year term. Draw a horizontal dotted line out to the vertical axis, such that the present value of the 10-year annuity is shown. Repeat the same operation for 5-year and 15-year annuities, respectively.

4. On the same diagram, indicate the value of the certainty equivalent of a lottery in which, depending on whether a flipped coin comes UD heads or tails, you win an annuity of either 5 or 15 years. (Assume risk neutrality.) Is this lottery worth more or less than the certainty of a 10-year annuity? Why?

5. How would the diagram, and the conclusions derived from it, change if the interest rate assumed were either higher or lower?

6. What analogy exists between the analysis in the preceding questions and the discussion in Chapter II above of the existence of risk aversion due to declining marginal utility of money?

c. Life Tenancy and Remainder Rights

A parcel of land has been condemned by the State and a hearing is being held by the Commissioners to award compensatory payments to those having interests in the land. A widow, under dower rights, enjoys a life tenancy on the land, whose annual rental value is presently estimated at $5,500. The Commissioners, after consulting standard mortality tables, find that the average remaining life span of a woman the widow’s age is 22 years. Hence, it is proposed to award her the present discounted value of a $5,500 stream for 22 years. The relevant interest rate is approximately 10%. The market value of the land in question, unencumbered by the life tenancy, has been set at $180,000. This sum, minus the computed value of the life tenancy, is to be given to the remainderman.

QUESTIONS

1. If you represented the remainderman, what argument would you raise to modify the proposed procedure for determining the respective compensatory payments to the life tenant and remainderman?

2. If you represented the life tenant, could you advance any argument to suggest that the proposed procedure will undercompensate your client? [Hint: Does there seem to be any inconsistency between the rental and the land value figures that have been proposed?]

3. The widow’s attorney proposes to introduce evidence that the woman in question is of exceptionally long-lived genetic stock and is in exceptional physical condition for a woman of her age. He contends that expert testimony will support a finding of an expected life span of 30 years for this particular woman. What considerations militate for and against hearing such evidence? Assess the “marginal benefits” versus the “marginal costs” of hearing such evidence. Are the issues in any way similar to Culver above?
2. AVERAGE VS. ACTUARILY EXPECTED LIFE EARNINGS

Exhibit 3.6 below provides a specific numerical example of a common error in computing lost future income damages, e.g., in wrongful death or disability cases. Under the incorrect procedure, the expected value of the victim’s remaining lifespan is first estimated, then the present discounted value of that income is computed and attributed to the victim. For instance, suppose that standard mortality tables indicate that for a 60 year old male the average number of years remaining is 17. It is then assumed as a certainty that the victim would have lived for precisely this average lifespan, neither more nor less. Then the projected annual income for each individual year within that period is discounted to its present value and cumulated. This procedure, corresponding to the leftmost eight columns of Exhibit 3.6\(^{32}\) (and similar to Exhibits 3.3 and 3.4 earlier in the chapter), indicates that the appropriate damage figure for a person who would have lived 17 years longer is approximately $219,013. (See the column 8 entry for the seventeenth year, or age 76.)

---

\(^{32}\) For spacing purposes, columns (3) through (5) are omitted from the Exhibit 3.6, but they are the same as the calculations in Exhibit 3.3 above, p.117.
Exhibit 3.6: Average Lifespan vs. Actuarial Calculations

<table>
<thead>
<tr>
<th>Year</th>
<th>Age</th>
<th>PV</th>
<th>Cumulative</th>
<th>Mortality Rate</th>
<th>Alive</th>
<th>Ex-Value</th>
<th>Cumulative Ex-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60</td>
<td>19,547.51</td>
<td>19,547.51</td>
<td>19.1</td>
<td>0.9809</td>
<td>19,174.15</td>
<td>19,174.15</td>
</tr>
<tr>
<td>2</td>
<td>61</td>
<td>18,899.46</td>
<td>38,446.97</td>
<td>21.1</td>
<td>0.9602</td>
<td>18,147.32</td>
<td>37,321.47</td>
</tr>
<tr>
<td>3</td>
<td>62</td>
<td>17,756.20</td>
<td>56,203.17</td>
<td>22.2</td>
<td>0.9389</td>
<td>16,671.06</td>
<td>53,992.53</td>
</tr>
<tr>
<td>4</td>
<td>63</td>
<td>16,413.48</td>
<td>72,616.66</td>
<td>25.0</td>
<td>0.9154</td>
<td>15,025.14</td>
<td>69,017.67</td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>15,262.46</td>
<td>87,879.12</td>
<td>26.9</td>
<td>0.8908</td>
<td>13,595.64</td>
<td>82,613.31</td>
</tr>
<tr>
<td>6</td>
<td>65</td>
<td>14,459.17</td>
<td>102,338.29</td>
<td>28.8</td>
<td>0.8651</td>
<td>12,509.13</td>
<td>95,122.45</td>
</tr>
<tr>
<td>7</td>
<td>66</td>
<td>13,698.16</td>
<td>116,036.45</td>
<td>30.8</td>
<td>0.8385</td>
<td>11,485.76</td>
<td>106,608.20</td>
</tr>
<tr>
<td>8</td>
<td>67</td>
<td>12,977.21</td>
<td>129,013.66</td>
<td>33.2</td>
<td>0.8107</td>
<td>10,519.99</td>
<td>117,128.19</td>
</tr>
<tr>
<td>9</td>
<td>68</td>
<td>12,294.20</td>
<td>141,307.86</td>
<td>36.0</td>
<td>0.7815</td>
<td>9,607.52</td>
<td>126,735.30</td>
</tr>
<tr>
<td>10</td>
<td>69</td>
<td>11,647.13</td>
<td>152,954.99</td>
<td>39.2</td>
<td>0.7508</td>
<td>8,745.06</td>
<td>135,480.77</td>
</tr>
<tr>
<td>11</td>
<td>70</td>
<td>11,034.13</td>
<td>163,989.12</td>
<td>42.6</td>
<td>0.7188</td>
<td>7,931.86</td>
<td>143,412.63</td>
</tr>
<tr>
<td>12</td>
<td>71</td>
<td>10,453.38</td>
<td>174,442.50</td>
<td>46.3</td>
<td>0.6856</td>
<td>7,166.48</td>
<td>150,579.11</td>
</tr>
<tr>
<td>13</td>
<td>72</td>
<td>9,903.20</td>
<td>184,345.70</td>
<td>50.3</td>
<td>0.6511</td>
<td>6,447.80</td>
<td>157,026.91</td>
</tr>
<tr>
<td>14</td>
<td>73</td>
<td>9,381.98</td>
<td>193,727.69</td>
<td>54.8</td>
<td>0.6154</td>
<td>5,773.70</td>
<td>162,800.61</td>
</tr>
<tr>
<td>15</td>
<td>74</td>
<td>8,888.19</td>
<td>202,615.88</td>
<td>59.8</td>
<td>0.5786</td>
<td>5,142.72</td>
<td>167,943.33</td>
</tr>
<tr>
<td>16</td>
<td>75</td>
<td>8,420.39</td>
<td>211,036.28</td>
<td>65.3</td>
<td>0.5408</td>
<td>4,553.91</td>
<td>172,497.24</td>
</tr>
<tr>
<td>17</td>
<td>76</td>
<td>7,977.22</td>
<td>219,013.49</td>
<td>71.1</td>
<td>0.5024</td>
<td>4,007.49</td>
<td>176,504.72</td>
</tr>
<tr>
<td>18</td>
<td>77</td>
<td>7,557.36</td>
<td>226,570.85</td>
<td>77.4</td>
<td>0.4635</td>
<td>3,502.71</td>
<td>180,007.44</td>
</tr>
<tr>
<td>19</td>
<td>78</td>
<td>7,159.61</td>
<td>233,730.46</td>
<td>84.4</td>
<td>0.4244</td>
<td>3,038.29</td>
<td>183,045.73</td>
</tr>
<tr>
<td>20</td>
<td>79</td>
<td>6,782.79</td>
<td>240,513.25</td>
<td>91.0</td>
<td>0.3857</td>
<td>2,616.45</td>
<td>185,662.17</td>
</tr>
<tr>
<td>21</td>
<td>80</td>
<td>6,425.80</td>
<td>246,939.04</td>
<td>98.3</td>
<td>0.3478</td>
<td>2,235.08</td>
<td>187,897.25</td>
</tr>
<tr>
<td>22</td>
<td>81</td>
<td>6,087.60</td>
<td>253,026.64</td>
<td>105.9</td>
<td>0.3110</td>
<td>1,893.21</td>
<td>189,790.46</td>
</tr>
<tr>
<td>23</td>
<td>82</td>
<td>5,767.20</td>
<td>258,793.84</td>
<td>113.7</td>
<td>0.2756</td>
<td>1,589.64</td>
<td>191,380.09</td>
</tr>
<tr>
<td>24</td>
<td>83</td>
<td>5,463.66</td>
<td>264,257.50</td>
<td>121.3</td>
<td>0.2422</td>
<td>1,323.30</td>
<td>192,703.39</td>
</tr>
<tr>
<td>25</td>
<td>84</td>
<td>5,176.10</td>
<td>269,433.60</td>
<td>128.4</td>
<td>0.2111</td>
<td>1,092.68</td>
<td>193,796.07</td>
</tr>
<tr>
<td>26</td>
<td>85</td>
<td>4,903.67</td>
<td>274,337.27</td>
<td>136.0</td>
<td>0.1824</td>
<td>894.39</td>
<td>194,690.46</td>
</tr>
<tr>
<td>27</td>
<td>86</td>
<td>4,645.58</td>
<td>278,982.86</td>
<td>144.1</td>
<td>0.1561</td>
<td>725.26</td>
<td>195,415.72</td>
</tr>
<tr>
<td>28</td>
<td>87</td>
<td>4,401.08</td>
<td>283,383.94</td>
<td>152.6</td>
<td>0.1323</td>
<td>582.25</td>
<td>195,997.97</td>
</tr>
<tr>
<td>29</td>
<td>88</td>
<td>4,169.44</td>
<td>287,553.38</td>
<td>161.6</td>
<td>0.1109</td>
<td>462.47</td>
<td>196,460.44</td>
</tr>
<tr>
<td>30</td>
<td>89</td>
<td>3,950.00</td>
<td>291,503.38</td>
<td>171.2</td>
<td>0.0919</td>
<td>363.12</td>
<td>196,823.56</td>
</tr>
<tr>
<td>31</td>
<td>90</td>
<td>3,742.11</td>
<td>295,245.49</td>
<td>181.3</td>
<td>0.0753</td>
<td>281.64</td>
<td>197,105.20</td>
</tr>
<tr>
<td>32</td>
<td>91</td>
<td>3,545.15</td>
<td>298,790.64</td>
<td>192.0</td>
<td>0.0608</td>
<td>215.59</td>
<td>197,320.78</td>
</tr>
<tr>
<td>33</td>
<td>92</td>
<td>3,358.57</td>
<td>302,149.20</td>
<td>203.4</td>
<td>0.0484</td>
<td>162.70</td>
<td>197,483.48</td>
</tr>
<tr>
<td>34</td>
<td>93</td>
<td>3,181.80</td>
<td>305,331.00</td>
<td>215.4</td>
<td>0.0380</td>
<td>120.93</td>
<td>197,604.42</td>
</tr>
<tr>
<td>35</td>
<td>94</td>
<td>3,014.34</td>
<td>308,345.34</td>
<td>228.2</td>
<td>0.0293</td>
<td>88.42</td>
<td>197,692.84</td>
</tr>
</tbody>
</table>

An alternative procedure is to recognize that the victim might have died either much earlier or later than the average time for a person having the same starting age. The probability of dying at each age can be found in standard mortality tables in the form of a “mortality rate” per thousand. These mortality rates have been entered as column (9) of Exhibit 3.6. Remember that, since the rates are in thousands, their further use requires that the decimal points be moved leftward three places, i.e., 19.9/1000 = .0199,
21.1/100 = .0211, etc. At each year, the cohort of current 60 year olds will be reduced by the proportion that dies. Hence, at age 61 the fraction \((1.00 - .0199) = .98090\) will still be alive, and that is the probability that a sixty year old will live to earn the full income from his 60th year. During the next year, the fraction .0211 of the .98090 still alive will die, yielding the \(.98090 (1 - .0211) = .9602\) of the original cohort of 60 year-olds alive to reap the second year’s income. These fractions left alive, as listed in column 10, are interpretable as the probabilities that the 60 year old victim would actually be alive to receive income in any particular year.

The next step is to use the probability for each year to “weight” its corresponding income figure from column 7. In effect, this method “credits” a person with a proportion of potential income that appropriately reflects the probability that the income will indeed be received. Of course, this probability-weighted figure is really what Chapter II identified as the “expected value” of the income in that year. We are interested in the expected value of the income for the victim’s lifetime, rather than for any single year. Therefore the expected values for each year, as listed in column 11, are cumulated in column 12. In principle, we should keep summing up until the column 11 figure becomes zero, but the reader can see that, because both the present value discounting and death factor operate so strongly in later years, little change in the “bottom line” number would occur if the table were carried out for additional periods. Hence, we can see that the expected value of the victim’s income stream is really less than $200,000. Contrast this result with the much higher figure achieved by using the average life expectancy.

In a real-world application, the income for each year should be weighted by realization probabilities much lower than those shown in the same. Obviously, being alive is a necessary but not a sufficient condition that income will be earned. A further adjustment should be made for the probability that a live individual is still an active participant in the workforce (i.e., not retired) and, if active, is not unemployed. Unemployment rates and labor force participation rates by age and sex are available to analysts, so that the appropriate joint probability of being alive, participating, and employed can be computed. Inclusion of these factors would make Exhibit 3.6 become somewhat more complicated, but the principle is the same.

There is no doubt that the expected value method is conceptually superior to the average life expectancy method and, as illustrated, the methods give rather different results. The specific numbers in Exhibit 3.6 and some of the queries in the last question block above should help you be able to explain in commonsense terms what the fallacy is in the average life expectancy method.

3. RISK SELECTION AND RISK COMPENSATION

Kay Glehn owns and manages a number of apartment units in a city which is the site of a large university. Most of her tenants are students and she has been plagued by a midyear “skipping” problem at around the end of first semester. Typically, about six percent of her tenants default after making only five monthly payments toward an annual lease. Of those remaining, four percent default after one additional payment. The rest stay through the year and fulfill their entire contractual obligation. Because of the seasonal nature of the rental market, a midyear default has the practical implication that an apartment will remain empty for the entire defaulted period. Al- though it is in theory possible to obtain a judgment for the lost rent, that alternative is impractical because of a combination of legal expenses and the low probability of recovery against persons who have left the jurisdiction, are impecunious, etc. Hence, although she attempts to create an impression otherwise, her actual policy is merely to absorb the loss from any default.

Since rental housing is essentially a competitive market, Kay’s rental schedule of $500 per month presumably suffices to give her a reasonable rate of return on invested capital and operating costs. Hence, she (and other landlords in the market) are “charging” renters for the default risk that landlords must bear. This charge, of course, manifests itself in the form of a contract rent that is slightly higher than it would otherwise be.
Experience suggests to Kay that the tenants are composed of three types: Alphas, who are sure that they will perform on the entire lease; Betas who regard themselves as sure “skippers” at mid-year; and Gammas, who regard the probability of performance as a fifty-fifty proposition. Unfortunately, she cannot tell in advance which prospective renter belongs to which category. A worksheet similar to Exhibit 3.7 can be found in Kay’s files.

Exhibit 3.7: Analysis of Rentals

<table>
<thead>
<tr>
<th>Payment</th>
<th>DiscFac</th>
<th>PresVal</th>
<th>Cumul</th>
<th>Default</th>
<th>OK</th>
<th>ExVal</th>
<th>CumulEx</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000000</td>
<td>500.00</td>
<td>50000</td>
<td>0.00%</td>
<td>100.00%</td>
<td>500.00</td>
<td>500.00</td>
</tr>
<tr>
<td>2</td>
<td>0.9900990</td>
<td>495.05</td>
<td>995.05</td>
<td>0.00%</td>
<td>100.00%</td>
<td>495.05</td>
<td>995.05</td>
</tr>
<tr>
<td>3</td>
<td>0.9802960</td>
<td>490.15</td>
<td>1485.20</td>
<td>0.00%</td>
<td>100.00%</td>
<td>490.15</td>
<td>1485.20</td>
</tr>
<tr>
<td>4</td>
<td>0.9705901</td>
<td>485.30</td>
<td>1970.49</td>
<td>0.00%</td>
<td>100.00%</td>
<td>485.30</td>
<td>1970.49</td>
</tr>
<tr>
<td>5</td>
<td>0.9609803</td>
<td>480.49</td>
<td>2450.98</td>
<td>0.00%</td>
<td>100.00%</td>
<td>480.49</td>
<td>2450.98</td>
</tr>
<tr>
<td>6</td>
<td>0.9514657</td>
<td>475.73</td>
<td>2926.72</td>
<td>4.00%</td>
<td>96.00%</td>
<td>456.70</td>
<td>2907.69</td>
</tr>
<tr>
<td>7</td>
<td>0.9420452</td>
<td>471.02</td>
<td>3397.74</td>
<td>6.00%</td>
<td>90.24%</td>
<td>425.05</td>
<td>3332.74</td>
</tr>
<tr>
<td>8</td>
<td>0.9327181</td>
<td>466.36</td>
<td>3864.10</td>
<td>0.00%</td>
<td>90.24%</td>
<td>420.84</td>
<td>3753.58</td>
</tr>
<tr>
<td>9</td>
<td>0.9234832</td>
<td>461.74</td>
<td>4325.84</td>
<td>0.00%</td>
<td>90.24%</td>
<td>416.68</td>
<td>4170.26</td>
</tr>
<tr>
<td>10</td>
<td>0.9143398</td>
<td>457.17</td>
<td>4783.01</td>
<td>0.00%</td>
<td>90.24%</td>
<td>412.55</td>
<td>4582.81</td>
</tr>
<tr>
<td>11</td>
<td>0.9052870</td>
<td>452.64</td>
<td>5235.65</td>
<td>0.00%</td>
<td>90.24%</td>
<td>408.47</td>
<td>4991.27</td>
</tr>
<tr>
<td>12</td>
<td>0.8963237</td>
<td>448.16</td>
<td>5683.81</td>
<td>0.00%</td>
<td>90.24%</td>
<td>404.42</td>
<td>5395.69</td>
</tr>
</tbody>
</table>

QUESTIONS

1. If the landlord could determine in advance who were Alphas, Betas, and Gammas, how might the rental terms to each group differ?

2. Since, under the facts of this hypothetical, the groups are not identifiable in advance, the rental terms offered to all prospective tenants would presumably have to be identical. If you were an Alpha, would you feel as though you were being “overcharged”? By approximately what dollar amount?

3. Given that the three classes of risks cannot be determined in advance, how can the terms of Ms. Glehn’s leases be manipulated so as to “smoke out” the bad-risk Betas? (If there is more than one way of doing this, explain the relative merits of each.) How would the Gammas react to such manipulations? Of what relevance is “least cost risk bearing”?

4. Obviously, finding a way to discriminate among the three classes of risk alters the relative welfare of the Alphas, Betas and Gammas. Do these effects merely cancel out? If you were behind the Rawlsian “veil of ignorance,” not knowing whether you were to be an Alpha, Beta or Gamma, what system of charging would you prefer?

5. People try to find proxy variables to discriminate between good and bad risks in their transactions. For instance, the seller of an almost-new used car may find a certain market resistance because buyers feel that the car is a “lemon.” [See Akerlof, “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism,” 84 Q.J.Econ. 488 (1970)] Are certain classes of buyers discriminated against because they are suspected of being “lemons” in some economic sense? Is this fair or not? In what ways can people give quality assurance when they are incorrectly impacted by the “lemon effect”?

6. Can the issues in any of the above questions be couched in the “externality” terminology introduced in Chapter I?
4. DEFAULT COST UNDER “BACKDOOR” INTEREST SCHEMES

If the market rate of interest for a particular class of creditor exceeds the legal maximum under a state usury statute, the continued availability of credit to that class may depend upon some way of circumventing the interest rate ceiling imposed by law. One way of doing this is by establishing a “cash price” and a higher “time price,” such that the latter embodies an implicit charge for credit in order to supplement the legally imposed interest. In jurisdictions where the time price subterfuge is not permitted by the courts, some sellers may establish a single inflated price and then make credit available at the legal rate of interest. Other sellers will charge a lower price but will decline to extend credit. Of course, purchasers who do not desire credit would be foolish to purchase at the establishments that offer the credit, since that would involve being implicitly charged for part of the price of credit that is not received. In any event, sellers will generally find a way to make credit available by adjusting price in one of the ways described.

Exhibit 3.8 describes in some detail exactly how the “time price” subterfuge works. Suppose that the cash price of the goods would normally be $1000. The maximum legal rate is 18% (1.5% per month), but the market rate for the risk class in question is 30% (2.5% per month). The market rate of 2.5% is the rate at which the vendor will discount the stream of payments made for the goods, but he-at least nominally-can only charge explicit interest at the rate of 1.5% per month on outstanding balances.

When the vendor provides $1000 worth of goods to the credit buyer, the unalterable reality is that the “real” extension of credit is only $1000. Setting a time price of $1225.74 essentially involves the creation of a fictional $225.74 worth of credit such that the stream of $36.01 payments on the fictional loan at the below-market nominal interest rate of 18% (1.5% monthly) is also adequate to amortize the “real” loan of $1000 at the market rate of 30% (2.5% monthly).

Check Exhibit 3.8 to confirm that an “inflation” of the real $1000 credit extension by $225.74 is exactly the right amount to accomplish the desired purpose. If the goods are sold for $1225.74, columns (2), (3), and (4) of Exhibit 3.8 show that 48 monthly payments of $36.01 will amortize the “time price” loan and reduce the balance to zero. Note that accrued interest on each month’s opening balance at the nominal 1.5% is first subtracted from the monthly payment, then any excess is applied to reduce the balance itself. Now examine columns (5), (6), and (7). These columns show what would be happening if the same monthly periodic payments of $36.01 were to be applied as amortization of the “true” loan of $1000 at the 2.5% monthly rate. In both cases, the “time price” loan and the real loan, the last payment is exactly enough to pay the accrued interest and reduce the outstanding balance to zero.

The repayment arrangement depicted in the example is a standard loan amortization procedure. The exact size of the equal periodic payments necessary to amortize any sum over any number of periods can be found by using a somewhat imposing-looking formula:

\[
A = V \left\{ \frac{r}{1 - (1 + r)^{-n}} \right\}
\]

where A is the periodic payment to be found, V is the value to be amortized, n is the number of payments, and r is the periodic interest rate. (Note well that r is the periodic interest rate, i.e., the rate as adjusted for the relevant period, not the annual interest rate.) For notational simplicity, we shall rename the collection of terms within the curly brackets of the formula as Z; that is, \(Z = \left\{ \frac{r}{1 - (1 + r)^{-n}} \right\}\). The calculation of an actual value for Z, which depends on the values assumed for the interest rate and the number of periods, would be tedious if done by hand. Not so long ago, people therefore kept books of tables that, in effect, provided values for Z under many alternative assumptions about the values of r and n, thus permitting the earlier formula for A to be solved conveniently. Now, the same information is quickly accessible to anyone who knows the formula and has a pocket calculator.

Exhibit 3.8: The “Time Price” Interest Subterfuge

<table>
<thead>
<tr>
<th>LegalRate</th>
<th>CashPrice</th>
<th>MktRate</th>
<th>TimePrice</th>
<th>Payment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.015</td>
<td>1000</td>
<td>0.025</td>
<td>1225.74</td>
<td>36.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Payment</td>
<td>Nominal</td>
<td>Principal</td>
<td>Nominal</td>
<td>Balance</td>
<td>TRUE</td>
<td>Principal</td>
<td>TRUE</td>
</tr>
<tr>
<td></td>
<td>Interest</td>
<td>Reduction</td>
<td>Interest</td>
<td>Reduction</td>
<td>Interest</td>
<td>Reduction</td>
<td>Interest</td>
</tr>
<tr>
<td>1</td>
<td>18.39</td>
<td>17.62</td>
<td>1208.12</td>
<td>25.00</td>
<td>11.01</td>
<td>988.99</td>
<td>219.13</td>
</tr>
<tr>
<td>2</td>
<td>18.12</td>
<td>17.89</td>
<td>1190.23</td>
<td>24.72</td>
<td>11.29</td>
<td>977.70</td>
<td>212.52</td>
</tr>
</tbody>
</table>
How does one then calculate the correct value, in this case $225.74, for the time price surcharge? The method is easy to follow if the underlying reality of the distinction between the real and fictional credit extensions is borne in mind. First, solve for the necessary payment $A$, using the real (i.e., market) rate of
interest and the actual credit extension of $1000. This reveals what the periodic payment $A$ must be in order to satisfy the seller under the realities of the marketplace, ignoring the legal constraints.

The next question is to ask what fictional credit extension would imply the very same repayment scheme if the artificial interest rate were the applicable one. Since the formula above told us that \( A = VZ \), a bit of inference from 9th grade algebra reveals that \( V = A/Z \). In other words, just as one can find the periodic payment flow necessary to amortize any given sum, it is equally possible to work backwards from the payment stream to the implied sum being amortized. There is only one trick to the present application: the value of \( Z \) to be plugged into the equation \( V = A/Z \) is now the value based on the legal rate rather than the real rate, since the legal rate is the interest rate being nominally applied to the time price transaction.

**QUESTIONS**

1. Explain the first four columns of the exhibit. At what interest rate is the interest being accrued for purposes of column (2)? Check the exhibit’s calculations for at least the first two payments.

2. Why is column (8) labeled an “Excess Balance”? Why does the amount of the overstatement vary over time as it does?

3. Suppose a purchaser were to make two years worth of payments and then default. What is the amount of the judgment that a vendor would be able to secure? Is this result arguably inequitable? [See Allen and Staaf, “The Nexus Between Usury, ‘Time Price’ and Unconscionability in Installment Sales,” 14 Uniform Commercial Code L.J. 219 (1982). Suppose that you had to defend the position that the result is equitable, or at least desirable on other public policy grounds. What argument(s) might support this position?]

4. Let \( V \) and \( Z \) stand for the values defined in the text above on the assumption that the “real” values of the credit extension and interest rates are used. Similarly, let \( V' \) and \( Z' \) stand for the analogous values when the time price and the legal rate of interest are involved. Can you show that the time price can be computed directly by the formula \( V' = V (Z/Z') \)? (Note that \( Z/Z' \) would then be interpretable as a “blowup factor” to be applied to the original cash price \( V \).)

5. In comparing the advantages of paying cash versus buying on credit, one national “consumer buying advice” magazine habitually compares the cash price to the sum of the monthly payments. In this example, for instance, one would be advised that he will have to pay $1728.29 for the merchandise rather than $1000. Why is this somewhat misleading? In your opinion, is such advice on balance more deceptive or more informative?
F. TAX TREATMENT OF TORT DAMAGES

We return now to the nagging problem of computing the lump sum equivalent of a stream of uncertain future damages. Earlier sections on this topic focused principally on discount rates and adjustments for the contingent nature of the income flows being replaced. The Liepelt decision, which allowed tax consequences to be considered in calculating lump sum damages, was only alluded to in passing. The following case raises the tax issue squarely and directly.

For many types of damage awards—such as personal injury and wrongful death—the lump-sum damage award does not become part of taxable income to the recipient, thus seemingly providing a tax benefit to the extent that the income being replaced would have been subject to taxation. On the other hand, as the Delucca opinion below recognizes, the interest rates used to discount future earnings are not generally reflective of the fact that earnings on the original lump sum are taxable and therefore do not constitute a net rate of return to the recipient. In other words, using the market interest rates over discounts future earnings. Note, however, that the tax exemption on the original lump sum and the over discounting are counterbalancing in the direction of their effects.

DELUCCA V. UNITED STATES
670 F.2d 843 (9th C., 1982)

SKOPIL, Circuit Judge:

The United States conceded liability in the death of plaintiff’s decedent. The court below awarded the wife of the deceased $483,380 for loss of her husband’s earnings and support. This amount was calculated as follows:

<table>
<thead>
<tr>
<th>Description</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Projected earnings for 1980</td>
<td>$ 21,468</td>
</tr>
<tr>
<td>Less federal income tax</td>
<td>-3,099</td>
</tr>
<tr>
<td>Less state income tax</td>
<td>-1,073</td>
</tr>
<tr>
<td>After tax income</td>
<td>$17,296</td>
</tr>
<tr>
<td>Less 30% for decedent’s consumption</td>
<td>-5,189</td>
</tr>
<tr>
<td>Annual Loss of Support</td>
<td>$12,107</td>
</tr>
</tbody>
</table>

The court then found that the rate of wage increases and the discount factor would “essentially balance off each other,” so it simply multiplied the $12,107 by 30.5 (the expected work life of the deceased). The award for loss of support came to $369,263.50. To this the court added $21,793 (loss of support from accident to date of trial) for a total of $391,056.50.

The district court then added an amount to compensate for income taxes on the investment earnings of the lump sum award. It is this part of the calculation the government challenges. The district court stated:

Since it is assumed that plaintiff wife will invest her award, a full compensation for loss of future support should also take into account the fact that income taxes must be paid on her investment earnings. Norfolk and Western Ry. Co. v. Liepelt, 444 U.S. 490 (February 1980). Cf. Sauers v. Alaska Barge, 600 F.2d 238, 247 (9th Cir. 1979). To provide plaintiff wife an income of $12,107.00 after paying state and federal income taxes reasonably approximated at a 20% effective rate, she must be given $15,134.00 ($15,134.00 X 80% = $12,107.00). To compensate for loss of support over 30.5 years, she should therefore receive $462,587.00 (30.5 x $15,134.00) as the present value of loss of support.

The total loss of support to plaintiff was therefore increased to $483,380.00 and judgment was entered accordingly.
The district court provided for interest on the total damage award at 4% per annum “from the date of judgment November 4, 1980, until this judgment is satisfied.”

The issues presented on appeal are:

(1) whether the district court erred in adding to the lump sum damage award an amount to compensate for taxes on income that will be earned by investing the award; and

(2) whether the district court erred in specifying that post-judgment interest must be paid beginning on the date of entry of the judgment until the judgment is satisfied.

Damages to be awarded under the Federal Tort Claims Act (“FTCA”) are governed by the law of the place of the wrongful act, but are limited to compensatory damages. California law, applicable in this case, also limits damages to compensatory damages. In re Paris Air Crash, 622 F.2d 1315 (9th Cir., 1980).

In United States v. English, 521 F.2d 63 (9th Cir. 1975), this court set specific guidelines for calculating wrongful death damages under the FTCA: (1) calculate the future earnings the decedent could have expected over his life work expectancy; (2) deduct taxes and other amounts the decedent would or could not have contributed; and (3) discount this amount to its present value.

In Sauers v. Alaska Barge, 600 F.2d 238 (9th Cir. 1979), the plaintiff argued that the district court erred in “failing to increase the lump sum awarded to account for the fact that the interest earned when that sum is invested would be taxable.” Id. at 247. The Sauers court rejected this contention. The court stated that

[m]athematically, the plaintiff is right. Had the judgment called for the defendant to make deferred payments rather than a lump sum award, the funding of such an award would have required either an investment large enough to allow for any income tax on the interest yield of the fund or management of the investment in a manner that would either eliminate, or compensate for, the income tax on its yield.

The court, notwithstanding the mathematical correctness of the plaintiff’s argument held that the plaintiff was not entitled to this increasing factor.

Shortly thereafter, the Supreme Court indicated that this tax on the income generated by the lump sum award should be included in the calculation. Norfolk and Western Ry. Co. v. Liepelt, 444 U.S. 490 (1980). In Liepelt, the plaintiff had contended that the decedent’s income taxes on his future earnings should not be deducted from the lump sum award. The court rejected this contention and noted that since the income taxes on the decedent’s future earnings should be deducted, properly offered estimates of the income taxes on the earnings generated by the lump sum damage award should correspondingly be added:

Logically, it would certainly seem correct that [the income to be earned by the damage award] should be estimated on an after-tax basis. But the fact that such an after-tax estimate, if offered in proper form, would also be admissible does not persuade us that it is wrong to use after-tax figures instead of gross earnings in projecting what the decedent’s financial contributions to his survivors would have been had this tragic accident not occurred.

This statement by the Supreme Court indicates the Court finds that a lump sum damage award should be increased by the amount of income tax that would have to be paid on the earnings of the award.

After Liepelt, this Circuit again addressed the identical issue in Hollinger v. United States, 651 F.2d 636 (9th Cir. 1981). This court, in remanding for a recalculation of damages, followed Liepelt and required the district court to adjust the lump sum award for the income taxes on the income generated by the award:

The interest that Hollinger will earn on the discounted principal in a safe investment, however, will also be taxable. The effective interest rate for an investor is equal to the stated interest rate reduced by the product of the stated interest rate and the investor’s tax rate. The principal amount necessary to produce a given after tax yield per month is the principal amount that would be necessary to produce
that amount if there were no tax divided by the percentage of his income that the investor retains after
taxes. Thus an award of $600,000, not taking the tax on interest into effect, would be increased to
$1,000,000 for a 40% tax bracket recipient. Clearly, then, the possible adjustments involved in taking
this aspect of taxes into account are significant. Although we have found no cases in our court where
such a calculation was made, other courts have properly made such an adjustment. McWeeney v. New
York, N.H. & H. R.R. Co., 282 F.2d 34, 37 (2d Cir. 1960). This Supreme Court has also recently
suggested that the tax on the discounted principal might properly be considered. Norfolk & Western

The government argues that the statement in Liepelt was dicta, and therefore is not controlling in this
court. However, the Liepelt statement gives a clear indication of how the Supreme Court feels damage
awards should be calculated. Even if the statement was not essential to its decision in Liepelt, the
direction given by the Supreme Court should be followed by this Circuit. This Circuit in Hollinger
expressly considered the Leipelt statement and followed it.

The government offers some calculations to show that even if plaintiff was awarded the amount of lost
earnings without the addition of the disputed tax factor, this amount invested at 7% would yield $27,000
the first year. Since the plaintiff lost only $12,000 this year, the government thus claims that even with
this smaller award the plaintiff would be awarded more than her annual loss of support. This calculation is
misleading and erroneous. The government neglects to explain that the lump sum award has been
discounted to present value, which means that a large portion of her earnings in the early years must be
reinvested in order to generate enough income in the later years to keep up with the inflation and wage
increases that will have occurred.

The government argues that the possibility of the plaintiff investing the award in tax-free bonds negates
the addition of a factor for income taxes on the earnings generated by the award. However, the district
court, to discount the award to present value, assumed that plaintiff would invest her award in a secure
investment. Tax-free bonds yield a lower interest rate than other secure investments available in the
market and thus are most advantageous to persons in a high tax bracket because the higher the tax bracket
the more the savings in taxes. An awardee not in a high tax bracket who invested in tax-free bonds would
receive less of a return than if he invested in a secure taxable investment, and thus the return would fail to
be sufficient to yield the yearly amounts determined by the district court before discounting to present
value. The district courts should not be prevented from estimating the effects of taxes of earnings from
awards simply because the method plaintiff will use to invest her award is speculative; “so are most
predictions courts make about future incomes [and] expenses.” United States v. English, supra at 75.
There was sufficient expert testimony about the effect of taxes and tax-free investments on the plaintiff’s
award such that the district court’s finding cannot be said to be clearly erroneous. In this situation! where
investment in tax-free bonds would not be advantageous, the possibility that plaintiff would invest in
tax-free bonds is low and should not negate the increase of her award for the amount of income taxes on
the earnings generated.

In light of the direction given by the Supreme Court in Leipelt, and followed by this Circuit in Hollinger, we affirm the district court.

***

QUESTIONS

1. For many years, courts have tended to hear evidence as to what the victim’s before-tax income would
have been, then discount that back to a present value without any allowance for taxes. This
procedure is implicit in the examples supplied earlier in the chapter. An alternative technique is to
estimate the lost income on an after-tax basis, then provide a lump-sum whose proceeds — the return
on which is taxable — will yield a flow equal to the lost after-tax income. How hard would it be to
do the latter in a substantially accurate manner? [Harder than you might imagine. Think carefully.]
INTERTEMPORAL COSTS AND BENEFITS

Chapter 3


2. How would a court cope with probabilistic factors that might have a material effect on one’s future tax liability, e.g., children, marital status, charitable giving, place of residence, etc.?

3. Although it is far too complicated to demonstrate rigorously here (as your reflections on question one above may have caused you to suspect), computer simulations for representative taxpayers show that the tax benefits from the non-taxability of the original lump sum tend to be dominated by the taxability of the proceeds, provided that the time period involved is more than a few years. In short, neglecting the tax consequences does lead to a slight under-recovery by plaintiffs, all other things being equal. You may obtain some intuitive grasp of why this is so by reflecting on the possible analogy with the loan amortization table earlier in the chapter as Exhibit 3.8. In the present context, the lump sum damage award is the principal amount of an initial asset (rather than liability) and the lost after-tax income to be replaced represents a withdrawal (rather than a payment). Accrued interest adds to the principal, but the IRS takes a bite of this, thus reducing the effective interest rate. To the extent that withdrawals for consumption come out of reduction in the original (“tax free”) principal, however, the beneficiary enjoys a species of tax shelter. The net effect is, as suggested above, generally a slight disadvantage to the victim or his beneficiary, varying directly with the interest rates. Can you see at least roughly why this might be? How small would the inaccuracy have to be before you would be inclined to ignore it? Why are there arguably counterbalancing inaccuracies?

G. FAIRNESS IN EMPLOYEE BENEFITS

1. CONTRIBUTIONS TO PENSION PLANS

LOS ANGELES DEPT. OF WATER AND POWER v. MANHART

435 U.S. 702 (1978)

Mr. Justice STEVENS delivered the opinion of the Court.

As a class, women live longer than men. For this reason, the Los Angeles Department of Water and Power required its female employees to make larger contributions to its pension fund than its male employees. We granted certiorari to decide whether this practice discriminated against individual female employees because of their sex in violation of § 703(a)(1) of the Civil Rights Act of 1964, as amended.

For many years the Department has administered retirement, disability, and death-benefit programs for its employees. Upon retirement each employee is eligible for a monthly retirement benefit computed as a fraction of his or her salary multiplied by years of service. The monthly benefits for men and women of the same age, seniority, and salary are equal. Benefits are funded entirely by contributions from the employees and the Department, augmented by the income earned on those contributions. No private insurance company is involved in the administration or payment of benefits.

Based on a study of mortality tables and its own experience, the Department determined that its 2,000 female employees, on the average, will live a few years longer than its 10,000 male employees. The cost of a pension for the average retired female is greater than for the average male retiree because more monthly payments must be made to the average woman. The Department therefore required female employees to make monthly contributions to the fund which were 14.84% higher than the contributions required of comparable male employees. Because employee contributions were withheld from paychecks, a female employee took home less pay than a male employee earning the same salary;

It is now well recognized that employment decisions cannot be predicated on mere “stereotyped” impressions about the characteristics of males or females. Myths and purely habitual assumptions about a woman’s inability to perform certain kinds of work are no longer acceptable reasons for refusing to employ qualified individuals, or for paying them less. This case does not, however, involve a fictional difference between men and women. It involves a generalization that the parties accept as unquestionably
true: Women, as a class, do live longer than men. The Department treated its women employees differently from its men employees because the two classes are in fact different. It is equally true, however, that all individuals in the respective classes do not share the characteristic that differentiates the average class representatives. Many women do not live as long as the average man and many men outlive the average woman. The question, therefore, is whether the existence or nonexistence of “discrimination” is to be determined by comparison of class characteristics or individual characteristics.

The statute makes it unlawful “to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual’s race, color, religion, sex, or national origin.” 42 U.S.C. § 2000e—2(a)(1). The statute’s focus on the individual is unambiguous. Even a true generalization about the class is an insufficient reason for disqualifying an individual to whom the generalization does not apply.

That proposition is of critical importance in this case because there is no assurance that any individual woman working for the Department will actually fit the generalization on which the Department’s policy is based.

It is true, of course, that while contributions are being collected from the employees, the Department cannot know which individuals will predecease the average woman. Therefore, unless women as a class are assessed an extra charge, they will be subsidized, to some extent, by the class of male employees. It follows, according to the Department, that fairness to its class of male employees justifies the extra assessment against all of its female employees.

But the question of fairness to various classes affected by the statute is essentially a matter of policy for the legislature to address. Congress has decided that classifications based on sex, like those based on national origin or race, are unlawful.

Even if the statutory language were less clear, the basic policy of the statute requires that we focus on fairness to individuals rather than fairness to classes.

Finally, there is no reason to believe that Congress intended a special definition of discrimination in the context of employee group insurance coverage. It is true that insurance is concerned with events that are individually unpredictable, but that is characteristic of many employment decisions. Individual risks, like individual performance, may not be predicted by resort to classifications proscribed by Title VII. Indeed, the fact that this case involves a group insurance program highlights a basic flaw in the Department’s fairness argument. For when insurance risks are grouped, the better risks always subsidize the poorer risks. Treating different classes of risks as though they were the same for purposes of group insurance is a common practice that has never been considered inherently unfair. To insure the flabby and the fit as though they were equivalent risks may be more common than treating men and women alike; but nothing more than habit makes one “subsidy” seem less fair than the other.

An employment practice that requires 2,000 individuals to contribute more money into a fund than 10,000 other employees simply because each of them is a woman, rather than a man, is in direct conflict with both the language and the policy of the Act. Such a practice does not pass the simple test of whether the evidence shows “treatment of a person in a manner which but for that person’s sex would be different.” It constitutes discrimination and is unlawful unless exempted by the Equal Pay Act of 1963 or some other affirmative justification.

The Department argues that the different contributions exacted from men and women were based on the factor of longevity rather than sex. It is plain, however, that any individual’s life expectancy is based on a number of factors, of which sex is only one.

In essence, the Department is arguing that the prima facie showing of discrimination based on evidence of different contributions for the respective sexes is rebutted by its demonstration that there is a like
difference in the cost of providing benefits for the respective classes. That argument might prevail if Title VII contained a cost-justification defense comparable to the affirmative defense available in a price discrimination suit. But neither Congress nor the courts have recognized such a defense under Title VII.

Although we conclude that the Department’s practice violated Title VII, we do not suggest that the statute was intended to revolutionize the insurance and pension industries. All that is at issue today is a requirement that men and women make unequal contributions to an employer-operated pension fund. Nothing in our holding implies that it would be unlawful for an employer to set aside equal retirement contributions for each employee and let each retiree purchase the largest benefit which his or her accumulated contributions could command in the open market. Nor does it call into question the insurance industry practice of considering the composition of an employer’s work force in determining the probable cost of a retirement or death benefit plan.

***

* * * [A]lthough we agree with the Court of Appeals’ analysis of the statute, we vacate its judgment and remand the case for further proceedings consistent with this opinion. It is so ordered.

* * *

Mr. Chief Justice BURGER, with whom Mr. Justice REHNQUIST joins, concurring in part and dissenting in part.

* * *

Gender-based actuarial tables have been in use since at least 1843, and their statistical validity has been repeatedly verified. The vast life insurance, annuity, and pension plan industry is based on these tables. As the Court recognizes, * * * it is a fact that “women, as a class, do live longer than men.” It is equally true that employers cannot know in advance when individual members of the class will die. * * * Yet, if they are to operate economically workable group pension programs, it is only rational to permit them to rely on statistically sound and proved disparities in longevity between men and women. Indeed, it seems to me irrational to assume Congress intended to outlaw use of the fact that, for whatever reasons or combination of reasons, women as a class outlive men.

* * *

The reality of differences in human mortality is what mortality experience tables reflect. The difference is the added longevity of women. All the reasons why women statistically outlive men are not clear. But categorizing people on the basis of sex, the one acknowledged immutable difference between men and women, is to take into account all of the unknown reasons, whether biologically or culturally based, or both, which give women a significantly greater life expectancy than men. It is therefore true as the Court says, “that any individual’s life expectancy is based on a number of factors, of which sex is only one.” * * * But it is not true that by seizing upon the only constant, “measurable” factor, no others were taken into account. All other factors, whether known but variable — or unknown — are the elements which automatically account for the actuarial disparity. And all are accounted for when the constant factor is used as a basis for determining the costs and benefits of a group pension plan.

* * * [S]ex is the umbrella-constant under which all of the elements leading to differences in longevity are grouped and assimilated, and the only objective feature upon which an employer — or anyone else, including insurance companies — reliably base a cost differential for the “risk” being insured.

This is in no sense a failure to treat women as “individuals” in violation of the statute, as the Court holds. It is to treat them as individually as it is possible to do in the face of the unknowable length of each individual life. Individually, every woman has the same statistical possibility of outliving men. This is the essence of basing decisions on reliable statistics when individual determinations are infeasible or, as here, impossible.

* * *
I find it anomalous, if not contradictory, that the Court’s opinion tells us, in effect, * * * that the holding is not really a barrier to responding to the complaints of men employees, as a group. The Court states that employers may give their employees precisely the same dollar amount and require them to secure their own annuities directly from an insurer, who of course, is under no compulsion to ignore 135 years of accumulated, recorded longevity experience.

QUESTIONS

1. Judge Kilkenny, dissenting in the Circuit court decision of Manhart, comments, “Unless and until unisex tables are developed, an employer, to comply with the EEOC regulations on equal benefits, may not charge any additional amount to his female employees. In thus forcing the employer himself to cover the added amount necessary to assure equal benefits, this makes the employment of females economically unattractive, a result clearly at odds with the thrust of Title VII.” 553 F.2d 581, 598 (9th Cir. 1976). Is this a valid argument? How might an employer react to the Manhart rule in making his hiring decisions? Does the rule benefit women employees in the manner intended by the Court?

2. The majority opinion states that “unless women as a class are assessed an extra charge, they will be subsidized, to some extent, by the class of male employees,” yet “when insurance risks are grouped, the better risks always subsidize the poorer risk.” What is the likely impact of this “subsidy” on the attitudes and employment decisions of male and female employees? Would the Manhart rule lead to a well-balanced composition of male and female employees in one particular company? What pressures might there be on the salaries paid to men and women?

3. “Nothing in our holding implies that it would be unlawful for an employer to set aside equal retirement contributions for each employee and let each retiree purchase the largest benefit which his or her accumulated contributions could command in the open market.” The dissent claims that this statement by the majority is “anomalous, if not contradictory”. Is there a valid distinction between employer-funded pension plans and those funded by third parties? Is there a difference between (a) equal contributions by employees and unequal benefits and (b) unequal contributions and equal benefits?

4. Some large employers are starting to create “flexible” benefits programs. Each employee is allotted a particular number of “points” that may be used to “purchase” various types of benefits from a “menu” of alternative group plan coverages. The menu may involve selecting various levels of medical coverage, retirement annu- ities, life insurance, child care, etc. The “points” of course are in effect shadow prices with some monetary equivalent and, by choosing to allocate them in different ways, individual employees can potentially end up with dramatically different benefit packages. The companies typically set the point prices of alternative benefit coverages in proportion to the cost incurred by the companies to support each type of group benefit. What would you predict would happen under such plans? Is this desirable or not? Why? Do you see any potential legal problems?

5. If the benefits from a pension plan could be assigned to third parties, a true market value of their worth could be arrived at. Should we attempt to determine the level of contributions required by an “assignment” concept? Or are there other considerations besides true market value?

2. CONSTRUING THE AGE DISCRIMINATION IN EMPLOYMENT ACT

The Department of Labor has interpreted the Age Discrimination in Employment Act (ADEA) to mean that workers in the protected 40-69 age group cannot be denied benefits extended to younger workers unless higher costs of providing the benefit to the older workers can be demonstrated actuarially and, perhaps most importantly, the modification of the benefits is “not a subterfuge” to evade the purposes of the statute. 29 CFR 860.120. Disputes have recently arisen with respect to the meaning of the ADEA as it applies to the group insurance benefits provided by certain employers. Consider the situation of a hypothetical employer, Beekins Manufacturing, Inc.
Under the group insurance plan provided by Beekins, life insurance proceeds are paid to the employee’s beneficiary at death but, in the event of Permanent Total Disability (PTD), the insured employee can elect to receive what would have been the death benefit in monthly installments. The same level of life insurance protection is provided to employees of all ages but, beginning at age 60, PTD coverage ceases. Many other firms are likely to be in the same position as Beekins. A study of benefits plans shows that, when disability coverage is linked to life insurance, the common practice is to drop disability coverage at age 60 or 65. See, Blostin, “Is Employer-Sponsored Life Insurance Declining Relative to Other Benefits,” Monthly Labor Rev., 3-13 (Sept. 1981).

Nobody contests the fact that disability insurance is more expensive for the older workers. But, suppose that the Equal Employment Opportunity Commission (EEOC) takes the position that the ADEA should be construed as forbidding the complete elimination of any benefit type to an age class. Specifically, it is argued that if a benefit e.g., coverage against a distinct risk such as PTD—is provided to any employees at all, then any cost differential may be used only to justify a reduction but not an abolition of protection against that same risk for older workers.

QUESTIONS

1. Disabling health problems do increase in frequency with advancing age, and this factor alone would create higher costs of providing PTD benefits for older workers. In addition, there is also a “moral hazard” problem that becomes increasingly significant. What is the moral hazard and what is its significance in this context?

2. Why do the facts suggest that, under the Beekins plan, the older workers are “overbenefitted” relative to younger workers in terms of life insurance coverage? Is it plausible to think that the “overbenefit” on the value of the mortality coverage might roughly compensate for the “underbenefit” from loss of PTD coverage? If not, should an age-related differential in the value of the total benefits package be regarded as a violation of the purposes of the Act?

3. If the EEOC view prevails, why would you predict that employers would reduce life insurance coverage to older workers in order to counterbalance the mandated increase in PTD coverage?

4. If you were writing an amicus brief for the Golden Age Workers Association, what position would be in the best interests of your clients?

5. The scene is a meeting of house counsels for other similarly situated employers. Another attendee says:

“Look, what are we worrying about? The Act probably requires us to spend roughly the same amount on benefits for all age groups anyway. Why should we care how our payments are located as long as their total amount isn’t going up?” You blurt out “Idiot! . . .” And what do you say after that?

6. How would you decide this issue as a federal judge trying to implement what you think was Congressional intent? (You may assume that there is no legislative history that is dispositive on this point.)
Economists use several different kinds of diagrammatic models. One advantage of diagrammatic models is their ability to depict the results of continuous adjustment of some variable. Those quantity adjustments, usually measured along the horizontal axis of the diagram, are then translated into a causal impact on benefits, costs, a probability of some happening, etc. The “result” variable is generally measured along the vertical scale of the diagram and the interaction between the two variables is captured in a “curve” that depicts the functional relationship between them. Note, however, that the number of variables involved is only two: one cause and one effect, corresponding to the horizontal and vertical dimensions of the diagram. Cause and effect relationships in the real world are frequently much more complex; for instance, two or more causes may act jointly to produce a particular effect. How can these more complex situations be modeled?

One response, of course, is to employ mathematical models. Professional economists frequently do employ mathematical models and such models are the most elegant and generalized method of dealing with many problems. But economists also have developed graphical models that retain the expository attractiveness of the “picture” approach while permitting at least a limited analysis of interactive causes. As we shall now see, certain widely used three-variable economic models (typically involving one effect and two interactive causes) are in essence topographic maps where the “height” dimension measures effect and the “longitude” and “latitude” dimensions measure the causes. This type of model will first be introduced in the context of an ordinary topographic map of real estate holdings where the questions dealt with are facially unrelated to economics. Nonetheless, the reasoning involved in interpreting such a map—which is familiar to most laymen—has almost exact analogues in economic models. Anyone who can understand the more familiar form of map should, with a modest amount of application, be able to make the transition to the more abstract context of the economic models.

As a prefatory matter, it should perhaps be indicated that the models dealt with in the present chapter are related most closely to what is known as “demand” economics. In the next chapter, very similar constructs will be reinterpreted as “production function” models that are useful in areas of law such as theories of tort liability.

A. TOPOGRAPHICAL MAP READING

1. DECISIONS BASED ON SPATIAL LOCATIONS

The topographical map in the accompanying Exhibit 4.1 shows land currently being subdivided and sold off by the Bomber Land Company. Although most of this tract is still held by the developer, two lots have already been sold, to McNabb and Zepp, respectively. The boundaries of the lots and the siting of the houses thereupon are as indicated on the attached map.

Bomber Land Company has entered into certain covenants with the purchasers of lots in this tract. One is that the company will, at its own expense, construct a community TV antenna. The agreement is that the antenna tower must be placed at the highest point within an area demarcated by the following:

1. a straight line connecting boundary markers 249-59 and 249-60;
2. a straight line due southeast from marker 249-60; and
3. a line running straight east from boundary marker 249-59.

The agreement calls for an antenna not less than 40 nor more than 60 feet in height, construction to be completed within 8 months after sale of the fourth lot in the development.
Bomber is also prepared to enter into other individualized agreements with subsets of property owners. One such agreement has already been completed. It involves an easement for a water pumping facility to be used jointly by the Zepp and McNabb families. Tentatively, the agreement entitles them to construct the pump at the point indicated on the map. However, the two property owners also have 12 months in which to agree upon and designate any other mutually agreeable site within Bomber Development’s holdings.

QUESTIONS

6. Prospective purchasers of other lots in this development would reasonably be concerned with the precise location of the TV-tower.
7. a. Draw the boundaries of the area within which the TV antenna must be placed.
8. b. Does the information on the topographical map permit you to say approximately where the antenna must be placed? Indicate on the map your best estimate of the antenna site.
9. c. A colleague sees your map work and exclaims: “Idiot! You have the whole business just backwards. You’re indicating one of the lowest sites in that tract, rather than the highest!” Can you prove her wrong?
10. 2.In connection with the water pump easement, you may assume that McNabb and Zepp each wish to minimize the distance of the pump from their residences. The reason for this, of course, is the high cost of piping from the pump to the residences, which must be borne by each party individually. Other than distance, the two parties have no preferences as to the location of the pump.
11. a. Prove that the tentative site designation is not an “efficient” one for the parties, i.e., that there are other sites that are mutually preferable.

12. b. What can you predict about the location of the alternative site that will be designated? What conditions must be met in order for the designated choice to be “efficient,” i.e., undominated by yet another site that is better for both parties?

2. INDIFFERENCE CURVE MODELS AND MARKET DECISIONS

In familiar topographic maps of land such as Exhibit 4.1, the “contour lines” on the map show points of equal height. Given a set of “coordinates” in terms of horizontal and vertical positioning, one can identify points on the map and say something about their height. Why is height important? In question one above, the best site for an antenna is presumably the highest one. Hence, any two points on the same contour line are, assuming that height alone is considered, equally good sites for an antenna. One would nonetheless prefer to be as high as possible on the “hill” or “mountain” depicted. The deed restrictions, however, constrain the choice of an antenna site to that subset of the land contained within well defined permissible boundaries. Except in the special case where there is a “peak” somewhere within the permissible boundaries, the highest achievable site will usually be found on one of the boundaries themselves, i.e., that on the “uphill” side.

Exhibit 4.2 is an analogous economic model. Now, the horizontal and vertical axes represent not distances but quantities of goods X and Y, respectively.

![Exhibit 4.2: Economic Topography and Constraints](image)

Point $a$ has “coordinates” of $8Y$ and $10X$, thus identifying it as a “bundle” of goods containing eight units of Y and ten units of X. The contour lines on this topographic map indicate equal levels of preference; a “higher” bundle will be preferred to a “lower” one. Since bundles on the same contour line, such as $a$ and $b$, are equally high, the preference relation between them is one of indifference and the
contour lines are therefore commonly called “indifference curves.” [Building on the Greek prefix “iso” meaning equal, the terms iso-utility or iso-preference curves are sometimes also encountered.] Hence, an “indifference curve model” is really a topographic map of an individual’s preferences expressed in terms of quantities of two “goods.” If allowed to do so, an individual will choose the bundle that is as high as possible on his or her “preference surface” or satisfaction mountain. The “constraints” involved in the first problem were property rights, presumably enforced in terms of a penalty for trespassing or some other sanction that makes location of the antenna outside of the designated boundaries impractical. In an economic model, the constraints on choices may be legal rather than narrowly economic. For instance, legal prohibitions on the use or possession of controlled drugs inhibit their consumption. The legal limitation may not be absolute, however, as exemplified by wartime “ration” allowances restricting consumers to certain maximum quantities of regulated commodities. Hence, it is quite appropriate to incorporate fundamentally legal constraints into economic-type decision models. The special form of constraint traditionally treated by economists is nonetheless based on narrowly economic parameters, namely prices and income.

Suppose Beyer has $100 and the prices of goods Y and X are $5 and $10 per unit, respectively. Is it possible for Beyer to purchase bundles a or b in Exhibit 4.2? What about any of the other possible bundles? It would be useful if we could partition the map, sometimes called a “commodity space”, into two areas, that containing bundles that Beyer does have enough money to buy and that containing other bundles whose purchase is impossible because it would require more money than is available. The area of real possibilities, from which Beyer must choose, is called the “opportunity set” to distinguish it from the set of other theoretically possible bundles that do not represent real possibilities under the assumed circumstances of income and prices.

It is worth suffering the pain of a little simple algebra to achieve some generalizable results. Let M stand for the money available, X and Y for the quantities of the two goods, and Pₓ and Pᵧ the respective prices of the goods. Then the combinations of X and Y purchases that exhaust all of the money are those that satisfy the equation

\[ Pₓ X + Pᵧ Y = M \]

where the left-hand side is the summation of the money (price times quantity) spent on X and on Y and the right hand side is the money available to be spent. Rearranging the terms just a little bit yields the following form of the same basic “budget equation” above:

\[ Y = \frac{M}{Pᵧ} - \frac{Pₓ}{Pᵧ}X \]

The above is the equation of the so-called “budget line” that forms the outer boundary of an individual’s opportunity set when constrained only by income and prices. The appropriate line is drawn in Exhibit 4.2 for the assumed $100 income and the $5 and $10 prices for the two goods. Bundles on the budget line or within the triangular shaped opportunity set are possible choices; those to the right of the budget line exceed Beyer’s resources. Note that the budget line is a straight line whose form corresponds to the algebraic equation above: if X were 0, then Y would be \( \frac{M}{Pᵧ} \) or \( \frac{100}{5} \), the height of the “intercept” of the straight line where it cuts the vertical axis. For each unit increase in the purchase of X, the amount of Y must fall by \( \frac{Pₓ}{Pᵧ} \) or \( \frac{10}{5} \), this magnitude corresponding to the downward “slope” of the budget line.

Economists traditionally eschew any analysis of why preference maps themselves change, i.e., why a set of indifference curves may shift. (That analysis is generally considered the domain of the sociologist or behavioral psychologist.) Rather, economic models concentrate on explaining observed changes in choices as due to shifts in constraints, such as the so-called budget line derived above. Exactly how the budget line shifts and what are the behavioral implications can have interesting legal ramifications.

The sections immediately following this introductory material provide illustrations of the use of indifference curve models in legal applications such as group decisionmaking, the theory of "efficient"
contract breach and the analysis of promissory reliance. The group decisionmaking analysis builds on the same principles as developed above for the location of an optimal location for the well. That analysis is a rather special case of the trading model called the Edgeworth Box by economists. The contract breach model is a more traditional and straightforward application of the same Edgeworth Box trading analysis. The promissory reliance model depicts a promise as knowledge that shifts a budget line. The next few problems deal with alternative forms of governmental grants. When you read the grant cases, attempt to relate the facts to shifts in budget lines. In order to do so, you may wish to refer explicitly to the equation above, reflecting on which of the parameters $M$, $P_x$, and $P_y$ have changed and what that implies in terms of a budget line shift.

B. GROUP DECISIONS UNDER MAJORITY RULE

1. “RATIONAL” GROUP PREFERENCES

Majority rule voting processes are mechanisms designed to take individuals’ preferences and aggregate them into a “group preference” in a reasonable way. Although earlier scholars had also pointed out some of the inherent peculiarities of voting as a rational decisionmaking process, Kenneth Arrow’s 1951 book on *Social Choice and Individual Values* effectively marked the birth of a new subdiscipline of the social sciences dedicated to the study of “public choice” and, specifically, the implications of alternative voting institutions. In many respects, only some of which will be exemplified in the problems below, the results of the public choice scholarship have been quite disturbing because they call into question the reasonableness of results achieved under any conceivable set of voting institutions.

What do we mean by a reasonable or rational group decision process? Two widely accepted desiderata for a rational collective decisionmaking process are what may be termed “consistency” and “path-irrelevance.” That is, there ought not to be anything inconsistent about the way in which different alternatives are ranked, nor ought the conclusion ultimately reached depend on the way in which the options are presented (i.e., the “path” a group follows in reaching the decision). Unfortunately, the mechanics of majority rule decision-processes seem to have certain inherent weaknesses. The demonstrations of this fact originally were couched in terms of the political process itself, but the results are applicable to all voting bodies. Consider why the models and questions presented below can be said to raise qualms about possible inherent inconsistency and arbitrariness in some familiar legal mechanisms that involve voting.

Two types of voting models are presented below. The first is a variant of the trading model developed above to deal with the possibility of a mutually agreeable relocation of the well site. Then, a similar situation is modelled in the “matrix” format that has been used several times in the preceding chapters. As will become apparent, each type of model has its own particular advantages.

---

2. MODELING OF VOTING DECISIONS

a. Topographic Model of 3-Person Majority Rule Decisions

A certain three-person local government commission has the authority to levy taxes up to 2% of assessed value on local real estate. Any funds raised through such taxation can then be allocated between use A and use B. Majority rule is used in making all decisions of the commission.

Each commissioner has different preferences about budgetary mix, as illustrated by Exhibit 4.3 below which shows the “ideal” outcomes of commissioners 1, 2, and 3 as points E1, E2, and E3, respectively. Of course, no single commissioner is in a position to impose any result unless the acquiescence of at least one other commissioner is also secured.

Exhibit 4.3 – Topographic Map of Budgetary Decisionmaking

QUESTIONS

1. Suppose that each commissioner regards all solutions equidistant from his ideal in Exhibit 4.3 as equally preferable; i.e., the relative preference ordering of different outcomes for any commissioner is inversely related to its distance from that commissioner’s ideal outcome E. Select one of the commissioners, then convert Exhibit 4.3 to a topographic map that shows at least three of the “contour lines” (i.e., the indifference curves) around that commissioner’s “peak.”

2. On the topographic map just created, select a point vertically above (“north” of) the peak and label it X. If result X is not ranked as “high” as the result at the peak, why is that true? Doesn’t result X correspond to more of both use A and use B? Then why isn’t it better? Similarly, pick a point
horizontally to the right of the peak and explain why the commissioner in question finds it inferior to the peak solution.

3. Now superimpose the preference maps of the other two commissioners on the same Exhibit. The interplay among these preferences will presumably determine the ultimate majority rule decision that emerges. Are there, however, any solutions shown on the map that are not possible choices for the commission? If so, indicate in some manner what constraint affects the possible choices. (If you believe that the information given is not sufficient, make a plausible assumption.)

4. Suppose that commissioners 1 and 2 were to form a majority coalition against commissioner 3. Show that all of the plausible “deals” involving a coalition between commissioners 1 and 2 ought to involve moving to some point along a line—commonly called the “contract curve”—connecting their two “peaks.” [What in the assumed facts causes this contract curve to be a straight line? What would determine its shape if that assumption were relaxed?] Pick a possible result from the coalition and label it A. What determines where along the contract curve for commissioners 1 and 2 that A will fall?

5. After result A has been reached through the operation of a coalition consisting of 1 and 2, commissioner 3 approaches 2 in hopes of creating a counter-coalition. Can 3 and 2 find any “improvements” on the tentative decision A such that they both would be better off? Identify these possibilities by using Exhibit 4.3, adding to it, if necessary, any other relevant curves or constructions.

6. Obviously, a majority coalition between any two voters is possible. If there are no procedural rules that limit counteroffers or the formation of alternative coalitions, what can be said about the final “majority choice” that will emerge from the voting process?

7. Assume that budget size and budget mix must be determined in separate decisions, i.e., first a vote is taken on the vertical adjustment of Exhibit 4.3, then a vote is taken on the horizontal adjustment. Does this specification of the procedural rules affect the answer to the previous question?

8. What in the original facts above required that the contour lines be perfect circles? Is this a plausible assumption? Why? Do any of the results in the questions above depend on the simplification embodied in the perfect circles?

9. The model developed here is a mere three-voter application. Can the same analytical tools be applied to more complex voting bodies? Give some thought to how the Exhibit 4.3 type of model would be used for the next step in generalization, a five-voter body. For an excellent and highly accessible discussion of this type of model, see Tullock, Toward a Mathematics of Politics (1967), Ch. II, esp. 33-36.

b. A MATRIX MODEL OF VOTE TRADING

A three-voter body is empowered to make a decision in which results A and B either may or may not be included in any combination. The three voters’ preferences about A and B and the four alternative outcomes involving possible combinations of them are summarized in Exhibit 4.4. The numbers can be understood as evaluations of the results in terms of dollars, utility, or any other index that you find appealing.
Voter | A Only | B Only | Both | Neither
--- | --- | --- | --- | ---
1 | 100 | -1 | +99 | 0
2 | -1 | -1 | -3 | 0
3 | -1 | +100 | +99 | 0

**Exhibit 4.4 – Voters’ Benefit Changes For Measures A & B**

**QUESTIONS**

1. How do the numerical entries in the last two columns of Exhibit 4.4 relate to the entries in the other columns?

2. If the inclusion of A and B were voted upon separately, e.g., as in a popular referendum, what result would be predictable?

3. Assume now that the result consisting of both A and B were offered directly. Does this command a majority over the result predicted in the previous answer?

4. “A result such as both A and B is achievable through either implicit or explicit logrolling that allows the intensities, and not just the directions, of the preferences on issues to count.” Explain. In light of this, what would you say about the desirability of logrolling? What is meant by “implicit” logrolling?

5. Suppose that, in a different context, one of these voters were given an opportunity to choose between single pieces of different kinds of fruit presented in a pairwise fashion. First, a Coconut is chosen over a Date. Then, the opportunity is taken to replace the Coconut with a Banana. Finally, the Banana is surrendered in favor of an Apple. At this point, the possibility emerges of reoffering the same person a Date in exchange for the Apple. What would you predict as to whether the Date would be accepted? Why? What would you say if, upon actual presentation of the choice, the person acted contrary to your prediction?

6. Relabel the last columns of the Exhibit 4.4 as options C and D, respectively. Let the symbol “>” represent the words “is preferred to.” Explain why the majority rule preferences of this voting body can then be summarized by the following relation:

   A > B > C > D > A

What is bizarre about this result? What relation is there to #5 above?

7. Suppose that A, B, C, and D in the table above represented amendments to a main motion. Roberts’ Rules of Order dictate that the last amendment introduced is the first to be voted upon. Is this likely to make a difference in the final result? Would you like your amendment to be first or last? How can these results be applied to the question of whether control of the agenda also controls the ultimate result? See the legal applications in Levine and Plott, “Agenda Influence and Its Implications,” 63 Va.L.Rev. 561 (1977).

8. Suppose that there are three voters who have the following sets of preference orderings over four alternatives A, B, C, and D: for Voter 1, C > B > A > D; for Voter 2, B > A > D > C; and for Voter 3, A > D > C > B. Assume that the alternatives are voted upon pairwise: first A against B, the winner of that first pairing against C, then the winner of the next pairing against D. Satisfy yourself that -D will come out the final winner. In addition to being an arbitrary result of the order of the voting (as the previous question suggests), ±, low-. what else seems unsatisfactory about this result?
Chapter 4  MAPPINGS AND “DEMAND” DECISIONS 155

9. How do the above models and series of questions apply to an appellate court panel or a regulatory agency such as the Federal Communications Commission? [See Spitzer, “Multicriteria Choice Processes: An Application of Public Choice Theory to Bakke, the FCC, and the Courts,” 88 Yale L.J. 717 (1979).] Does your answer to question 4 above depend on whether a court, an agency, or a legislature is involved? Why?

10. How can the “topographic map” trading model of the commissioners voting on a budget size and mix be interpreted as a generalization of this present matrix model? Explain. What would the preference maps of the individual voters have to look like in order for the problems reflected in the questions above not to arise?

C. DECISIONMAKING IN THE COURTS

The question sets in the immediately preceding section were designed to highlight inherent problems in decisions that are made by voting. If the voting process is unconstrained, so that results previously rejected may be re-raised, there generally will be no “final” decision because there is no result that commands a majority against all possible alternatives. This is called the “cyclical majority” problem. Typically, procedural rules of some kind are used to constrain the process such that majority cycles are avoided. For instance, Roberts’ Rules require a special procedure to “reconsider” a motion that has already been defeated. In the courts, a similar effect is exerted by the doctrine of stare decisis, which precludes the reassertion of arguments rejected in an earlier consideration. Any such limiting rule, however, produces “path dependence” wherein the ultimate results depend on seemingly arbitrary factors such as the order of consideration, the number and identity of other options considered at each stage, etc. In short, neatness and consistency can be imposed, but at the price of arbitrariness. For an interesting application of these notions to the decisions of the Supreme Court, see Easterbrook, “Ways of Criticizing the Court,” 95 Harv.L.Rev. 802 (1982).

The previous problem set only hinted, however, at the issues raised by the possibility of vote-trading and strategic voting by judicial panels, juries, regulatory panels, etc. Once one understands the workings of voting processes, it becomes apparent that there are frequently practical gains to be had from voting for results that are, at least in part, strongly disfavored. There is, in fact, no possibility of designing a voting system that will not be subject to manipulation by strategic voting.*

Read the following criminal cases for possible practical applications of logrolling, strategic voting, or any of the other theoretical problems inherent in voting processes. How many of these can you identify?

STATE OF MISSOURI v. BARTON
602 S.W.2d 479 (Mo. App., 1980)

CRIST, Judge.

Defendant was charged by information with assault with the intent to kill with malice aforethought, but was convicted by the jury of a lesser included offense of assault with intent to do great bodily harm without malice. Under the Second Offender Act, his punishment was assessed at five years imprisonment. We affirm.

* * *  [Here the court briefly recounted the events leading up to the shotgun shooting by the defendant of the victim, Thomas Wells, outside a tavern where the two men had quarreled earlier in the evening.]

---

* * * [Defendant’s] argument indicated he believes the evidence was insufficient to support a conviction under the original charge, to-wit: assault with intent to kill with malice aforethought. Accordingly, he reasons that the trial court committed error when it failed to sustain his motion for acquittal and to dismiss the case in its entirety at the close of all the evidence. Defendant also argues that the court committed contiguous error in that it permitted the jury to be instructed as to the original charge.

Defendant’s argument points to evidence which indicates that defendant was some 25-30 feet from the victim when he fired. Defendant states that had he actually intended to shoot Wells, from such a distance he would have inflicted considerably more damage than the victim actually suffered. Defendant relies on State v. Harty, 569 SOIT 783 (Mo.App.1978) which requires proof of intent in order to convict a defendant with either assault with intent to kill with malice aforethought or assault to do great bodily harm without malice. Defendant feels that without the requisite demonstration of intent, he is guilty of no more than common assault.

* * *

Contrary to defendant’s ill-conceived supposition, it is not prejudicial error to submit the greater offense for the jury’s consideration, even where the facts would not support a conviction therefor, when the verdict was for the lesser offense and that verdict is sustained by the evidence. State v. Brooks, 567 SOIT 348, 352 (Mo.App.1978). And, we specifically note that the charge of assault with intent to kill with malice aforethought did not expire from lack of evidence, but rather, from lack of a jury verdict.

Our foregoing, particularized recitation of the evidence purveys an adequate factual basis for defendant’s conviction in that there was substantial evidence that defendant shot at Wells. We cannot oblige defendant by ignoring the evidence adduced by the state or in reweighing the testimony.

Judgment affirmed.

HARRIS V. RIVERA

PER CURIAM.

The questions presented by the certiorari petition concern the constitutionality of inconsistent verdicts in a nonjury criminal trial. Certiorari is granted and the judgment of the United States Court of Appeals for the Second Circuit is reversed.

* * *

Inconsistency in a verdict is not a sufficient reason for setting it aside. We have so held with respect to inconsistency between verdicts on separate charges against one defendant, Dunn v. United States, 284 U.S. 390, and also with respect to verdicts that treat codefendants in a joint trial inconsistently, United States v. Dotterweich, 320 U.S. 270, 279.\(^\text{14}\) Those cases, however, involved jury trials; as the Court of Appeals correctly recognized, both of those opinions stressed the unreviewable power of a jury to return a

\(^{14}\) “Equally baseless is the claim of Dotterweich that, having failed to find the corporation guilty, the jury could not find him guilty. Whether the jury’s verdict was the result of carelessness or compromise or a belief that the responsible individual should suffer the penalty instead of merely increasing, as it were, the cost of running the business of the corporation, is immaterial. Juries may indulge in precisely such motive or vagaries. Dunn v. United States, 284 U.S. 390.”
verdict of not guilty for impermissible reasons. It is argued that a different rule should be applied to cases in which a judge is the factfinder.

We are not persuaded that an apparent inconsistency in a trial judge’s verdict gives rise to an inference of irregularity in his finding of guilt that is sufficiently strong to overcome the well-established presumption that the judge adhered to basic rules of procedure.

The question that respondent has standing to raise is whether his trial was fairly conducted. The trial judge, the New York appellate courts, the Federal District Court, and the United States Court of Appeals all agreed that the record contains adequate evidence of his guilt. These courts also agreed that the proceedings leading up to respondent’s conviction were conducted fairly. Apart from the acquittal of Robinson, this record discloses no constitutional error. Even assuming that this acquittal was logically inconsistent with the conviction of respondent, respondent, who was found guilty beyond a reasonable doubt after a fair trial, has no constitutional ground to complain that Robinson was acquitted.

QUESTIONS

1. The court in *Missouri v. Barton* held that it was not prejudicial error to submit a greater offense for the jury’s consideration, even where the facts would not support a conviction on this offense, when the verdict was for the lesser offense and that verdict was sustained by the evidence. One interpretation is that the court is saying that, even conceding arguendo that the submission of the more serious-charge was error, the error was obviously harmless because the jury rejected the impermissible result. Is that a fair interpretation, or do you think that something else was meant? Assuming that the interpretation is correct, do you agree with the inference as to the harmlessness of the error?

2. *Harris v. Rivera* was not a jury trial. If it had been, what additional explanations of the “inconsistent” verdict (besides those potentially attributable to a judge) would be possible?

3. As a general matter, would defendants be better off if inconsistent verdicts were impermissible? From the perspective of society as a whole, “should” inconsistent verdicts be acceptable? [Note that, in the excerpted case, the Supreme Court does not directly decide this normative issue. Its holding is the more limited one that an inconsistent verdict does not ipso facto violate constitutional guarantees of a fair trial.]

---

15 Justice Holmes’ opinion in *Dunn*, his last for the Court, characteristically was brief and to the point. He quoted the following passage from *Steckler v. United States*, 7 F.2d 59, 60 (CA2 1925):

“The most that can be said in such cases is that the verdict shows that either in the acquittal or the conviction the jury did not speak their real conclusions, but that does not show that they were not convinced of the defendant’s guilt. We interpret the acquittal as no more than their assumption of a power which they had no right to exercise, but to which they were disposed through lenity.” *Dunn v. United States*, 284 U.S. 390, 393.

After citing *Horning v. District of Columbia*, 254 U.S. 135, he added:

“That the verdict may have been the result of compromise, or of a mistake on the part of the jury, is possible. But verdicts cannot be upset by speculation or inquiry into such matters.” 284 U.S., at 393-394.

20 “The question whether the evidence is constitutionally sufficient is of course wholly unrelated to the question of how rationally the verdict was actually reached. Just as the standard announced today does not permit a court to make its own subjective determination of guilt or innocence, it does not require scrutiny of the reasoning process actually used by the factfinder—if known.” *Jackson v. Virginia*, 443 U.S. 307, 319 n. 13.
D. EDGECOURTH BOX ANALYSIS OF CONTRACTS

We earlier explored the application of “mapping” models to explore possibilities of trade and adjustment: the well-relocation problem in the map-reading section. Very similar trading models may be constructed for a host of other legal applications such as out-of-court settlement, plea bargaining, etc. Although the models can thus be applied in contexts that seemingly have little or no relationship to market-type transactions, the general class of mapping-based trade model used above is nonetheless closely related to the “Edgeworth Box” model used widely in traditional economic applications wherein “goods” are being traded for money. It will be useful at this point to present the Edgeworth Box model explicitly and exemplify its application in a market-like trading context.

One area of legal analysis that—like traditional economics—deals with goods-for-money transactions is contract law. Hence, it should not be surprising that the Edgeworth Box model can be used for contract analysis in more or less its original form, as it might appear in an economics textbook. An example of the application of an Edgeworth Box to contract analysis is therefore provided immediately below. This analysis is adapted from Goetz and Scott, “Liquidated Damages, Penalties and the Just Compensation Principle: Some Notes on an Enforcement Model and a Theory of Efficient Breach,” 77 Col.L.Rev. 554, 563-565 (1977). In this context, the same Edgeworth Box construction is used to address the problem of why the law sometimes seems to encourage breach of contract (together with payment of damages) rather than honoring of the original contractual obligation.

CONTRACTUAL REALLOCATION THROUGH “EFFICIENT BREACH”

Important elements of the rationale underlying modern contract theory can usefully be articulated in terms of the “indifference curve” analysis commonly used in economic theory. Basically, indifference curve maps may be thought of as mapping a “preferredness mountain” wherein higher “elevations” represent the greater preferredness of outcomes. The points in any single indifference curve constitute outcomes having equal preferredness, i.e., results among which the individual in question would profess indifference or a state of equal satisfaction. When indifference curves are used in mapping a preference “mountain,” their significance is exactly analogous to the equal-elevation “contour lines” on a standard topographical map. The indifference curve labeled $I_{B1}$ is a plausibly drawn contour line for B because his more preferred results (closer to his mountain’s “peak”) are presumably in the northeast direction where B’s stocks of money and goods are increasing. Exactly the opposite reasoning applies for S, whose “peak” of preferredness is approached by moving toward the southwest, where S’s money and goods stocks are maximized. Conceptually, an indifference curve can be drawn through any point on the map of possible results in an Edgeworth box. However, in order to avoid excessive clutter of the diagrams, it is customary to provide only the indifference curve segments actually necessary to understand the general contours of the individual’s preferences about outcome. In this application, it will be necessary to remember only that:

1. for B, any point southwest of his indifference curve $I_{B1}$ is inferior to the promised performance of $R_c$, and
2. for S, any point southwest of his indifference curve $I_{S1}$ through $R_c$ is preferred to performance while any point northwest is inferior.

Exhibit 4.5 is an adaptation of the economic trading model familiarly known as the “Edgeworth Box.” The potential trade partners, B and S, are assumed to have started with money holdings of $80,000 and $20,000 respectively, for an aggregate of $100,000. Also, B is assumed to have a zero stock of “goods,” the product being traded, while S has 100,000 units of goods. By measuring B’s holdings from an origin at the southwest corner of the box and S’s holdings from the northwest corner, we can represent, as a point in the box, any possible division of the aggregates of money and goods between B and S. For
instance, the initial situation (which is sometimes called an “endowment point”) is represented by $R_0$, where B has $O$ goods (bottom scale) and $80,000$ (left-hand scale).\(^{35}\)

A contract between B and S would be representable as an agreement on some particular movement from the original endowment point or status quo ante result $R_0$ to a point such as $R_c$. The motivation for breach will ordinarily be that the perceived advantages of an agreement at the time of contract have been modified by changed conditions. Hence, if a potential breach is to be analyzed, we should expect to find that one of the parties now “regrets” his promise to move from $R_0$ to $R_c$.

Exhibit 4.5 – Indifference Curve Model of Efficient Breach

As Exhibit 4.5 is drawn, changed conditions have shifted the seller’s indifference curve ($I_{s2}$) back toward the northeast origin.\(^{36}\) The seller now perceives the agreed upon post-performance result ($R_c$) as

\(^{35}\) Note that if we look at the situation in terms of S, the point $R_0$ is unchanged. It represents S’s holding of $20,000$ and $100,000$ units of goods.

\(^{36}\) It should be emphasized that conditions inducing breach need not always be represented by a “shiftback” in seller’s indifference curve. It could equally occur where the buyer’s curve shifted. The essential condition to breach
an inferior outcome when compared to the status quo ante (R_o) and thus an incentive to breach exists. In terms of Exhibit 4.5, the just compensation principle of modern contract theory is representable by B’s indifference curve I_b1. This curve not only indicates solutions equivalent (equally preferred by B) to full performance at R_c, but also identifies end results southwest of I_b1 which would fail to provide the buyer with an outcome meeting his original “expectations” of performance at R_c. Since I_b1 does identify results which are, from B’s standpoint, interchangeable with full performance (R_c), this important indifference contour will be termed the “quasi-performance” curve. The concept of quasi-performance expresses the fact that any point on the curve places the buyer in the same preferredness situation as if performance had taken place.

As Exhibit 4.5 is constructed, full breach by the seller would require that S pay B $15,000 in damages in order to move the buyer from the non-performance solution R_o to the appropriate point R_f on the quasi-performance curve. Point R_f is the appropriate solution precisely because it supplies B with the equivalent of performance when no goods at all are delivered. Exhibit 4.5 is also drawn to reflect the fact that S can do even better by tendering partial performance of 15,000 units, thus creating the post-compensation result of R_b. Result R_b, permitting the seller to move to a new higher indifference curve 193, is preferable to both R_c and R_f. It is the most preferred available to S since the law guarantees to B at least some point on the quasi-performance curve. Not only is solution R_b unambiguously superior to performance at R_c, but it also represents an “efficient” end result in the sense that there is no movement from R_b possible without making at least one party worse off. In reaching R_b, the seller would collect the original contract price of $1 per unit for the 15,000 unit partial performance, but would then be required to remit to the buyer $7,500 in damages in order to provide B with his quasi-performance expectation at R_b rather than the no compensation result R_j.

This efficient damage model indicates that upon total or partial non-performance the seller’s obligation is to provide “just compensation” sufficient to place the buyer on his quasi-performance curve. Since the points along this curve represent the buyer’s subjective preferences between performance and compensation, “just compensation” may require consideration of nonobjeetable elements in determining appropriate compensation to the non-breaching buyer.

QUESTIONS

1. What is meant by an “efficient breach” or an “efficient contract”? What are the “error costs” if an efficient breach is prevented? Does the term “inefficient,” as used in this context, mean about the same thing as “wasteful”?

2. The analysis in the text should suggest to you that the law actually encourages promisors not to perform their promises under certain circumstances. Explain why, under the standard contract rules of compensatory damages, the breacher gets all of the so-called efficiency gains when an inefficient contract is not performed.

3. Think about some of the alternatives to the standard system of contract damages. For instance, compare the normal system of compensation for “expectation” damages to one wherein the promisee had the right to specific performance. Would the latter result in the performance of inefficient contracts? How about a system wherein the breacher and breachee “split” the gains fifty-fifty? Why

---

37 The movement from R_o to R_c is a movement down the preferredness surface for the seller. It places S on an indifference curve (I_c2) which is northwest of R_o and thus inferior to R_c in S’s eyes.

38 The point R_b is superior as Exhibit 4.5 is drawn because it makes the seller better off by putting him on indifference curve I_b3 while retaining the buyer on his indifference plateau I_b1. The result is also “optimal” in the sense that any movement away from R_b would necessarily produce an inferior result for one of the parties.
do you think the standard legal rule is for compensatory damages rather than any of these other possible rules?

4. As the example was developed in the text, it was the breacher’s indifference curves that shifted, representing a post-contract change in his assessment of the relative values of goods and-money. Suppose that the breachee’s preferences had also shifted. Would the quasi-contract curve for “expectation” damages be based on the “old” indifference curve running through $R_c$ or on the “new” indifference curve existing at the time of breach? Explain.

E. CONSUMER CHOICE: MODELING PROMISSORY RELIANCE

In the initial “map-reading” section, one question dealt with the mutually agreeable relocation of a well site. That part of the problem involved what should now be recognizable as a variation of the Edgeworth Box trading model, several additional examples of which have been provided in the intervening sections. The other part of the map-reading section, however, involved a somewhat different type of model, one that traditional economics would label a “constrained maximization” model of consumer choice. In such models, the focus is on the adjustment of a single decision-maker to the opportunities confronting him, rather than two-party trade. In the antenna application, the constraint facing the decision-maker was a legal one, the boundaries within which the antenna could be placed, and the “maximand” (thing to be maximized) was the height of the ground as mapped by the contour lines. Almost identical constructions are used in traditional economics to model consumer choice. In most traditional applications, the constraints or boundaries are determined by the individual’s income and the rate at which the things open to choice must be substituted for one another in order to remain within the “budget constraint.” The maximand is the level of utility or “preferredness” as represented by the indifference curve map. This type of indifference curve model of consumer choice was explained briefly as Exhibit 4.2 above and will now be the focus of the next few application sections.

Indifference curve models of consumer choice have a great many legal applications. The following examples involve constructions that might be used in any undergraduate economic theory course, since they explicitly cast the analysis in terms of income and prices. Bear in mind, however, that the “goods” chosen do not have to be standard economic entities nor does the opportunity set open to the chooser have to be defined in terms of income and prices. In short, the antenna-location application is just as legitimate a use of the model as these more standard applications that follow.


A. The Function of Promises: Adaptation by the Promisee

In analyzing the promisee’s reaction to a promise, it is critically important to bear in mind the conceptual distinction between the promise itself and the future benefit that it foretells. By communicating a promise, the promisor informs the promisee about the proposed future receipt of a benefit. The promise itself is merely the production of a piece of information about the future. Normally, advance knowledge of a future transfer will increase the benefit to the promisee because he can more perfectly adapt his consumption decisions to the impending change in wealth. For instance, a person informed of a $25,000 bequest to be made one year hence may revise some of the plans that he otherwise would have followed in the intervening twelve months. Because of the revisions in his plans, the individual can achieve a higher level of satisfaction than if the wealth were transferred without any advance notice. Such adaptive gain from the information embodied in a promise may appropriately be termed “beneficial reliance.” The problem occurs, however, when the transfer foretold by the promise is not actually performed. In this case, the information conveyed by the promise turns out to have been
misleading and the promisee’s induced adaptation in behavior makes him worse off than he would have been without the erroneous expectation of a future benefit. Losses incurred by ill-premised adaptive behavior are commonly termed “detrimental reliance.” Because the role of promises as units of information is so fundamental to the entire analysis developed below, we will use an economic indifference curve model to give more rigorous content to such key legal concepts as reliance and the reasonableness of the promisee’s adaptation process.

1. Reliance Reactions of a Promisee

Exhibit 4.6 will be utilized to develop a very simple intertemporal allocation model, one in which a person must allocate his income between two periods, present and future. A, the potential promisee, begins with $100, which he can divide between consumption now and consumption in the future. In Exhibit 4.6, his possible choices are represented by the straight line budget constraint indicating all combinations of present and future consumption that sum to $100. His preferences about alternative combinations of present and future consumption are summarized by the indifference curves, which define a kind of topographic map of the desirability of different present-future consumption patterns. On these assumptions, the highest preference level consistent with the scarce resources is point e, where indifference curve II is tangent to the budget constraint. This consumption allocation involves spending $50 now together with a planned expenditure of $50 in the future.

Exhibit 4.6 – Intertemporal Model of Promissory Reliance

Suppose now that B promises A a transfer of $50 to be made in the future period and that A believes the promise. Even neglecting the possibility of borrowing against his future wealth, A’s budget constraint
will shift out to the right in Exhibit 4.6. The new constraint indicates that, although no more than $100 can be spent in the present, the two-period consumption levels may now total $150 rather than $100. Based on the new information, A’s best two-period plan would be point e2. A is thus led to revise his current consumption upward to $75 and to project $75 worth of future consumption.

Because detrimental reliance is widely regarded as a basis for damage computation, an advantage of the Exhibit 4.6 model is that it clearly illustrates why it is a mistake to use A’s observable action in reliance on a promise as the measure of his damages. Under the facts assumed, A’s observable reliance is the $25 extra he spends in period 1. If, however, he were awarded this $25 as damages for breach, his ultimate position would be at e5 on indifference curve I5. He would have spent $75 in the first period and, including the $25 damages, would have available $50 for the final period. But since return to the status quo ante requires only that indifference curve II be achieved, the true reliance damages are equal only to the lesser amount indicated by the horizontal distance between e3 and e6 in Exhibit 4.6. The common-sense explanation for this, of course, is that detrimental adaptation in behavior is usually only a partial rather than a total loss. In this case, A did get some benefit from the excessive $25 consumption even though not as much as he would have if the consumption had been postponed to the optimal time. Reliance is, simply, the opportunity cost of the broken promise. Thus, true damages are measured by the difference between the value of the stream of consumption choices not taken—indifference curve I1—and those choices induced by the promise—indifference curve I3. Because it is important to distinguish compensation based on observable reliance from true reliance damages, we shall refer to the former as “reimbursement damages.”

Even where the precise meaning is imperfectly captured, the notion of detrimental reliance tends to be better understood than that of beneficial reliance. This is unfortunate because the production of beneficial reliance is perhaps the principal social rationale of promising; the risk of detrimental reliance is merely the unavoidable concomitant cost. Exhibit 4.6 aptly illustrates the beneficial consequences of promising when the promise is performed. Assume, for instance, that the $50 in our hypothetical is merely transferred to A in period two without any advance warning. Not knowing about the wealth increase, A will have committed himself to plan e,. When period two arrives, A will unexpectedly find himself with $100 to spend At this point, the best available choice is at e4, which is on a lower indifference curve than e2, the point that would have been achievable had A obtained advance knowledge of the transfer. In commonsense terms, the difference between I2 and I4 illustrates the benefits to A of being able to adjust, because of the promise, to revised expectations about the future.

QUESTIONS

1. How is an indifference curve defined? Why are indifference curves useful devices in analyzing damages that have a restitutio ad integrum or compensatory purpose?

2. Draw an additional budget constraint, with the same slope as the other two, such that it is tangent to indifference curve I4. Explain the relevance of that line to the following problem:

I had intended to give A $100, but was unable to tell him about it in advance so that he could rely on it. Now I find it possible to inform him of my intent and give assurances of my performance.

---

24 For simplicity, we assume that the interest rate on money is zero in order to produce a one-to-one tradeoff between present and future consumption. The introduction of interest is irrelevant to this analysis because it merely alters the rate of tradeoff, reflected in the slopes of the budget lines.

25 The promisee’s consumption of all goods with nonzero income elasticities will be modified. [See the section on “elasticity” measures later in this chapter.] His intertemporal consumption stream will be adjusted to the higher wealth level created by the gift.
But it occurs to me that I can now make a promise of somewhat less than the originally intended $100 and achieve the same effect. How much less can I give?

Does this suggest to you that the promisor as well as the promisee may benefit from having promises be enforceable? See Posner, “Gratuitous Promises in Economics and Law,” 6 J.Leg.Studies 411 (1977).

3. Suppose the promisee thought that there were only a 50% chance that the promise would be performed. He would have to pick his position on the vertical axis on the assumption that one of two possible budget lines would determine his eventual position on the horizontal axis. Why? The exact solution to this choice problem is difficult to derive graphically from an indifference curve model (although an analogous mathematical model handles it quite easily). Consider the problem in an approximate way, however, and take a guess at what decision the promisee would make. Why is this inevitably a “compromise” decision that is always somewhat “wrong” as viewed ex post? And what is the loss from not knowing whether the promise is completely reliable or not?

F. PERMISSIBLE AND IMPERMISSIBLE EFFECTS OF GOVERNMENT GRANTS

MATHEWS ET AL. v. MASSELL


RICHARD C. FREEMAN, District Judge.

This is an action brought by plaintiffs, citizens and taxpayers of the City of Atlanta on behalf of themselves and all others similarly situated against defendants, the mayor and members of the Board of Aldermen of the City of Atlanta. Plaintiffs challenge the way in which defendants have utilized a portion of the funds received by the City pursuant to the “State and Local Fiscal Assistance Act of 1972” popularly known as the Revenue Sharing Act.

On February 22, 1973, a hearing was held on plaintiffs’ motion for preliminary injunction. A stipulation of the relevant facts was entered into by the parties prior to the hearing.

* * *

The relevant facts as stipulated by the parties are here summarized. Pursuant to the Revenue Sharing Act, which went into effect on October 20, 1972, the City of Atlanta received approximately $5,962,169 in revenue sharing funds for the year 1972. It has been anticipated that the City will receive approximately an additional $1.5 million in Revenue Sharing funds per quarter for the first three quarters of 1973.

* * *

The effect of the * * * ordinances [being challenged in this case] is as follows: The Revenue Sharing funds received for 1972 are to be utilized to raise firemen’s salaries, an approved use under the Revenue Sharing Act. The $4.5 million in Revenue Sharing funds anticipated for the first three quarters of 1973 is likewise designated as being used to pay firemen’s salaries, thereby relieving the general fund of this obligation. The $4.5 million from the general fund originally slated for the payment of firemen’s salaries is then to be transferred to the City’s Water/Sewer Fund to reimburse it for the $4.5 million reduction in the Water/Sewer Fund resulting from the rebate conferred upon firms and individuals having water accounts with the City. It is the Revenue Sharing Funds anticipated for the first three quarters of 1973 (a total of $4.5 million) which are to be used to release the $4.5 million in general funds which defendants propose to disburse in the form of a water/sewer rebate.

In his State of the City Annual Message, delivered on Tuesday, January 2, 1973, defendant [Mayor] Massell stated that he intended to return directly to the citizenry of Atlanta a portion of the Revenue Sharing funds in the amount of $4.5 million. On February 12, 1973, in a news release, defendant Massell
announced that he was going to use the Revenue Sharing funds at least in part to “create some relief for the monthly budget of the average Atlanta householder * * * [T]his has made it possible for us to reduce the water/sewer rates on over 1,000,000 accounts by $4 per month per meter * * * a total benefit to the public of $4.5 million.” Affidavits from three defendants, members of the Board of Aldermen, state that the series of ordinances described above was designed to carry out a plan to return $4.5 million of Federal Revenue Sharing funds to firms and individuals having active water/sewer accounts with the City of Atlanta.

Plaintiffs claim that the actions of the defendants, as set forth above, constitute a violation of Sec. 103(a) of the Revenue Sharing Act, which provides as follows:

(a) In general Funds received by units of local government under this sub[title] may be used only for priority expenditures. For purposes of this [title], the term “priority expenditures” means only--

(1) ordinary and necessary maintenance and operating expenses for

(A) public safety (including law enforcement, fire protection, and building code enforcement)
(B) environmental protection (including sewage disposal, sanitation, and pollution abatement)
(C) public transportation (including transit systems and streets and roads)
(D) health
(E) recreation
(F) libraries
(G) social services for the poor or aged, and
(H) financial administration; and

(2) ordinary and necessary capital expenditures authorized by law.

It is plaintiffs’ contention that defendants have violated the Act by, in effect, spending $4.5 million for water sewer rebates, which is not one of the priority uses set forth in Sec. 103(a).

* * *

The requisite financial harm to the taxpayers of the City of Atlanta is found in the Revenue Sharing Act itself. Section 123(a) of the Act provides * * * for a return of any Revenue Sharing funds spent in violation of the priority use requirements set forth in Sec. 103(a), plus a penalty of 10%. If defendants were to carry out their proposed plan to rebate $4.5 million of the Revenue Sharing funds and if, as plaintiffs allege, such use is in violation of Sec. 103(a), the City would be liable to the Secretary of the Treasury for $4.5 million plus 10%, a total of $4.95 million.

* * *

The court must now proceed to a consideration of the merits of plaintiffs’ claim that the proposed plan of defendants to distribute a water/sewer rebate is in violation of Sec. 103(a) of the Revenue Sharing Act. It is clear from Sec. 103(a), and defendants do not contend otherwise, that the defendants could not have directly used any portion of the Revenue Sharing funds to make a rebate to those with water sewer accounts. A water/sewer rebate does not fall within one of the “priority” uses permitted by Sec. 103(a).

Defendants do contend, however, that they have fully satisfied their duties under the Act by placing the Revenue Sharing funds in a trust account for the payment of firemen’s salaries, a priority use. It is their position that the Revenue Sharing Act imposes restrictions only on the specific funds which are received by operation of the Act. They contend that the Act imposes no restrictions upon the City’s general funds
which are freed up by the influx of Revenue Sharing funds. Thus in the present case they state that $10.5 million of general funds would have been used for the payment of firemen’s salaries but for the receipt of the Revenue Sharing funds. Defendants contend that use of the Revenue Sharing funds for firemen’s salaries allows them to make any use whatsoever of the general funds which would otherwise have been put to such use.

It is true that the Revenue Sharing Act does not specifically impose any restrictions upon the use of legitimately freed-up funds. Thus the Act seems clearly to have contemplated that the infusion of Revenue Sharing funds into state and local governments would permit future tax relief to the hard-pressed taxpayers of those governments. Further, there is no requirement that a local government maintain at pre-Revenue Sharing levels its spending on “priority expenditures.” There is a clear difference, however, between funds which are legitimately freed up by the designation of federal Revenue Sharing funds to provide municipal services which otherwise would have to have been paid for out of general City funds, and funds which are transferred from one account to another simply to avoid the restrictions imposed by Sec. 103(a) of the Act. The actions of defendants, the public statements made by defendant Mayor Massell and the affidavits of three of the defendant members of the Board of Aldermen, show clearly that the steps taken by defendants were designed to carry out a plan to return $4.5 million in Revenue Sharing funds to certain taxpayers, the defendants having decided to confer such tax relief by way of rebates on the water/sewer accounts.

Such an attempt to avoid the clear restrictions of federal statute cannot be accepted. While there is as yet no case law on the Revenue Sharing Act, in the interpretation of federal statutes generally, the courts have long made it clear that Congressional intent cannot be overridden by sham transactions. In the areas of federal income taxation, securities law, and antitrust statutes, the courts have consistently refused to exalt artifice over reality or to ignore the actual substance of a particular set of transactions. Thus the court must recognize that the defendants have merely transferred funds from one account to another in an effort to disguise the fact that they plan to distribute $4.5 million of Revenue Sharing funds to the holders of water/sewer accounts.

Defendants would seem to argue, however, that their proposed plan does not conflict with Congressional intent. Defendants, while recognizing the restrictions set forth in Sec. 103(a), contend that such restrictions were not meant to have any force and that thus defendants’ proposal to circumvent those provisions would not actually run counter to the intent of Congress. Defendants’ contention that Congress intended no effective restrictions on the funds is clearly belied by an examination of the legislative history of the Act.

* * *

One final argument by defendants is that Sec. 103(a) has no effect because it is so difficult to enforce. They draw support for this argument from the Senate report on the Senate version of the bill in which the committee justified the Senate’s deletion of the restrictions on a local government’s use of the funds by the argument that enforcement of the restrictions would be impossible and that therefore any restrictions would be illusory. And facially this argument is persuasive, but * * * it was in effect overruled by the action of the House-Senate conference and by the passage of the Act in its present form by both houses of Congress. Moreover, the Act itself provides an enforcement mechanism by requiring the Secretary of the Treasury to establish accounting and auditing procedures (Sec. 123c) and also by providing for imposition of a penalty if funds are spent in violation of Sec. 103(a).

Defendants seek support from a recent article on Revenue Sharing published in the Harvard Journal on Legislation. “The Revenue Sharing Act of 1972: Untied and Untraceable Dollars from Washington”, 10 Harv.J.Legis. 276 (1973). In discussing the priority expenditures provision of the Act, the article makes the point that violations of Sec. 103(a) will be extremely difficult to discover and prove, for Revenue Sharing funds will be commingled in fact with other local funds, even if the books of account are kept separately. Such problems of proof will undoubtedly arise: however, in the present case the use of
Revenue Sharing funds in violation of Sec. 103(a) has been clearly proved by plaintiffs in large part by the statements of defendants themselves.

In conclusion then, the court has looked to the substance of defendants’ actions and determined that defendants’ proposed plan would entail the expenditure of Federal Revenue Sharing funds for other than one of the priority expenditures set forth in Sec. 103(a) In light of this determination the court need not consider plaintiffs’ other claims.

Accordingly, judgment is hereby entered for the plaintiffs. Defendants, their agents, employees, successors in office and attorneys are hereby permanently, enjoined from utilizing $4.5 million in Revenue Sharing funds in the manner proposed, that is, to make a reduction in water/sewer rates or to give a rebate to all those with active water/sewer accounts.

It is so ordered.

QUESTIONS

1. Assume that a grantee government has a set of preferences representable as a set of indifference curves. [The section above on group decisionmaking under majority rule should suggest that this assumption about the existence of a well ordered preference mapping for a voting group (in this case, the City Council) is a debatable simplification. Do you understand why? When you have completed this current question set, ask yourself to what extent the making of the assumption alters the points being discussed in the present exercise.] Construct a model showing quantity of Priority Expenditure goods on the horizontal axis and quantity of “Other” goods on the vertical axis. Show a pre-grant budget constraint and a hypothetical set of indifference curves.

2. After the receipt of the revenue-sharing grant, the grantee government’s new economic constraint will lie to the right of and parallel to the original budget constraint. Why?

3. On this same model, use a distinctive line to distinguish the set of permissible choices under Sec. 103(a) from those which would violate the Act. This line should, in effect, represent the proverbial “bright yellow line” that separates legal and illegal budgetary decisions. If you think there is more than one possible interpretation of where this legal constraint lies, indicate each possible answer and explain the underlying reasoning.

4. On completion of the last question, you should have what is really just a variant of the standard consumer choice model, except that it has two constraints, one economic and one legal. What is the “price” for violating the statutory constraint? Is that price probabilistic (i.e., a “prospect”)? Have you incorporated this element in your model?

5. Defendants in this case allegedly supplied the proof against themselves. Absent such admissions, what possible evidence would permit an inference that the funds had been used improperly? For a discussion of the “tracing” problem, see Comment, “The Revenue Sharing Act of 1972: Untied and Untraceable Dollars from Washington,” 10 Harv.J.Legis. 276 (1973). An empirical analysis of the extent to which funds were used for substitutional purposes rather than new spending can be found in Nathan, “The Uses of Shared Revenue,” 30 J.Finance 557 (1975).

TILTON V. RICHARDSON

403 U.S. 672 (1971)

[Mr. Chief Justice Burger.]

This appeal presents important constitutional questions as to federal aid for church-related colleges and universities under Title I of the Higher Education Facilities Act of 1963, 77 Stat. 364, as amended, 20
USC Secs. 711-721 (1964 ed. and Supp. V), which provides construction grants for buildings and facilities used exclusively for secular educational purposes. We must determine first whether the Act authorizes aid to such church-related institutions and, if so, whether the Act violates either the Establishment or Free Exercise Clauses of the First Amendment.

* * *

**The Act authorizes federal grants and loans to “institutions of higher education” for the construction of a wide variety of “academic facilities.” But Sec. 751(a)(2) (1964 ed., Supp. V) expressly excludes “any facility used or to be used for sectarian instruction or as a place for religious worship, or * * * any facility which * * * is used or to be used primarily in connection with any part of the program of a school or department of divinity * * *.”**

* * *

We are satisfied that Congress intended the Act to include all colleges and universities regardless of any affiliation with or sponsorship by a religious body. Congress defined “institutions of higher education,” which are eligible to receive aid under the Act, in broad and inclusive terms. Certain institutions, for example, institutions that are neither public nor nonprofit, are expressly excluded, and the Act expressly prohibits use of the facilities for religious purposes. But the Act makes no reference to religious affiliation or nonaffiliation. Under these circumstances “institutions of higher education” must be taken to include church-related colleges and universities.

Numerous cases considered by the Court have noted the internal tension in the First Amendment between the Establishment Clause and the Free Exercise Clause. *Walz v. Tax Comm’n*, 397 U.S. 664 (1970), is the most recent decision seeking to define the boundaries of the neutral area within which the legislature may legitimately act. There, as in other decisions, the Court treated the three main concerns against which the Establishment Clause sought to protect: “sponsorship, financial support, and active involvement of the sovereign in religious activity.” Id., at 668.

* * * [W]e consider four questions: First, does the Act reflect a secular legislative purpose? Second, is the primary effect of the Act to advance or inhibit religion? Third, does the Administration of the Act foster an excessive government entanglement with religion? Fourth, does the implementation of the Act inhibit the free exercise of religion?

The stated legislative purpose appears in the preamble where Congress found and declared that “the security and welfare of the United States require that this and future generations of American youth be afforded ample opportunity for the fullest development of their intellectual capacities, and that this opportunity will be jeopardized unless the Nation’s colleges and universities are encouraged and assisted in their efforts to accommodate rapidly growing numbers of youth who aspire to a higher education.” 20 USC Sec. 701. This expresses a legitimate secular objective entirely appropriate for governmental action.

The simplistic argument that every form of financial aid to church-sponsored activity violates the Religion Clause was rejected long ago in *Bradfield v. Roberts*, 175 U.S. 291 (1899). There a federal construction grant to a hospital operated by a religious order was upheld. Here the Act is challenged on the ground that its primary effect is to aid the religious purposes of church-related colleges and universities. Construction grants surely aid these institutions in the sense that the construction of buildings will assist them to perform their various functions. But bus transportation, textbooks, and tax exemptions all gave aid in the sense that religious bodies would otherwise have been forced to find other sources from which to finance these services. Yet all of these forms of governmental assistance have been upheld. [Citations omitted.] The crucial question is not whether some benefit accrues to a religious institution as a consequence of the legislative program, but whether its principal or primary effect advances religion.

The Act itself was carefully drafted to ensure that the federally subsidized facilities would be devoted to the secular and not the religious function of the recipient institutions. It authorizes grants and loans
only for academic facilities that will be used for defined secular purposes and expressly prohibits their use for religious instruction, training, or worship. These restrictions have been enforced in the Act’s actual administration, and the record shows that some church related institutions have been required to disgorge benefits for failure to obey them.

** * ***

Appellants * *** rely on the argument that government may not subsidize any activities of an institution of higher learning that in some of its programs teaches religious doctrines. This argument rests on Everson where the majority stated that the Establishment Clause barred any “tax * *** levied to support any religious * *** institutions * *** whatever form they may adopt to teach or practice religion.” 330 U.S. at 16, 67 S.Ct. at 511. In Allen, however, it was recognized that the Court had fashioned criteria under which an analysis of a statute’s purpose and effect was determinative as to whether religion was being advanced by government action. 392 U.S. at 243; Abington School District v. Schempp, 374 U.S. 203, 222 (1963).

Under this concept appellants’ position depends on the validity of the proposition that religion so permeates the secular education provided by church-related colleges and universities that their religious and secular functions are in fact inseparable. The argument that government grants would thus inevitably advance religion did not escape the notice of Congress. It was carefully and thoughtfully debated, 109 Cong.Rec. 19474-19475, but was found unpersuasive. It was also considered by this Court in Allen. There the Court refused to assume that religiosity in parochial elementary and secondary schools necessarily permeates the secular education that they provide.

The record, similarly, provides no basis for any such assumptions here. * * * There is no evidence that religion creeps into the use of any of these facilities.

** * ***

We next turn to the question of whether excessive entanglements characterize the relationship between government and church under the Act. Walz v. Tax Comm’n, supra, at 674-676. Our decision today in Lemon v. Kurtzman and Robinson v. DiCenzo has discussed and applied this independent measure of constitutionality under the Religion Clauses. There we concluded that excessive entanglements between government and religion were fostered by Pennsylvania and Rhode Island statutory programs under which state aid was provided to parochial elementary and secondary schools. Here, however, three factors substantially diminish the extent and the potential danger of the entanglement.

** * ***

There are generally significant differences between the religious aspects of church-related institutions of higher learning and parochial elementary and secondary schools. The “affirmative if not dominant policy” of the instruction in pre-college church schools is “to assure future adherents to a particular faith by having total control of their education at an early age.” Walz v. Tax Comm’n, supra, at 671. There is substance to the contention that college students are less impressionable and less susceptible to religious indoctrination. Common observation would seem to support that view, and Congress may well have entertained it. The skepticism of the college student is not an inconsiderable barrier to any attempt or tendency to subvert the congressional objectives and limitations. Furthermore, by their very nature, college and postgraduate courses tend to limit the opportunities for sectarian influence by virtue of their own internal disciplines. Many church-related colleges and universities are characterized by a high degree of academic freedom and seek to evoke free and critical responses from their students.

Since religious indoctrination is not a substantial purpose or activity of these church-related colleges and universities, there is less likelihood than in primary and secondary schools that religion will permeate the area of secular education. This reduces the risk that government aid will in fact serve to support
religious activities. Correspondingly, the necessity for intensive government surveillance is diminished and the resulting entanglements between government and religion lessened. Such inspection as may be necessary to ascertain that the facilities are devoted to secular education is minimal and indeed hardly more than the inspections that States impose over all private schools within the reach of compulsory education laws.

The entanglement between church and state is also lessened here by the nonideological character of the aid that the Government provides. Our cases from Everson to Allen have permitted church-related schools to receive government aid in the form of secular, neutral, or nonideological services, facilities, or materials that are supplied to all students regardless of the affiliation of the school that they attend. In Lemon and DiCenso, however, the state programs subsidized teachers, either directly or indirectly. Since teachers are not necessarily religiously neutral, greater governmental surveillance would be required to guarantee that state salary aid would not in fact subsidize religious instruction. There we found the resulting entanglement excessive. Here, on the other hand, the Government provides facilities that are themselves religiously neutral. The risks of Government aid to religion and the corresponding need for surveillance are therefore reduced.

Finally, government entanglements with religion are reduced by the circumstance that, unlike the direct and continuing payments under the Pennsylvania program, and all the incidents of regulation and surveillance, the Government aid here is a one-time, single-purpose construction grant. There are no continuing financial relationships or dependencies, no annual audits, and no government analysis of an institution’s expenditures on secular as distinguished from religious activity. Inspection as to use is a minimal contact.

No one of these three factors standing alone is necessarily controlling; cumulatively all of them shape a narrow and limited relationship with government which involves fewer and less significant contacts than the two state schemes before us in Lemon and DiCenso. The relationship therefore has less potential for realizing the substantive evils against which the Religion Clauses were intended to protect.

We think that cumulatively these three factors also substantially lessen the potential for divisive religious fragmentation in the political arena. This conclusion is admittedly difficult to document, but neither have appellants pointed to any continuing religious aggravation on this matter in the political processes. Possibly this can be explained by the character and diversity of the recipient colleges and universities and the absence of any intimate continuing relationship or dependency between government and religiously affiliated institutions. The potential for divisiveness inherent in the essentially local problems of primary and secondary schools is significantly less with respect to a college or university whose student constituency is not local but diverse and widely dispersed.

Finally, we must consider whether the implementation of the Act inhibits the free exercise of religion in violation of the First Amendment. Appellants claim that the Free Exercise Clause is violated because they are compelled to pay taxes, the proceeds of which in part finance grants under the Act. Appellants, however, are unable to identify any coercion directed at the practice or exercise of their religious beliefs. Board of Education v. Allen, supra, 392 U.S., at 248-249. Their share of the cost of the grants under the Act is not fundamentally distinguishable from the impact of the tax exemption sustained in Walz or the provision of textbooks upheld in Allen.

We conclude that the Act does not violate the Religion Clauses of the First Amendment except that of Sec. 754(b)(2) providing a 20-year limitation on the religious use restrictions contained in Sec. 751(a)(2). We remand to the District Court with directions to enter a judgment consistent with this opinion.

Vacated and remanded.

Mr. Justice DOUGLAS, with whom Mr. Justice BLACK and Mr. Justice MARSHALL concur, dissenting in part.

* * *
The fact that money is given once at the beginning of a program rather than apportioned annually as in Lemon and DiCenso is without constitutional significance. Thus it is hardly impressive that rather than giving a smaller amount of money annually over a long period of years, Congress instead gives a large amount all at once. The plurality’s distinction is in effect that small violations of the First Amendment over a period of years are unconstitutional (see Lemon and DiCenso) while a huge violation occurring only once is de minimis. I cannot agree with such sophistry.

What I have said in Lemon and in the DiCenso cases decided today is relevant here. The facilities financed by taxpayers’ funds are not to be used for “sectarian” purposes. Religious teaching and secular teaching are so enmeshed in parochial school that only the strictest supervision and surveillance would insure compliance with the condition. Parochial schools may require religious exercises, even in the classroom. A parochial school operates on one budget. Money not spent for one purpose becomes available for other purposes. Thus the fact that there are no religious observances in federally financed facilities is not controlling because required religious observances will take place in other buildings. Our decision in Engel v. Vitale, 370 U.S. 421, held that a requirement of a prayer in public schools violated the Establishment Clause. Once these schools become federally funded they become bound by federal standards [Ivanhoe Irrig. Dist. v. McCracken, 357 U.S. 275, 296; Rosado v. Wyman, 397 U.S. 397, 427 (concurring opinion): Simkins v. Moses H. Cone Memorial Hosp., 4 Cir., 323 F.2d 959] and accordingly adherence to Engel would require an end to required religious exercises. That kind of surveillance and control will certainly be obnoxious to the church authorities and if done will radically change the character of the parochial school. Yet if that surveillance is not searching and continuous, this federal financing is obnoxious under the Establishment and Free Exercise Clauses for the reasons stated in the companion cases.

In other words, surveillance creates an entanglement of government and religion which the First Amendment was designed to avoid. Yet after today’s decision there will be a requirement of surveillance which will last for the useful life of the building and as we have previously noted, “[i]t is hardly lack of due process for the Government to regulate that which it subsidizes.” Wickard v. Filburn, 317 U.S. 111, 131. The price of the subsidy the Act is violation of the Free Exercise Clause. Could a course in the History of Methodism be taught in a federally financed building? Would a religiously slanted version of the Reformation of Quebec politics under Duplessis be permissible? How can the Government know what is taught in the federally financed building without a continuous auditing of classroom instruction? Yet both the Free Exercise Clause and academic freedom are violated when the Government agent must be present to determine whether the course content is satisfactory.

As I said in the Lemon and DiCenso cases, a parochial school is a unitary institution with subtle blending of sectarian and secular instruction. Thus the practices of religious schools are in no way affected by the minimal requirement that the government financed facility may not “be used for sectarian instruction or as a place for religious worship.” Money saved from one item in the budget is free to be used elsewhere. By conducting religious services in another building, the school has--rent free--a building for nonsectarian use. This is not called Establishment simply because the government retains a continuing interest in the building for its useful life, even though the religious schools need never pay a cent for the use of the building.

* * *

I would reverse the judgment below.

QUESTIONS

1. Create a model similar to the one used for Mathews v. Massel above. Place quantities of “Secular Buildings” and “Religious Uses” on the vertical and horizontal axes, respectively. Assume that the
federal grant program paid for exactly 50% of the cost of secular buildings. What are the pre- and post-grant economic constraints faced by a church-related university? Why is the shift in the budget constraint not a parallel shift as illustrated by the revenue sharing grants?

2. If the grant moves the church-related institution onto a higher “indifference curve,” is that fact relevant for an analysis of the Act’s constitutionality? Will such a movement to a higher indifference curve always take place?

3. “The proper constitutional test for this program is whether it has the predictable effect of increasing, or stimulating, the religious activities of the grantee.” On your indifference curve model, depict in some appropriate manner the implicit constitutional constraint represented by the quoted statement. Do you agree with the quote? If not, depict the constitutional constraint implied by the test you deem proper. Explain your answer.

4. Does one’s attitude towards the issues in *Tilton v. Richardson* depend on empirical data or can the problem be framed adequately in theoretical terms?

5. Viewing both *Mathews v. Massell* and *Tilton v. Richardson* against the background of the indifference curve models, what comments do you have on the decisions? If any grant is given at all, is some degree of unintended effect inevitable? Does the magnitude of the unintended effects depend on the form of the grant?

Hypothetical: Christian Friars Wine Sales to Military

By the year 2035, the previously upper middle-class penchant for quiche and white wine has made deep inroads into the traditional American preference for pizza and beer. Nowhere is this more apparent than in federal procurement contracts for the military services, where wine consumption has recently outstripped beer in mess halls and enlisted personnel’s recreational clubs.

Low bidder on a recent government contract for military forces wine supply was the Christian Friars Vinery, a west-coast wine producer run by a religious order that devotes all of the profits from its wine sales to subsidize the religiously-oriented secondary schools in which its members teach. A ruling of the General Services Administration required that the Department of Defense reject the Vinery’s bid and award the contract to the next lowest qualified bidder, reasoning that the transaction with the Christian Friars would involve a constitutionally impermissible relationship between church and state. The Friars have filed suit in federal court alleging that the refusal to accept its bid is an unconstitutional discrimination on the basis of religion. “Religion shouldn’t be any part of this,” the Brother Vintmaster, stated in a press conference. “The government wants wine and we have wine to sell. What difference does it make how we use our profits?”

On what principles should this case be decided? Does it make a difference if the labels of the wine bottles bear a trademark consisting of a Friar pausing to meditate while he harvests a basket of wine grapes? To what extent is the position of the Christian Friars Vinery analogous to that of Albertus Magnus College in the *Tilton* case?

G. BRIBES VERSUS THREATS

If one takes seriously the notion of opportunity cost, there is a dismaying similarity between the impact of bribes and threats: a bribe foregone is as much a cost as a tax or penalty paid out. To the extent that certain activities are deemed to be protected from interference by the State, the issue may arise as to whether the State may achieve with the carrot what it would be forbidden to achieve via the stick or, alternatively, to achieve by the back door what it cannot at the front. The next few real cases and hypotheticals should be read with at least two major questions in mind. One is the distinction, if any,
between using bribes or threats to modify behavior. The other is whether the “bribe” given by the State is real or in some sense only apparent.

UNITED STATES V. BUTLER

297 U.S. 1 (1936)

[The Agricultural Adjustment Act of 1933, in order to raise farm -prices by reducing production, authorized the secretary of agriculture- -to contract with farmers to reduce productive acreage. Payments were made out of the revenues of a processing tax exacted from the processors of the agricultural products. A cotton processor challenged the tax as beyond Congress’ constitutional authority.

Mr. Justice ROBERTS delivered the opinion of the Court.

The Government asserts that * * * Article I, Section 8 of the Constitution authorizes the contemplated expenditure of the funds raised by the tax. This contention presents the great and the controlling question in the case.

* * *

The clause thought to authorize the legislation, the first, confers upon the Congress power “to lay and collect Taxes, Duties, Imposts and Excises, to pay the Debts and provide for the common Defense and general Welfare of the United States. * * *” It is not contended that this provision grants power to regulate agricultural production upon the theory that such legislation would promote the general welfare. The Government concedes that the phrase “to provide for the general welfare” qualifies the power to “to lay and collect taxes.” The view that the clause grants power to provide for the general welfare, independently of the taxing power, has never been authoritatively accepted. Mr. Justice Story points out that if it were adopted “it is obvious that under color of the generality of the words, ‘to provide for the common defence and general welfare,’ the government of the United States is, in reality, a government of general and unlimited powers, notwithstanding the subsequent enumeration of specific powers.” The true construction undoubtedly is that the only thing granted is the power to tax for the purpose of providing funds for payment of the nation’s debts and making provision for the general welfare.

* * *

We are not now required to ascertain the scope of the phrase “general welfare of the United States” or to determine whether an appropriation in aid of agriculture falls within it. Wholly apart from that question, another principle embedded in our Constitution prohibits the enforcement of the Agricultural Adjustment Act. The act invades the reserved rights of the states. It is a statutory plan to regulate and control agricultural production, a matter beyond the powers delegated to the federal government. The tax, the appropriation of the funds raised, and the direction for their disbursement, are but parts of the plan. They are but means to an unconstitutional end.

From the accepted doctrine that the United States is a government of delegated powers, it follows that those not expressly granted, or reasonably to be implied from such as are conferred, are reserved to the states or to the people. To forestall any suggestion to the contrary, the Tenth Amendment was adopted. The same proposition, otherwise stated, is that powers not granted are prohibited. None to regulate agricultural production is given, and therefore legislation by Congress for that purpose is forbidden.

It is an established principle that the attainment of a prohibited end may not be accomplished under the pretext of the exertion of powers which are granted.

* * *
The power of taxation, which is expressly granted, may, of course, be adopted as a means to carry into operation another power also expressly granted. But resort to the taxing power to effectuate an end which is not legitimate, not within the scope of the Constitution, is obviously inadmissible.

* * *

If the taxing power may not be used as the instrument to enforce a regulation of matters of state concern with respect to which the Congress has no authority to interfere, may it, as in the present case, be employed to raise the money necessary to purchase a compliance which the Congress is powerless to command? The Government asserts that whatever might be said against the validity of the plan if compulsory, it is constitutionally sound because the end is accomplished by voluntary cooperation. There are two sufficient answers to the contention. The regulation is not in fact voluntary. The farmer, of course, may refuse to comply, but the price of such refusal is the loss of benefits. The amount offered is intended to be sufficient to exert pressure on him to agree to the proposed regulation.19

The power to confer or withhold unlimited benefits is the power to coerce or destroy. If the cotton grower elects not to accept the benefits, he will receive less for his crops; those who receive payments will be able to undersell him. The result may well be financial ruin. * * *

* * *

But if the plan were one for purely voluntary co-operation it would stand no better so far as federal power is concerned. At best it is a scheme for purchasing with federal funds submission to federal regulation of a subject reserved to the states.

It is said that Congress has the undoubted right to appropriate money to executive officers for expenditure under contracts between the government and individuals; that much of the total expenditures is so made. But appropriations and expenditures under contracts for proper governmental purposes cannot justify contracts which are not within federal power. And contracts for the reduction of acreage and the control of production are outside the range of that power. An appropriation to be expended by the United States under contracts calling for violation of a state law clearly would offend the Constitution. Is a statute less objectionable which authorizes expenditure of federal moneys to induce action in a field in which the United States has no power to intermeddle? The Congress cannot invade state jurisdiction to compel individual action; no more can it purchase such action.

* * *

Since, as we have pointed out, there was no power in the Congress to impose the contested exaction, it could not lawfully ratify or confirm what an executive officer had done in that regard. Consequently the Act of 1935 does not affect the rights of the parties.

The judgment is affirmed.

Mr. Justice STONE, dissenting.

I think the judgment should be reversed.

* * *

---

19 U.S. Dept. of Agriculture, Agriculture Adjustment, p. 9. “Experience of cooperative associations and other groups has shown that without such Government support, the efforts of the farmers to band together to control the amount of their product sent to nothing. Almost always, under such circumstances, there has been a noncooperating minority, which, refusing to go along with the rest, has stayed on the outside and tried to benefit from the sacrifices the majority has made. * * * It is to keep this noncooperating minority in line, or at least prevent it from doing harm to the majority, that the power of the Government has been marshaled behind the adjustment.
Here regulation, if any there be, is accomplished not by the tax but by the method by which its proceeds are expended, and would equally be accomplished by any like use of public funds, regardless of their source.

The method may be simply stated. Out of the available fund payments are made to such farmers as are willing to curtail their productive acreage, who in fact do so and who in advance have filed their written undertaking to do so with the Secretary of Agriculture. In saying that this method of spending public moneys is an invasion of the reserved powers of the states, the Court does not assert that the expenditure of public funds to promote the general welfare is not a substantive power specifically delegated to the national government, as Hamilton and Story pronounced it to be. It does not deny that the expenditure of funds for the benefit of farmers and in aid of a program of curtailment of production of agricultural products, and thus of a supposedly better ordered national economy, is within the specifically granted power. But it is declared that state power is nevertheless infringed by the expenditure of the proceeds of the tax to compensate farmers for the curtailment of their cotton acreage. Although the farmer is placed under no legal compulsion to reduce acreage, it is said that the mere offer of compensation for so doing is a species of economic coercion which operates with the same legal force and effect as though the curtailment were made mandatory by Act of Congress. In any event it is insisted that even though not coercive the expenditure of public funds to induce the recipients to curtail production is itself an infringement of state power, since the federal government cannot invade the domain of the states by the “purchase” of performance of acts which it has no power to compel.

Of the assertion that the payments to farmers are coercive, it is enough to say that no such contention is pressed by the taxpayer, and no such consequences were to be anticipated or appear to have resulted from the administration of the Act. The suggestion of coercion finds no support in the record or in any data showing the actual operation of the Act. Threat of loss, not hope of gain, is the essence of economic coercion. Members of a long depressed industry have undoubtedly been tempted to curtail acreage by the hope of resulting better prices and by the proffered opportunity to obtain needed ready money. But there is nothing to indicate that those who accepted benefits were impelled by fear of lower prices if they did not accept, or that at any stage in the operation of the plan a farmer could say whether, apart from the certainty of cash payments at specified times, the advantage would lie with curtailment of production plus compensation, rather than with the same or increased acreage plus the expected rise in prices which actually occurred.

The Constitution requires that public funds shall be spent for a defined purpose, the promotion of the general welfare. Their expenditure usually involves payment on terms which will insure use by the selected recipients within the limits of the constitutional purpose. Expenditures would fail of their purpose and thus lose their constitutional sanction if the terms of payment were not such that by their influence on the action of the recipients the permitted end would be attained. The power of Congress to spend is inseparable from persuasion to action over which Congress has no legislative control. Congress may not command that the science of agriculture be taught in state universities. But if it would aid the teaching of that science by grants to state institutions, it is appropriate, if not necessary, that the grant be on the condition, incorporated in the Morrill Act, 12 Stat. 503, 26 Stat. 417, that it be used for the intended purpose. Similarly it would seem to be compliance with the Constitution, not violation of it, for the government to take and the university to give a contract that the grant would be so used. It makes no difference that there is a promise to do an act which the condition is calculated to induce. Condition and promise are alike valid since both are in furtherance of the national purpose for which the money is appropriated.
The limitation now sanctioned must lead to absurd consequences. The government may give seeds to farmers, but may not condition the gift upon their being planted in places where they are most needed or even planted at all. The government may give money to the unemployed, but may not ask that those who get it shall give labor in return, or even use it to support their families. It may give money to sufferers from earthquake, fire, tornado, pestilence or flood, but may not impose conditions—health precautions designed to prevent the spread of disease, or induce the movement of population to safer or more sanitary areas. All that, because it is purchased regulation infringing state powers, must be left for the states, who are unable or unwilling to supply the necessary relief. The government may spend its money for vocational rehabilitation, 48 Stat. 389, but it may not, with the consent of all concerned, supervise the process which it undertakes to aid. It may spend its money for the suppression of the boll weevil, but may not compensate the farmers for suspending the growth of cotton in the infected areas. It may aid state reforestation and forest fire prevention agencies, 43 Stat. 653, but may not be permitted to supervise their conduct. It may support rural schools, 39 Stat. 929, 45 Stat. 1151, 48 Stat. 792, but may not condition its grant by the requirement that certain standards be maintained. It may appropriate moneys to be expended by the Reconstruction Finance Corporation “to aid in financing agriculture, commerce, and industry,” and to facilitate “the exportation of agricultural and other products.” Do all its activities collapse because, in order to effect the permissible purpose, in myriad ways the money is paid out upon terms and conditions which influence action of the recipients with the states, which Congress cannot command? The answer would seem plain. If the expenditure is for a national—public purpose, that purpose will not be thwarted because payment is an on condition which will advance that purpose. The action which Congress induces by payments of money to promote the general welfare, but which it does not command or coerce, is but an incident to a specifically granted power, but a permissible means to a legitimate end. If appropriation in aid of a program of curtailment of agricultural production is constitutional, and it is not denied that it is, payment to farmers on condition that they reduce their crop acreage is constitutional. It is not any the less so because the farmer at his own option promises to fulfill the condition.

STEWARD MACHINE CO. V. DAVIS
301 U.S. 548 (1937)

[A tax was imposed on employers, based on their employees’ wages. The proceeds went into general fund revenues and were not earmarked for any purpose; however, a credit of up to 90 percent of the federal tax was allowed to the extent the employer contributed to a state unemployment fund that met detailed statutory requirements and was also approved by the Social Security Board.]

The excise is not void as involving the coercion of the States in contravention of the Tenth Amendment or of restrictions implicit in our federal form of government.

* * *

[Petitioner argues] that the tax and the credit in combination are weapons of coercion, destroying or impairing the autonomy of the states. * * * Who then is coerced through the operation of this statute? Not the taxpayer. He pays in fulfillment of the mandate of the local legislature. Not the state. * * * For all that appears she is satisfied with her choice, and would be sorely disappointed if it were to be annulled. * * * [E]very rebate from a tax when conditioned upon conduct is in some measure a temptation. But to hold that * * * temptation is equivalent to coercion is to plunge the law in endless difficulties. The outcome of such a doctrine is the acceptance of a philosophical determinism by which choice becomes impossible. Till now the law has been guided by a robust common sense which assumes the freedom of the will as a working hypothesis in the solution of its problems. The wisdom of the hypothesis has illustration in this case. Nothing in the case suggests the exertion of a power akin to undue influence, if we assume that such a concept can ever be applied with fitness to the relations between state and nation. * * * We cannot say that [Alabama] was acting, not of her unfettered will, but under the strain of a persuasion equivalent to undue influence, when she chose to have relief administered under laws of her
own making, by agents of her own selection, instead of under federal laws, administered by federal
officers, with all the ensuing evils, at least to many minds, of federal patronage and power. * * *
In ruling as we do, we leave many questions open. We do not say that a tax is valid, when imposed by act
of Congress, if it is laid upon the condition that a state may escape its operation through the adoption of a
statute unrelated in subject matter to activities fairly within the scope of national policy and power. No
such question is before us. In the tender of this credit Congress does not intrude upon fields foreign to its
function. * * *
United States v. Butler is cited by petitioner as a decision to the contrary. * * * The decision was by a
divided court, a minority taking the view that the objections were untenable. None of them is applicable
to the situation here developed.
(a) The proceeds of the tax in controversy are not earmarked for a special group.
(b) The unemployment compensation law which is a condition of the credit has had the approval of the
state and could not be a law without it.
(c) The condition is not linked to an irrevocable agreement, for the state at its pleasure may repeal its
unemployment law, terminate the credit, and place itself where it was before the credit was accepted.
(d) The condition is not directed to the attainment of an unlawful end, but to an end, the relief of
unemployment, for which nation and state may lawfully cooperate.
The statute does not call for a surrender by the state of powers essential to their quasi-sovereign
existence.
* * * A wide range of judgment is given to the several states as to the particular type of statute to be
spread upon their books. * * * What they may not do, if they would earn the credit, is to depart from
those standards which in the judgment of Congress must have the benefit of a fair margin of discretion.
One cannot say with reason that this margin has been exceeded, or that the basic standards have been
determined in any arbitrary fashion. * * *
We are to keep in mind steadily that the conditions to be approved by the Board as the basis for a credit
are not provisions of a contract, but terms of a statute, which may be altered or repealed. * * * The state
does not bind itself to keep the law in force. It does not even bind itself that the moneys paid into the
federal fund will be kept there indefinitely or for any stated time. On the contrary, the Secretary of the
Treasury will honor a requisition for the whole or any part of the deposit in the fund whenever one is
made by the appropriate officials. * * *

**CARROTS vs. STICKS**

The underlying theme of the above two cases continues to arise in varying contexts many decades
later. For instance, the salient issue in Brown v. Environmental Protection Agency, 521 F.2d 827 (9th
Cir. 1975) was whether the Administrator of EPA may, as part of a State Implementation Plan issued
pursuant to the Clean Air Amendments of 1970, direct the State of California to undertake certain
air-improvement programs under penalty of fines or other civil sanctions. The D.C. Circuit held that the

---

* The following are examples of the numerous cases in which courts have held that a state’s action was voluntary
because it could have rejected a national grant that imposed obligations on the state: Florida Department of Health
and Rehabilitative Service v. Califano, 449 F.Supp. 274, 284 (N.D. Fla. 1978), affirmed per curiam on opinion of
court below, 585 F.2d 150 (5th Cir.), certiorari denied 441 U.S. 931 (1979); Walker Field, Colorado, Public Airport
Authority v. Adams, 606 F.2d 290, 297-98 (10th Cir. 1979); Oklahoma v. Harris, 480 F.Supp. 581, 588 (D.D.C.
1979).
State might decline to undertake the EPA-mandated programs without becoming liable to sanctions. However, it noted that

Our interpretation is in no way inconsistent with the recognition that Congress has the power to authorize the Administrator to obtain the consent of a reluctant state by conditioning certain federal expenditures within that state on the granting of such consent.

In this connection, the Court cited *Steward Machine Co. v. Davis* (1937) in favor of the “conditioning” position. But the *Steward* decision might well be compared with the seemingly contrary view in *United States v. Butler* (1936). Which of the lines of argument is most valid?

In cases such as *National League of Cities v. Usery*, 426 U.S. 833 (1976), the Supreme Court continues to show a concern for the Federal system by limiting the power of the national government over state-local governments. On the other hand, the growing importance of conditional grants may be viewed as a de facto erosion of the states’ independent spheres of authority. This position is expounded trenchantly in a law review note:

When tested against a standard of uninfluenced decisionmaking, state acceptance of grants whose conditions invade state powers fails. In practice, state legislatures often go beyond allowing national grants to influence their decisions on such policy matters. The states, in effect, cede the responsibility for making some decisions to the national government by accepting conditions that are subject to subsequent determination and revision by national agencies. * * * Such an abdication is a dangerous threat to Federalism and should be found unconstitutional.

Note, “Taking Federalism Seriously: Limiting State Acceptance of National Grants,” 90 Yale L.J. 1694, 1713 (1981) [footnotes omitted]. But would not an outright ban on all such “deals” between governmental units risk losing mutually advantageous “gains from trade”? Is a principled distinction possible?

Finally, consider the hypothetical question involving interpretation of the United World League Charter in the year 2053. The charter contains a provision which forbids any member state from “making any agreements or taking any actions which infringe the sovereign powers of any other nation.” Middleclasia proposes a treaty which offers to pay 50% of the annual GNP of Lesser Developia to the latter’s Treasury in return for the right to enact and enforce “public health laws,” specifically including compulsory sterilization programs and environmental regulations, in Lesser Developia. Since both nations are UWL member states, the Global Court is asked to construe the charter on this point. The justices of that august body issue the following opinion, 597 G.C. 214, at 223 (2053):

* * * [Through this treaty] Middleclasia confronts Lesser Developia with a situation which, while it retains the trappings of a consensual arrangement, is nonetheless recognizable as one devoid of any meaningful choice. When the illusory veil of voluntariness is pierced, the impact on Lesser Developia’s sovereignty can be adjudged just as predictable and irresistible as a threat of armed attack. In the last century, behavioral scientists were already cognizant of the essential symmetry between bribes and threats. We therefore hold that the proposed treaty is inconsistent with Article 24 of the UWL charter.

By way of brief historical digression only, it may be worth noting that the reaction of Middleclasia was then to switch to a strategy of nuclear threat, with the comment “Well, it sure is a whole lot cheaper—and apparently not any more contrary to the UWL charter—than our illegal and invalidated treaty.”

**QUESTIONS**

1. What is the theoretical argument that the impact of bribes—and threats is symmetrical? Of what relevance are such economic concepts as opportunity cost, budget line slopes, etc.? Sketch out, if possible, the opportunity sets of the taxpayers or other relevant economic units before and after the programs being considered by the courts in the above cases.
2. Are there any important distinctions among the Brown, Butler, and Middleclasia cases? Or do they each embody essentially the same situation?

3. In your opinion, what is the “right” ruling on the three cases? Can you suggest any useful compromises between strict constitutional logic as to “noninterference” and the demands of practicality? Or does such a conflict not even really exist?

4. Briefly, suggest one or two other real world contexts where the issues embodied in the above section on “bribes” might be raised in the same or closely related forms.

H. MEASURING RESPONSIVENESS: ELASTICITIES

In many legal applications, it is of great interest to know the sensitivity of response on the part of decisionmakers to changes in key economic parameters. Often, these key parameters are changes in the costs (prices) of goods or activities or in the expendable resources (income) available to the decisionmaker. Why this may be important is illustrated by the Mathews and Tilton cases above.

In Mathews, the economic impact of the block grant is equivalent to an income increase, a parallel outward shift of the recipient’s budget constraint. The grantor is—if the language of the statute is taken at face value—interested in stimulating “Priority Expenditures” and not those other things upon which the grant recipient may prefer to spend the money. Although this intent of the grantor is manifested in the form of a nominal legal constraint, our analysis of the Mathews case should have suggested that proof of a violation of the legal constraint will, barring highly exceptional circumstances, be difficult to establish. In sum, the legal constraint is of little or no practical consequence as an influence on the results attributable to the grant. Hence, a realistic appraisal of the grant’s impact would focus on the grant being, in effect, a mere income transfer to the recipient government. Then, we might ask with what sensitivity the quantities of the Priority Expenditure goods will vary as the municipal budgetary authorities perceive their unit as being “richer” after the grant is received. The exact response of any municipality depends partially on its own preferences, i.e., the configuration of its particular indifference curve map for public goods. But Congress would presumably still be interested in what is expected to be the typical or average responsiveness to the revenue sharing grants.

One straightforward way of measuring responsiveness would be with a simple rate of change concept:

\[
\text{Rate of Change} = \frac{\text{Change in Quantity Chosen}}{\text{Change in Income}}
\]

For some purposes, this simple concept might be suitable, but it generally objectional as being ambiguous, especially if the quantity responsiveness rates of two different goods are to be compared. In short, the problem is that the actual number yielded by the formula above depends on the units of measurement chosen for both the quantity and the income changes. For instance, the quantity of roads may be measured in feet or yards, linearly or areally, while income changes may be measured in dollars, cents, thousands of dollars, etc. Hence, as a sensitivity index the simple rate of change concept is in a fundamental sense arbitrary and of limited use for comparative purposes.

The economist’s response is to use an “elasticity” concept rather than rates of change. For an income change, the relevant elasticity formula is the “income elasticity”:

\[
\text{Elasticity} = \frac{\text{Change in Quantity Chosen}}{\text{Change in Income}} \times \frac{\text{X}}{\text{Quantity}}
\]
or, by rearranging the terms,

\[
\text{Income Elasticity} = \frac{\text{Change in Quantity Chosen}}{\text{Quantity}} \times \frac{\text{Change in Income}}{\text{Income}}
\]

where income and quantity are the starting income and quantity. Because the “rate of change” term is multiplied by a new term, income divided by quantity, any change in the units that affects the numerator of one of the fractional terms will similarly affect the denominator, thus negating any effect on the size of the elasticity index. Hence, a group of economists might discuss a Mathews-type block grant in terms of the expected income elasticity that recipient governments would have for the quantities of priority and non-priority goods. The grantor presumably would like to see a highly positive income elasticity for the priority goods. To the extent that the non-priority goods had a highly positive income elasticity, an unintended “diversion” of the grant funds would occur.

In Tilton, the grant program is a so-called “matching” grant scheme wherein the effective price of one category of goods is lowered while the cost of other goods remains the same. Here, a “price elasticity” rather than an income elasticity is relevant. What is usually meant by a price elasticity is the index measuring the responsiveness of a good’s quantity to a change in its own price:

\[
\text{Price Elasticity} = \frac{\text{Change in Quantity Chosen}}{\text{Change in Own Price}} \times \frac{\text{Quantity}}{\text{Price}}
\]

An own-price elasticity is always expected to have negative algebraic sign because any change in price produces a change in quantity of opposite sign (i.e., price increases stimulate quantity reductions, and vice versa).

In addition to its own price, the quantity of a good chosen may also be influenced by the price of another good to which it is “related.” The relevant index of sensitivity is called a “cross elasticity”:

\[
\text{Cross Elasticity} = \frac{\text{Change in Quantity of A}}{\text{Price of B}} \times \frac{\text{Change in Price of B}}{\text{Quantity of A}}
\]
where A and B are two different goods. The absolute magnitude (i.e., size without regard to algebraic sign) of the cross elasticity between two goods tells to what extent they are “related.” If goods are related, the terms on which one good is offered will affect the quantity purchased of the other. The kind of relationship between the goods is indicated by the algebraic sign of their cross elasticity: positive sign indicates substitutes and negative sign is evidence of complements. Examine the first term of the cross-elasticity formula above to see why it makes sense to interpret the sign of the cross elasticity as evidence about substitutability or complementarily.

There are many legal applications of cross elasticities. For instance, courts examine the cross elasticities among sets of goods to determine which should be grouped as forming part of the same “market” for antitrust analysis. In the Tilton case, the concept of “entanglement” between religious and secular expenditures of colleges may be expressed in terms of cross elasticities. Examine the abortion case immediately following with a view to the possibility of expressing the issue in terms of cross elasticities.

**HARRIS v. MCRAE**

448 U.S. 297 (1980)

[The majority opinion, by Justice Stewart, held that a state that participates in medicaid is not obligated under Title XIX of the Social Security Act to continue to fund medically necessary abortions nor are the funding restrictions of the Hyde Amendment itself in violation of either the Fifth Amendment or the establishment clause of the First Amendment.]

Mr. Justice BRENNAN, with whom Mr. Justice MARSHALL and Mr. Justice BLACKMUN join, dissenting.

I agree entirely with my Brother Stevens that the State’s interest in protecting the potential life of the fetus cannot justify the exclusion of financially and medically needy women from the benefits to which they would otherwise be entitled solely because the treatment that a doctor has concluded is medically necessary involves an abortion. I write separately to express my continuing disagreement with the Court’s mischaracterization of the nature of the fundamental right recognized in *Roe v. Wade*, 410 U.S. 113 (1973), and its misconception of the manner in which that right is infringed by federal and state legislation withdrawing all funding for medically necessary abortions.

*Roe v. Wade* held that the constitutional right to personal privacy encompasses a woman’s decision whether or not to terminate her pregnancy. *Roe* and its progeny established that the pregnant woman has a right to be free from state interference with her choice to have an abortion—a right which, at least prior to the end of the first trimester, absolutely prohibits any governmental regulation of that highly personal decision. The proposition for which these cases stand thus is not that the State is under an affirmative obligation to ensure access to abortions for all who may desire them; it is that the State must refrain from wielding its enormous power and influence in a manner that might burden the pregnant woman’s freedom to choose whether to have an abortion. The Hyde Amendment’s denial of public funds for medically necessary abortions plainly intrudes upon this constitutionally protected decision, for both by design and in effect it serves to coerce indigent women to bear children that they otherwise would elect not to have.

After the first trimester, the State, in promoting its interest in the mother’s health may regulate the abortion procedure in ways that are reasonably related to that end. And even after the point of viability is reached, State regulation in furtherance of its interest in the potentiality of human life may not go so far as to proscribe abortions that are necessary to preserve the life or health of the mother. See *Roe v. Wade*, 410 U.S. 113 164-165 (1973).

My focus throughout this opinion is upon the coercive impact of the congressional decision to fund one outcome of pregnancy—childbirth—while not funding the other—abortion. Because I believe this alone renders the Hyde Amendment unconstitutional, I do not dwell upon the other disparities that the Amendment produces in the
When viewed in the context of the Medicaid program to which it is appended, it is obvious that the Hyde Amendment is nothing less than an attempt by Congress to circumvent the dictates of the Constitution and achieve indirectly what Roe v. Wade said it could not do directly. Though it may not be this Court’s mission “to decide whether the balance of competing interests reflected in the Hyde Amendment is wise social policy,” it most assuredly is our responsibility to vindicate the pregnant woman’s constitutional right to decide to bear children free from Governmental intrusion.

UNITED STATES v. E. I. DU PONT DE NEMOURS & CO.

In a civil action under Sec. 4 of the Sherman Act, the Government charged that appellee had monopolized interstate commerce in cellophane in violation of Sec. 2 of the Act. During the relevant period, appellee produced almost 75% of the cellophane sold in the United States; but cellophane constituted less than 20% of all flexible packaging materials sold in the United States. The trial court found that the relevant market for determining the extent of appellee’s market control was the market for flexible packaging materials and that competition from other materials in that market prevented appellee from possessing monopoly powers in its sales of cellophane. Accordingly, it dismissed the complaint.

The Government contends that, by so dominating cellophane production, du Pont monopolized a “part of the trade or commerce” in violation of Sec. 2. Respondent agrees that cellophane is a product which constitutes “a ‘part’ of commerce within the meaning of Section 2.” Du Pont brief, pp. 16, 79. But it contends that the prohibition of Sec. 2 against monopolization is not violated because it does not have the power to control the price of cellophane or to exclude competitors from the market in which cellophane is sold. The court below found that the “relevant market for determining the extent of du Pont’s market control is the market for flexible packaging materials,” and that competition from those other materials prevented du Pont from possessing monopoly powers in its sales of cellophane. Finding 37.

The Government asserts that cellophane and other wrapping materials are neither substantially fungible nor like priced. For these reasons, it argues that the market for other wrappings is distinct from the market for cellophane and that the competition afforded cellophane by other wrappings is not strong enough to be considered in determining whether du Pont has monopoly powers. Market delimitation is necessary under du Pont’s theory to determine whether an alleged treatment of rich and poor, pregnant and non-pregnant. I concur completely, however, in my Brother Stevens’ discussion of those disparities. Specifically, I agree that the congressional decision to fund all medically necessary procedures except for those that require an abortion is entirely irrational either as a means of allocating health-care resources or otherwise serving legitimate social welfare goals. And that irrationality in turn exposes the Amendment for what it really is—a deliberate effort to discourage the exercise of a constitutionally protected right.

It is important to put this congressional decision in human terms. Nonpregnant women may be reimbursed for all medically necessary treatments. Pregnant women with analogous ailments, however, will be reimbursed only if the treatment involved does not happen to include an abortion. Since the refusal to fund will in some significant number of cases force the patient to forego medical assistance, the result is to refuse treatment for some genuine maladies not because they need not be treated, cannot be treated, or are too expensive to treat and not because they relate to a deliberate decision to abort a pregnancy, but merely because treating them would as a practical matter require termination of that pregnancy. Even were one of the view that legislative hostility to abortions could justify a decision to fund obstetrics and child delivery services while refusing to fund nontherapeutic abortions, the present statutory scheme could not be saved. For here, that hostility has gone a good deal farther. Its consequence is to leave indigent sick women without treatment simply because of the medical fortuity that their illness cannot be treated unless their pregnancy is terminated. Antipathy to abortion, in short, has been permitted not only to ride roughshod over a woman’s constitutional right to terminate her pregnancy in the fashion she chooses, but also to distort our Nation’s health care programs. As a means of delivering health services, then, the Hyde Amendment is completely irrational. As a means of preventing abortions, it is concededly rational—brutally so. But this latter goal is constitutionally forbidden.
monopolist violates Sec. 2. The ultimate consideration in such a determination is whether the defendants control the price and competition in the market for such part of trade or commerce as they are charged with monopolizing. Every manufacturer is the sole producer of the particular commodity it makes but its control in the above sense of the relevant market depends upon the availability of alternative commodities for buyers: i.e., whether there is a cross-elasticity of demand between cellophane and the other wrappings. This interchangeability is largely gauged by the purchase of competing products for similar uses considering the price characteristics and adaptability of the competing commodities. The court below found that the flexible wrappings afforded such alternatives. This Court must determine whether the trial court erred in its estimate of the competition afforded cellophane by other materials.

* * *

If cellophane is the “market” that du Pont is found to dominate, it may be assumed it does not have monopoly power to control prices or exclude competition. It seems apparent that du Pont’s power to set the price of cellophane has been limited only by the competition afforded by other flexible packaging materials. Moreover, it may be practically impossible for anyone to commence manufacturing cellophane without full access to du Pont’s technique. However, du Pont has no power to prevent competition from other wrapping materials. The trial court consequently had to determine whether competition from the other wrappings prevented du Pont from possessing monopoly power in violation of Sec. 2. Price and competition are so intimately entwined that any discussion of theory must treat them as one. It is inconceivable that price could be controlled without power over competition or vice versa. This approach to the determination of monopoly power is strengthened by this Court’s conclusion in prior cases that, when an alleged monopolist has power over price and competition, an intention to monopolize in a proper case may be assumed. 19

If a large number of buyers and sellers deal freely in a standardized product, such as salt or wheat, we have complete or pure competition. Patents, on the other hand, furnish the most familiar type of classic monopoly. As the producers of a standardized product bring about significant differentiations of quality, design, or packaging in the product that permit differences of use, competition becomes to a greater or less degree incomplete and the producer’s power over price and competition greater over his article and its use, according to the differentiation he is able to create and maintain. A retail seller may have in one sense a monopoly on certain trade because of location, as an isolated country store or filling station, or because no one else makes a product of just the quality or attractiveness of his product, as for example in cigarettes. Thus one can theorize that we have monopolistic competition in every nonstandardized commodity with each manufacturer having power over the price and production of his own product. However, this power that, let us say, automobile or soft-drink manufacturers have over their trademarked products is not the power that makes an illegal monopoly. Illegal power must be appraised in terms of the competitive market for the product.

* * *

IV. The Relevant Market.—When a product is controlled by one interest, without substitutes available in the market, there is monopoly power. Because most products have possible substitutes, we cannot, as we said in Times-Picayune Co. v. United States, 345 U.S. 594, 612 give “that infinite range” to the definition of substitutes. Nor is it a proper interpretation of the Sherman Act to require that products be fungible to be considered in the relevant market.

---

The Government argues:

“We do not here urge that in no circumstances may competition of substitutes negative possession of monopolistic power over trade in a product. The decisions make it clear at the least that the courts will not consider substitutes other than those which are substantially fungible with the monopolized product and sell at substantially the same price.”

But where there are market alternatives that buyers may readily use for their purposes, illegal monopoly does not exist merely because the product said to be monopolistic differs from others. If it were not so, only physically identical products would be a part of the market. To accept the Government’s argument, we would have to conclude that the manufacturers of plain as well as moistureproof cellophane were monopolists, and so with films such a Pliofilm, foil, glassine, polyethylene, and Saran, for each of these wrapping materials is distinguishable. These were all exhibits in the case. New wrappings appear, generally similar to cellophane: is each a monopoly? What is called for is an appraisal of the “cross-elasticity” of demand in the trade. See Note, 54 Col.L.Rev. 580.

* * *

An element for consideration as to cross-elasticity of demand between products is the responsiveness of the sales of one product to price changes of the other. If a slight decrease in the price of cellophane causes a considerable number of customers of other flexible wrappings to switch to cellophane, it would be an indication that a high cross-elasticity of demand exists between them; that the products compete in the same market. The court below held that the “[g]reat sensitivity of customers in the flexible packaging markets to price or quality changes” prevented du Pont from possessing monopoly control over price. 118 F.Supp., at 207. The record sustains these findings. See references made by the trial court in Findings 123-149.

We conclude that cellophane’s interchangeability with the other materials mentioned suffices to make it a part of the flexible packaging material market.

* * *

QUESTIONS

1. In this case, “flexible packaging material” would include products that have necessarily similar physical characteristics. Suppose that the issue involved the definition of a market for the provision of “off-hours theft protection” to business establishments. Some alternative approaches would involve watchmen, guard dogs, local alarms, alarms connected to a central response station, special locks, etc. Would all of these be part of the same market? See United States v. Grinnell Corp., 384 U.S. 563 (1966).

2. The Court’s use of the cross elasticity measure in this case is generally regarded by economists as a case of erroneous inference Can you tell why? [Hint: Will the cross elasticity among products be the same at all prices?]

3. LEGAL APPLICATIONS FROM DEMAND THEORY

As developed above, elasticities are the economist’s measures for the responsiveness of choices to changes in economic variables such as income, prices, etc. The magnitudes of elasticities may be important for various types of legal analysis, as exemplified in the questions that appear below.

QUESTIONS

1. “One view of constitutional protections is this: the State must exercise great caution when it affects the choices that have high cross-elasticities with protected goods or activities.” Explain what you think is meant by the quoted statement. Do you agree?
2. A recent New York State statute provided for heavy mandatory penalties on drug pushers. By raising the cost of purchasing drugs, the statute was intended indirectly to curb the type of street crime and burglary that is commonly engaged in by addicts seeking to support a drug habit. How does the rationale of this law depend on an implicit premise about the [own price] elasticity of demand for illegal drugs? Do you regard this premise as plausible?

3. Your client, Zenavox Corp., manufactures consumer electronics equipment including television receivers. Suit has been filed against Zenavox by a group of its own retailers seeking treble damages under the Sherman Antitrust Act. The conduct alleged is, inter alia, that the corporation conspired with a group of “favored” retailers to monopolize sales of Zenavox TV sets in the southern part of the State of North Dakota. The time period is 1969. Your economic consultant advises that a study be conducted to determine the cross elasticity of demand between Zenavox TVs and other brands. Normally, what results would you be looking for? Why? (This part is easy.)

4. Plaintiffs also allege that Zenavox’s semi-annual factory sales of TV sets to its own employees make it a retailer. If the court so holds, Zenavox’s Fair Trade agreements with its franchisees would not be protected by the Miller-Tydings amendment to Section 1 of the Sherman Act. In this time period (1969), the Miller-Tydings provision protected “vertical” price-fixing agreements among entities competing at the same market level but not “horizontal” agreements as, for example, among several retailers. Unless so protected, price-fixing agreements would be per se illegal, i.e., not subject to any defense such as lack of market power. Under these circumstances, does it still seem prudent to generate evidence that there is a high cross elasticity of demand between Zenavox products and those of competing brands? [Hint: Under this type of resale price maintenance agreement, dealers were required to charge no less than certain minimum prices. What kind of damage does a retailer suffer by not being able to lower his price?]
The notion of legal rules as behavioral modifiers is hardly a new one. In this chapter, we take a focused look at some important liability rules and their behavioral implications. The discussion leads off with familiar notions from tort theory.

A. TORT LIABILITY MODELS

A finding of negligence on the part of a person who accidentally injures another has for almost two centuries been the most important standard that triggers civil liability in tort. But exactly what is meant by the term “negligence”? Perhaps the most famous attempt to give exact content to the concept was that of Judge Learned Hand in *United States v. Carroll Towing Co.*, 159 F.2d 169 (2d Cir. 1947). Hand stated that the factfinder must consider three factors: the magnitude of the loss if the accident should occur, the probability of occurrence, and the cost of taking precautions that would avert it. The product of the first two factors—what an economist might term the expected value of the harm—should be balanced against the cost of precautions, with failure to engage in cost-effective precautions constituting negligence. This formulation is widely known as the Learned Hand Rule.

Obviously, the Learned Hand Rule has an economic flavor to it, but economists tend to criticize the original formulation as ambiguous or at least imprecise. The complaint is that the original formulation does not make it sufficiently clear that the rule must, in order to convey the sense that Hand almost certainly intended, be interpreted as applying to *marginal adjustments* of care and not to the total amount of care exercised. In most situations, many different levels of care are possible, each of which yield different answers for the three factors that Hand emphasized. Except perhaps for classificatory purposes, it is unimportant for the Hand approach whether precautions affect the probability of harm, the magnitude of the harm if it occurs, or both. At base, what is really involved is reduction of the expected value of the harm. Remember also that “care” is a convenient summary or portmanteau term; in general, care consists of heterogeneous activities rather than adjustment of any single dimension of activity.

1. THE LEARNED HAND RULE: A MARGINAL COST-BENEFIT FORMULATION

What Hand presumably wanted potential tortfeasors to do is to act reasonably in the light of the cost of the precautions. The marginal benefits of each “dose” of additional care (these benefits being composed of reductions in the expected value of harm) should be compared to the marginal cost of taking the precautions. Exhibit 5.1 depicts such a marginal analysis. The vertical axis is used to measure costs and benefits in dollars, while the variable T on the horizontal axis represents levels of care taken by the tortfeasor. For simplicity, the marginal cost of precautions is represented as constant, but increasing or decreasing marginal cost would not materially affect the analysis. Also for simplicity, the “injurer” in the remainder of these materials will generally be referred to loosely and somewhat inaccurately as the “tortfeasor” even when his conduct conforms to the legal standard so that technically the injury is not a “tort.”
1. In Exhibit 5.1, draw a line (bright yellow if you have it!) perpendicular to the horizontal axis such that levels of care to the left would be negligent under the Learned Hand Rule. On the horizontal axis, label the level of care that just satisfies the rule as $T^*$.

2. Absent any legal rule, what level of $T$ would you expect the tortfeasor to pick? Why? If a tortfeasor must compensate the victim only when negligent, what level of $T$ would be predictable? Why? (Assume, for the moment, that courts always apply the standard with accuracy and precision. This assumption will be relaxed later.)

3. It is sometimes said that legal liabilities are used to “internalize” external costs. What is meant by that assertion? And of what relevance is it to the negligence standard?

4. Assume that all potential tortfeasors adopt care level $T^*$. Some positive level of expected costs will be imposed by the class of potential tortfeasors on the class of potential victims. How can you tell that from Exhibit 5.1? Does Exhibit 5.1 permit you to say what the magnitude of those costs would be?

5. Suppose that liability could be escaped only if “all possible effective precautions” were exercised. Label that level of care $T'$ in Exhibit 5.1. Assume that $T'$ were adopted by all potential tortfeasors. In what way would that assumption affect what is said in question 4 immediately preceding?

6. You were just asked to assume a precaution level of $T'$. Under the liability rule posited, is precaution level $T''$ the behavior that you would predict?

7. What do you understand to be meant by the standard of “strict liability.” Explain, in terms of Exhibit 5.1, the level of $T$ that would be likely under that standard. How is question 4 above affected if recast to apply to the strict liability rule?

8. From a normative perspective, do you find anything particularly attractive about the negligence standard? If not, can you articulate what might be the position of its adherents?
2. THE HYPOTHETICAL MODEL UNDERLYING THE DIAGRAMS

Careful completion of the question set immediately preceding should have led you to what is to many persons a surprising and counterintuitive conclusion: under a perfect adjudication system, the level of care taken by a potential tortfeasor is the same under a strict liability system as under a negligence rule. The reason why this is so, and why the conclusion changes when there are “errors” in the adjudication system, can be further explored with some additional simple models.

In order to make the illustrations more concrete, the diagrammatic models provided as Exhibits 5.1 through 5.4 in the beginning of this chapter are all based upon a specific numerical example. This example assumes that the cost of a unit of care, \( T \), is $100. A specific assumption is also made about the efficacy of care; namely, that the expected value of the harm, \( H \), from a dangerous activity varies according to the formula:

\[
H = 22000 - [200T - .5T^2] \quad \text{if} \quad T < 200
\]

where \( T \) is the level of precautionary care taken by the tortfeasor and the square-bracketed term indicates the reduction in harm that has taken place as a result of the care exercised. Hence, if no care is taken, the level of harm will be $22,000. Thereafter, the value of \( H \) declines as \( T \) increases within the range from zero to 200. At the 200 level, additional care will have no further effect, as the “side condition” to the right of the formula indicates by constraining \( T \) to values of less than 200. By inference from the formula for the total level of \( H \), the marginal effect of \( T \) in reducing harm can be expressed by the formula 200 - \( T \); that is, the first dose of \( T \) reduces \( H \) at a rate of almost 200, but the efficacy of care declines and reaches zero when 200 units of care have been applied. (Readers who have had a little calculus will recognize 200 - \( T \) as the first derivative of the formula given above for \( H \), i.e., the value of \( dH/dT \).) Note that the marginal benefits curve in Exhibit 5.1 above follows the factual assumptions just laid out: it incorporates (1) a marginal benefits curve for care having an intercept of 200 and a negative slope of one and (2) a marginal cost curve of care which is constant at the level 100.

Although economic models usually focus almost exclusively on marginal effects, the present topic provides an instance when it may be helpful to one’s understanding to see the problem in total terms. Exhibit 5.2 plots out the total magnitudes underlying the problem presented in marginal terms in Exhibit 5.1. The straight line is the total cost of care taken. If we symbolize the total cost of care as \( C \), then Exhibit 5.2 shows the relationship

\[
C = 100T
\]

as a straight line. The “harm reduction” curve is the square-bracketed term in the formula for \( H \) above. It is, in essence, the “total benefits” curve from the societal standpoint. Note that the greatest excess of benefits over costs occurs at the 100 level of \( T \), just as the marginal analysis in Exhibit 5.1 suggested. Indeed, the slope of any total curve is diagrammatically equivalent to the marginal effect.
Verify the optimality of the 100 level by constructing the tangent to the harm reduction curve of Exhibit 5.2 at the 100 level. Since the tangent indicates the slope of a curve at the point where it touches, the slope should be the same as the straight line that represents total precautionary costs. Do you see the common-sense of the result that one should stop incurring costs at the point where benefits begin to rise more slowly (i.e., have a lesser slope) than the associated rate of increase in costs? Finally, the heavy top curve shows the net result at different levels of care, i.e., the value $HC$. Note that the minimum point of this curve lies, as expected, at the 100 level of $T$, but the absolute height of the curve is quite high; $H$ has not been reduced to zero, even in the range that clearly meets the negligence standard. Take a colored pencil and trace out a potential tortfeasor’s costs for different levels of care under a perfectly administered negligence standard. It follows the top curve up to the point where $T = 100$, then “jumps” down to the straight C curve. Do you see why? By contrast, what curve or curves are relevant under strict liability? Why is the expected conduct the same?

3. ERRORS IN THE APPLICATION OF TORT LAW

Thus far, the results from Exhibit 5.2 are just a different way of expressing the analysis of Exhibit 5.1. Let us now, however, begin to introduce “errors” into the adjudication system.

a. Incorrect Assessment Of Damages

Let us start by assuming that courts will correctly set the negligence standard at 100 units of care, but will underassess damages, setting them at only 75% of $H$. Under the alternative rule of strict liability, the tortfeasor will always have to pay damages, but the same degree of underassessment will occur. This form
of error will, of course, undercompensate victims and may also be expected to affect behavior with respect to the level of care taken.

In effect, the form of error just hypothesized converts the original formula for $H$ into:

$$H = 0.75 \{22000 - [200T - 0.5T^2]\} \quad T < 200$$

where only 0.75 of the true harm is “charged” to the tortfeasor. The question set below will explore the implications of this situation.

**QUESTIONS**

1. Which curves in Exhibits 5.1 and 5.2 are based on $H$? Is it true that these curves now lie exactly three-fourths as high as they did originally? Draw them in on your copies of the respective exhibits.

2. The dotted curve in Exhibit 5.2 is the new “net result” or $H + C$ curve. About where does its minimum point lie? Is this the same level of $T$ as suggested by the marginal analysis in Figure 5.1?

3. Imagine that you were a potential tortfeasor. If the standard were one of negligence, it would be rational for you still to invest in 100 units of care. Why? What is your cost-minimizing level of care under strict liability? What explains the difference?

4. Suppose that the error took the form of an overassessment of damages by 25%, rather than an underassessment. How are the results affected?

5. If $T = 100$ is the desired behavior under the Learned Hand Rule, then the preceding analysis suggests that the negligence standard is “robust” in that damage miscalculations—even of high relative magnitude—may not distort the desired behavior. What would the degree of misassessment have to be to affect the results under the conditions hypothesized in Exhibits 5.1 and 5.2? How could the facts of the hypothetical have been chosen so as to weaken the finding of no behavioral effect?

6. Would most people regard the results developed above as counterintuitive or contrary to expectation? If so, are they the result of any “sleight of hand” or are they a fair representation of reality? (You may wish to defer a final judgment on this question until after the completion of this chapter.)

7. Robert Cooter developed models almost identical to those in this section in his interesting analysis of the rationale for punitive damages. See Cooter, “Economic Analysis of Punitive Damages,” 56 S.Cal.L.Rev. 79 (1982). Cooter suggests that most injurers are at fault only “unintentionally” in the sense of what economists would regard as mistaken or “disequilibrium” behavior. Gross fault, he argues, is likely to be intentional and attributable to an idiosyncratic cost or returns function possessed by the individual. Punitive damages would then be necessary to secure the individual’s compliance with a normal standard. Would a persistent pattern of modest fault fit his hypothesis equally well?

**4. MISTAKES IN APPLYING THE LIABILITY STANDARD**

In contrast to the last section, let us now assume that damages are assessed correctly (equal to $H$) when liability is found, but that judges and juries sometimes set the standard of care either above or below the 100 level called for by a precise Learned Hand Rule application. Figure 5.3 shows a familiar bell-shaped Gaussian or “Normal” curve centered on the “correct” level of 100, but dispersed for a considerable range around this midpoint. (For those really interested, the “standard deviation” of the probability distribution assumed in Exhibits 5.3 and 5.4 is 15.) The height of the bell-shaped curve indicates the relative probability that a court will find that a particular level of $T$ fails to satisfy the minimum requirements of
the negligence standard. Hence there is almost no chance that a court would accept a level as low as 55 or demand one as high as 145, but the intervening values are all possible. The area around 100 is, as the diagram suggests, the most likely standard to be adopted by a judge or jury.

The heavier curve of cumulative probabilities is derived from the bell-shaped one. Note that the area under the bell-shaped curve must account for 100% of the probabilities. The heavy downward sloping curve is the fraction of the area under the bell-shaped curve that, at any level of T, lies to the right. Therefore, it can be interpreted as telling us the probability that a court would set its liability standard higher than any given level of T, thus finding the tortfeasor who chooses that level of care to be negligent.

Exhibit 5.3: Probability of Jury Requiring Care Level

For expository simplicity, assume that a potential tortfeasor is risk-neutral, so that he wishes to minimize the expected value of his costs. Under the circumstances posited, the sum of the costs under the negligence system may be expressed as:

\[ L = pH + C \]

where \( L \) signifies the costs in question, \( p \) is the probability of being held liable, \( H \) is the damages (assumed now to be accurately levied as compensatory damages), and \( C \) is the cost of the care taken. In the specific numeral example being used in these sections

\[ L = p \{22000 - [200T + .5T^2] \} + 100T. \]

The care level \( T \) appears explicitly in part of this formula, but it also is involved implicitly as a determinant of the value taken on by \( p \). (See Exhibit 5.3.)
An orthodox marginal adjustment model can be constructed by examining the effect of this situation on marginal benefits and costs from increasing care. In general terms, the marginal effect of T on the expected value of the damages to be paid will be:

\[ p(\text{marginal harm reduction}) \] + \[(\text{marginal reduction of } p)H\]

Readers who have had calculus will recognize the words above as the first derivative of pH with respect to T, i.e., d(pH)/dT = p dH/dT + dp/dT H. The “common sense” of the formulation may, in any event, be perceived by giving a practical interpretation to its components. The first square-bracketed term above is what we shall call the “harm reduction effect,” since it measures the amount by which one’s costs will be reduced due to the necessity to pay smaller damages upon any judgment of liability. By contrast, the second square-bracketed term is the “liability risk reduction effect,” since it measures the impact on cost of a diminished probability that one will indeed be held liable at all.

Whenever p is less than one, the harm reduction effect is, of course, attenuated because the ex ante prospect is that one will be held somewhat less than fully responsible for any harm done. Thus, in Exhibit 5.4, the heavy straight line shows the true marginal harm reduction effect (when p= 1.00) while the curve below it indicates how the importance of harm reduction is attenuated when the probability of liability begins to fall below 1.00, starting at about 55 on the care axis. The liability risk reduction effect from the second part of the formula works in the opposite direction. It has the effect of making care adjustments have a greater impact to the extent that they produce a change in p. Under plausible conditions, this latter effect greatly overbalances the attenuation factor, producing a net change (sum of both effects) that distorts the true marginal benefits curve upward in the manner indicated by the humped curve in Exhibit 5.4. Under the assumptions of our specific numerical example, the formula for that net curve is:

\[ p (200 - T) + \Delta p (22000 - 200T + .5T^2) \]

where \( \Delta p \) is the reduction in p that occurs as a result of a marginal increase in T. Values for p and \( \Delta p \) can be inferred approximately from Exhibit 5.3 or looked up more precisely in a reference book containing tables of values for the Gaussian Normal probability distribution. One can, however, merely inspect the formula and see that even relatively small values for \( \Delta p \), the change occurring in p, will cause the second part of the formula to have a value sufficiently high to overbalance the effect of a small value of p in otherwise diminishing the marginal benefit curve from its “true” value of (200 - T).
Exhibit 5.4: Behavior Effects of Errors Regarding Care Standard

Add the appropriate marginal cost of care curve to Exhibit 5.4. Then find the intersection point that identifies the level of $T$ at which the marginal costs of care and the marginal effect on legal exposure are properly balanced so that the tortfeasor’s prospective costs are minimized in the face of this inaccurately applied negligence standard. How does this compare with what would happen under a strict liability rule?

QUESTIONS

1. Explain in as non-technical terms as possible why high—and arguably “excessive”—levels of care might be very cost-effective for tortfeasors in the context just laid out. Why do the symmetric errors of the courts (equal incidence of too-strict and too-lenient standards) not just “cancel out” in terms of behavioral incentives?

2. How could the facts of the example in the text be altered to change the results? Under what circumstances would strict liability produce higher levels of care than negligence?

3. What do you think the shapes of the actual probability distributions of standard-setting errors look like? Why do the shapes matter?

4. Suppose that the liability standard were perfectly applied in all cases. Draw a graph that shows the level of $p$ as a function of the level of $T$. If $T^*$ is the level of care that satisfies the negligence standard, what value does $\Delta p$ have when $T^*$ care is taken? This exercise may improve your understanding of the analysis in the previous section, where it was suggested that the levels of care under negligence standards may not be much affected by incorrect calibration of the damages. Try to explain that earlier result in terms of the factors introduced in this present section.

5. What if the level of care actually chosen either exceeds or falls short of that called for in the Learned Hand Rule? Who is better or worse off? Why should we care?
6. “The negligence rule, properly understood, is really a right of private eminent domain given to performers of dangerous acts. What’s more, it doesn’t even require them to pay full compensation.”

Comment.

B. MULTI-PARTY HARM PRODUCTION FUNCTIONS

Up to this point, all of the discussion in this chapter has been couched in terms of the impact that the tortfeasor’s conduct has on the probability and magnitude of an accident. But a very large proportion of potential torts involve situations in which the conduct of the victim can influence either the probability or magnitude of the harm. Indeed, this fact is recognized in the legal doctrine of contributory negligence, which imposes upon the victim a duty to take cost-effective precautions in a manner similar to that which we have just discussed for the tortfeasor. If the victim is found contributorily negligent, he cannot recover damages even if the injurer is found negligent.

1. PRECAUTIONS THAT MINIMIZE NET SOCIAL COST OF ACCIDENTS

When both parties have a duty to take care, and both parties have an impact on the result, the analysis becomes a bit more complicated. Frequently, that the effect of the victim’s and the tortfeasor’s care is interdependent, the efficacy of each party’s acts depending to some extent on what the other party does. Sometimes care taken by one party is merely a substitute for the other party’s care, so that the more care the victim takes the less efficacious is the marginal care by the tortfeasor, and vice versa. Alternatively, the levels of care may be complementary or synergistic; more care by one makes the conduct of the other more effective. When any form of interdependency exists, one cannot draw the kind of marginal benefits curve used in Exhibit 5.1 without first knowing the level of care that will be taken by the other party; the other party’s care will shift the marginal benefits curve up and down.

From the discussion of the Learned Hand Rule above, it may fairly be hypothesized that the combination of the negligence and contributory negligence doctrines is designed to create incentives for parties to minimize harms in the most cost-effective way possible. We can adapt the “mapping” tools employed in the opening section of these materials to a consideration of this interaction between the “inputs” of care by the two parties and the resultant “output” of a level of social cost. The model yielded is similar to “production function” models in more traditional economics.

In what follows, we will let T stand for the level of care taken by the tortfeasor and V the level of care taken by the victim. In its more abstract and generalizable form, a two-party production function for harm could then be written as \( H = H(V, T) \) where the \( H(V, T) \) is understood as some mathematical function that has \( V \) and \( T \) as two of its variables. Similarly, cost could be written abstractly as \( C = C(V, T) \), where \( C(V, T) \) simply says that the total cost of care is determined in some way by the levels of \( V \) and \( T \). If one were interested in minimizing the net social cost attributable to the dangerous activity in question, then the goal could be formalized as one of minimizing \( S \) where

\[
S = H(V, T) + C(V, T)
\]

Exhibit 5.5 is an example of what a “map” of the production function of \( S \) might look like. Think of it as a topographic map of a valley or “bowl” where the contour lines show equal levels of \( S \). Social cost is minimized by getting to the bottom of the bowl, point a in Exhibit 5.5. As one moves around on the map, by varying the levels of \( V \) and \( T \), \( H \) and \( C \) undergo changes in opposite directions.

---

Exhibit 5.5 also shows two lines that are clearly distinguishable from the contour lines. These are “lines of optima” that denote the S-minimizing reaction level of each party to any level of the other party’s care. Follow line tt’ to where it intersects the middle contour line. Now draw in a horizontal line at that point. When referred to the vertical axis, the horizontal line can be interpreted as a fixed or predetermined level of V. The tortfeasor can move back and forth along this line, because he controls the value of T, but he cannot move up or down in a vertical direction since V is controlled by the other party. Where, then, is the “optimal” response of the tortfeasor in terms of a level of T that achieves the lowest achievable level of S? It is at the tangency point with the contour line. Do you see why? You should satisfy yourself that every point on the tt’ line is the “optimal” S-minimizing response of the tortfeasor on the assumption that V has been fixed at the corresponding level of the vertical axis. Line vv’ is, of course, an analogous construct showing the optimal cost-minimizing response of the victim on alternative assumptions as to the care taken by the tortfeasor.
QUESTIONS

1. What relationship does a point on $tt'$ in Exhibit 5.5 have to the marginal analysis in Exhibit 5.1?

2. To what extent is it accurate to interpret $tt'$ and $vv'$ as, respectively, conceptualizations of the negligence and contributory negligence standards?

3. Absent any legally imposed incentives, does minimizing $S$ represent the best result for the tortfeasor? For the victim? If not, to what location on the “map” would you expect the parties to adjust if there were no legal system at all?

4. Do the relevant legal doctrines give the parties incentives that will lead to a solution at point $a$, the minimum level of $S$? Why?

5. In the absence of the contributory negligence doctrine, what solution would you predict?

6. What kinds of changes in the facts of the world would cause the “topography” represented by the map to shift?

7. What, if anything, does Exhibit 5.5 tell you about behavior under a strict liability rule rather than a negligence standard?

2. COMPARATIVE NEGLIGENCE

The doctrine of comparative negligence is now applied in many jurisdictions in lieu of a strict rule of no recovery when the victim is contributorily negligent. Perhaps the increasing popularity of comparative negligence reflects a sense that there exists a certain harshness and apparent injustice when, although both parties fail in their duties of care, the consequences are borne solely by the contributorily negligent victim.

Exactly what does comparative negligence mean? Loosely speaking, nobody seems to have any difficulty with the notion that comparative negligence involves allocating the damages between the negligent tortfeasor and the contributorily negligent victim in a manner proportional to their relative degrees of negligence. Surprisingly enough, however, the concept is somewhat ambiguous if one attempts to give it more precise operational content. This ambiguity can be illustrated effectively with a mapping model.

Exhibit 5.6 is another production function model similar to that introduced in the previous section as Exhibit 5.5. The contours of these two harm production function maps are dissimilar because they represent two different activities or products, each of which involves its own unique “technology” of harm reduction and diverse costs of care. Otherwise, the contour lines and lines of optima have the same significance as we have already discussed.
If both parties complied with the negligence standard, the tortfeasor would supply $T^*$ of care and the victim $V^*$, thus achieving the global minimum of social cost $S$ at point $a$, the "bottom of the bowl." At that optimal result, the level of social costs (the cost of the harm plus the cost of the care taken to minimize the harm) would be only 50. Assume, however, that tortfeasor takes only $T^*$ of care and victim $V$ of care, so that the actual result is at point $x$. Exhibit 5.6 tells us exactly what the shortfalls of care are by both parties and exactly what the consequences of those shortfalls are; we have none of the uncertainties or measurement problems of the real world; Consequently, it is a fair question to ask how the exact degrees of comparative negligence should be determined in light of the full information available.

The actual result $x$ is on a contour line labeled 90, thus indicating that there are 90 dollars (or hundred or thousands or millions of dollars, if you prefer) in social costs to be allocated between the two parties. One plausible approach would be to look to the shortfalls in inputs of care as the basis for measuring relative degrees of negligence. The victim’s shortfall in care inputs is equal to $(V^* - V)\text{ and the tortfeasor’s to } (T^* - T)$. If those two shortfalls were somehow added together, we would have an index of total negligence and could express each party’s negligence as a proportion of the total. A threshold problem is that $V$ and $T$ are not necessarily the same kinds of care; i.e., $T$ may be a pedestrian’s care in looking carefully before crossing the street and $V$ the restraint with which an automobile is driven. One way (perhaps not the only or the best way) of making the inputs comparable for aggregation purposes is to measure them in cost terms. For expository simplicity, assume that the prices of each kind of care input are identical, so that we can avoid the necessity to weight the physical units by their costs. Then we can allocate the victim’s share of the social costs as

$$(V^* - V) / [(V^* - V_x) + (T^* - T_x)]$$
and the tortfeasor’s as
\[
\frac{(T^* - T_x)}{(V^* - V_x) + (T^* - T_x)}
\]

The following values for the relevant variables are consistent with Exhibit 5.6: \(T^* = 48\); \(T_x = 22\); \(V^* = 50\); and \(V_x = 30\). (You may wish to measure the axes to confirm that the relevant distances are indeed in the proportions indicated by the numbers just suggested.) By inserting the numbers in the formulas, the comparative negligence of the tortfeasor will be computed as 56.5% and that of the victim as 43.5%. Those percentages would be used to apportion the damages.

The method just described may be termed the “inputs approach,” since it focuses on the inputs of care. An alternative, however, is to focus on the impact of the care shortfalls. The latter is the “outputs approach” and proceeds as follows. The optimal result at point a would have involved costs of 50. If only the tortfeasor had been negligent, the “difference” from the optimal result would have been the increase in cost between point a and point d. Point d is on a contour line labeled 77, so this impact of negligence by the tortfeasor alone is (77 - 50) or 27. Analogously, contributory negligence alone would have moved the result from a to f where the contour line indicates a level of 70. Thus, the output differences caused by each party’s shortfall of care might be measured as 27 and 20 for the tortfeasor and victim respectively. On the basis of this outputs approach, the tortfeasor would find that, instead of 56.5%, his comparative negligence would be based on the proportion 27/(27+20) or 57.0% Only a small difference, to be sure.

Now, consider the fact that the dotted contour line with point d on it might be relabeled with any number between 70 and 80. This relabeling would correspond to a change in the “steepness” of the topography in that band of the production function map, a change whose economic significance is that the efficacy of care inputs in that range has been raised or lowered. Using a value of 71, the tortfeasor’s comparative negligence would be 21/(21+20) or 51.2%, while assuming a level of 79 for d’s contour would yield 29/(29+20) or 59.2%. The perceptive reader will soon see the trick: if the “bowl” has unequal degrees of steepness in different parts of the map, then the effect on output will not necessarily be proportional to the relative sizes of the input shortfalls.

QUESTIONS

1. The text demonstrates that different results can be achieved through the inputs approach and the outputs approach. Therefore, they cannot both be right. What are the arguments for and against each approach? Which is to be preferred? If you were a judge, how would you charge a jury on this question?

2. Just to muddy the waters further, one of the parties claims that the procedure of the outputs approach computation as described in the text above is exactly backwards. “You are asking the wrong question in starting from point a. You should start from point x and ask what difference it would have made if each party had not been negligent. For the tortfeasor, the relevant difference is between the result at x and the result at f, while for the victim it is the difference between x and d.” Which party would benefit from adopting this approach rather than the one in the text above? Which is more correct?

3. Note that the map was originally defined in terms of social cost S, which includes the cost of care taken by each party. In discussing the comparative negligence standard, should a map of only the harm H been used? If so, the contour lines of a map of H could have been superimposed on the map of S. Which, if any, of the difficulties discussed in the text above would have been affected?

4. Now that you have been put in the mood, can you think of any other defensible interpretations of how comparative negligence should be measured? See also, Schwartz, “Contributory and Comparative Negligence: A Reappraisal,” 87 Yale L.J. 697 (1978).

5. In Exhibit 5.6, the lines of optimal response indicate a “substitutional” relationship between the care inputs of the two parties. That is, the lower the level of V the higher T should be in order to minimize
social cost $S$, and vice versa. How do you account for the lines of optimal response in Exhibit 5.5? Is one of these exhibits constructed unreasonably? If not, give examples of real activities or products that you believe might be exemplified by the different forms of production function. What relationship, if any, is there to the type of marginal analysis done in Exhibit 5.1?

6. The text in this section suggests that comparative negligence is not a very precise legal concept, a thesis that may be buttressed by the cases and other material below. Whatever its meaning, however, comparative negligence means that each of the negligent parties will bear only partial responsibility for the damages. On these grounds, the doctrine has been criticized as not providing the proper incentives for the parties to take care. Elaborate on the form that you think the criticism might take. Is the criticism persuasive?

**BENTZLER v. BRAUN**
149 N.W.2d 626 (Supreme Ct. Wisc., 1967)

The question, therefore, is not whether the guest’s negligence contributed to the cause of the accident, but, rather, whether it contributed to the injuries. In view of the Wisconsin statutes that the legislative mandate in regard to seat belts applies merely to installation and not to use, the failure to use available seat belts is a question for determination by the jury as in the case of any ordinary negligence, i.e., was the conduct a substantial factor in producing a result.

We therefore conclude that, in those cases where seat belts are available and there is evidence before the jury indicating causal relationship between the injuries sustained and the failure to use seat belts, it is proper and necessary to instruct the jury in that regard. A jury in such case could conclude that an occupant of an automobile is negligent in failing to use seat belts. In the instant case, however, because of the lack of any evidence of causation, the trial judge properly refused the requested instruction. There was proof that seat belts were available and were not used, but that fact alone does not prove causation, for the driver of the vehicle also failed to use the available seat belts, but his injuries were minimal. In motions after verdict the trial judge stated that there was a complete failure of proof in that regard:

“There is no evidence, by expert witnesses or otherwise, that Janet Bentzler’s injuries would have been reduced or minimized had she been wearing a seat belt. The jury could only speculate as to whether her injuries would have been less severe.”

The only witness offered was an orthopedic surgeon, who, although qualified in his chosen profession, did not purport to be able to testify what effect the use of seat belts might have had in this particular case. The record supports the trial court's determination that there was no proof whatsoever to show that Janet Bentzler’s injuries were caused or aggravated by the failure to use the seatbelts. In the absence of credible evidence by one qualified to express the opinion of how the use or nonuse of seat belts would have affected the particular injuries, it is improper for the court to permit the jury to speculate on the effect that seat belts would have had.

While in this case there was evidence that could have sustained the court’s finding that Janet Bentzler was negligent in failing to wear the seat belts, that negligence could only have been related to her negligence in using ordinary care to protect herself from injuries and could not have been negligence that was a factor in producing the crash. In the absence of any proof of causation, the court properly refused the requested instruction on the question of Janet Bentzler’s negligence in the use of seat belts.

**VICTOR SCHWARTZ, COMPARATIVE NEGLIGENCE, p. 297**

The process is not allocation of physical causation, which could be scientifically apportioned, but rather of allocating fault; which cannot be scientifically measured. For example, suppose an intoxicated motorcyclist speeds at eighty miles an hour down a highway with a 55-mile speed limit. He loses control of his vehicle and crosses over into the opposite lane where he collides with a large truck traveling 65
miles per hour. The point of collision is the left fender of the truck. As a result of the impact, the motorcyclist is killed and his vehicle is a total loss. The truck, however, is only slightly damaged and the truck driver is not hurt at all.

In terms of pure physical causation, perhaps an expert could testify that the truck supplied 95% of the force that killed the motorcyclist, based on formulae combining the relative weights, speeds and directions of the vehicles. Even without expert assistance, the jurors might, by instinct, regard the truck as the more substantial cause. Nevertheless, the jury’s line of inquiry under comparative negligence does not focus on physical causation; rather, it considers and weighs culpability. In Nebraska and South Dakota, the question is whether plaintiff’s negligence was “slight” and defendant’s “gross” in comparison. In many other states, the jury must determine the percentage of fault to be allocated to the plaintiff when the accident is considered as a whole.

**STATE V. KAATZ**

572 P.2d. 775 (Supr.Ct. Alaska, 1977)

* * * What is to be compared is negligence—conduct, fault, culpability—not causation, either physical or legal. There is no requirement that the party who is negligent in more ways, or for whom a greater number of individual negligent acts can be found, must be assigned the higher percentage. [Citations omitted].

Briefly summarized, the facts indicating Kaatz’ negligence are these: The Mitkof Highway was icy and very slippery, which condition was apparent to the casual observer. Kaatz had driven the section of the Mitkof Highway in question shortly before riding in the loader. Loaders are not designed for travel on a highway, but they are often driven over highways under their own power. Kaatz was familiar with the characteristics of the loader. He also knew the condition of the road. The loader is designed to carry only a driver, not a passenger, and the policy of Kaatz’ employer forbade riding on loaders, with exceptions not relevant to this case.* * * On previous trips, Kaatz had ridden in a pickup truck instead of on the loader. In those circumstances, Kaatz was negligent in riding in the loader on the Mitkof Highway.

The facts from which the state was held negligent in brief, are: The state highway department had a standard operating procedure that called for sanding of roads in inclement weather, even in outlying areas and even if it required overtime or night work. Cecil Donohue, the state maintenance foreman in the Petersburg sector, knew of the slippery condition of the Mitkof Highway before the accident occurred. His sanding machine did not work, but repairs were being undertaken. His stockpile of sand was frozen and covered with snow. He drove the stretch of road in question a few hours before the accident but decided not to sand it until the next morning, because he thought it might snow and cover the sand, even though he thought the road probably needed sanding that evening. Later that evening, after the accident, he and his men sanded the road. In those circumstances, the state was negligent in failing to sand the road in the late afternoon.

The state’s foreman testified that the road needed sanding for regular traffic, not for front-end loaders. Two nearby residents testified the road was worse that evening than they had ever seen it before. If he had considered only the testimony of Mr. Schouweiler, one of these nearby residents, the trial judge said he would have found Kaatz at least fifty percent negligent. On the other hand, he pointed out, other witnesses testified that they had negotiated the highway without inordinate difficulty. Kaatz himself drove the highway in a pickup truck without apparent incident shortly before riding the loader. Kaatz suggests another aggravating factor concerning the state’s negligence: it exposed large numbers of persons—everyone who traveled the highway—to danger, whereas Kaatz’ negligence had the potential to harm no one but himself.
Kaatz and the driver of the loader were killed when the loader slipped off the road and plummeted down an embankment.

QUESTIONS

1. How would you fit the facts of the above cases into the production function paradigm? Is the court in Bentzler rejecting the “fault” criterion advocated in Schwartz’s treatise? In State v. Kaatz, how does it seem to be contemplated that one would compare the state’s negligence or fault with that of Mr. Kaatz?

2. In Butand v. Suburban Marine & Sport Goods Inc., the Alaska Supreme Court applied its version of comparative negligence to a strict liability case. Consider the concurrence by Justice Rabinowitz:

   I generally agree with the court’s treatment of the issues which arise in the situation where a plaintiff who is himself negligent asserts a claim founded upon strict liability. Perhaps it is only a semantic difference rather than reflective of a true functional distinction but I prefer adoption of a comparative causation analysis in strict liability cases. Thus, I would require the trier of fact to compare the harm caused by the product’s defect with the harm caused by the claimant’s own negligence. See Solet v. M/V Capt. H.V. Dufrene, 303 F.Supp. 980, 986 (E.D.La.1969). Adoption of a comparative causation approach would avoid the theoretical problems inherent in any attempt to compare relative degrees of fault where the defendant’s negligence, or fault, is determined by principles of strict liability. 555 P.2d. 42 (1976) at 47.

   Can what Justice Rabinowitz seems to have in mind be translated into the production function approach? Are there any ambiguities in what he seems to be saying?

3. If you are still wondering what courts “really” have in mind, consider the statement made by the Wisconsin Supreme Court in Kohler v. Dumke, 13 Wisc.2d 211, 108 N.W.2d 581 (1961):

   * * * This court has never attempted to lay down any formula for determining how much weight is to be accorded to the element of negligence and how much to that of causation in comparing causal negligence. Neither do we think it advisable to attempt to do so. This is something that had best be left to the common sense of juries.

   Do you think that, given differing arguments from counsel, the typical jury might come up with wildly variant results in the face of precisely the same facts? Indeed, what is your own—presumably sophisticated—view of the best “common sense” treatment of these issues?

C. THE “LEVEL” OR QUANTITATIVE PROBLEM

One way to understand liability rules is to hypothesize that they are designed to allow the balancing of the benefits and harms from potentially injurious activities. This is the “marginal cost vs. marginal benefits” theme that has been present in the preceding sections. To the extent that this hypothesis is true, however, it means that there are two conceptually distinct dimensions to the balancing. One is that the “quality” of a potentially injurious act should be properly adjusted in the sense that all cost-beneficial care should be exercised in connection with any act taken. In addition, the “quantity” of acts must also be adjusted such that only acts whose benefits exceed their costs are taken. Our discussion thus far has not been very explicit about those distinct dimensions, but the perspective has in fact been principally from the qualitative or “caretaking” perspective. It is now time to complete the picture.

As Chapter I suggested, individuals have the proper incentives to engage in maximization behavior in the light of their own internal costs and benefits. The problem arises when there are third-party effects, i.e., “externalities” associated with an act. Unless legal liabilities are imposed to alter the incentives, a person will engage in what is arguably a “wrong” quantity and/or quality of the act in
question. Since we have already discussed the quality-adjusting implications of liability rules at some length, let us now elaborate some simple models in which the quantitative adjustment is the focus. Absent legal liability of any kind, the individual’s cost-benefit calculus might be represented as:

\[ \Delta B(Q) = \Delta K(Q) \]

where \( B(Q) \) is a benefit function that presumably rises monotonically as the quantity \( Q \) of the act increases, \( K(Q) \) is whatever private costs there are to the activity, and the symbol \( \Delta \) indicates that we are dealing with the marginal changes as \( Q \) is adjusted. In a no-liability world, the individual would choose a value of \( Q \) that satisfies equation \([1]\).

Now let us go to the other extreme and describe an “ideal” incentive system that balances all relevant costs and benefits, including those to third parties. The notation and terminology of earlier sections in this chapter will be modified only slightly. Care levels \( T^* \) and \( V^* \) are the tortfeasor and victim responses that satisfy the “qualitative” considerations discussed above, while \( C_t(Q,T^*) \) and \( C_v(Q,V^*) \) are the respective costs of achieving those care levels for activity level \( Q \). The “harm function” now also explicitly has \( Q \) added to it in addition to \( T \) and \( V \). Then the ideal marginal balance is where the following condition is satisfied:

\[ \Delta B(Q) = \Delta K + \Delta C_t(Q,T^*) + \Delta C_v(Q,V^*) + \Delta H(Q,T^*,V^*) \]

The last two terms in \([2]\) are the externalities or third party effects that do not appear in \([1]\) above because their consequences are ignored by the potential injurer. Similarly, \( \Delta C_t(Q,T^*) \) did not appear in \([1]\) because there would be no incentive for the injurer to incur care costs when no compensatory benefits accrue.

It will now be of interest to produce equations that represent the behavioral incentives under each of several major liability principles. The terms of the individual equation are written with the same positions and spacing as equation \([2]\) in order to facilitate comparison. Assuming (in line with the discussion earlier in the chapter) that potential injurers will rationally choose to adopt care level \( T^* \), negligence may be represented as:

\[ \Delta B(Q) = \Delta K + \Delta C_t(Q,T^*) \]

In words, the marginal cost of additional quantity is increased only by the incremental cost of engaging in that quantity of activity non-negligently; the injurer’s choice of activity level is not deterred by actual harm occurrence. By contrast, strict liability involves the condition:

\[ \Delta B(Q) = \Delta K + \Delta C_t(Q,T^*) + \Delta H(Q,T^*,V^*) \]

if contributory negligence is permitted as a defense, but.

\[ \Delta B(Q) = \Delta K + \Delta C_t(Q,T^*) + \Delta H(Q,T^*,V^*) \]

when contributory negligence is not a defense and the victim will engage in some unknown level of self-defensive care \( V' \). (The value of \( V' \) will normally be less than \( V^* \) and may be zero.) If the victim is not at \( V^* \), then the tortfeasor may find it advantageous to adjust his behavior to \( T' \), a care level different from the optimal \( T^* \).

Conditions \([3]\) and \([4]\) seem plainly to suggest underdeterrence and an excessive level of \( Q \), since one or more of the elements of external cost are not imposed on the injurer. Condition \([5]\) is ambiguous with respect to the optimal level \( Q^* \): the victim’s care costs (if any) are omitted, but the marginal legal liability \( \Delta H(Q,T^*,V) \) may be excessive due to the victim’s presumptive failure to exercise \( V^* \) of care and the consequent upward bias introduced into the resultant damages. Once raised explicitly, this “level problem” emerges as an extremely thorny barrier to the construction of simple liability rules that have all of the ideal behavior-adjusting characteristics. The reader may, however, also wish to reflect on...
whether any of the equation formulations are arguably too simplistic. For instance, should the definition of “negligence” include the choice of Q as well as T?

There is a very noteworthy special class of situations that affect the conclusions drawn from the simple models just discussed. Suppose that an activity results in a substantial amount of reciprocal externalities being imposed by persons who engage in it. For instance, when a person drives, he imposes costs on other motorists, both in the form of victim avoidance costs $C_v(V)$ and possible uncompensated harms. But the other motorists, in turn, may impose roughly comparable costs on him. What implications does this have for the level problem? Are these implications compensatory or behavioral?

QUESTIONS

1. Create a diagrammatic exhibit that illustrates the potential tortfeasor’s marginal choice calculus under the different situations of equations [1] through [5], labelling the quantities as $Q_1$ through $Q_5$. By shading or other markings, indicate the magnitude of what might be regarded as the social “error costs” in each situation.

2. Is there a possible “level problem” in defining contributory negligence?

3. The text of this section was implicitly based upon a situation in which the rules of liability are accurately applied. Earlier in the chapter, the influence of various types of adjudication errors was discussed: How do those types of errors affect the results in this section? And, equally important, how are the earlier results affected by the necessity to consider the level problem?

GREENLAWN NEIGHBORHOOD ASSOCIATION v. FIRST GUARANTY BANK

The president of the First Guaranty Bank likes to fly helicopters, perhaps because his license to drive automobiles has been revoked upon conviction for excessive vehicular violations. When a new building was constructed for the bank, a rooftop helicopter pad was incorporated into the design. The bank uses the helicopter to transport its checks to a Federal Reserve clearinghouse about 75 miles away, thus avoiding an additional day’s “float” on the checks being presented for collection. The alternative would be to transport the checks to an airport about 8 minutes away by automobile and then to use a fixed-wing aircraft. Fixed-wing aircraft have much better accident rates than helicopters.

QUESTIONS

1. If an accident occurs, what should be the standard of negligence? Is it just the appropriately careful operation of the helicopter? Or does a comparison with alternative transport modes have to be factored into the calculus called for under the Learned Hand rule?

2. What difference might it make if the applicable standard were one of strict liability rather than negligence?

3. The bank is in an area, Greenlawn Heights, which is now quite congested. Does it make any difference if the accident involves a crash into an area that did not develop until after the bank had instituted a regular practice of using the helicopter?

4. What are some other activities or scenarios that are analogous to the helicopter case?

5. If a strict liability standard is used, what are the arguments for and against coupling it with a contributory negligence defense?
D. TAKING COST-MINIMIZATION SERIOUSLY(?)

The cost-minimization principles that seemingly underlie the negligence and contributory negligence doctrines have at least a defensible if not outright attractive-rationale. Pushing a principle to what are apparently its logical implications nevertheless often poses interesting questions. This is, of course, the classical method of *reductio ad absurdum*. Arguably, the following scenarios push cost-minimization to an unacceptable extreme. Or are they so “absurd” after all???

*The Case Of Innocente v. Vinho*

Innocente lives in a house about 50 yards from a sharp turn on a mountain road. He has observed that Vinho makes a regular practice of getting drunk after work on paydays, which occur on alternate Fridays. Vinho then drives down the mountain in a dangerous manner and appears likely, sooner or later, to be incapable of negotiating the curve in the road. The most convenient place for Innocente to park his car is on the curving driveway in front of his house. Unfortunately, the driveway is in the likely path of Vinho’s vehicle if it leaves the road. At the cost of some modest inconvenience, Innocente could park his car on the utility driveway at the side of the house, thus avoiding all prospect of damage due to Vinho’s binges, except for the insubstantial matter of possible tire tracks across the lawn. Assume that Innocente nonetheless continues to park his car in front of the house, Vinho smashes into it, and the damage to the Innocente automobile is estimated at $8500.

In answering the following questions, you should attempt to utilize the model incorporated in Exhibit 5.6 above. Vinho’s predicted level of care would be the level indicated as $T_n$ in Exhibit 5.6. Normally, Innocente would be held only to the $V^*$ level of care, on the assumption that Vinho is non-negligent.

**QUESTIONS**

1. If you were counsel for Vinho, how would you try to argue that Innocente should not collect $8500?
2. If you were the attorney for Innocente, how would you respond to the arguments made?
3. Would it make any difference in the previous two questions if Innocente were able, at a time when he still had time to move his car, to look 5 miles up the mountain and see the Vinho vehicle approaching in a manner that made an accident seem almost certain?
4. What does the doctrine of “Last Clear Chance” require? Does it *always* apply to one who has “last chance” to avoid a harm? Why?
5. On these facts, how do you think a real court would find? What do you think is the “right” rule? Why?
6. Assume that the jurisdiction in question employs a comparative negligence liability rule. Does this affect your answer to any of the first three questions above?
7. Suppose that Innocente came to you before the accident, described the situation, and asked whether he had to move his car. What advice would you have given?

The above hypothetical was created purely to dramatize a pedagogical point and was first used in the author’s classes in 1977. Many students may have regarded it as infected with more than a hint of unreality. That explains the inclusion of the following item, based on press accounts in August, 1982.
Drunks Won’t Let The Preacher Alone

Chattanooga, Tenn.—A Baptist preacher says he’s fed up with the flying bodies and flaming cars of drunken drivers who repeatedly crash into his house. He’s so angry he’s threatening to erect a brick wall in his front yard even though “it’s going to kill everybody who hits it.”

“If the city doesn’t put up a guardrail, we’re going to put up a brick wall,” the Rev. Martin Pinion said Wednesday. “It’s either deflect them with a guardrail or kill them with a brick wall.” “I’m not real concerned about the motorists because they’re just about always drunk. It’s just like war,” he said. Seven drivers in the past six years have slammed into the house after careening off the busy highway that curves past.

“They’re always drinking,” said Pinion’s wife, Debra. “It seems like they always hit in the middle of the night when you’re sleeping. Then it’s up to us to go out and pull them out of their burning cars.”

“These wrecks move the whole house,” Mrs. Pinion said. “Things get shifted around, pushed one way and the other. After an accident the doors won’t close and the cabinets are crooked. Everything in the house is out of keel.”

E. PROMISE-BREAKING ANALYZED AS A TORT

Review, if necessary, the indifference curve model on p. 161ff. above where the concepts of beneficial and detrimental reliance were illustrated. To the extent that a broken promise has created detrimental reliance, the promisee has suffered a harm caused by the promisor. An extraterrestrial visitor upon first encountering our legal system might inquire what is the difference between the treatment of a harm that is labelled as a contract breach and one that carries the name of a “tort.”

The extraterrestrial might well be puzzled as to why the harms flowing from some broken promises are penalized—or, if you prefer, are the subject of compensation—while others are ignored as being the product of “naked” or “gratuitous” promises. The uncompensated breaches, he might argue, appear to the otherworldly eye to be no less than unrestrained torts. On the other hand, if a promise is enforceable because it is for some reason classified as a “contract,” then we seem to go to the opposite extreme of over-restraining; after all, the only true social loss is the detrimental reliance and not the full value of the expectation interest usually given as contract damages.

1. ECONOMICALLY OPTIMAL DAMAGES FOR BROKEN PROMISES

The analysis that follows is worked out more thoroughly in Goetz and Scott, “Enforcing Promises: An Examination of the Basis of Contract,” 89 Yale L.J. 1261, 1273-1286 (1980). This edited excerpt, however, may suggest how one would approach the enforcement of promises from a perspective very similar to tort theory. You may find that, with due allowance for practical considerations, the principles governing the enforcement of torts and contracts are quite congruent.

* * *

* * * When a promise is made in good faith, the promisor presumably believes that he is likely to perform. Still, many good-faith promisors would acknowledge the possibility that events may arise that cause them to regret having made the promise. Thereafter, if it were costless to do so, they would indeed breach the promise. Such contingencies may involve a wide range of factors, from changes in personal conditions to disappointment about external considerations that originally made the promise seem desirable. The term “regret contingency” will be used to denote the future occurrence of a condition that would motivate breach if breach were a costless option for the promisor. Assuming any reliance, the occurrence of a regret contingency necessarily implies that either the promisor or promisee must bear a cost.
When a regret contingency arises, the promisor’s options are either to bear the loss attributable to performance, which now costs more than it is worth, or to breach and accept the cost of any corresponding sanction. Presumably the promisor would adopt the cheaper of these regret costs. In any event, someone will suffer a net loss whenever a regret contingency arises, whether in the form of regret costs to the promisor, uncompensated detrimental reliance to the promisee, or both.

By what means does the promisor adapt to the prospective costs of promising? The promisor can substantially influence the probability of a regret contingency, and thus its prospective costs, by adjusting his behavior ex ante. One means of mitigating potential costs is by altering the form of the promise. For instance, the promisor may condition performance on the proviso that certain circumstances—potential regret contingencies—not arise. Alterations in the form of the promise will generally entail a cost to the promisor either in terms of direct resource cost—time and trouble—or in the possibility that the benefit of the promise-making to the promisor will be diminished. The second means of avoiding regret costs is simply to make fewer promises. The costs of this option are forgone benefits from unmade promises.

The costs that result from restrictions in the scope or number of one’s promises can be termed “precautionary costs.” It is useful to distinguish these further as either quality precautions or quantity precautions. Quality precautions involve adjustments restricting the scope of promises and impose a cost of decreased reliability. Quantity precautions, which consist of reductions in the number of promises made, result in a loss of benefits from promising. A rational promisor will pursue precautionary adjustments up to the point at which marginal precautionary costs are exactly balanced by marginal reductions in regret costs.

Precautionary adjustments by the promisor decrease the value of the promise. Conversely, when the promise is worth more to the promisor than its prospective cost, the promisor may engage in “reassurance.” Reassurance includes such actions as the offer of guarantees, verbal persuasions, and the development of a reliable reputation, designed to convince the promisee that the promise is valuable. Reassurances increase the value of the promise to the promisee. Indeed, promisees may regard voluntary reassurance measures as substitutes for sanctions. Reassurance usually entails some cost to the promisor and, hence, will be pursued up to the point at which marginal reassurance costs are exactly balanced by increases in resulting benefits to the promisor.

Precautionary and reassurance reactions by promisors are triggered by variations in the cost of promising. It should be apparent, therefore, that an additional legal sanction will raise prospective costs, thereby precipitating adjustments of the scope and volume of promises.

Optimizing Promisor-Promisee Interaction

To what extent do legal sanctions optimize the interactions between promisor and promisee? In the present context, optimization is defined as maximizing the net social benefits of promissory activity—that is, the benefits of promises minus their costs. This approach is equivalent to the balancing of prospective costs and benefits under the widely accepted Learned Hand test for the required duty of care in potential tort-producing activities. Indeed, there are strong theoretical parallels between the production of dangerous, but useful, products and the making of promises.

The role of damages or sanctions in generating socially optimal behavior can be focused more sharply by observing the distinction between internal and external effects. Because self-interested maximizing behavior entails consideration of only internal costs and benefits, unfettered individual behavior is incompatible with social optimization in circumstances in which significant external costs or benefits are present. Individuals will oversupply activities with external costs and undersupply those with external benefits. By imposing costs and creating incentives, the law can cause individuals to consider external effects in their decisionmaking and thus “internalize” them.
Inducing optimal promise-making therefore requires that the promisor’s costs of promising be adjusted to reflect any external effects on the promisee. But this adjustment process is complex. Changes in the costs and benefits of promising are highly interactive in two senses. First, an individual’s adjustments may substitute one category of his costs for another. Second, the actions of one party may produce reactions by the other and, in turn, feedback responses to the first party.

1. Nonreciprocal Promises

Consider the case of a gratuitous promisor who has adjusted his promise-making to an arbitrarily assumed level of extra-legal sanction so that he cannot further improve his situation. In addition, assume that social considerations effectively prevent the promisee from influencing the promisor’s calculations through bargaining. Under these conditions, when does the intervention of the law lead to optimal results?

The effect of a decision to enforce legally any particular class of nonreciprocal promises depends upon the nature of the sanction imposed for breach. Promisors will respond to higher levels of sanction by increasing both the reliability of a given volume of promises and the number of promises actually made.

A necessary starting point in determining an optimal damage rule is to specify the external effects of a nonreciprocal promise as the supply of such promises is increased by one marginal unit. The external effects are the prospective detrimental reliance incurred if the promise is broken and the prospective beneficial reliance enjoyed if the promise is performed. Proper reflection of external effects therefore requires not only that the promisor be charged for the harm expected from broken promises, but also that he be rewarded for the prospective benefits of performance. It is helpful to state this condition symbolically. Let $p$ be the promisor’s reasonable, subjective assessment of the probability that he will perform a promise under an existing legal rule calling for damages of $D$ in the event of breach. For the damage rule to deter all promises with net social costs and encourage those with net benefits, the amount of damages awarded must satisfy the following equation:

$$(1-p)D = (1-p)R - pB$$

where $R$ and $B$ are the values of detrimental and beneficial reliances, respectively. Assuming that all broken promises are litigated, the left-hand side of the equation represents the expected value of the prospective legal sanction. Because only broken promises are affected by the law, the probability $(1-p)$ of the promise being broken is used to “discount” the damages $D$. The values for $R$ and $B$ on the right-hand side of the equation should be understood as those resulting from optimal self-protection by the promisee. Thus, promisees will appropriately minimize the value of the right-hand term, which is the net social cost of the promise. In calculating this prospective net reliance, the magnitudes of the potential detrimental and beneficial reliances are each discounted by their probabilities. When the equation is satisfied through the imposition of optimal damages $D$, the promisor’s internal cost-benefit calculus will reflect the external effects of his promise-making. If the external effects are thus accounted for, the promisor’s maximization of his internal net benefits is consistent with supply of the socially optimal quantity and quality of promises. We call this damage rule the “prospective net reliance” formulation.

In some cases, the prospective beneficial reliance from a promise will exceed its prospective detrimental reliance. Because the net external effect of such a promise is beneficial, it would be optimal to reward the making of such promises. However, in the nonreciprocal setting no practical legal mechanism exists for rewarding promises. This limitation renders true optimization impossible; the situation is necessarily second-best. At minimum, promises with prospects of net beneficial reliance should not be the subject of damages if breached. Only promises with prospective net detrimental external effects should be enforceable. The prospective net reliance formulation developed above can be used to analyze the optimal level of enforcement. By dividing both sides of the original equation by the probability of breach $(1-p)$, the following damage rule emerges:
D = R – [p/(1-p)] R

The optimal damage rule thus subtracts from the promisee’s reliance cost a fraction of his potential beneficial reliance. This fraction is the ratio of the ex ante subjective probability of performance to that of nonperformance. It determines the extent to which the prospect of beneficial reliance when the promise was made is credited against the promisee’s prospective detrimental reliance. Because this ratio may be thought of as an index of the promisor’s good faith, we call it the “good-faith ratio.” A damage offset based on the good-faith ratio and on the amount of potential beneficial reliance will encourage the optimal quantity and quality of promises by reflecting in the promisor’s decision calculus both the harmful and beneficial effects of his promise-making. This optimal legal sanction is likely to be unattainable in an environment of costly legal process and imperfect information. But specifying an optimal sanction permits more rigorous evaluation of the error produced by any practical adjustment attributable to process costs.

The rule suggested by the formula above will admittedly result in a large quantity of uncompensated damages from broken nonreciprocal promises. Although greater damages would deter many injurious transactions, a stricter standard also would deter beneficial transactions in even greater magnitudes. Viewing an already-broken promise, the affected promisee would always prefer the highest possible damage award. But from the ex ante standpoint, a promisee would not wish to discourage a promise that creates a prospect of gain out weighing the risk of uncompensated loss. Such promises are “good bets” for the promisee over the long run, even though some of the promises will result in uncompensated harm. The penalty formulation developed above awards damages both to protect promisees from “bad bet” promises and to avoid deterrence of promises that are “good bets.”

Under this prospective net reliance damage formula, the gratuitous promisor also has an incentive to undertake cost-effective qualitative precautions to modify prospective reliance induced by any promise actually made. A properly calibrated legal sanction will induce the promisor to convey to the promisee socially beneficial information about the risk of regret contingencies. In essence, such a promisor is encouraged to make cost-effective adjustments in both the quality and the quantity of his promises because the legal rule converts social benefits into savings for him. If, in addition, the law recognizes only the amount of damages that constitutes a “reasonable”—i.e., cost-effective reliance by the promisee, then the promisee will also have an incentive to minimize net social costs. Thus, the rule penalizes each party for failing to take cost-effective steps to minimize the social costs of promising. Damages exceeding those described above will tend to induce the promisor to invest too much in precautionary adjustments. This phenomenon is analogous to the excessive level of prudence anticipated if tort victims were awarded a multiple of their true damages.

2. Reciprocal Promises

Promising is reciprocal when the parties can adjust interactively to the nature and amount of promise-making. The prospective net reliance formulation is equally applicable to reciprocal as well as to nonreciprocal promises. But the net reliance damage rule seems in sharp conflict with accepted legal doctrine in the reciprocal promise context, in which damages for breach are typically based on the promisee’s full-performance expectation rather than on his detrimental reliance. Upon analysis, the apparent conflict can be dissipated; moreover, reciprocal promises are easier than nonreciprocal promises for the law to address.

This conclusion is buttressed by two independent lines of argument. First, in the case of reciprocal promises, a plausible empirical generalization is that a promisee’s acceptance of one promise frequently requires his foregoing a potential substitute promise. The forgone value of the best substitute promise available—the opportunity cost—is key in determining the promisee’s detrimental reliance when an
accepted promise is subsequently broken. In a well-organized market, alternative promises will be close, if not perfect, substitutes. In that case, detrimental reliance is equal to the full performance value of the breached promise. Similarly, beneficial reliance will be small, because the promisor’s pledge, even if performed, will not constitute a very substantial improvement over the potential beneficial reliance from substitute promises. This empirical generalization implies that, in the damage formula developed above, full performance expectation \( E \) can be substituted for detrimental reliance \( R \) because \( E \approx R \). Furthermore, because \( B \approx O \), the term of \( \frac{p}{1-p}B \) drops out. We are left with \( D \approx E \); thus, expectation damages are a good proxy for the prospective net reliance damage formulation developed above. Nothing in the logic of this argument limits its applicability to reciprocal promises. Although the argument’s empirical premise tends to have greater validity with respect to bargained-for promises, classes of nonreciprocal promises may exist for which forgone substitute promises also yield a close convergence between reliance and expectation.

Second, a fundamental theoretical difference exists between reciprocal and nonreciprocal promises. In the case of a reciprocal promise, the principal objective of a promisor is to obtain consideration in the form of a return promise. The value of the return promise elicited is the main element of the promisor’s benefit. Therefore, changes in the qualitative aspects of the promise are reflected in commensurate shifts of benefits to the promisor; a higher quality promise motivates a more valuable return promise, and vice versa. In contrast to the case of nonreciprocal promises, qualitative adjustments are internalized in the promisor’s cost-benefit calculus by generating a more or less valuable consideration for his promise. Hence, the bargaining process accomplishes an important part of the behavioral regulation that, for nonreciprocal promises, must be performed by the legal system.

Furthermore, the bargaining process, not available by definition in the nonreciprocal context, can facilitate the optimal allocation of risk for reciprocal promises. Precautionary action is subject to a test of the ability of the promisee to bribe the promisor to make an unconditioned promise. Within any scheme of enforcement, then, the parties can reallocate the risks of regretted promises by buying or selling protection through the terms of their agreement. The least-cost bearer of any risk will presumably agree to absorb that risk in exchange for an enhanced return promise.

For much the same reasons, the consequences of excessive damages for breaching reciprocal promises are also mitigated, as long as the rule providing for excessive damages is understood in advance. The parties can always bargain out from the rule, for instance by a limited damages agreement. Thus, when transactions costs are zero, the particular damage rule selected for reciprocal promises is irrelevant. Although the existence of transactions costs renders bargaining over damage rules costly in practice, the feedback adjustment of the return promise markedly reduces the potentially inefficient effects of legal rules. While in the nonreciprocal case, excessive damage overdeter the promisor from promissory activity, in the reciprocal relationship, the promisee will regard the excessive damages as a quality improvement and will offer an enhanced return promise. The enhanced return promise will tend to offset the deterrence effect of the damages. However, some inefficiencies remain. Legally mandated “overinsurance” induces a moral hazard because the promisee will not exercise optimal self-protection. Furthermore, there is a cheaper allocation of risk than the legally mandated level of reassurance provided to—and paid for by—the promisee.

The result may be even more costly, however, when the law provides a suboptimal level of enforcement. The extreme case of a complete refusal to enforce reciprocal promises provides an instructive illustration. The initial impact of the rule, which is to underdeter promisors, will be counterbalanced by reductions in the value of return promises. Promisors may then substitute extra-legal forms of reassurance for legal sanctions. Creating adequate extra-legal enforcement mechanisms is likely to be less efficient than legally sanctioned reassurance. If so, the inefficiency consequences of underenforcement may be more serious than those of overenforcement.
In sum, the theoretical damages principles developed in connection with nonreciprocal promises apply to reciprocal promises as well. The difference in their legal treatment may be due to a close empirical identity of reliance and expectation in reciprocal promises. However, modification of the return promise is a powerful additional adjustment mechanism, which exists, by definition, only for reciprocal promises. By internalizing many of the promissory interactions among contracting parties, the return promise reduces the stress placed on legal rules for optimally influencing the behavior of promisors and promisees.

**QUESTIONS**

1. In what respect was the present section concerned with what was discussed earlier as a “level problem”? Why is the problem more complicated here?

2. Under what circumstances would promisees prefer that a certain class of promise not be enforced?

3. “In the realm of bargained-for promises, the breach damages can be considered as a form of ‘insurance policy’ that each party sells to the other as protection against his own breach.” Explain. Does this perspective suggest that parties would never want more than refinance damages?

4. Suppose that the law mandates higher damages than the parties themselves would prefer. Why is this less troublesome with bargained-for promises than with gratuitous ones? In what form do the resultant “error costs” occur?

5. Buyer contracts with Seller for the delivery of one gadget at a price of $200. Seller is the exclusive source of gadgets and can produce a single gadget per period at a cost of $150. Buyer’s potential benefits from the gadget depend on how many hours of gadget-using instructions he takes at a cost of $10 per hour: $230 of gross benefits with one hour of instruction, $242 with 2 hours of instruction, $258 with three hours of instruction. Since gadget expertise is quickly forgotten, any hours of instruction purchased by Buyer will be worthless if the promised gadget is not delivered at the agreed time. Both Buyer and Seller know that there is one chance in three that Otherbody will, at the time set for delivery, also desire a gadget and, if so, would bid $260 for it. No matter which of the common contract remedies is assumed, Otherbody will probably wind up with the gadget. Why? Would a reliance measure of damages cause Buyer to over-rely? Does the same answer apply in the case of expectation damages?

6. To what extent is “reasonable reliance” in contract analogous to the contributory negligence doctrine in tort?

7. If contract damages are in any sense similar to insurance payments, does it make sense to distinguish between exogenous risks not controllable by either party and endogenous risks that are subject to a significant degree of control? What kinds of things would fit into each category? [If this question is not clear at this point, return to it after the section on relational contracting below.]

2. **SEC. 90, RESTATEMENT OF CONTRACTS**

Section 90 of the Restatement of Contracts is one basis upon which gratuitous promises have traditionally been enforced. The original wording of Sec. 90 was altered in Restatement (Second) by deleting the bracketed phrase below and adding the underlined sentence.

A promise which the promisor should reasonably expect to induce action or forbearance [of a definite and substantial character] on the part of the promisee and which does induce such action or forbearance is binding if injustice can be avoided only by enforcement of the promise. The remedy granted for breach may be limited as justice requires.
QUESTIONS

1. What reasoning might account for the deletion of the requirement for reliance to be “definite and substantial”?

2. What factors would justify a limitation of the remedy?

3. Why are the words “if injustice can be avoided only by enforcement of the promise” included in Section 90? After all, are there any circumstances when an uncompensated detrimental reliance would not imply an “injustice”?

4. Which version of Section 90 do you find preferable? Why?

CENCO, INC. v. SEIDMAN & SEIDMAN
686 F.2d 449 (7th Cir., 1982).

[Between 1970 and 1975, managerial employees of Cenco, Inc. engaged in a massive fraud which Cenco’s independent auditor throughout the period, the accounting firm of Seidman & Seidman, either never discovered or, if it did, failed to report. The fraud involved the inflating of inventories far above their actual value, thus exaggerating the profitability of Cenco and leading to a higher market price for its stock. The unmasking of this practice, followed by a 75% fall in the price of Cenco stock, led to the filing of a class action in federal district court against Cenco, its corrupt managers, and Seidman & Seidman. The class was composed of purchasers of Cenco stock during the period of the fraud, when the stock price was inflated. Cenco filed a cross-claim against Seidman & Seidman for having failed to prevent the fraud by Cenco’s managers. Seidman then filed its own cross-claim against Cenco, alleging that Seidman had been one of the victims of Cenco’s fraud and was thus entitled to damages from Cenco. Cenco’s and Seidman’s cross-claims against each other were set for trial together. Several aspects of the cross claims were settled by directed verdict or dismissal. The case went to the jury, however, on three counts in Cenco’s cross-claim, alleging respectively breach of contract, professional malpractice (negligence), and fraud. Cenco’s evidence claimed to show that in the early stages of the fraud Seidman had been careless in checking Cenco’s inventory figures and its carelessness had prevented the fraud from being nipped in the bud. Seidman’s evidence attempted to show that Seidman had diligently attempted to follow up all signs of fraud but had been thwarted by the efforts of the large group of managers at all levels of Cenco who collaborated to prevent Seidman from learning about the fraud. The jury found for Seidman on all three counts. Cenco appeals from the judgment, contending that the jury’s verdict was based on erroneous instructions.]

POSNER, Circuit Judge.

This brings us to the main issue in the case—whether the district judge gave erroneous instructions to the jury. The challenged instructions relate to the question whether Seidman was entitled to use the wrongdoing of Cenco’s managers as a defense against the charges of breach of contract, professional malpractice, and fraud. Despite the plurality of charges it is one question because breach of contract, negligence, and fraud, when committed by auditors, are a single form of wrongdoing under different names. The contract in question here * * * consists of the letters between Seidman and Cenco outlining the terms of Seidman’s annual retention to audit Cenco’s books. The material part of the letters is the incorporation by reference of general accounting standards which require the auditor to use his professional skill to follow up any signs of fraud that he discovers in the audit. The tort of negligence in the context of auditing is likewise a failure to use professional care and skill in carrying out an audit. And if such care and skill are not used, then the audit reports to the client will contain misrepresentations, either negligent or, if the auditor knows that the representations in the reports are untruthful or is indifferent to whether or not they are truthful, fraudulent.
Because these theories of auditors’ misconduct are so alike, the defenses based on misconduct of the audited firm or its employees are also alike, though verbalized differently. A breach of contract is excused if the promisee’s hindrance or failure to cooperate prevented the promisor from performing the contract. See Restatement (Second) of Contracts Sec. 245 (1979). The corresponding defense in the case of negligence is, of course, contributory negligence.

Negligence is not a defense to an intentional tort such as fraud. E.g. Mother Earth, Ltd. v. Strawberry Camel, Ltd., 72 Ill.App.3d 37, 52, 28 Ill.Dec. 226, 238, 390 N.E.2d 393, 405 (1979). But a participant in a fraud cannot also be a victim entitled to recover damages, for he cannot have relied on the truth of the fraudulent representations, and such reliance is an essential element in a case of fraud. See, e.g. Broberg v. Mann, 66 Ill.App.2d 134, 139, 213 N.E.2d 89, 92-93 (1965). If the misrepresentation is negligent rather than intentional, contributory negligence plays the same role it would play in an ordinary negligence case. See Prosser, supra, at 706.

The jury instructions in this case stated these defenses accurately, but Cenco contends that the instructions should not have been given, because they related not to Cenco’s conduct but to that of its managers. The judge was aware of the distinction but instructed the jury that the acts of a corporation’s employees are the acts of the corporation itself if the employees were acting on the corporation’s behalf. If this instruction was correct, then the instructions which allowed the jury to consider Cenco’s misconduct as a defense to Seidman’s alleged wrongdoing were proper.

To determine the correctness of the instruction requires us to decide in what circumstances, if any, fraud by corporate employees is a defense in a suit by the corporation against its auditors for failure to prevent the fraud. Illinois precedent allows us to reject one extreme position on this question, which is that the employee’s fraud is always attributed to the corporation by the principle of respondeat superior. This position * * * was rejected in Cereal Byproducts Co. v. Hall, 8 Ill.App.2d 331, 132 N.E.2d 27 (1956), where a company’s independent auditors were held liable for negligently failing to detect embezzlement by the company’s bookkeeper. Auditors are not detectives hired to ferret out fraud, but if they chance on signs of fraud they may not avert their eyes-they must investigate. The references to keeping an eye out for fraud that appear in the accounting standards incorporated (by reference) in the retention letters between Cenco and Seidman would have little point if not interpreted to impose a duty on auditors to follow up any signs of fraud that come to their attention.

But this does not tell us what the result should be if the fraud permeates the top management of the company and if, moreover, the managers are not stealing from the company—that is, from its current stockholders—but instead are turning the company into an engine of theft against outsiders creditors, prospective stockholders, insurers, etc. On this question the Illinois cases on auditors’ liability provide no guidance. In fact, to our knowledge the question has never been the subject of a reported case. Leeds Estate, Bldg. & Inv. Co. v. Shepherd, 36 (L.R.) Ch. D. 787, 802, 809 (1887), described by Cenco in its main brief as “the one squarely applicable common law decision on accountants’ liability,” is nothing of the sort. The auditor in that case had failed to discover that the company’s manager, by misrepresenting the profits of the company, had caused the company to pay out dividends, directors’ fees, and bonuses for himself—all in violation of the charter—as a result of which the company went broke. This was stealing from, not for, the company. In predicting how the Illinois courts might decide the present case, we assume they would be guided by the underlying objectives of tort liability. Those objectives are to compensate the victims of wrongdoing and to deter future wrongdoing. With regard to the first, we must refine our earlier statement that the “victim” of Seidman’s alleged laxity was Cenco Incorporated. A corporation is a legal fiction. The people who will receive the benefits of any judgment rendered in favor of Cenco on its cross-claim against Seidman are Cenco’s stockholders, comprising people who bought stock in Cenco before the fraud began, people who bought during the fraud period and either sold
afterwards when the stock price fell or continue to hold the stock at a loss, and people who bought after the fraud was unmasked. A judgment in favor of Cenco on its claim against Seidman would not differentiate among these classes, but would benefit every stockholder as of the date of the judgment (or the date when a judgment was anticipated with some precision) in proportion to the number of shares he owned.

Once the real beneficiaries of any judgment in favor of Cenco are identified, it is apparent that such a judgment would be perverse from the standpoint of compensating the victims of wrongdoing. Among the people who bought stock in Cenco before the fraud began are the corrupt officers themselves. To the extent they are still-stockholders in the company, they would benefit pro rata from a judgment in favor of Cenco. The other stockholders in this class are innocent in a sense, but of course it is they who elected the board of directors that managed Cenco during the fraud. The people who bought during the fraud period and either sold at a loss or continue to hold at a loss are the plaintiffs in the recently settled class action in which both Cenco and Seidman were defendants. Seidman has already paid $3.5 million to them. Those who continue to own stock in Cenco (as distinct from those who sold at a loss) would receive additional compensation if Cenco prevailed in this action against Seidman. This is not to say they would be overcompensated; but it seems odd that the same shareholders should be able to recover damages from Seidman twice for the same wrong-once directly and once, in this suit, indirectly. Finally, the shareholders who bought after the fraud was unmasked lost nothing. The unmasking of the fraud caused the price of Cenco’s stock to be bid down to reflect not only the true value of its inventories but also any anticipated injury to the company as a result of the fraud.

Because of shareholder turnover, there is always a potential mismatch between the recovery of damages by a corporation and the compensation of the shareholders actually injured by the wrong for which the damages were awarded. It is simply a more dramatic mismatch in this case than usual.

From the standpoint of deterrence, the question is whether the type of fraud that engulfed Cenco between 1970 and 1975 will be deterred more effectively if Cenco can shift the entire cost of the fraud from itself (which is to say, from its stockholders’ pockets) to the independent auditor who failed to prevent the fraud. We think not. Cenco’s owners—the stockholders—hired managers (directly, in the case of the president and chairman, who were both members of the board of directors, indirectly in the case of the others) who turned out to be thoroughly corrupt and to corrupt the corporation so thoroughly that it caused widespread harm to outsiders. If Seidman had been a more diligent auditor, conceivably if it had been a more honest auditor, the fraud might have been nipped in the bud; and liability to Cenco will make Seidman, and firms like it, more diligent and honest in the future. But if the owners of the corrupt enterprise are allowed to shift the costs of its wrongdoing entirely to the auditor, their incentives to hire honest managers and monitor their behavior will be reduced. While it is true that in a publicly held corporation such as Cenco most shareholders do not have a large enough stake to want to play an active role in hiring and supervising managers, the shareholders delegate this role to a board of directors, which in this case failed in its responsibility. And not all of Cenco’s shareholders were “little people.” During the period of the fraud Curtiss-Wright Corporation owned between 5 and 16 percent of Cenco’s common stock and had its own accounting firm conduct a study of Cenco’s operations—without discovering the fraud.

Thus, not only were some of Cenco’s owners dishonest *, *, but the honest owners, and their delegates ** were slipshod in their oversight and so share responsibility for the fraud that Seidman also failed to detect. In addition, the scale of the fraud—the number and high rank of the managers involved—both complicated the task of discovery for Seidman and makes the failure of oversight by Cenco’s shareholders and board of directors harder to condone.

** [I]f Cenco may be divorced from its corrupt managers, so may Seidman from the members and employees of the firm who suspected the fraud. If Seidman failed to police its people, Cenco failed as or more dramatically to police its own.
Furthermore, we must assume that Cenco’s corrupt managers were acting for the benefit of the company, not against it as in the Cereal Byproducts case. The jury was instructed that it could attribute the fraud of Cenco’s managers to Cenco only if it found that the managers had been acting on Cenco’s behalf, and the verdict for Seidman implies that the jury either so found or found that Seidman had not even committed a prima facie breach of duty to Cenco. The former assumption is more favorable to Cenco.

Fraud on behalf of a corporation is not the same thing as fraud against it. Fraud against the corporation usually hurts just the corporation; the stockholders are the principal if not only victims, their equities vis-a-vis a careless or reckless auditor are therefore strong. But the stockholders of a corporation whose officers commit fraud for the benefit of the corporation are beneficiaries of the fraud. Maybe not net beneficiaries, after the fraud is unmasked and the corporation is sued—that is a question of damages, and is not before us. But the primary costs of a fraud on the corporation’s behalf are borne not by the stockholders but by outsiders to the corporation, and the stockholders should not be allowed to escape all responsibility for such a fraud, as they are trying to do in this case.

We need not go so far as to predict that the Illinois courts would hold that in any action by a corporation against its auditors an employee’s fraud intended to benefit the company rather than the employee at the company’s expense will be attributed to the corporation, however lowly the employee. It is true that the lower down the employee is in the company hierarchy, the less likely he is to commit fraud for rather than against the company. But there are overzealous employees at every level—many a corporation has paid heavy damages for antitrust violations committed by low-level sales managers who thought they were acting in the company’s best interests as well as their own—and we think it premature as well as unnecessary to decide that an auditor is never liable for the frauds of loyal but misguided company employees that he could have prevented by taking care. But here the uncontested facts show fraud permeating the top management of Cenco. In such a case the corporation should not be allowed to shift the entire responsibility for the fraud to its auditors.

Hence the challenged jury instruction was proper in the circumstances * * *

**HYPOTHETICAL: PERILS OF A TAXJACQUES**

Marie Taxjacques is a 1973 graduate of Prestige University Law School with degrees earned through the joint MBA/JD program. Her undergraduate major was in accounting and she has since qualified as a CPA. Presently, she is a partner in the firm of Taxjacques and Lehgerscan, a medium-sized company that provides a broad range of legal and financial services on a contract basis.

a. Last year, Taxjacques’ services were much in demand as an advisor on certain arcane tax shelters arising out of the latest revisions in the tax code. Unfortunately, Taxjacques tendered advice that now appears clearly to have been not only erroneous but the result of professional negligence on her part. In addition to her own direct clients who relied to their detriment on her advice, two other groups of plaintiffs have also filed suit against Taxjacques.

The first of these are clients of a local branch of S.N. Arbloch, a local financial advisor who, in turn, had retained Taxjacques as his own source of expert information to be passed on to his clients. It has been established that Arbloch was not in any way himself negligent and he is, in any event, not possessed of financial resources adequate to cover the claimed damages of his clients. Consequently, Arbloch’s clients are seeking to recover against Taxjacques, the “real source” of the incorrect advice.
A second group of suits has been filed against Taxjacques by friends of Edith M. Goodolgal. It seems that Goodolgal sought Taxjacques’ advice on her own behalf but also passed on this advice to her wealthy friends at the next local meeting of the Dames of the American Insurrection.

Should Taxjacques’ be liable to any of these “third party” plaintiffs who have been injured as a result of her negligence? Why? If the answer depends in part on facts that are not specified, say what they are and why they are relevant.

b. It has truly not been a good year for Taxjacques. She was also the principal in a longstanding contract to perform the annual audits of the books of the Pseudo Corporation. It now turns out that for the past three years the officers of the Pseudo Corp. had been engaging in sham accounting transactions designed to over-state the company’s profits. The emergence of the truth has now seriously damaged the company in the financial markets. A stockholders suit was filed against Taxjacques on the grounds that she negligently failed to discover the financial manipulations during her periodic audits.

You represent the stockholders. A colleague has just in-formed you, albeit somewhat sketchily, of the gist of a recent Federal Appeals Court opinion (by Judge Richard Poser of Law and Economics fame) that holds the stockholders primarily responsible for monitoring the management that they have, after all, themselves selected and perpetuated in office. Fortunately, the opinion in question is not binding precedent in your District, but another judge may be persuaded by similar reasoning. Do you have any good counter argument(s)?

3. PROMISES WITH INDEFINITE CONTENT

PERFORMANCE DEFINITION AND RELATIONAL CONTRACTING

At first glance, it seems as if another difference between contract and tort law is the way in which the obligations of parties are defined. In contracts, the promisee’s obligation is frequently defined in rather precise terms, and the nature of this obligation does not vary with external circumstances such as the cost of performance; there is usually the rough equivalent of what in tort would be strict liability. By contrast, tort law frequently depends on a broad “standard” of behavior. This is exemplified by the negligence rule, which calls for a level of care that depends upon what is in essence a highly circumstantial cost-benefit analysis. On further reflection, however, it is not difficult to cite many tort duties that are defined in rather specific and absolute terms, as well as to understand why these departures from the ordinary rule occur. Less well understood is the phenomenon of contracts wherein the rules are defined in extraordinarily general ways. We therefore now turn our attention to contracts wherein the standard of liability may be expressed in terms of a standard or other variable obligation rather than a fixed performance.

The class of contracts wherein flexible standards of performance are most likely to be found is that of so-called “relational contracts.” This section and the one that follows are adapted from Goetz and Scott, “Principles of Relational Contracts,” 67 Va.L.Rev. 1089 (1981) where the themes presented below are developed in further detail and related to other legal literature.

* * *

Many private contracts deviate significantly from the presuppositions of classical legal analysis. One reason for this deviation is the pivotal role played in conventional legal theory by the concept of the complete contingent contract (or a “fully specified” contract). Parties in a bargaining situation are presumed able, at minimal cost, to allocate explicitly the risks that future contingencies may cause one or the other to regret the agreement. Under these conditions, contract rules serve as standard or common risk allocations that can be varied by the individual agreement of particular parties, thus saving most bargainers the cost of negotiating a tailor-made arrangement. If the basic risk allocation provided by a legal rule fails to suit the purposes of particular parties, then bargainers are free to negotiate an alternative allocation of risks. All relevant risks can thus be optimally assigned—either by legal rule or
through individualized agreement—because future contingencies are not only known and understood at the time the bargain is struck, but can also be addressed by contractual responses that are efficacious in coping with the risks presented by the future contingencies.

In a complex society, parties frequently enter into continuing, highly interactive contractual arrangements where a complete contingent contract may not be feasible. When future contingencies are peculiarly intricate or uncertain, practical difficulties arise that impede the contracting parties from allocating all risks at the time of contracting. Not surprisingly, therefore, parties who enter into such exchange relationships seek specially adapted contractual devices. The resulting “relational contracts” encompass most generic agency relationships, including distributorships, franchises, joint ventures, and employment contracts.

Although a certain ambiguity has always existed, there has been a tendency to equate the term “relational contract” with long-term contractual involvements. The discussion below adopts a very specific construction of the term based more explicitly on the characteristics that distinguish the relational contract from the classical contingent contract. A contract is relational to the extent that the parties are incapable of reducing important terms of the arrangement to well defined obligations. Such definitive obligations may be impractical because of inability to identify uncertain future conditions or because of inability adequately to characterize complex adaptations even when the contingencies themselves can be identified in advance. As the discussion below illustrates, long-term contracts are more likely than short-term agreements to fit this conceptualization, but temporal extension per se is not the defining characteristic. Most actual contracts are, of course, neither perfectly contingent nor entirely relational. Legal theory has merely tended to focus on agreements that fall close to the one polar extreme. The focus in the following analysis is deliberately directed toward the other, relatively neglected, end of the continuum.

A key step in sharing the benefits of contractual exchange involves specifying the performance standard of each party and then selecting a mechanism to ensure compliance with the agreed-upon standard. In conventional contracts, the parties are generally able to reduce performance standards to rather specific obligations. By contrast, relational contracts create unique, interdependent relationships, wherein unknown contingencies or the intricacy of the required responses may prevent the specification of precise performance standards. Complexity and uncertainty each play conceptually distinct roles, although they frequently operate in combination. An example of this is provided if one thinks of attempting to write a contract providing for the care of a fine home garden during a summer out of town. Uncertainty is represented by the difficulty of determining in advance the climatic conditions, incursions of the gypsy moth, wind-borne powdery mildew, etc. Quite another level of difficulty is involved in specifying to the gardener exactly what responses should be made in each case: how much to spend on sprays, whether to water, when a diseased plant should be cut down to prevent infection of adjacent ones.

A typical response to this problem of complexity and uncertainty is to define the performance standard in unusually general terms. Some responses involving generalized performance standards have already been hinted at in earlier chapters: these include the ethical standards of attorneys, brokers and other agents; and the implied fiduciary obligation that attaches to certain special classes of transactions. As the Blood v. Falstaff case below indicates, another common response to the articulation of performance obligations in relational contexts is the “best efforts” clause.

Because these standards are typically described in general terms, it is difficult to apply them in any specific context. Relational contracts therefore require more creative control mechanisms than do conventional contingent contracts. In any cooperative contract where performance obligations remain imprecise, there are inevitable costs in insuring that any particular level of performance is achieved. Parties will bear this cost in a number of ways. They may grant the principal the right to monitor the
agent’s efforts. Performance can thus be controlled by direct supervision or by indirect incentive systems designed to encourage the agent to consider fully the principal’s interests. Alternatively, in cases where monitoring is relatively costly, the agent may seek to reassure the principal by a “bonding” agreement. Liquidated damages provisions, covenants not to compete, and unilateral termination clauses are common examples of agent bonding. Ideally, the parties will select that combination of monitoring and bonding arrangements that optimizes the costs of governing the standard of performance.

Parties will enmesh themselves in the inevitable vagaries of relational contracts only after considering alternative methods of achieving their objectives. One obvious alternative is vertical integration of potentially separable activities, such as manufacturing and distribution, into a single firm. An integrated firm would presumably take all the relevant cost and benefit interactions of the two activities into consideration and would provide the optimal level of manufacturing and distribution inputs such that overall profits are maximized. The vertically integrated firm thus provides a benchmark against which various alternative contractual arrangements can be measured. Vertical integration will be the preferred mechanism if the cost of performing the activity within the firm is less than the costs of accomplishing the same goal by “contracting out” the function. Part of the cost of contracting out a function is, unfortunately, the imperfection of the contractual mechanisms available to the parties. Consider, for example, an industry in which manufacturing and distribution are specialized activities and are, at least potentially, much more efficiently performed by separate firms. Where a single integrated firm is no longer the exclusive decision maker, the respective parties need to engage in a form of explicit agreement in order to ensure that the interactions between the manufacturing and distributing activities will be properly considered. How, then can the parties trade in order to exploit the potential technical efficiencies of separate firms while preserving an optimal level of combined efforts analogous to that achieved by the integrated firm?

An analysis of this question can usefully begin by identifying some conditions that produce variations in the nature and form of contracting. For the sake of concreteness and ease of exposition, the discussion will be carried on in terms of a “manufacturer” and a “distributor.” These are merely convenient labels to describe relative positions of the parties in a production chain. In any particular real world situation the appropriate terminology may vary: e.g., franchiser-franchisee, supplier-fabricator, client-broker.

Consider the common situation where one party is the distributor of a product supplied, at least in part, by the manufacturer. As a simple hypothetical case, assume that adjustments in distribution efforts are the only dimension of production influencing output on the margin and that “distribution efforts” can be regarded as units of product distributed. (Although product volume provides us with a straightforward concretization, the same analysis would apply if efforts were reinterpreted in the form of any other volume or quantitative adjustment, including, for example, quality level, advertising, or any other activity that affects joint profits.) The parties are assumed each to know their own costs (but not necessarily the other party’s costs) and to know the external market conditions during the effective period of the proposed contract.

In Exhibit 5.7, IMC represents the aggregate of the marginal costs of “manufacturing” (MM) and of the marginal “distributing” costs (the vertical distance between IMV and MM) where a firm carries on both manufacturing and distribution as an integrated process. If MR is the marginal revenue curve, then the intersection of MR and IMC at Q represents the output volume that maximizes profits. In other words, Q and its associated profit level is the best that the integrated firm can do by both manufacturing and distributing. The agent presumably enters the picture because he enjoys a productive advantage over the manufacturer, such as, lower marginal distribution costs as exemplified by MD in Exhibit 5.7. The joint marginal cost curve JMC would result if the parties were able to treat their separate costs as a single entity.
The potential savings from separate performance of their respective functions now provide manufacturer and distributor with an economic incentive to enter into a distributorship agreement. Two predictions can be made as to the terms of an agreement designed to exploit that potential. First, as Exhibit 5.7 suggests, the parties will not have exploited all of the potential gains from trade in the situation unless their agreement somehow calls for the manufacture and distribution of quantity \( Q^* \) in Exhibit 5.7, where the sum of the marginal costs to the joint producers equals the marginal revenues from sales. From any output other than \( Q^* \), a movement to \( Q^* \) will increase the combined profits of the parties. If, therefore, there are no special impediments in the form of bargaining or other transaction costs, one would expect to find contract terms facilitating this “joint maximization” quantity outcome.

Second, certain limits can be placed on the minimum and maximum amounts that the distributor will pay to the manufacturer for the predicted \( Q^* \) units of the product. Each party must be at least as well off under the contractual arrangement as it would have been otherwise. Consequently, the manufacturer must receive at least its additional manufacturing costs under the higher-volume distributorship agreement plus the “go-it-alone” profits that it would otherwise earn as an integrated firm. Graphically, these sums correspond, respectively, to area ghij and triangle abc in Exhibit 5.7. The distributor would, in turn, be willing to pay a maximum of the total revenues from sales minus its distribution costs (area kcdm). A
range of indeterminacy exists because the gains from trade (cost saving fa!z+ profits on expanded output zbd) must be divided through bargaining between the parties. These potential net gains from the distributorship arrangement are represented by the shaded area in Exhibit 5.7.

On the bare facts presented above, one would not predict a best efforts term or another flexible performance standard in the contract. There simply is no need for it. Where the optimal output Q * can be predetermined, a fixed quantity term would be a direct and perfectly suitable mechanism for specifying volume. Furthermore, fixed quantity terms and other precisely stated contractual obligations are generally the most efficient instruments for measuring subsequent performance. Thus, under these assumed conditions, the parties will predictably require the manufacturer to provide and the distributor to sell Q * units of production. Suppose, however, that the curves depicted in Exhibit 5.7 are known only probabilistically at the time of contract formation. How is the parties’ behavior altered when they are uncertain as to future conditions?

**Fixed-Quantity Terms.** Even under uncertain conditions there is still a determinate output Q * that optimizes the parties’ contractual relationship based on essentially the same considerations discussed above. Assume, however, that this output must now be calculated with future cost and revenue information known only imperfectly by the parties at the time of contract formation. One approach would be to retain the fixed-quantity term in the contract, specifying the volume that maximizes the expected value of the joint profits based on an ex ante calculation of the risks and their associated probabilities. Such a determinate volume will, however, always turn out to be “wrong” at the time set for performance.

In Exhibit 5.8, for example, a distribution contract is negotiated in the context of but a single contingency—the future imposition of a particular governmental regulation. Assuming that governmental regulation increases costs and that the probabilities of imposition and non-imposition are equal, quantity E will represent the fixed quantity term that maximizes the expected value of the exchange at the time of contracting. Nevertheless, output E will always be inferior to some other output level. If the regulation is imposed, quantity Q 1 will be the optimal output and the actual profits will diverge from the optimal profits by the shaded area A in Exhibit 5.8. On the other hand, if the regulation is not imposed, optimal profits at volume Q 2 will exceed actual profits by the amount represented by area B. The foregone profits constitute “error costs.” Under any conditions of uncertainty, an obligation designed in advance to be optimal on average will tend always be wrong in the particular situation that ultimately pertains. In either case, therefore, the difference between actual and optimal profits is the error cost incurred by the parties from couching their agreement in terms of a fixed contractual quantity.
Exhibit 5.8

Of course, the parties are not required to set a single output term. In Exhibit 5.8, if the uncertainties about costs and revenues are tied to a set of objectively verifiable contingencies that are few in number, the volume called for in the contract may be “keyed” to those uncertain future events. Thus, the parties can agree to quantity $Q_1$ if the regulation is imposed and, in the alternative, specify quantity $Q_2$ in the absence of regulation. However, any increase in the complexity of the risk factors or any greater uncertainty about the future will substantially increase the risk that any fixed volume contract will specify the “wrong” output. Such contingent volume agreements thus represent an intermediate point between the complete contingent contract and more complex forms of relational exchange.

Sequential Contingent Contracts: Responses to Strategic Behavior. One response to increasing levels of complexity or uncertainty is to limit the temporal scope of the contract. By negotiating a series of recurring contingent contracts, the parties can reduce the error costs of specifying fixed obligations. But in many cases the performance of the initial agreement will produce specialized, idiosyncratic skills. Assume, for example, that the parties agree to a one-year distribution contract for a determinate quantity of the manufactured product. Once the distributor learns the skills peculiar to the distribution of the manufactured product in the market area, he enjoys a comparative advantage over the market of unskilled agents. The distributor then has an incentive, when the contract is renegotiated the following year, to secure a larger compensation for similar efforts in order to exploit the advantage gained by his “on the job” training. On the other hand, his newly acquired skills can be fully exploited only if he remains as the
manufacturer’s representative. Thus, as specialization occurs, each party becomes more vulnerable to strategic demands from the other. When the contract is renegotiated, the bargaining stakes are greater and both have incentives to use strategic or opportunistic behavior in order to secure a larger slice of the enhanced contractual “pie.” The essence of the problem here is that, even where perfectly substitutable trading parties are initially available in a competitive market, the increasing specialization of the parties vis-à-vis each other produces a species of bilateral monopoly. Continuance of the original relationship becomes increasingly desirable in order to exploit the accrued specialization advantages, but the division of those gains must be bargained out in a non-competitive environment.

Requirements Contracts. As another alternative, the parties may enter into a requirements contract, thus confronting the problem of strategic behavior by encouraging competition in the market for manufacturers and distributors. The manufacturer is obligated to supply each distributor with the product quantity it “requires” but does not offer any particular distributor the exclusive right to a marketing area. The requirements contract permits marginal adjustments in distribution effects by the distributors since the quantity “required” will vary as conditions change over time. The advantage of this arrangement lies in the use of a competitive distribution market to monitor any individual distributor’s “requirements.” The availability of close market substitutes for any particular distribution contract prevents a distributor from using contractually-based discretion to extort special advantages. In addition, where there are substitutes for the manufacturer’s product, the ability of distributors to switch products may curtail strategic behavior by the manufacturer.

Exclusive Dealing Arrangements. Frequently, however, the profit-maximizing level of distribution activity cannot be achieved in an unrestrainedly competitive market. The conditions that characterize a natural monopoly—substantial fixed costs coupled with relatively modest marginal costs, or restricted market size—occur in many relational contexts, especially when the distribution takes place in spatially limited sub-markets. When classical natural monopoly conditions arise, the economies of scale associated with an activity are not fully exploited unless the number of firms is limited, possibly even to a single distributor. Under these conditions, ordinary requirements contracts are inefficient as well as unstable. The manufacturer will examine the market and conclude, either because economies of scale are large or market size is limited, that a single distributor is the most profitable arrangement. If he licenses more than one distributor, the potential returns from the distribution arrangement will be diminished because competition among distributor will produce an inefficient investment in distribution activities.

As an alternative, the parties may agree to an “exclusive dealing” contract in which the manufacturer will offer the distributor the exclusive rights within a defined market. The exclusivity of these rights creates a relationship of dependence and vulnerability for the manufacturer during the life of the contract. In response, the manufacturer will generally attempt to limit his vulnerability from the exclusive arrangement by securing the agent’s promise to use his “best efforts” to promote sales. The tensions inherent in such a situation are obvious. If a sequential contingent contract is negotiated, the parties once again expose themselves to future strategic demands for increased compensation. Alternatively, if the compensation agreement extends for the life of the relationship, the parties must either specify an “erroneous” fixed-volume term or the manufacturer will be vulnerable to a failure by the distributor to extend those “best efforts” that were paid for in the original compensation agreement.

In the simple model described above, where a manufacturer licenses a single distributor, the exclusive licensing agreement is adequate to induce an optimal investment in distribution activities. More typically, however, the manufacturer will license a number of distributors, granting each a form of access to a limited sub-market. Once a network is established, the parties face a second-order “free rider” problem. Optimal distribution activity will frequently include pre-sales advertising and promotion. Yet, any individual distributor will be reluctant to invest in such services where the prospective customers can learn about the product through one distributor’s promotion efforts and then purchase the product from another distributor. To protect the integrity of the distribution network, therefore, parties to an exclusive
dealings arrangement will typically bargain for territorial restrictions on the market each distributor is entitled to serve. These restrictions serve to prevent a distributor from “poaching” on either the customers or the distribution efforts of other distributors in the network.

**Indirect Monitoring Mechanisms: Price Adjustments.** As we have suggested above, contractually fixed quantity terms are not always effective means of achieving joint profit maximization in an uncertain environment; some system of monitoring the level of effort undertaken by the distributor or agent is needed. Direct monitoring is feasible in contexts such as franchising where operating procedures are standardized and one agent’s efforts can be compared to those of similar agents. Indeed, franchisors may retain vertically integrated distributional outlets in representative locations in order to provide benchmarks against which franchisee operations can be compared. Such direct forms of monitoring may not always be practical. In such circumstances, joint profit maximization will be better achieved by indirect means.

One indirect instrument for adapting volume to uncertain future conditions is the proper calibration of the marginal price between the manufacturer and the distributor. Since the distributor has de facto control of the maximum volume sold by the parties, the objective should be to set pricing terms that will induce the distributor always to choose a close approximation of the optimal quantity \( Q^* \) even when he is nominally free to choose any volume level at all. Marginal cost to the distributor is the sum of his own marginal distribution costs plus the marginal contractual price payable to the manufacturer. In practice, this “price” will typically be termed a royalty, license fee, franchise fee, or some other term appropriate to the business context. Since the distributor will choose the output that equates marginal cost with marginal revenue \( (MC = MR) \) in order to maximize his own profits, a proper calibration of price by the manufacturer can induce the \( Q^* \) optimal output to be chosen. The optimal output will be chosen voluntarily by the distributor only if his marginal costs can be made to approximate the aggregate marginal costs of distribution and manufacturing \( (MD + MM) \), i.e., what we have above termed the joint marginal costs.

One useful scheme is for the contract to call for sales to the distributor at marginal manufacturing costs \( (MM) \), but pay over to the manufacturer some share \( (K) \) of the distributor’s net profits in the form of a franchise fee, license, or royalty. Marginal cost to the distributor then becomes

\[
MC = MD + MM + K [MR - (MD +MM)]
\]

and automatically leads to the selection of \( Q^* \). This result will obtain because the distributor’s own profit maximization calculus will lead him to expand volume until

\[
(i-K) [MR - (MD + MM)] = 0
\]

where the term in square brackets is the increment in joint profits and \( (I-K) \) is the fractional share of those joint profits that the distributor gets to keep. Indeed, \( (I-K) \) may be reinterpreted as a fractional “commission” payment paid over to the distributor or agent out of the proceeds of any sales. The above formulation directly reflects the common sense purpose of this pricing scheme, because it shows how the distributor cannot maximize his own profits without also fully exploiting any opportunities for joint profits as represented by the square-bracketed term. In addition, a simple algebraic manipulation demonstrates that the formulation implicitly yields a result wherein \( MR = MD + MM \), which is the underlying general condition for joint profit maximization: marginal revenue equals marginal cost where both costs and revenues are summed over all affected parties.

Unfortunately, there are numerous practical drawbacks to the implementation of a profit-sharing arrangement such as that just described. One limitation is that the pricing scheme requires information on marginal costs of the manufacturer. Since accurate incremental cost information is much harder to acquire than total cost information, the parties may not wish to tie the contract to magnitudes that are costly to
ascertain or otherwise impractical to monitor. This limitation is less important when marginal manufacturing costs are either insignificant in magnitude or closely approximated by average cost, a more readily ascertainable value. For instance, marginal “manufacturing” costs are frequently negligible when the “manufacturer” is providing only a license rather than some actual product.

Even though the distributor’s costs and revenue need—at least in theory—to be measured only in totals in order to arrive at the profits figure upon which the contractual payments are based, the distributor’s treatment of costs may create problems. The divergence between accounting costs and economic costs may confuse profit measurement. One example of this is the exclusion of “fixed” costs as economic costs. Perhaps most important is the problem of monitoring and segregating costs and revenues within multi-product firms where it becomes increasingly difficult to monitor the allocation of joint costs—overhead, etc.—which may be shifted strategically among various contracts. A multi-product distributor has an incentive to impute as many costs as possible to those activities where his share of net profits is lowest.

Similarly, the distributor may “pad” costs, disguising his own returns and denying profits to the manufacturer. For instance, the distributor may provide himself with unnecessary amenities such as a “company car” or place relatives on the payroll at unwarranted salaries. Gross abuses will, of course, be easy to detect, but cost padding can take a variety of subtle forms that are extremely difficult to discover and perhaps impractical to prove.

Given the increased monitoring burdens of profit-based pricing, it is not surprising that relational contractors frequently choose alternative pricing arrangements. Commonly observed, for example, are initial flat fee payments coupled with royalty payments tied to a percentage of gross sales. These arrangements reduce the costs of monitoring the distributor’s expenditures but, in turn, also reduce the congruence of interests concerning optimal output between the manufacturer and the distributor. Once the distributor’s return is reflected in a sales-based pricing mechanism, there is an inherent conflict of interests between the contracting parties over profit-maximizing output: the distributor’s self-interest will induce him to produce a lesser quantity than the manufacturer’s interests would demand. Thus, these more common pricing arrangements reduce the burden of monitoring the agent’s costs and revenues but increase the burden of controlling the agent’s level of effort. Presumably, firms choose sales-based pricing because the cost of monitoring the level of effort is perceived as less than the cost of monitoring operating expenditures.

**BEST EFFORTS PROVISIONS**

The previous section on the general problems of relational contracting provides a background to consideration of what is perhaps the most poorly understood class of relational contracts: that involving agreements wherein one party explicitly, or even implicitly, undertakes the contractual duty of using its “best efforts” to carry on an activity beneficial to the other. Notwithstanding the frequency with which such terms are observed empirically, the precise legal meaning to be attached to a best efforts requirement is not at all clear, either from a consideration of the case law or from theoretical discussions in standard legal scholarship. Nevertheless, there appears to be a relatively straightforward and persuasive definition that emerges from the preceding economic conceptualization of the problem faced by two parties who are attempting to set a contractual volume in which they have joint interests.

Best efforts cases hinge on two factors: strategic adaptation to the conflict of interest between the parties and the problem of managerial incompetence. These elements may, of course, coexist in a single case. Since the courts appear not to have sharply distinguished the factors, it is not surprising that existing case law has been unhelpful in working out a consensus about the meaning of the legal rule.

Exhibit 5.9 illustrates a contract in which the distributor faces a marginal cost curve MC composed of his own marginal distribution costs MD plus the marginal “price” R negotiated with the manufacturer. In addition, it is assumed that the payment takes the form of a 50% royalty on gross sales. Absent any other
information, one might expect that the distributor would then be legally entitled to choose volume Q where his marginal costs ($MC = MD + R$) are equal to marginal revenues ($MR$). This is the point at which the distributor’s own profits are maximized. If he were required to sell an additional quantity beyond Q, the distributor’s profits would be reduced, as exemplified by the shaded triangle A in Exhibit 5.9.

Suppose, however, there were some way in which he could oblige himself to adjust to the joint maximization output that we have previously identified at $Q^*$. In exchange for such an undertaking, which at the time it is accepted represents a loss to the distributor, the manufacturer should be willing to agree in advance to a compensatory contractual concession through which the two parties can split the additional profits generated by the higher volume. These profits are represented by the cross-hatched triangle B in Exhibit 5.9.

*Exhibit 5.9*

*Optimal Volume Definition.* The obligation to produce at the joint maximization volume is the meaning that is proposed here for the best efforts term in commercial contracts. This interpretation of the best efforts provision has a great deal of theoretical attractiveness since, absent the specification of an alternative construction by the parties, it directs the outcome that maximizes the net gains that parties could achieve from their contractual relationship. It is a plausible means of identifying a goal presumably desired by most parties, albeit not always well articulated.
In addition, the definition suggested above is consistent with a “fairness” obligation as might be formulated by distributive justice theorists. The distributor is required, under this conception, to treat the manufacturer “fairly,” giving the manufacturer’s interests (profits) equal weight with his own when output decisions are made. Moreover, such special consideration has presumably been anticipatorily paid for by the manufacturer in the form of some compensatory concession.

Unfortunately, a best efforts obligation, as defined above, inherently implies a serious monitoring problem. This is illustrated in Exhibit 5.9 where shaded triangle A represents the reduction in profits suffered by the distributor because he is obligated to produce at Q* rather than Q. Hence, the best efforts promisor will generally have a strong economic incentive to “chisel” on the obligation. In a world of cost-free information, such breaches of the best efforts requirement would be easily detected, and the behavior restrained through the legal damages imposed for breach of contract. In a real-world situation, however, the requisite information for proof of liability or quantification of damages may be chillingly costly to obtain, especially when the plaintiff bears the burden of such proof. Hence the standard legal mechanism may not be a viable one for enforcement of this type of contract provision.

Where recourse to the courts is not an attractive option, these economic considerations suggest that a best efforts promisee such as the manufacturer-will attempt to contract for other means of controlling the standard of performance. Presumably, the self interest of both contracting parties will induce them to seek out that combination of monitoring or bonding arrangements that represents the optimal trade off between expected costs of contractual governance and profits foregone because the ideal output Q* is not perfectly enforced. As the cost of contract-specific monitoring strategies increases, the price of contracting to the best efforts promisor similarly increases. The best efforts promisor has an incentive, therefore, to propose cost-effective bonding agreements which, by reducing the costs of contractual control, reduce the contract “price” paid by him. The “price” reductions might take the form of a reduction in the initial license payment required by the manufacturer, or a reduction in the royalty paid to the manufacturer on the contractual product or, indeed, any other adjustment in contract terms favorable to the distributor.

A commonly observed form of bonding is a termination privilege that could be invoked by the manufacturer if he detects a breach of the best efforts obligation. Moreover, the parties might be expected to negotiate a termination clause that granted the manufacturer considerable discretion as to the circumstances under which termination would be permitted. If, instead, the termination clause were only triggered by specific events, any attempt to exercise it might create precisely those problems of proof that the clause was originally designed to circumvent. A limited right of termination, by reducing the value to the manufacturer of the reassurance of contractual performance, would presumably induce some compensatory increase in the contract “price” paid by the distributor. A discretionary termination clause is not an ideal safeguard. Considering the available alternatives, however, it frequently permits a mutually beneficial adaptation to the inevitable conflict of interest generated by a best efforts agreement.

While the best efforts result Q* is in theory a clearly optimal result for the parties, the realities of enforcement, especially coupled with the inherent chiseling incentive, may dim the practical attractiveness of such agreements. On the other hand, the problems arising in legal regulation of such agreements should not be viewed as dispositive. Many contractual provisions are honored even where there is no effective legal sanction for their breach. In some circumstances this is due to the existence of informal, extra-legal sanctions, including a sense of commercial ethics. Notwithstanding practical difficulties of securing legal enforcement, therefore, a contractual provision also has value as a mere communication of understanding between the parties as to their mutual rights and duties. Hence, the inclusion of a best efforts term may, at a minimum, serve as a signal alerting good faith bargainers that the proposed contractual relationship is one in which special concerns are to be considered.

Where courts are compelled to attach a meaning to otherwise ambiguous terms of agreements, it makes sense to look to the likely intent of the parties or the goal the parties might reasonably be deemed
to have sought. The “joint-maximization” definition of best efforts is, we argue, the single most plausible interpretation of the underlying economic motivations involved. This proposed meaning of best efforts need not be seen as describing in any particular case what the contracting parties actually intended. Rather, it is designed to offer a plausible way in which the legal rule can allocate unknown risks in advance of individual bargaining, thus reducing the uncertainty costs of an imprecise legal standard.

Best Efforts as Diligence Insurance. Although the joint-maximization output interpretation may be the single most persuasive meaning for best efforts, one alternative plausible definition merits discussion: best efforts requires the exercise of due diligence and ordinary business prudence. Under this alternative conception, a breach of the best efforts obligation would exist where the distributor’s efforts diverge from the standard of diligent or reasonably prudent business conduct.

In terms of Exhibit 5.9, our previous analysis has focused on the losses from a shortfall between the joint maximization volume $Q^*$ and the distributor’s profit-maximizing output $Q$. However, it is interesting to consider in addition the case where the distributor chooses an even lower output than $Q$, for example $Q$, in Exhibit 5.9.

Why would the distributor ever choose an output of less than $Q$ if, under the cost and revenue conditions for the product he sells, the quantity $Q$ represents the point of maximum profitability? One of the simplest explanations is to reinterpret Exhibit 5.9 as containing “objective” cost and revenue curves as they would exist for a typical distributor. The cost and revenue functions underlying the actual output calculus of any particular distributor may diverge greatly from the objective, either because of truly subjective elements such as misapprehension of the market or due to more concrete factors such as carelessness or incompetence in restraining production costs. The distributor, either through misapprehension or incompetence in restraining costs, perceives marginal cost to be $MC'$ and selects volume $Q_1$. Whatever the reason for the distributor’s failure to serve even his own interests competently, the manufacturer will understandably be distressed if the original agreement was predicated on his perfectly reasonable expectation that the objective circumstances experienced “should” have motivated a volume of at least $Q$.

For at least two distinct reasons, it makes good economic sense that the distributor would be the efficient bearer of the risk of both his own and the manufacturer’s lost profits from sales foregone due to business mistakes on the part of the distributor. First, the distributor is the party who has effective control of his own level of care invested in undertaking business activities and, hence, the opportunity to adjust that level of care to the cost-effective extent. Second, the distributor is in a better position to assess ex ante his own capability to achieve the ordinary or expectable level of business performance. Consequently, inclusion of the best efforts term might be construed as an explicit allocation of all risks from the possible business blunders of the party who promises his “best” efforts.

For expository ease, we shall refer to this interpretation of best efforts as the “diligence insurance” definition. Such a conception of the meaning of best efforts is not an unattractive one and provides at least a minimal standard for the term. One can argue persuasively, however, that the diligence standard is properly subsumed within the joint-maximization output definition. This result is suggested when one attempts to give rigorous content to the question of precisely how much diligence is required. The response is that the obligation is to use a cost-effective amount, to, in lay terms, “do the prudent thing, taking into account the interests of both parties” in a manner similar to that mandated by the negligence standard in tort law. But this is only another way of expressing the original fundamental insight that the parties can mutually benefit when the distributor acts as a joint maximizer.

Once it is granted that the parties are motivated by a concern to maximize the joint contractual product, it seems odd to restrict that type of reasoning to a single aspect of the business relationship. One can
be diligent and produce an “erroneous” or nonoptimal output such as the distributor’s profit-maximizing quantity $Q$. On the other hand, the obligation to produce the jointly optimal output is easily understood as an all-encompassing optimality condition, directly mandating the “correct” result in objective terms. Since the two interpretations spring from essentially the same underlying principles, the one with greater generality seems preferable.

**HYPOTHETICAL: INTERWOOD v. FORREST LUMBER**

In July 1969, Interwood Products and Forrest Lumber Co. entered into a joint venture, the Forinter Corporation. As its name suggests, Interwood Products is a large, vertically integrated wood products company with interests extending from raw lumber production and wholesale wood supply to the fabrication of finished wood products such as office furniture, parquet flooring, etc. Forrest Lumber Co. has traditionally operated in the “wood farming” business. It plants and cultivates forests suitable for commercial wood production. Forrest also is well known for engaging in research and development of highly innovative techniques of wood production, ranging from experimental tree varieties to new methods of planting and cultivation.

The Forinter joint venture was conceived of as harnessing Forrest’s greater expertise in the raw product end of the business with Interwood’s marketing ability and superior access to business capital. Forrest has filed suit against Interwood for losses arising from alleged abuse of the fiduciary relationship created by their joint venture. The existence of a fiduciary relationship between joint venturers is widely acknowledged:

“The relationship between joint venturers, like that existing between partners is fiduciary in character and imposes on all the participants the obligation of loyalty to the joint concern and of the utmost good faith, fairness and honesty in their dealings with each other with respect to matters pertaining to the enterprise.” 46 Am.Jur.2d 69.

Forest cites three specific examples of conduct by which Underweight is thought to have breached its fiduciary duty.

1. In 1970, Forinter agreed to a supply contract with Interwood whereby Forinter undertook to furnish the raw wood components for a future delivery contract between Interwood and Maxim Construction Co. Under this contract, Maxim Construction Co. acquired an option on a large quantity of oak paneling to be delivered by Interwood in 1973. The agreed price to Forinter under this contract was $2,220,000. Forinter delivered the required raw oak to Interwood’s fabrication division in early 1973, although the current market price of the wood was at that time $3,020,000. It has subsequently been conceded, however, that Interwood waived a 6-month prepayment provision in its contract with Maxim, due to Maxim’s “tight cash position such that it would otherwise have been unable to exercise its option on the paneling.” Forrest contends that Interwood should not have waived the prepayment provision.

2. Interwood outbid Forinter on a valuable forestland property, the Brush Mountain Tract. Forrest asserts that the value of the Brush Mountain Tract was recognized by Interwood only because of information and expertise developed through its Forinter joint venture.

3. Interwood outbid Forinter on another valuable tract, the Cleve Valley Property. Interwood’s willingness to outbid Forinter was based on its own large new orders for finished wood products, which it knew would tend to drive up the value of all wood resources. Forrest asserts that this special knowledge should have been shared with Forinter, thus placing Forinter in an equal bidding position.

Consider yourself to be a member of a state appellate court panel. There are no directly pertinent precedents in your state. How would you decide the case? In preparing your opinion, you would presumably have to cover each of the following points, even if the considerations in (b) below did not appear explicitly in the opinion. (You may wish to reflect on whether or not they should appear.)
QUESTIONS

(a) Which, if any, of the examples of conduct by Interwood should be actionable? What are the underlying reasons for your conclusions with respect to each charge? (If you need to assume additional facts, do so, but develop a clear rationale as to why the additional facts are relevant and necessary.) [Hint: In the author’s opinion, one of the charges should fairly plainly be actionable, one should not, and another is uncertain.]

(b) What are the implications for future contractual relations of the precedents that you would create? Specifically, what would be the likely alternative effects of a decision either restricting or expanding the scope of the duty owed to a joint venturer? How would you predict future contracting parties would adjust their behavior to any specific resolution of the legal issues confronted in this case?

BLOOR v. FALSTAFF BREWING CO.
454 F.Supp. 258

[The following is perhaps the leading recent case dealing with an important term of contractual liability, the obligation to use “best efforts.” Before reading the case, try to define rigorously your own understanding of that common legal term of art.]

BRIEANT, District Judge.

This action was filed July 21, 1976 to recover monetary damages for breach of contract. Plaintiff James Bloor is the Reorganization Trustee of Balco Properties Corporation, formerly named P. Ballantine & Sons (“Ballantine”). Defendant is the Falstaff Brewing Corporation (“Falstaff”), which on March 31, 1972 bought from Investors Funding Corporation (“IFC”) the Ballantine brewing labels, trademarks, accounts receivable, distribution systems and other property, excepting only the Ballantine brewery.

The purchase agreement called for an immediate payment to Ballantine of $4,000,000 and royalty payments thereafter of $.50 to be paid on each barrel (31 gallons) of the Ballantine brands sold between April 1, 1972 and March 31, 1978. The contract contained a liquidated damages clause, calling for payments of $1,100,000 a year which were to be made in the event Falstaff “substantially discontinue[d] the distribution of beer under the brand name ‘Ballantine’.” The contract also required that Falstaff “use its best efforts to promote and maintain a high volume of sales” of the Ballantine brands.

This action arises out of defendant’s alleged breach of these covenants by its substantial discontinuance of the Ballantine brands or its failure to use best efforts in their promotion, and also by its failure to pay any royalties whatsoever on sales of those brands after December 1975 and by its alleged underpayment of royalties before that date.

* * *

By contract dated March 3, 1972, Falstaff acquired Ballantine’s assets. The closing took place on March 31, 1972. At that time Falstaff was the fifth ranking brewer in the United States but had failed to acquire a substantial foothold in New York, New Jersey and Pennsylvania, the highest beer consuming region in the country. Its purchase of Ballantine was apparently prompted by the desire to acquire a ready-made distribution system in the New York area, and by its need to utilize the excess capacity of its own breweries. Falstaff did not buy the Ballantine brewery in Newark, New Jersey. At some undetermined date during its ownership of Ballantine, Falstaff began using its own beer, without formula alteration, to fill Ballantine Beer containers. At the present time, the only difference between Ballantine and Falstaff Beer is the label. Falstaff was also motivated in its purchase by the fact that Ballantine Beer
was (and is) a “price” beer in the Northeast and would in the normal course of events not compete with Falstaff’s own so-called “premium” beer.

Shortly after acquiring the Ballantine assets, Falstaff moved the “retail” distribution operation from Newark, New Jersey (the location of the Ballantine brewery) to North Bergen, New Jersey, where it continued until 1975 to service generally the same accounts. Between 1972 and 1975, Falstaff also continued the former policy of substantial advertising of Ballantine products, spending more than $1 Million a year for that purpose. Falstaff continued in every way the former pricing policies used by IFC, including substantial “post-offs” from listed prices. During this period, Falstaff claims to have lost $22 Million in its Ballantine brands operations.

A very significant change in the modus operandi of Falstaff-Ballantine took place in 1975. In March of that year, Mr. Paul Kalmanovitz, an entrepreneur with 40 years experience in the brewing industry who had owned a small position in Falstaff stock prior to that time, advanced $3 million to Falstaff to enable it to meet its payroll and other pressing debts then due. On March 10, 1975, in return for some $10 million cash advanced, and additional loan guaranties, Mr. Kalmanovitz was issued new convertible preferred shares in amounts equal to about 35% of the company’s outstanding shares. A voting trust arrangement was made to give him control of the Board of Directors. The shareholders of Falstaff approved the agreement on April 28, 1975. At present, Mr. Kalmanovitz is Chairman of the Board of Falstaff.

* * *

Since Mr. Kalmanovitz’s assumption of control of Falstaff the advertising budget for Ballantine products has decreased from about $1 Million a year to a point near non-existence (about $115,000 since the beginning of 1976). Substantial cuts have also been made in sales and management personnel. In late 1975 four of the six “retail” distribution centers, including the North Bergen depot, were substantially closed or phased out. The North Bergen depot was replaced by two independent distributors who together service substantially fewer accounts. In mid-1976 Mr. Kalmanovitz also discontinued the price cutting policies of former management, ordering that no beer was to be “given away” in the future.

Concomitant with these changes, and plaintiff argues, causally related to them, there has been a precipitous decline in the sales of Ballantine products, and a slightly less precipitous diminution in the sale of Falstaff products. In December 1975, Falstaff unilaterally discontinued royalty payments on sales of Ballantine products.

Plaintiffs Claims Pleaded

Plaintiff is suing to recover money damages for Falstaff’s breach of contract in three separate instances: (1) Falstaff’s “substantial discontinuance” of distribution of Ballantine products, or its failure to use best efforts to keep sales of them high; (2) Falstaff’s underpayment of royalties prior to December 1975; and (3) Falstaff’s discontinuance of royalty payments on Ballantine products sold after December 1975. We consider these claims separately.

[After trial, plaintiff abandoned its claims for substantial discontinuance and lack of best efforts for the period prior to May 1975, the date Mr. Kalmanovitz assumed operating control of Falstaff.]

I find that plaintiff has failed to prove “substantial discontinuance” under the contract. Falstaff now and at all times has distributed beer under the Ballantine name to all who would purchase it at a price above Falstaff’s costs. Whether the totality of its merchandising efforts complies with the obligation to use “best efforts” presents a more difficult question.

Ballantine has, however, proved its claim that Falstaff failed to use its best efforts to promote sales of Ballantine products after May 1975.

The parties differ with respect to the meaning and effect of “best efforts”. Defendant has urged on the Court the argument that the interpretation of such a clause requires the application of a subjective
standard: Falstaff contracted to use “its” best efforts, and any determination of those efforts must include considerations of Falstaff’s own allegedly precarious financial position. Plaintiff, on the contrary, cites substantial precedent holding that financial difficulty and economic hardship do not excuse performance of a contract, and argues for the application of an objective standard, that of the “average, prudent comparable” brewer. Arnold Productions, Inc. v. Favorite Films Corp., 176 F.Supp. 862, 866 (S.D.N.Y.1959), aff’d, 298 F.2d 540 (2d Cir. 1962).

The point of the argument appears to be defendant’s contention that it promoted Ballantine to the full extent that its own straitened financial abilities permitted, and could do no more.

“Best efforts” is a term “which necessarily takes its meaning from the circumstances.” Perma Research & Development v. Singer Co., 308 F.Supp. 743, 748 (S.D.N.Y.1970). In a patent assignment case, for example, in which the court found an implied obligation to use “best efforts,” Judge Duffy examined the assignee’s research capability and net earnings, and stated: “To properly evaluate Singer’s performance under the patent assignment it is necessary to look first at Singer’s capability.” Id., 402 F.Supp. 881, 890 (S.D.N.Y.1975), aff’d, 542 F.2d 111 (2d Cir.), cert. denied, 429 U.S. 987, 97 S.Ct. 507, 50 L.Ed.2d 598 (1976). “Capability,” however, is a far broader term than “financial ability.” It includes the marketing expertise and experience attributable to the “average, prudent, comparable” brewer. It is obvious that any determination of the best efforts achievable by Falstaff must take into account Falstaff’s abilities and the opportunities which it created or faced; Falstaff did not contract to promote Ballantine at the level or to the degree that Anheuser-Busch or Schlitz might have done. It did, however, contract to merchandise Ballantine products in good faith and to the extent of its own total capabilities, and this it failed to do.

The evidence in the case also shows that Falstaff’s financial position during most of the period in litigation was far from that depicted by defendant, and that Falstaff failed to use even its own temporarily circumscribed abilities and resources to promote the sale of Ballantine products.

As a consequence of this precipitous departure from the New York City market area, Ballantine registered no sales there for the month of September 1975. In October and November 1975, Falstaff, on the initiative of several independent distributors, selected Thomas Fatato, Inc. to distribute them in New Jersey. The performance of these two has been less than impressive, and, indeed, at the time of their appointment, strong concern about their ability to cover the metropolitan area adequately was expressed by Mr. Griesedieck, President of Falstaff. Fatato, for example, services only about 1,500 to 2,000 of the approximately 15,000 to 17,000 retail accounts formerly serviced from North Bergen. In the first month of his operation (October 1975), sales dropped to 2,500 barrels from some 7,000 barrels sold in August. There was convincing testimony from defendant’s own employees that Mr. Fatato’s ownership of a competing brand, Milwaukee Beer, on which he earns both the distributor’s and manufacturer’s profit would naturally result in his giving preference to his own brand in the competition for shelf space, rather than aggressively marketing Ballantine. Since Fatato has been distributing Ballantine in the New York City area, sales of Ballantine draught beer have declined more than 90 percent. Despite all this, Mr. Kalmanovitz testified that Fatato’s performance as a distributor was better than expected.

Under its union contract with its employees at the North Bergen depot, Falstaff considered itself legally bound not to appoint an official distributor for the tristate area to whom Falstaff would make shipments from its brewery in Cranston, Rhode Island. Both Fatato and Piccirrillo bought Ballantine Beer in Rhode Island and transported it themselves into the New York City area. This union contract would have expired in May 1976, but Falstaff’s failure to terminate it with proper notice resulted in an arbitration award extending the contract until May 1977. Falstaff urges the existence of this contract as the excuse for its failure to take advantage of a very significant opportunity in the fall of 1975 which would have cost it nothing to exploit, but Ballantine may not be charged with Falstaff’s negligence in failing to terminate its labor contracts properly.
In September or October 1975, Mr. John Molyneux, Vice President and General Manager of the Guinness-Harp Corporation, an importer of foreign beers, and distributor of various American brands including Schlitz, Schmidt’s, Pabst and Piels, journeyed to San Francisco to discuss with Mr. Kalmanovitz the possibility of undertaking the distribution of Ballantine products. Mr. Molyneux testified that his company would be able to handle all of the Ballantine products (except Munich Beer) at the prices quoted by Mr. Kalmanovitz and would be able to distribute them profitably through the Metrobeer Division of Guinness-Harp, to about 13,000 of the 15,000 retail licensees in New York City. On October 20, 1975, however, Falstaff wrote to Mr. Molyneux that even that offer was withdrawn because Falstaff wanted to give Mr. Fatato a chance in New York and did not want to “invite competition” by other wholesale operations.

I find that Guinness-Harp presented Falstaff with an opportunity to promote Ballantine products that an “average, prudent, comparable brewer” situated as was Falstaff reasonably would and should have explored and sought to exploit. Existence of the union contract was not an obstacle to such an arrangement. There is no credible evidence that Mr. Molyneux demanded an exclusive “appointment” in the New York City area. In any event, after May 1976, any inability to appoint such an exclusive distributor in the New York area was caused by the fault or negligence of Falstaff. To the extent such fault or negligence prevented it from using effective marketing methods in the area which Molyneux proposed to serve, it is answerable in damages to Ballantine.

Defendant argues that it had the right, and even the duty to use in good faith its best business judgement in all matters arising under the contract, and also had the right to look to its own interest. This argument has some validity, but Falstaff has gone beyond that point. As the New York Court of Appeals said in a similar situation: “Although a publisher has a general right to act on its own interests in a way that may incidentally lessen an author’s royalties, there may be a point where that activity is so manifestly harmful to the author, and must have been seen by the publisher so to be harmful, as to justify the court in saying there was a breach of the covenant to promote the author’s works.” Van Valkenburgh v. Hayden Publishing Co., 30 N.Y.2d 34, 46, 330 N.Y.S. 329, 334, 281 N.E.2d 142, 145 (1972).

Some of this apparent callousness towards Ballantine sales is undoubtedly caused by the fact that even though the liquid in a can of Ballantine Beer and in a can of Falstaff Beer is identical, and accordingly costs exactly the same amount to produce, sale of Falstaff Beer produces a greater profit for Falstaff. In part this is the result of the fact that Falstaff is a “premium” beer and nets Falstaff about $4.20 more a barrel than does Ballantine, even before the $.50 Ballantine royalty is subtracted from the latter. Also, Mr. Kalmanovitz testified, the process of changing a production line over from Falstaff to Ballantine Beer “cost[s] a lot of money.”

In addition to Falstaff’s refusal to entertain Mr. Molyneux’s proposals despite the apparent inadequacy of the Fatato-Piccirillo distributorships, Falstaff has also failed in several other notable respects to promote Ballantine products with its best efforts. After Mr. Kalmanovitz’s assumption of control of Falstaff, the company severely cut back personnel in distribution, sales, marketing, administrative and warehousing areas. It virtually eliminated its promotion and advertising of Ballantine Beer and closed its advertising department, all this despite the repeated assertions by experts for both sides (for example, Mr. Weinstein for plaintiff and Mr. Dependahl, former Vice-President of Falstaff, for defendant) that personal contact, merchandising and advertising were essential to the marketing of any beer. While some cutbacks may have been temporarily justified in late 1975, the situation was quite different in 1976 and 1977 when Falstaff continued its stated policy of simply making the beer available, and when solicited by distributors, funding one-half of “co-op” promotional costs. After Mr. Kalmanovitz assumed control of Falstaff, the company even discontinued the establishing of goals for its remaining salesmen, an essential
step in any marketing effort, and one which would have cost the company nothing to implement. This is not “best efforts.”

Defendant has stressed the argument that Falstaff’s policies after 1975 were evenhanded, that is, that Falstaff’s cutbacks and marketing strategy affected Falstaff and Ballantine performance equally. Even if this were the case, which it is not, Falstaff’s relationship to Ballantine is essentially different from its relationship to its own products. In the latter case, it may promote, continue or discontinue its products as it wills subject to its duty to shareholders; in the former case it is bound by a contractual duty to the promisee. As the court said in a case cited by the defendant here:

“‘[B]est energies’ meant such effort as in the exercise of sound judgment would be likely to produce the most profitable results to the promisee in view of the nature of the business and the extent of the territory over which it was to be conducted.” Randall v. Peerless Motor Car Co., 212 Mass. 352, 374, 99 N.E. 221, 226 (1912).

Additionally, Falstaff has not treated both products equally. Robert Thibaut, a Falstaff Vice-President, testified that Falstaff but no Ballantine had been advertised extensively in Texas and Missouri. In the same areas Falstaff, although a “premium” beer, was sold for extended periods below the price of Ballantine. These actions and failures to act were not consistent with those of the average, prudent, comparable brewer in the same circumstances seeking to promote a beer to the extent of his best efforts. The “business judgement” argument is of no help to Falstaff in this regard.

* * *

Accordingly, I find that from September 1975 until the present, Falstaff has breached the covenant in its agreement with plaintiff’s predecessor, in that it failed without justification to use its best efforts to promote the sale of Ballantine products and has neglected to act in the manner required of the average, prudent, comparable brewer in marketing his product.

* * *

**Damages**

Falstaff has breached its contract with Ballantine to use its best efforts to promote sales of Ballantine products and thus assure fair royalty payments to Ballantine. The period of this breach extends from September 1975 when Mr. Kalmanovitz’s policies began to take effect, through the date of trial, and anticipatorily, through March 31, 1978, the end of the royalty period. Since December 1975, moreover, Falstaff has, as a set-off against its claims against Ballantine, withheld royalty payments admittedly due on sales of Ballantine products. In the light of our discussion of defendant’s counterclaims, these setoffs were not justified. Finally, it is undisputed that royalties to Ballantine for the period from April 1, 1972 through November 30, 1975 were underpaid in the amount of $41,399.50. During that period Falstaff sold 5,108,908 barrels of Ballantine brands, for which the royalty payments should have been $2,554,454.00, rather than the $2,513,054.50 actually paid.

Damages for Breach of Contract

“In simple terms, the measure of the damage [for failure to use best efforts] is the amount necessary to put the injured party in the exact position he would have been if the contract had not been breached.” Perma Research & Development v. Singer Co., 542 F.2d 111, 116 (2d Cir.), cert. denied, 429 U.S. 987, 97 S.Ct. 507, 50 L.Ed.2d 598 (1976). In this case the position Ballantine would have enjoyed but for the breach is established by reference to the sales achieved by “comparable” brewers between mid-1975 and the present. This method of computing damages has often been employed by the courts. In Autowest, Inc. v. Peugeot, Inc., 434 F.2d 556 (2d Cir. 1970), which involved a claim of wrongful termination of a franchise to sell Peugeot automobiles, plaintiff proved his damages by, among other
methods, comparison of the sales of Volvo automobiles, which had increased from 5,000 to 10,000 between 1965 and 1968, with sales of Peugeot automobiles which only rose from 700 to 900 during the same period. Contemporary Music, Inc. v. Famous Music Corp., 557 F.2d 918 (2d Cir. 1977), was an action to recover damages for defendant’s failure to use reasonable efforts to promote Contemporary’s records, specifically the song “Fear No Evil”, which had reached No. 61 on the Hit Parade at the time it was wrongfully withdrawn by defendant. After stating the general rule that a plaintiff who has proved injury “need only show a ‘stable foundation for a reasonable estimate of royalties he would have earned had defendant not breached,’” id. at 926, quoting Freund v. Washington Square Press, Inc., 34 N.Y.2d 379, 383, 357 N.Y.S.2d 857, 314 N.E.2d 419 (1974), the Court approved a statistical method of comparison which tended to show that 51% of songs reaching No. 61 on the Hit Parade went on to reach No. 20 and consequently to achieve greater sales. Finally, in Bigelow v. RKO Radio Pictures, Inc., 327 U.S. 251, 260, 66 S.Ct. 574, 90 L.Ed. 652 (1946), the Supreme Court approved a method of computing damages based on a comparison of the profits of plaintiff’s theater, which had been deprived of first-run motion pictures by defendant, with the profits of a theater of comparable size which had access to such films.

[Here the court engaged in a lengthy discussion of the evidence presented by Messrs. Weinberg and Horowitz, expert witnesses for plaintiffs and defendants, respectively, as to the performance of “comparable” brands during the contract period. Both experts considered Schaefer and Rheingold to be products with market opportunities comparable to those of Ballantine. The court therefore found that the experiences of these two brands reveal what Ballantine’s performance would have been had it not been for Falstaff’s breach. Specifically, the court found that Ballantine’s sales in each of the years 1975-1977 should have been the same percentage of base-year 1974 sales as experienced by the “comparable” brands. Actual Ballantine sales were then compared to this benchmark level for each year and the shortfall used to compute the lost royalties.]

Accordingly, the damages for Falstaff’s breach amount to $628,528.10. These damages are reasonable under all the circumstances presented here. As our Court of Appeals stated in Contemporary Mission, Inc. v. Famous Music Corp., 557 F.2d 918, 926-27 (2d Cir. 1977):

“[T]he burden of uncertainty as to the amount of damages is upon the wrongdoer, Perma Research & Dev. v. Singer, supra, 542 F.2d at 116, and the test for admissibility of evidence concerning prospective damages is whether the evidence has any tendency to show their probable amount. Duane Jones Co. v. Burke, 306 N.Y. 172, 192, 117 N.E.2d 237, 247-48 (1954). The plaintiff need only show a ‘stable foundation for a reasonable estimate of royalties he would have earned had defendant not breached.’ Freund v. Washington Square Press, Inc., supra, 34 N.Y.2d at 383, 357 N.Y.S.2d at 861, 314 N.E.2d at 421. ‘Such an estimate necessarily requires some improvisation, and the party who has caused the loss may not insist on theoretical perfection.’ Entis v. Atlantic Wire & Cable Corp., 335 F.2d 759, 763 (2d Cir. 1964). ‘[T]he law will make the best appraisal that it can, summoning to its service whatever aids it can command.’ Sinclair Refug. Co. v. Jenkins Co., 289 U.S. 689, 697, 53 S.Ct. 736, 739, 77 L.Ed. 1449 (1933).”

Underpaid Royalties

It is undisputed that between April 1, 1972 and November 30, 1975 Falstaff sold 5,108,908 barrels of Ballantine products, on which a royalty of $2,554,454.00 was due. It has been stipulated that during that period Falstaff paid to Ballantine only $2,513,054.50. The underpaid royalties due for that period thus amount to $41,399.50.

Withheld Royalties

Plaintiff is also entitled to receive its $.50 per barrel royalty on all sales of Ballantine products after November 30, 1975. These were unjustifiably withheld by Falstaff in anticipation of its counterclaims against Ballantine, all of which have been found lacking in merit. In the period from December 1, 1975
to December 31, 1977, Falstaff sold 1,159,241 barrels of Ballantine products, and the royalty due is $579,620.50. Using the method described above, Falstaff would have sold 105,525 barrels of Ballantine products (30% of Ballantine’s 1974 volume) during the first quarter of 1978. The royalties due for this period amount to $52,762.50.

The sum total due to Balco is thus $1,302,310.60. Prejudgement interest at 6% per annum is awarded, with the computation to be made in accordance with New York CPLR Section 5001(b), with costs to be taxed.

QUESTIONS

1. Did Falstaff fail in what was described earlier in the chapter as its “diligence insurance” obligation? In what way? Would it surprise you to know that most of the litigated cases involve a diligence failure? Does that constitute strong evidence in favor of the “diligence” construction of “best efforts” as opposed to the “joint maximization” interpretation?

2. What is the meaning of the discussion of “evenhandedness” by the court? Why would Falstaff not have been evenhanded?

3. In a contract case, the method of computation of the damages usually permits one to infer the performance standard. Why? In the Bloor case, the method of damage computation is consistent with the joint maximization criterion. Why? Do you think that this consistency is deliberate, or could it have been inadvertent?

4. If Bloor had wanted to protect itself against this type of situation, how could the contract with Falstaff been rewritten?
This chapter deals in large part with “systemic” interactions among firms or other kinds of economic units. The traditional economic environment of “markets” is the principal system analyzed below, but reference is also made to non-market systems, such as intercommunity migration (see the Petaluma case below). When individual units are linked in some form of interdependent system, it is necessary to trace the process whereby a large number of individualistic decisions become aggregated in a system result such as a market price, an industry-wide output, etc. Conversely, it may also be of interest to ferret out the impact on individual units arising as a consequence of some change in the structure of the system itself.

In order to understand the supply side of markets, one must first have some minimal familiarity with the standard “theory of the firm” and the way in which individual producers are understood to make their production decisions. Considerations of cost are, obviously, paramount in production planning. The first few sections below are therefore designed to raise some preliminary issues about the nature of cost and its relevance to several different kinds of legal problems. Inter alia, the sections on costs should demonstrate how important and illuminating it is to consistently adhere to an “opportunity cost” concept of cost.

The preliminary materials on cost are followed by some extremely simple models of market adjustment where individual decisions are aggregated into an industry-wide outcome. One should beware of taking these models too seriously as a description of the real world. They are offered here for the more limited purpose of supplying stripped-down constructs whose mechanics are hopefully simple enough to be intelligible while at the same time retaining some analogical relevance to the forces that operate in real situations.

Although there has been nothing really exotic about the economic concepts used in earlier sections of these materials, they were imbedded in a legal context that is a bit removed from the traditional focus of economic literature. The underlying economics of this last segment is, by contrast, precisely the area in which standard economic texts have lavished a great deal of attention and sophistication. Be assured, therefore, that we are not about to undertake the task of re-inventing the wheel.

The appropriately abbreviated and relatively non-technical view provided below will nonetheless permit access to a number of stimulating legal issues. The more interested reader may find it useful to follow up by consulting any of a number of excellent economic theory texts and applied works that deal more or less directly with the topics touched upon below.

A. PRODUCTION COSTS AND CONTRACT BREACH

1. COST CONDITIONS FOR A FIRM PRODUCING STANDBPRODS

Firm S is a producer of standprods. It has contracted to deliver 6 lots of three units apiece to six different buyers, B1, B2, B3, B4, B5, and B6. All orders are to be delivered on June 1. Normally, production commences one working day before delivery is due. On April 29, before S has produced any of the standprods involved in these orders, market conditions undergo a dramatic change. Standprods become much less desirable to buyers and a series of breaches by S’s buyers takes place. First B1 breaches his order of three units, followed by B2 on an order of three, etc. After the final breach by B6, all of S’s orders have been lost.

We shall assume that S will file suit against each of the six breathers separately. Our problem will be to determine the “proper” recovery from each breacher. The costs that the firm would have had to incur in order to perform the contracts is an important determinant of that recovery. Exhibit 6.1 is a schedule of production costs for the firm in question. (We will also use this schedule for a subsequent problem.) One issue will be the appropriate treatment of what are sometimes called “overhead” costs.
Economists classify production costs in several ways that are relevant here. One important distinction is between “fixed” and “variable” costs. Fixed costs are whatever costs that remain even if output is shut down to zero. They are exemplified by items such as rental payments on a term lease of the plant, contractual obligation to pay the wages of the overseer, etc. In our Exhibit 6.1 example, such costs are $100. Variable costs are avoidable by adjusting output. Marginal costs are equivalent to the rate of change in these variable costs as output is varied. “Total” costs are, of course, the sum of both fixed and variable costs.

Variable costs are sometimes further subdivided to distinguish a category of “unallocable” costs. Such costs may be avoidable by shutting down completely, but for positive levels of production, it may be difficult to impute them to any particular item of production. In the example, it is assumed that a $15 license fee per period is necessary if any production takes place. Hence, the cost of the first unit of production is listed as $20: $15 for the license and $5 for variable inputs specifically attributable to the first unit of product.

### Exhibit 6.1: Costs at Alternative Outputs

<table>
<thead>
<tr>
<th>Output</th>
<th>Variable Cost</th>
<th>Total Cost</th>
<th>AVC</th>
<th>ATC</th>
<th>MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>NA</td>
<td>100</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>20.0</td>
<td>120.0</td>
<td>20.0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>12.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>101</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>113</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>126</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>155</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>171</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### QUESTIONS

1. Fill in the column of Exhibit 6.1 labelled “Total Cost.” The column labelled “ATC” provides space for Average Total Cost, computed by dividing Total Cost by the relevant output quantity. “AVC” signifies Average Variable Cost, a concept analogous to ATC. Fill in however much of this “average” cost information as you deem useful or relevant.

2. Determine the incremental cost of each unit and place the resultant information in the column headed “MC”.

3. Suppose that each of the contracts was for a price of $15.95 and the current market price has dropped to $6.95. How much does S need to collect in total. in order to be as well off as if the contracts had been performed? Should each breacher pay the same amount? Why?
4. What do you understand to be meant by “efficient breach”? How is this concept relevant to the answer in the question immediately preceding?

5. Same facts as above, but the market price has fallen only to $7.95. What answer now?

6. If the fixed costs had been $150 rather than $100, how would the answers to the last two questions have changed? Of what relevance is the ATC column of Exhibit 6.1?

7. “Fixed costs are not really costs at all. They don’t matter.” Comment.

2. CASES ABOUT DAMAGES FOR BREACHES OF DUTY

JERICHO SASH AND DOOR CO. v. BUILDING ERECTORS, INC.

The defendant appeals from a final decree for the plaintiff on this bill in equity to reach and apply assets of the defendant to the plaintiff’s claim of damages for breach of contract. The sole question presented is whether the trial judge erred in allowing damages for “profit (including reasonable overhead)” under G.L. c. 106, Section 2-708(2), in the absence of evidence showing separate figures for profit and for overhead. The plaintiff delivered 1,420 pairs of assorted sizes of window sash, for which the defendant admitted liability, and the defendant then repudiated the undelivered balance of 5,580. The plaintiff introduced evidence showing the “weighted average sales price per pair” and the “weighted average direct cost per pair” of the delivered sash. Subtraction of cost from price gave “lost profit and overhead per unit,” and multiplication by the number of undelivered units gave “total lost profit and overhead,” and the judge awarded more than $21,000 on that account. There was no error.

The judge followed the statutory injunction that the remedy “be liberally administered to the end’ that the aggrieved party had fully performed.” G.L. c. 106, Section 1-106(1), as appearing in St.1957, c. 765, Section 1. Damages need not “be calculable with mathematical accuracy. Compensatory damages are often at best approximate: they have to be proved with whatever definiteness and accuracy the facts permit, but no more.” Comment 1 to Section 1-106(1) of the Uniform Commercial Code, 1 U.L.A. (Master Ed.). See Dyecraftsmen, Inc. v. Feinberg, Mass., 269 N.E.2d 693; Coyne Industrial Laundry of Schenectady, Inc. v. Gould, Mass., 268 N.E.2d 848. There is no requirement that “overhead” be separated from “net profit” in the computation. As the plaintiff’s witness and the judge clearly understood, “profit (including reasonable overhead)” in the statute is the equivalent of “gross profit,” including fixed costs but not costs saved as a result of the breach.

Decree affirmed with costs of appeal.

QUESTIONS

1. Even if you have never heard the accounting term “direct cost” before, what can you infer that it must mean?

2. What do you understand to be meant by the phrase in Sec. 2-708(2) of the Uniform Commercial Code allowing the seller to recover “profit (including reasonable overhead)” lost as a result of a buyer’s breach? Is the discussion in the case unnecessarily muddy? How would you explain the point?

NOBS CHEMICAL, U.S.A., INC. v. KOPPERS CO., INC.
616 F.2d 212. (5th C., 1980)

Before CHARLES CLARK, RONEY and HENDERSON, Circuit Judges.

HENDERSON, Circuit Judge:

[Koppers contracted with the plaintiffs, Nobs Chemical and Calmon-Hill Trading Corp., to purchase 1000 metric tons of cumene, an oily additive for high-octane motor fuel, at a contract price of $540,000. Plaintiffs had arranged to purchase the cumene in Brazil for $400 a ton and to expend $45 per ton for the
cost of transporting the cumene to the defendant, for a total expense of $445,000. The district court applied Tex.Bus. & Com.Code Sec. 2.708(b) and determined that the plaintiffs were entitled to recover lost profits of $95,000 ($540,000 minus $445,000). The district court ruled that the plaintiffs could not recover the extra $25 per ton they allegedly were forced to pay their Brazilian supplier when the price per ton increased because their total order with the supplier was reduced from 4,000 metric tons to 3,000 metric tons because of Kopper’s breach. The court decided this lost quantity discount amounted to consequential damages and was, therefore, not recoverable.]

* * *

We first turn to the issue of whether the district court was correct in applying the lost profits measure of damages to the plaintiff’s loss. According to Tex.Bus. & Com.Code Ann. Sec. 2.708 (Vernon)

(a) * * * the measure of damages for non-acceptance or repudiation by the buyer is the difference between the market price at the time and place for tender and the unpaid contract price together with any incidental damages provided in this chapter (Section 2.710), but less expenses saved in consequence of the buyer’s breach.

(b) if the measure of damages provided in Subsection (a) is inadequate to put the seller in as good a position as performance would have done then the measure of damages is the profit (including reasonable overhead) which the seller would have made from full performance by the buyer, together with any incidental damages provided in this chapter (Section 2.710), due allowance for costs reasonably incurred and due credit for -payments or proceeds of resale.

The plaintiffs urge that subsection (a) should govern in this case. Because the market value of cumene dropped to between $220.40 and $264.48 a metric ton at the time of the breach, the plaintiffs contend that they should recover the difference between the contract price ($540,000) and the market price (between $220,400 and $264,480), substantially more than the $95,000 awarded them under subsection (b).

* * *

Because there does not appear to be any law directly on point, we take the liberty of looking to those more learned on the subject of the Uniform Commercial Code. Professors White and Summers, recognizing that Sec. 2.708(b) is not the most lucid or best-drafted of the sales article sections, decided that the drafters of the Uniform Commercial Code intended subsection (b) to apply to certain sellers whose losses would rarely be compensated by the subsection (a) market price-contract price measure of damages, and for these sellers the lost profit formula was added in subsection (b). One such type of seller is a “jobber,” who, according to the treatise writers, must satisfy two conditions: “[f]irst, he is a seller who never acquires the contract goods. Second, his decision not to acquire those goods after learning of the breach was not commercially unreasonable. * * * “J. White & R. Summers, Uniform Commercial Code Sec. 7-10, at 228 (1972). Nobs and Calmon-Hill clearly fit this description. The plaintiffs never acquired the goods from their Brazilian supplier, and, as White and Summers point out, an action for the purchase price or resale was therefore unavailable. See also, American Metal Climax, Inc. v. Essex International, Inc., 16 U.C.C.Rep. 101, 115 (S.D.N.Y. 1974) (“[C]ompensatory damages as provided in the contract-market formula of Sec. 2-708(1) [Sec. 2.708(a) ] are realistic only where the seller continues to be in a position to sell the product to other customers in the market.”).

The plaintiffs argue, however, that in this case the measure of damages under subsection (a) would adequately compensate them and therefore, according to the terms of subsection (a), subsection (b) does not control. This is an intriguing argument. It appears that the drafters of Sec. 2.708(a) did not consider the possibility that recovery under that section may be more than adequate. White & Summers, supra, Secs. 7-12, at 232-233.

It is possible that the code drafters intended subsection (a) as a liquidated damage clause available to a plaintiff-seller regardless of his actual damages. There have been some commentators who agree with this philosophy. [Citations omitted.] But, this construction is inconsistent with the code’s basic philosophy, announced in Tex.Bus. & Com.Code Ann. Sec. 1.106(a) (Vernon), which provides “that the aggrieved
party may be put in as good a position as if the other party had fully performed” but not in a better posture. * * * No one insists, and we do not think they could, that the difference between the fallen market price and the contract price is necessary to compensate the plaintiffs for the breach. Had the transaction been completed, their “benefit of the bargain” would not have been affected by the fall in market price, and they would not have experienced the windfall they otherwise would receive if the market price-contract price rule contained in Sec. 2.708(a) is followed. Thus the premise contained in Sec. 1.106 and Texas case law is a strong factor weighing against application of Sec. 2.708(a).

Our conclusion that the district court was correct in applying Sec. 2.708(b) brings us to the second issue—was it error for the district court to refuse to award the plaintiffs the additional $75,000, which they were required to pay when they lost their quantity discount?

We believe the trial court was correct in declining to award the plaintiffs the extra $75,000. Under Sec. 2.708(b), in addition to profit, the seller may recover “incidental damages” and “due allowance for costs reasonably incurred.” The code does not provide for the recovery of consequential damages by a seller. Tex.Bus. & Com.Code Ann. Sec. 1.106(a) (Vernon); Petroleo Brasileiro, S.A. v. Ameropan Oil Corp., 372 F.Supp. 503, 508 (E.D.N.Y.1974); Cf. Tex.Bus. & Com.CodeAnn. Sec. 2.715 (Vernon) (buyer’s remedies). “Incidental damages” are defined in Tex.Bus. & Com.Code Ann. Sec. 2.710 (Vernon) as “any commercially reasonable charges, expenses or commissions incurred in stopping delivery, in the transportation, care and custody of goods after the buyer’s breach, in connection with return or resale of the goods or otherwise resulting from the breach.” The draftsmen’s comment to the section states that the purpose is to “authorize reimbursement of the seller for expenses reasonably incurred by him as a result of the buyer’s breach.” We think it is clear that Sec. 2.710 was intended to cover only those expenses contracted by the seller after the breach and occasioned by such things as the seller’s need to care for, and, if necessary, dispose of the goods in a commercially reasonable manner. See, Guy H. James Construction Co. v. L.B. Foster Co., No. 17,473 (Tex.Civ.App.1979) (cost of replacing material in stock for resale recoverable as incidental damages).

QUESTIONS

1. Of what relevance is the sum that Nobs “expected” to make when it entered into the contract? Would giving it this sum be the same as expectation damages?

2. A careful reading of the case will suggest that Nobs was not contractually bound to purchase any fixed amount from the Brazilian supplier and had not yet acquired the cumene at the time of breach. If it had been able to perform on the contract, at what price do you think it would have been able to buy the necessary cumene?

3. Suppose that Nobs had already purchased the cumene for $400 per ton as originally contemplated and was holding it in inventory at the time of breach. What costs does it save if it does not have to deliver the cumene?

4. What is the “correct” recovery due to the plaintiff? Is it properly recoverable under subsection (a) of 2-708?

5. Could the plaintiff have plausibly argued that it had “purchased” the right to sell cumene at $540? What consideration did it give in exchange?


CURT’S TRUCKING CO. v. CITY OF ANCHORAGE

Supreme Court of Alaska, 1978.
578 P.2d 975.

RABINOWITZ, Justice.
An overhead telephone cable belonging to the Anchorage Telephone Utility was severed when struck by the raised body of a dump truck operated by Gerald Curt on Minnesota Drive in Anchorage. The Anchorage Telephone Utility is a public utility owned by the City of Anchorage; and City employees repaired the damaged cable. The City then brought an action in superior court alleging that $45,349.23 in damages resulted from the negligence of Gerald R. Curt and Curt’s Trucking Company.

The parties stipulated that Curt’s Trucking was liable for the direct costs of repairing the damaged cable. The direct costs of material and labor amounted to $4,457.72. However, Curt’s Trucking disputed the City’s claim to an additional 20% for “overhead” and the trial in superior court was directed only to that issue.

The City introduced evidence indicating that its 20% overhead figure was derived from two separate sources: (1) general administrative overhead applicable to all reimbursable City services, comprising 13% of the 20% overhead figure, and (2) expenses incurred by the City’s Risk Management Office in processing claims against third parties who damage City property. The superior court concluded that the entire 20% “overhead” charge should be allowed and awarded total damage in the amount of $5,349.23, i.e., $4,457.72 for direct costs of labor, materials and equipment plus $891.54 for indirect costs.

Curt’s Trucking argues that the 13% charge for administrative expenses applicable to all reimbursable City services was not a proper element of damages because it is too remote or too speculative. Both parties have cited opinions of this court explaining that the general principle underlying the assessment of damages in tort cases is that an injured person is entitled to be placed as nearly as possible in the position he would have occupied had it not been for the defendant’s tort. This principle also applies where the injured party performs the repairs itself. Many courts considering this problem have concluded that indirect expenses are proper items of damage and are not too remote or speculative when the evidence establishes a sufficient connection between the overhead expense and the necessary repairs.

In Baltimore and Ohio Railroad Co. v. Commercial Transport, Inc., 273 F.2d 447 (7th Cir. 1960), the court affirmed the inclusion of overhead in a damage award arising out of a collision between plaintiff’s diesel locomotive and defendants’ tractor-trailer. The railroad had applied various overhead percentages to its costs of materials and labor; the specific percentages had been taken from formulas established by 25 major railroads for use, in part, in billing each other for self-repair of their own track or equipment damaged by another railroad. The Seventh Circuit concluded that the reasonable cost of repairs made by the railroad’s own employees should not be limited to the dollar amount actually paid workmen and suppliers because “[s]uch a limitation would ignore the facts of business life.” The court also determined that the percentage formulas were properly submitted to the jury:

On the basis of the facts adduced in the instant case as to the establishment of the percentage formulae contained in the Rules of the General Managers Association, and their long use and continued acceptance in the industry, we are of the opinion that the district court did not err in admitting the testimony and exhibit in question. This evidence was material and relevant in proof of plaintiff’s damages. It is infrequent that damages are measurable with mathematical certainty or that a plaintiff can prove all elements from his own records or experience. The factors here employed and submitted for the jury’s consideration were not designed to produce a profit. Methods accepted in everyday business affairs as fixing fair measures of value are not to be excluded from a jury’s consideration in arriving at damages. [Idem.]

Other federal courts have also permitted recovery of properly calculated overhead expenses. In Bultema Dock & Dredge Co. v. Steamship David P. Thompson, 252 F.Supp. 881 (W. D. Mich. 1966), plaintiff recovered overhead expenses in connection with repair of its underwater construction project, which had been struck by defendant’s vessel. The district court explained, in part:

Had it been necessary for an outside company to repair the damage, the cost of such repairs would include similar charges. There is no reason to free respondent of these costs simply because the injured party is in a position to make the repairs itself.
However, in *Crain Brothers, Inc. v. Duquesne Slag Products Co.*, 273 F.2d 948, 952-53 (3d Cir., 1959), the Third Circuit explained that the reason a party which performs its own repairs may recover an amount including such overhead elements is that its actual loss would be the full amount charged by the independent contractor, as if it paid that amount out-of-pocket. Thus, where plaintiff has carried out the repairs itself, the losses and expenses actually incurred as a result of the accident should be included in a damage award. Costs which would have been recoverable had plaintiff hired someone else to do the work are a useful indicator of reasonable costs of repair only if plaintiff actually expends such an amount.

Although the state courts are not unanimous in awarding overhead expenses, many have concluded that such costs are properly included in damage awards. Courts permitting recovery generally have required plaintiffs to show that the claimed indirect costs were actually incurred and that the formula used to calculate overhead was accurate.

Sufficient accuracy has been found where the overhead calculation was based on sound accounting principles; where the accounting methods used to allocate indirect expenses had been approved by government regulatory bodies; or where the formulas were widely accepted in plaintiff’s trade or business as correct measures of overhead when self-repairs are performed.

We are persuaded that a party which repairs property tortiously damaged by another should be permitted to recover indirect expenses incurred in making such repairs if its calculation of overhead is shown to be a fair and reasonable measure of actual costs. A plaintiff need not identify the particular dollar amount of each overhead component which was expended in connection with specific direct costs; it is sufficient if reasonably accurate formula allocations or percentages are used to calculate the overhead. Whether such estimates are sufficiently accurate must be determined from all the facts and circumstances of the particular case.

In the case at bar, the City introduced the deposition of its internal auditor, Lawrence Campbell, a certified public accountant who explained how the 13% charge for administrative costs of reimbursable City services had been calculated. Campbell stated that the City performs a variety of services for which it recovers its costs. These services are known as “reimbursable City services,” and they range from thawing water pipes for private citizens to maintaining streets for the state highway department. Included in “reimbursable City services” are repairs to property damaged by third parties. Campbell conducted a study and determined that the average overhead attributable to all City departments which perform reimbursable services was approximately 13%.

Curt’s Trucking called its own expert witness, Frederick M. Strand, a certified public accountant who was present at the taking of Campbell’s deposition. Strand observed that the 13% overhead figure had been computed in accordance with sound accounting principles. Curt’s Trucking Company introduced no other evidence suggesting that the overhead figure was not an accurate measure of the indirect costs actually incurred in connection with repairing the damaged telephone cable.

As noted previously, we have determined that overhead expenses incurred by a party in repairing its damaged property are proper items of damages if such costs actually have been expended and have been allocated accurately to the repair work. In the case at bar, the superior court applied the correct principles of law in permitting recovery of these costs. In addition, the record contains nothing which indicates that the 13% overhead calculation was inaccurate. Thus, we conclude that the award of 13% of the direct costs was not clearly erroneous.

Curt’s Trucking Company next asserts that the additional 7% charge representing expenses of the City’s Risk Management Office are not proper items of damage because they are in the nature of costs incurred in preparing for litigation. The records shows that the 7% figure was derived in a different manner than the 13% charge associated with the cost of making repairs. The City’s internal auditor, Lawrence Campbell, testified that his administrative cost study which established the 13% figure did not consider the Risk Management Department’s expenses because the office had not yet been created. Both Campbell and Frederick Strand, the expert called by Curt’s Trucking, stated that the functions presently performed by the Risk Management Office were formerly spread among the City’s various departments.
Thus, to the extent these activities were carried on by departments which performed reimbursable City services, the costs associated with claims preparation may have been reflected in the 13% overhead calculation. Edward Hite, head of the City’s Risk Management Office, testified that to the extent claims processing occurred prior to creation of his position in 1973 it was conducted by personnel within the utility which had suffered damage. However, Hite stated that many of his office’s functions had never been performed prior to 1973.

Hite also described the tasks of the Risk Management Office in some detail. According to Hite, his office collects information about the incident, takes statements of employees involved, coordinates with any police investigation, determines whether the other party was negligent, evaluates the validity of potential claims, consolidates possible claims of different departments arising out of the same incident, estimates damages or costs of repair, and sends the claim to an adjuster for collection or to the City’s insurance company. Hite explained that the Risk Management Office does not send out investigators and rarely corresponds with the people who have damaged City property or with their insurers. On cross-examination, counsel for Curt’s Trucking asked Hite what his office actually does with respect to a claim; Hite stated, in part:

The same thing you as an individual would do if someone came along and smashed into your car. You’re going to take that car down and get yourself an estimate on it. You may have to go to three different places. You’re probably going to have to arrange for a rental car. You’re going to sit down and try to figure out how much money you have been out and you’re probably going to end up arguing with an adjusting firm on how much money you have really got coming. And you’re going to spend an hour or two and probably a lot of tears in trying to get your money. * * * Someone has come along and damaged our goodies and we’re going to try to get our money back for the damage that is done. Someone from the municipal point of view has got to keep track of that money. Someone’s got to keep track of who owes whom and what we’re going to pay and why we’re going to pay it. And if we don’t, we’re not handling your tax money right.

Even if the testimony could be understood as suggesting that some costs of preparing claims possibly were included in the 13% overhead figure as computed in 1971, we are not persuaded that the superior court’s allowance of the full 13% was clearly erroneous. Both parties’ accounting witnesses agreed that a 13% charge accurately reflected the indirect costs to the City of making repairs; no evidence was introduced to indicate the unrelated costs had been included. In addition, Campbell testified that his 1974 reevaluation of the 13% figure showed it was still accurate; in March, 1973 the City’s risk manager had taken over claims processing functions which previously had been performed by the individual utilities.

QUESTIONS

1. Is all overhead fixed cost? What does it mean to require that “such costs actually have been expended and have been allocated accurately to the repair work”?

2. Should the treatment of “overhead” costs differ in any way because this is a tort and not a contractual relationship? Consider this issue also in regard to the Lund case immediately below which involves a criminal theft of services.

LUND v. COMMONWEALTH

Supreme Court of Virginia, 1977.

Defendant, Charles Walter Lund, was charged in an indictment with the theft of keys, computer cards, computer print-outs and using “without authority computer operation time and services of Computer Center Personnel at Virginia Polytechnic and State University [V.P.I. or University] * * * with intent to
defraud, such property and services having a value of one hundred dollars or more.” Code Sec. 18.1-118 were referred to in the indictment as the applicable statutes.

The computer used by the defendant was leased on an annual basis by V.P.I. from the IBM Corporation. The rental was paid by V.P.I. which allocates the cost of the computer center to various departments within the University by charging it to the budget of that department. This is a bookkeeping entry, and no real money is involved. The departments are then issued “computer credits [in dollars] for their use [on] a proportional basis of their [budgetary] allotments.” Each department manager receives a monthly statement showing the allotments used and the running balance in each account of his department.

An account is established when a duly authorized administrator or “department head” fills out a form allocating funds to a department of the University and an individual. When such form is received, the computer center assigns an account number to this allocation and provides a key to a locked post office box which is also numbered to the authorized individual and department. The account number and the post office box are the access code which must be provided with each request before the computer will process a “deck of cards” prepared by the user and delivered to computer center personnel. The computer print-outs are usually returned to the locked post office box.

Defendant came under surveillance on October 12, 1974, because of complaints from various departments that unauthorized charges were being made to one or more of their accounts. When confronted by the University’s investigator, defendant initially denied that he had used the computer service, but later admitted that he had. He gave to the investigator seven keys for boxes assigned to other persons.

The director of the computer center testified that the unauthorized sum spent out of the accounts associated with the seven post office box keys, amounted to $5,065. He estimated that on the basis of the computer cards and print-outs obtained from the defendant, as much as $26,384.16 in unauthorized computer time had been used by the defendant. He said, however, that the value of the cards and print-outs obtained from the defendant was “whatever scrap paper is worth.”

The defendant contends that his conviction of grand larceny of the keys, computer cards, and computer print-outs cannot be upheld under the provisions of Code Sec. 18.1-100 because (1) there was no evidence that the articles were stolen, or that they had a value of $100 or more, and (2) computer time and services are not the subject of larceny under the provisions of Code Secs. 18.1-100 or 18.1-118.

The Commonwealth concedes that the defendant could not be convicted of grand larceny of the keys because there was no evidence that those articles were stolen and that they had a market value of $100 or more. The Commonwealth argues, however, that the evidence shows the defendant violated the provisions of Sec. 18.1-118 when he obtained by false pretense or token, with intent to defraud, the computer print-outs which had a value of over $5,000.

Under the provisions of Code Sec. 18.1-118, for one to be guilty of the crime of larceny by false pretense, he must make a false representation of an existing fact with knowledge of its falsity and, on that basis, obtain from another person money or other property which may be the subject of larceny, with the intent to defraud. See Hubbard v. Commonwealth, 201 Va. 61, 66, 109 S.E.2d 100, 104 (1959).

At common law, larceny is the taking and carrying away of the goods and chattels of another with intent to deprive the owner of the possession thereof permanently. Code Sec. 18.1-100 defines grand larceny as a taking from the person of another money or other thing of value of five dollars or more, or the taking not from the person of another goods and chattels of the value of $100 or more. The phrase “goods and chattels” cannot be interpreted to include computer time and services in light of the often repeated mandate that criminal statutes must be strictly construed. See Commonwealth v. McCray, 430 Pa. 130, 133, 242 A.2d 229, 230 (1968).

At common law, labor or services could not be the subject of the crime of false pretense because neither time nor services may be taken and carried away. It has been generally held that, in the absence of a clearly expressed legislative intent, labor or services could not be the subject of the statutory crime of
false pretense. McCray, supra; 2 Wharton, Criminal Law and Procedure Sec. 604 at 369 (Anderson ed. 1957). Some jurisdictions have amended their criminal codes specifically to make it a crime to obtain labor or services by means of false pretense. We have no such provision in our statutes.

Furthermore, the unauthorized use of the computer is not the subject of larceny. Nowhere in Code Sec. 18.1-100 or Sec. 18.1-11.8 do we find the word “use.” The language of the statutes connotes more than just the unauthorized use of the property of another. It refers to a taking and carrying away of a certain concrete article of personal property. See People v. Ashworth, 220 App. 498, 222 N.Y.S. 24, 27 (1927). There it was held that the unauthorized use of machinery and spinning facilities of another to process wool did not constitute larceny under New York’s false pretense statute.

We hold that labor and services and the unauthorized use of the University’s computer cannot be construed to be subjects of larceny under the provisions of Code Secs. 18.1-100 and 18.1-118.

The Commonwealth argues that even though the computer printouts had no market value, their value can be determined by the cost of the labor and services that produced them. We do not agree.

The cost of producing the printouts is not the proper criterion of value for the purpose here. Where there is no market value of an article that has been stolen, the better rule is that its actual value should be proved. See 50 Am.Jur.2d, Larceny Sec. 45 at 211 and Sec. 148 at 329. Compare, Hancock v. State, 402 S.W.2d 906, 18 A.L.R.3d 1113 (Tex.Cr.App.1966), where the prosecution was based on a statute and there was evidence of the monetary value of the computer programs.

Here the evidence shows that the print-outs had no ascertainable monetary value to the University or the computer center. Indeed, the director of the computer center stated that the print-outs had no more value than scrap paper. Nor is there any evidence of their value to the defendant, and the value to him could only be based on pure speculation and surmise. Hence, the evidence was insufficient to convict the defendant of grand larceny under either Code Sec. 18.1-100 or Sec. 18.1-118.

For the reasons stated, the judgment of the trial court is reversed, and the indictment is quashed.

Note:
Following the decision in Lund supra, the Code of Virginia was amended as follows:

18.2-98.1. Computer time or services or data processing services or information or data stored in connection therewith is hereby defined to be property which may be the subject of larceny under Sec. 18.2-95 or 18.2-96 or embezzlement under Sec. 18.2-111, or false pretenses under Sec. 18.2-178.

QUESTIONS

1. Suppose that the defendant had stolen special purpose goods, such as an orthotic shoe device made by a podiatrist to a mold of defendant’s foot and therefore of no use to anyone else. The materials of the device are about three dollars worth of plastic composition and an eighty-cent leather cover. Such devices are regularly sold by the podiatrist for $145. Defendant is charged with grand larceny under a state statute that requires a value of $100 or more for the goods illegally obtained. Would the fact that tangible goods are involved avoid the problem raised in Lund above? Does such an orthotic device involve both tangible goods and a service? If so, how are the values to be separated for purposes of the minimum valuation called for in the statute?

2. Is the Lund court correct in rejecting cost of production as an index of value? If cost of production were to be used, what elements (e.g., pro rata share of computer rental charges paid by university, labor of computer center personnel, etc.) should be included in the cost?

3. If “cost” to the University is construed as how much worse off it was made by not being able to exclude Lund from obtaining services illicitly, then what was the cost of the services taken?

4. At various times of the year, such as the ends of terms, most University computer facilities become overwhelmed with excess demand for the limited amount of processing time available on the system. This results in long waiting times, sometimes as long as several days, between the submission of a “job” and the reception of finished output. Assume that a significant portion of Lund’s use occurred
during such periods. What “costs” did he impose? Could they be quantified in any way? Do such costs have any legal relevance? If not, should they?

B. ADJUSTMENT OF A COMPETITIVE INDUSTRY

Up to this point, we have dealt exclusively in terms of the adjustments made by single decisionmakers or small groups; there has been no attempt to aggregate the behavior of individuals into “market” results. This section takes up the manner in which individual firms’ decisions are translated into market prices and aggregate supply of a product. The treatment here is sketchy since the behavior of markets is the principal topic of traditional economics books, many of which deal with the subject very elegantly and in great detail. The material supplied below is, however, designed to provide the essential “flavor” of the process of market adjustment.

We present the traditional textbook case of a competitive market. The firms in this industry produce a homogeneous product and are “price-takers,” i.e., respond to market prices but do not individually have the power to set prices. The indirect manner in which individual firms affect prices will be illustrated through a problem that has been used in economics classes for many years.

1. A SIMONS MODEL OF COMPETITIVE EQUILIBRIUM

The problem is loosely based on a simple model of competitive market adjustment attributed to Professor Henry Simons. This model is explicitly a pedagogical one and, in order to maintain its simplicity, necessarily incorporates certain artificialities. Anyone who has reason to understand the process of market adjustments should nevertheless take the time to work through an exercise of this type at least once.

FACTS. At the beginning of the period under analysis, the industry which produces “standprods” (which are perfectly standardized goods) contains exactly 500 firms, each of which operates only a single plant. The per-period cost of producing standprods consists of “overhead” costs of $100 imputable to the plant itself, plus certain variable or “operating” costs imputable to different levels of output. The costs are identical for each firm and are the same as indicated above in Exhibit 6.1 used for the buyers’ breach hypothetical problem.

The type of plant being used deteriorates completely in 500 weeks, regardless of how much it is used. At the beginning of the scenario, there is a uniform distribution of plant ages, i.e., one plant one week old, one plant two weeks old, etc. Plant construction takes ten weeks and, once initiated, is an irreversible process. Hence, there are initially 10 replacement plants “in process” which cannot be cancelled.

The demand curve faced by the industry can be represented by the equation price = 65 - (Q/170) where Q is total industry output. This implies that price will fall a little more than half a cent for every unit produced. But, the number of competitors is sufficiently great that no single producer feels that his own output will affect price at all.

a. Part I. Initial Equilibrium State

1. Fill in any remaining blank columns of Exhibit 6.1. Using graph paper, plot the approximate shapes of the marginal cost, average variable cost, and average total cost curves as Exhibit 6.2.

2. (The abbreviations MC, AVC, and ATC are commonly used.)

3. Using a smaller scale than for Exhibit 6.2, construct Exhibit 6.3 to show the supply curve of the present 500-firm industry. This will be the horizontal sum of each firm’s marginal cost curve, except

...
for the part of MC which lies below the AVC curve. [Why?] Superimpose the industry’s demand curve onto Exhibit 6.3 and ascertain the equilibrium price where the supply and demand curves intersect. Why is this an “equilibrium” price? What will happen at a higher or lower price?

Exhibit 6.2: Graph of Firm’s Cost Curves

---

38 A convenient format for Exhibit 6.3 will involve vertical (price) values in the range from 7 to 24 and horizontal (output supplied or quantity demanded) values between 4,000 and 13,000. An accurate version of Exhibit 6.3 is supplied as Exhibit 6.3A in the Appendix to this chapter.
4. Suppose the world rolls on with the original situation basically unchanged. Will there be any tendency for changes to occur in price, the number of firms, etc.? Why?

**b. Partial Displacement And Short-Run Equilibrium Adjustment**

Evidence is presented to Congress that the use of standprods causes significant environmental damage. A statute is subsequently passed requiring that standprods be rendered “environmentally harmless.” Fortuitously, the AZD Corporation reveals that it has a patented process which will eliminate the environmental damages otherwise produced due to the use of standprods. The process in question is essentially a clever improvement in fabrication technique and does not in any way change the real resource costs of production. After estimating the market value of the technique, AZD Corporation fixes a $4 per standprod royalty for the use of its patented process.

1. Other than what is in effect a legal requirement to use the AZD process, no other changes in the conditions facing the industry occur. Fill in Exhibit 6.4 as the revised version of Exhibit 6.1, reflecting the cost of the royalty for using the AZD process.
2. Add to Exhibit 6.2, in red or some distinctive manner, the new cost curves for a firm, based on Exhibit 6.4.
3. Add to Exhibit 6.3 the new industry supply curve and determine the new (short-run) equilibrium price immediately after imposition of the necessity to use the AZD process.
4. What is your answer now to the same question as asked in Part I-c above?
Exhibit 6.4: Costs to Firm Paying ASD Royalty

<table>
<thead>
<tr>
<th>Output</th>
<th>Variable Cost</th>
<th>Total Cost</th>
<th>AVC</th>
<th>ATC</th>
<th>MC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>100</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**c. Progress Toward New Equilibrium**

1. When (after how many weeks) will firms begin to leave the industry? Why? Draw the industry supply curve and determine the price when the time elapsed after enactment of the new law is 10 weeks, 30 weeks, 50 weeks. Why would these prices be regarded as only short run equilibrium prices”? Draw the short run supply curve that would exist if firms kept exiting for 70 weeks. How many firms would there be? Would firms keep leaving for 70 weeks? Why?
   
   What is happening to the output of a single firm as the short run equilibrium price adjusts?

2. When do firms stop leaving and things “settle down” to a long run equilibrium? What makes this an “equilibrium”?

**d. Additional Implications**

After the industry has reached a new long-run equilibrium it is revealed by a *Washington Pillar-and-Post* investigative reporter that AZD’s patent was fraudulently obtained and is certainly invalid. Assuming this to be the case, AZD may be liable for treble damages under the antitrust statutes.

1. Who should have the cause of action to sue? The standprods producers? The standprods consumers? Someone else? Explain the implications of the relevant alternatives. At this point, try not to be influenced in your response by what you know about the present stance of the courts in cases such as Hanover Shoe, Inc. v. United Shoe Machinery Corp., 392 U.S. 481 (1968) and Illinois Brick Co. v.
Illinois, 431 U.S. 720 (1977). These cases are presented below. You may wish to reassess your answer after reading them.

2. Suppose that ultimate consumers of standprods (the indirect users of AZD process) are to be plaintiffs. What is the measure of their damages? In what way, if at all, do such damages differ from those that a producer might ask if he were the antitrust plaintiff?*

**e. A Simple Mathematical Model  [Optional]**

*This is an optional section.* Readers who are absolutely appalled by the slightest use of mathematics may wish to skip this section. Indeed, those who will be more confused than enlightened by the mathematical formulation can be assured that the results of this section are not crucial to what follows. On the other hand, the modest effort necessary to read this material will be well repaid in terms of increased understanding for many people, especially those who have never studied any type of mathematical economic model.

Even a vague recollection of 9th grade elementary algebra will suffice to follow the rudimentary mathematical model of a competitive market about to be presented. Indeed, this model will be merely a translation into simple algebra of the graphical version of the Henry Simons model already discussed above.

Recall that the facts of the Simons model contained a stipulation of the market demand curve in the form of an equation:

\[ P = 65 - (1/170)Q \]

where \( P \) is the maximum market price that consumers will pay if \( Q \) units are to be sold. Since this relationship is really the equation of the demand curve, the \( Q \) should be identified as quantity demanded by appending an identifying subscript \( d \). Then we can merely manipulate equation \([1]\) in order to yield its more useful form:

\[ Q_d = 170 (65 - P) \]

Equation \([1a]\) says that quantity demanded, \( Q_d \), is inversely related to price.

Now we need to model the supply side of the market. We know that each individual firm will produce out to the quantity where marginal cost equals the marginal revenue it receives. For competitive firms, price equals marginal revenue, so that the profit-maximizing output of a single firm will be where \( MC = P \). In the relevant range of outputs (where price exceeds average variable cost), the behavior of marginal cost can be summarized by the following equation:

\[ MC = k + q \]

where the lowercase \( q \) is the single firm’s output. The parameter \( k \) is equal to -2 in the situation before the AZD royalty fee was necessitated; i.e., \( MC = -2 + q \). If the marginal costs are affected by the $4 royalty fee, they become \( MC = -2 + 4 + q \), so that \( k \) would then be \((-2 + 4) \) or \(+2\). If we wish to identify the profit-maximizing \( q \), we must set \( MC = P \) by substituting the solution for \( MC \) from \([2]\)

\[ k + q = P \]

and then solving for \( q \) on the left:

\[ q = P - k \]

Since we are working with a simple “identical firms” model, the industry supply is the quantity supplied by any individual firm multiplied by the number of firms operating:

\[ Q_s = nq \]
where $Q_s$ is the quantity supplied. We can also substitute equation [3] for the $q$ term in equation [4], yielding

$$[4a] \quad Q_s = n \,(P-k)$$

so that [4a] represents a short-run “supply curve” of the kind used in the graphical models.

Now the demand and supply relations can be put together to express the industry equilibrium condition that price must adjust until consumers desire to buy exactly the quantity supplied:

$$[5] \quad Q_d = Q_s$$

Of course, it will be necessary to substitute the demand curve and supply curve equations derived in [1a] and [4a] for $Q_d$ and $Q_s$, respectively, in order to solve for the equilibrium value of the price $P$. That substitution yields:

$$[5a] \quad n(P - k) = 170 \,(65 - P)$$

as the equilibrium condition of the market. Finally, solving for $P$, we have the “solution” of the model in terms of an equilibrium price:

$$[6] \quad P = \frac{(11050 + nk)}{(n + 170)}.$$

By substituting +2 for $K$, as indicated for the with-fee situation, we can use equation [6] to follow the market adjustment process.

It would be desirable to express the number of firms, $n$, in terms of the time period involved. Thus, for the first week $t = 1$, for the second week, $t = 2$, etc. Because 10 new plants have irrevocably been ordered, $n = 500$ for the first ten weeks after the AZD fee is imposed. After that, $n$ will decrease by one for each week as long as price remains lower than average total cost for the producers. Hence, from week eleven until a new equilibrium is established, $n$ will be representable as $500 - (t - 10)$, or $n = 500 - t + 10$. What will be the situation in week 12, for instance? If we substitute the right values for $k$ and $n$ in equation [6] above, we will have:

$$P = \frac{[11050 + 499(2)]}{(499 + 170)} = 18.00.$$

At an $18$ price, what will a single firm’s output be? The answer can be gained from equation [3] above.

The model just created can be used to address the issue of when a new equilibrium will occur. In essence, this question really asks how many firms must enter or leave in order for price to become exactly equal to the minimum average total cost (ATC) of production for an individual firm. Do you understand why? Those interested in doing so may attempt to formulate this question as an algebraic equilibrium condition. [Hint: In the relevant range of outputs, Total Cost = Fixed Cost + 32.875 + .5(q + 2.5)².] Those less interested in such an exercise should at least examine Exhibit 6.4b below and attempt to understand how it is possible to fill in the blanks. This exhibit uses the simple model just developed, together with the original facts of the Simons problem, to indicate the values of critical market parameters at various numbers of weeks after the AZD fee was imposed.

---

1 Here, in some detail for those whose algebra is rusty, are the series of intermediate steps that might be followed in getting from [5a] to [6]:

- $n(P - k) = 170(65) - 170P$
- $nP + 170P = 11050 + nk$
- $P(n + 170) = 11050 + nk$
- $nP - nk = 11050 - 170P$

and thence to the result in the text.
Exhibit 6.4b: Solutions of Market Model

<table>
<thead>
<tr>
<th>Week</th>
<th>Firms</th>
<th>Firm Output</th>
<th>Firm Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>499</td>
<td>16.0</td>
<td>______</td>
</tr>
<tr>
<td>31</td>
<td>479</td>
<td>16.5</td>
<td>______</td>
</tr>
<tr>
<td>50</td>
<td>460</td>
<td>______</td>
<td>______</td>
</tr>
</tbody>
</table>

Note that Exhibit 6.4b demonstrates the possibility that output solutions (for \( q \) in equation [3]) may take on non-integer values. Is there any interpretation of the original cost data that could make sense out of fractional outputs? Or must these models impose an additional mathematical condition that the solutions be integers?

2. **BANKRUPTCY LAWS AND FIXED COSTS**

There is a famous “quotation” that economists are fond of giving as a “comment” question on economic theory exams:

Economists are crazy. Obviously, recovering fixed costs is very important and necessary to businessmen. Economists, though, say that in the short run fixed costs don’t matter and in the long run they don’t exist.

The position attributed to economists is not so crazy at all, of course, as long as one understands costs in the opportunity cost sense. Past decisions to which one is irrevocably committed may require outlays, but they can never represent opportunity costs imputable to current activities.

There is, nevertheless, at least one important respect in which it would be incorrect to say that fixed costs do not matter at all. Fixed costs are often legal obligations to third parties. If the firm is unable to pay such obligations, judgments may be secured against it, attaching its working capital, physical assets, inventory, etc. Even though the firm would, through continued operation, be able to pay all future costs, it may be put out of business through its inability to satisfy creditors who have covered past costs. Bankruptcy laws play an important role in coping with this situation. Consider, for instance, the following scenario, one whose facts should have a familiar ring.

Universal Standprods, Inc., has just brought on line a new physical plant financed by a loan from Westhamden Guaranty & Trust and secured by a first mortgage on the building used for standprod production. Machinery in the plant was financed by various unsecured loans from five other financial institutions. A three month inventory of the raw materials and chemicals used in the production process is stored in Universal’s factory warehouse. Most of this inventory was acquired from suppliers on 60 to 90-day open book credit. Standprods had been selling until last month at a price of $15, an amount exactly sufficient to provide Universal with the cashflow necessary to cover its costs and provide a normal return on equity capital. Unfortunately, a new EPA regulation has just increased unit costs by $4.00., Market price has risen in reaction, but is hovering around the $18 level. At any price below $19, Universal will have insufficient cash flow to cover production costs and the payments on its obligations to creditors. Universal has only a very small amount of liquid assets. It does possess wholly-owned standprod metering equipment purchased new for $500,000 shortly before the EPA regulation was announced. Creditors are beginning to become very nervous, since it appears that the present discounted value of Universal’s liabilities exceeds that of its expected operating income.

**QUESTIONS**

1. Do you think the metering equipment could be sold for anything near its $500,000 cost? What factors determine its current value?
2. Under what circumstances could it be said that the creditors are in a difficult Prisoner’s Dilemma? (Be as specific as possible in explaining the dilemma, using the concept of expected value where appropriate.) How do bankruptcy laws respond to this situation?

3. Who is made better or worse off by bankruptcy laws?

C. INCIDENCE ANALYSIS: SHIFTING AND PASSING ON

As presented above, the simple graphical and mathematical models of a competitive market demonstrate an interesting point: a cost initially levied on producers turns out, after the passage of some time for adjustment, to be “shifted” entirely to consumers. The tracing out of this process in an attempt to discover the ultimate consequences of a market change is called incidence analysis. Traditionally, incidence analysis has been used by economists primarily to investigate the burdens imposed by various forms of taxes. The same analytic principles are, however, applicable to other kinds of costs and benefits, notably including those created by the legal system.

If it were always the case that costs are ultimately shifted 100% to the consumer of the affected product, the world would be a simpler place. In that respect, however, the models used above are based upon an assumption that will not always be satisfied in the real world.

In order to understand what actuates-and limits-shifting, one needs to understand the fundamental equilibrium conditions of competitive markets. Two time periods are relevant. In what is called the “short run,” the firm is “stuck” in the industry and outlays on so-called fixed costs are not economically relevant. Subject to the proviso that average variable costs are covered, the single firm’s “supply curve” is its marginal cost curve and the market supply curve is the aggregation of the single-firm supply curves. When cost variations cause the firms’ marginal cost curves to shift upward or downward, the effect is directly translated into a shift of the short run market supply curve, and therefore a change in price.

In what is called the “long run,” firms are free to exit from or enter the industry. The actuating principle here is the relationship between price and average total cost. If price is greater than average total cost, a firm can do better in the industry than anywhere else, thus motivating new entrants whose increased supply will move price downward. If price is less than average cost, exit will take place, supply will fall, and price will start to rise. Equilibrium exists when price is exactly equal to average total cost. (Remember that cost is in terms of opportunity cost, so that when revenues and costs are equal the firm is doing just as well in the present industry as in the next best opportunity.) In sum, long-run equilibrium requires that price rise or fall, as may be required, to the level of average total cost.

In our earlier models, the average cost curves of the firms were assumed to be affected only by the change in the royalty fee. Obviously, then, price would in the long run have to track the fee variations exactly. The long run supply curve would be a horizontal line at the level of the low point on the ATC curve; at that price, firms will enter without limit to accommodate demand, while at any lower price nobody will continue to produce in the long run.

But what if the cost curves of the individual firms were also affected by the volume of output produced by the industry? For instance, as industry output expands, certain inputs may become scarce or have to be lured away from other uses. As the best inputs become fully employed, recourse to inferior inputs may be necessary. Such conditions would mean that costs to the individual firm will shift upward as industry output grows and, similarly, decline when it contracts. Graphically, this condition would be manifested by an upward sloping long-run supply curve. The degree of the slope depends on the strength of the scarcity factor as industry output fluctuates. A high slope of the supply curve is likely to be associated with the use of specialized inputs in an industry; when the industry output expands, these inputs can command a scarcity premium, but when output declines, their very specialization causes them to accept lower returns rather than shift to some other use or employment.

Of course, the degree to which industry output does expand or contract is in part determined by the demand conditions. If the slope of the demand curve is relatively flat, it means that consumers will adjust
to any supply cutback principally by switching to some other form of consumption rather than raising their bids for the product and hence its price.

In sum, the incidence of a change in market conditions depends on the relative slopes of the demand and supply curves. In turn, the slopes of these curves depend upon the substitutability of other goods for consumers and the mobility of productive inputs, respectively. In general, we might expect that part of a cost increase will be passed on to consumers and some part may also be borne by the producing sector. Unfortunately, the relative degrees of incidence borne by particular parties are strictly empirical questions. For a more detailed discussion of the theory and practice of incidence analysis, see any standard public finance textbook.

The difficulty of doing accurate incidence analysis is an issue that troubles the Supreme Court in the following cases.

**HANOVER SHOE, INC. v. UNITED SHOE MACHINERY CORP.**

Supreme Court of the United States, 1968. 392 U.S. 481

Mr. Justice WHITE delivered the opinion of the Court.

Hanover Shoe, Inc. (hereafter Hanover) is a manufacturer of shoes and a customer of United Shoe Machinery Corporation (hereafter United), a manufacturer and distributor of shoe machinery.

Hanover’s action against United alleged that United had monopolized the shoe machinery industry in violation of Sec. 2 of the Sherman Act; that United’s practice of leasing and refusing to sell its more complicated and important shoe machinery had been an instrument of the unlawful monopolization; and that therefore Hanover should recover from United the difference between what it paid United in shoe machine rentals and what it would have paid had United been willing during the relevant period to sell those machines.

* * *

The District Court found that Hanover would have bought rather than leased from United had it been given the opportunity to do so. The District Court determined that if United had sold its important machines, the cost to Hanover would have been less than the rental paid for leasing these same machines. The difference in cost, trebled, is the judgment awarded to Hanover in the District Court. United claims, however, that Hanover suffered no legally cognizable injury, contending that the illegal overcharge during the damage period was reflected in the price charged for shoes sold by Hanover to its customers and that Hanover, if it had bought machines at lower prices, would have charged less and made no more profit than it made by leasing. At the very least, United urges, the District Court rejected this assertion of the so-called “passing-on” defense, and we affirm that judgment.

If in the face of the overcharge the buyer does nothing and absorbs the loss, he is entitled to treble damages. This much seems conceded. The reason is that he has paid more than he should and his property has been illegally diminished, for had the price paid been lower his profits would have been higher. It is also clear that if the buyer, responding to the illegal price, maintains his own price but takes steps to increase his volume or to decrease other costs, his right to damages is not destroyed. Though he may manage to maintain his profit level, he would have made more if his purchases from the defendant had cost him less. We hold that the buyer is equally entitled to damages if he raises the price for his own product. As long as the seller continues to charge the illegal price, he takes from the buyer more than the law allows. At whatever price the buyer sells, the price he pays the seller remains illegally high, and his profits would be greater were his costs lower.

* * *

* * * The rule, United argues, should be the subject to the defense that economic circumstances were such that the overcharged buyer could only charge his customers a higher price because the price to him was higher. It is argued that in such circumstances the buyer suffers no loss from the overcharge. This situation might be present, it is said, where the overcharge is imposed equally on all of a buyer’s
competitors and where the demand for the buyer’s product is so inelastic that the buyer and his competitors could all increase their prices by the amount of the cost increase without suffering a consequent decline in sales.

We are not impressed with the argument that sound laws of economics require recognizing this defense. A wide range of factors influence a company’s pricing policies. Normally the impact of a single change in the relevant conditions cannot be measured after the fact; indeed a businessman may be unable to state whether, had one fact been different (a single supply less expensive, general economic conditions more buoyant, or the labor market tighter, for example), he would have chosen a different price. Equally difficult to determine, in the real economic world rather than an economist’s hypothetical model, is what effect a change in a company’s price will have on its total sales. Finally, costs per unit for a different volume of total sales are hard to estimate. Even if it could be shown that the buyer raised his price in response to, and in the amount of, the overcharge and that his margin of profit and total sales had not thereafter declined, there would remain the nearly insuperable difficulty of demonstrating that the particular plaintiff could not or would not have raised his prices absent the overcharge or maintained the higher price had the overcharge been discontinued. Since establishing the applicability of the passing-on defense would require a convincing showing of each of these virtually unascertainable figures, the task would normally provide insurmountable. On the other hand, it is not unlikely that if the existence of the defense is generally confirmed, antitrust defendants will frequently seek to establish its applicability. Treble damage actions would often require additional long and complicated proceedings involving massive evidence and complicated theories.

In addition, if buyers are subjected to the passing-on defense, those who buy from them would have to meet the challenge that they passed on the higher price to their customers. These ultimate consumers, in today’s case the buyer of single pairs of shoes, would have only a tiny stake in a lawsuit and little interest in attempting a class action. In consequence, those who violate the antitrust laws by price fixing or monopolizing would retain the fruits of their illegality because no one was available who would bring suit against them.

***

*** We recognize that there might be situations—for instance, when an overcharged buyer has a pre-existing “cost-plus” contract, thus making it easy to prove that he has not been damaged—where the considerations requiring that the passing-on defense would not be permitted in this case would not be present. We also recognize that where no differential can be proved between the price unlawfully charged and some price that the seller was required by law to charge, establishing damages might require a showing of loss or profits to the buyer.

ILLINOIS BRICK CO. v. ILLINOIS
Supreme Court of the United States, 431 U.S. 720 (1977)

[Suppliers of concrete block in Illinois pleaded nolo contendere to an indictment charging them with price-fixing. When a number of private suits were subsequently filed, the suppliers settled out of court with all of the contractors who had directly purchased the block. There were, however, also other suits by indirect purchasers, such as the purchasers of the buildings that the contractors had constructed. The Court of Appeals held that these indirect purchasers could recover for any of their provable damages.]

Mr. Justice WHITE delivered the opinion of the Court.

The parties in this case agree that however Sec. 4 is construed with respect to the pass-on issue, the rule should apply equally to plaintiffs and defendants—that an indirect purchaser should not be allowed to use a pass-on theory to recover damages from a defendant unless the defendant would be allowed to use a pass-on defense in a suit by a direct purchaser. Respondents, in arguing that they should be allowed to recover by showing pass-on in this case, have conceded that petitioners should be allowed to assert a pass-on defense against direct purchasers of concrete block; they ask this Court to limit Hanover Shoe’s bar on
pass-on defenses to its “particular factual context” of overcharges for capital goods used to manufacture new products.

Before turning to this request to limit Hanover Shoe, we consider the substantially contrary position, adopted by our dissenting Brethren and lower courts that have allowed offensive use of pass-on, that the unavailability of a pass-on theory to a defendant should not necessarily preclude its use by plaintiffs seeking treble damages against that defendant. Under this view, Hanover Shoe’s rejection of pass-on would continue to apply to defendants unless direct and indirect purchasers were both suing the defendant in the same action, but it would not bar indirect purchasers from attempting to show that the overcharge had been passed on to them. We reject this position for two reasons.

First, allowing offensive but not defensive use of pass-on would create a serious risk of multiple liability for defendants. Even though an indirect purchaser had already recovered for all or part of an overcharge passed on to it, the direct purchaser would still recover automatically the full amount of the overcharge that the indirect purchaser had shown to be passed-on; similarly, following an automatic recovery of the full overcharge by the direct purchaser, the indirect purchaser could sue to recover the same amount. The risk of duplicative recoveries created by unequal application of the Hanover Shoe rule is much more substantial than in the more usual situation where the defendant is sued in two different lawsuits by plaintiffs asserting conflicting claims to the same fund. A one-sided application of Hanover Shoe substantially increases the possibility of inconsistent adjudications and therefore of unwarranted multiple liability for the defendant—by presuming that one plaintiff (the direct purchaser) is entitled to full recovery while preventing the defendant from using that presumption against the other plaintiff; overlapping recoveries are certain to result from the two lawsuits unless the indirect purchaser is unable to establish any pass-on whatsoever. As in Hawaii v. Standard Oil Co., we are unwilling to “open the door to duplicative recoveries” under Sec. 4.

Second, the reasoning of Hanover Shoe cannot justify unequal treatment of plaintiffs and defendants with respect to the permissibility of pass-on arguments. The principal basis for the decision in Hanover Shoe was the Court’s perception of the uncertainties and difficulties in analyzing price and output decisions “in the real economic world rather than an economists’ hypothetical model, and on the costs to the judicial system and the efficient enforcement of the antitrust laws of attempting to reconstruct those decisions in the courtroom. This perception that the attempt to trace the complex economic adjustments to a change in the cost of a particular factor of production would greatly complicate and reduce the effectiveness of already protracted treble-damage proceedings applies with no less force to the assertion of pass-on theories by plaintiffs than to the assertion by defendants. However, “long and complicated” the proceedings would be when defendants sought to prove pass-on, they would be equally so when the same evidence was introduced by plaintiffs. Indeed, the evidentiary complexities and uncertainties involved in the defensive use of pass-on against a direct purchaser are multiplied in the offensive use of pass-on by a plaintiff several steps removed from the defendant in the chain of distribution. The demonstration of how much of the overcharge was passed on by the first purchaser must be repeated at each point at which the price-fixed goods changed hands before they reached the plaintiff.

It is argued, however, that Hanover Shoe rests on a policy of ensuring that a treble-damage plaintiff is available to deprive antitrust violators of “the fruits of their illegality,” a policy that would be furthered by allowing plaintiffs but not defendants to use pass-on theories. We do not read the Court’s concern in Hanover Shoe for the effectiveness of the treble-damage remedy as countenancing unequal application of the Court’s pass-on rule. Rather, we understand Hanover Shoe as resting on the judgment that the antitrust laws will be more effectively enforced by concentrating the full recovery for the overcharge in the direct purchasers than by allowing every plaintiff potentially affected by the overcharge to sue only for the amount it could show was absorbed by it.

Permitting the use of pass-on theories under Sec. 4 essentially would transform treble-damage actions into massive efforts to apportion the recovery among all potential plaintiffs that could have absorbed part of the overcharge—from direct purchasers to middlemen to ultimate consumers. However appealing this
attempt to allocate the overcharge might seem in theory, it would add whole new dimensions of complexity to treble-damage suits and seriously undermine their effectiveness.

As we have indicated, potential plaintiffs at each level in the distribution chain are in a position to assert conflicting claims to a common fund—the amount of the alleged overcharge—contending that the entire overcharge was absorbed at that particular level in the chain. A treble-damage action brought by one of these potential plaintiffs (or one class of potential plaintiffs) to recover the overcharge implicates all three of the interests that have traditionally been thought to support compulsory joinder of absent and potentially adverse claimants: the interest of the defendant in avoiding multiple liability for the fund; the interest of the absent potential plaintiffs in protecting their right to recover for the portion of the fund allocable to them; and the social interest in the efficient administration of justice and the avoidance of multiple litigation.

Opponents of the Hanover Shoe rule have recognized this need for compulsory joinder in suggesting that the defendant could interplead potential claimants under 28 U.S.C.A. Sec. 1335. But if the defendant, for any of a variety of reasons, does not choose to interplead the absent potential claimants, there would be a strong argument for joining them as “persons needed for just adjudication” under Fed. Rule Civ.Proc. 19(a).

It is unlikely, of course, that all potential plaintiffs could or would be joined. Some may not wish to assert claims to the overcharge; others may be unmanageable as a class; and still others may be beyond the personal jurisdiction of the court. We can assume that ordinarily the action would still proceed, the absent parties not being deemed “indispensable” under Fed. Rule Civ.Proc. 19(b). But allowing indirect purchasers to recover using pass-on theories, even under the optimistic assumption that joinder of potential plaintiffs will deal satisfactorily with problems of multiple litigation and liability, would transform treble-damage actions into massive multiparty litigations involving many levels of distribution and including large classes of ultimate consumers remote from the defendant. In treble-damage actions by ultimate consumers, the overcharge would have to be apportioned among the relevant wholesalers, retailers, and other middlemen, whose representatives presumably should be joined. And in suits by direct purchasers or middlemen, the interests of ultimate consumers are similarly implicated.

There is thus a strong possibility that indirect purchasers remote from the defendant would be parties to virtually every treble-damage action (apart from those brought against defendants at the retail level). The Court’s concern in Hanover Shoe to avoid weighing down

More generally, the process of classifying various market situations according to the amount of pass-on likely to be involved and its susceptibility to proof in a judicial forum would entail the very problems that the Hanover Shoe rule was meant to avoid. The litigation over where the line should be drawn in a particular class of cases would inject the same “massive evidence and complicated theories” into treble-damage proceedings, albeit at a somewhat higher level of generality. Hanover Shoe itself implicitly discouraged the creation of exceptions to its rule barring pass-on defenses, and we adhere to the narrow scope of exemption indicated by our decision there.

QUESTIONS

1. For the simple case of the “standprod” industry developed above, review who are the injured parties and what is the quantum of their damages. In the real world, which of the facts would be difficult to know, and how would they affect the damage calculations?

2. Recall that the standprod hypothetical depicts a “constant cost” industry, the cost of whose product is not affected by the magnitude of industry output. How would the presence of scarce or specialized inputs complicate matters?

3. The Illinois Brick rule substantially undercuts the compensatory function of antitrust damages, does it not? What about the deterrent function? Does a party who has not really been injured have enough incentive to sue?

5. Might limitation of recovery to direct purchasers result in underdeterrence of antitrust violators?

D. ANTICOMPETITIVE SUPPLY RESTRICTION

The antitrust cases in the previous section provide the occasion to comment at least briefly on non-competitive markets, i.e., those in which producers’ own decisions are capable of materially influencing the prices that they obtain. An assumption of our simple competitive model was, after all, that the producer acts as a “price taker” rather than as a “price maker.” It is worth exploring why a producer who is one of many in a market is, in general, neither willing nor able to engage in output-restricting activities that will raise price. In sum, our message will be that a producer who is only one among many does not have the necessary incentives to try to influence price, nor would he be likely to have much impact if he did try to do so.

It is quite true that even a producer in a competitive market can put any price tag he wants on his products; the problem is that the price at which he can actually sell the product is set by the demand conditions of the market, not by any decision controlled by the producer himself. As the Henry Simons model illustrated, the market demand curve reflects the maximum price that consumers are willing to pay for different quantities placed on the market. For each possible quantity supplied, the corresponding price is really determined by the demand curve alone. The power to control price, therefore, can effectively be exercised only by controlling quantity supplied.

Still, it may seem that a competitive producer could control the price he gets, at least to some extent. As a competitive firm contracts or expands its own production, does not the aggregate quantity supplied on the market change, moving the price up or down the demand curve within the range of the producer’s own quantity variations? Why does the competitive producer not take this price-changing effect into consideration? One facile answer is “Well, a competitive producer is too small to have much effect on price.” Bright principles of economics students are sometimes troubled by this response, and rightly so. “Wait a minute,” they reason. “These economics books insist that the monopolist is sensitive to the effect that even a one-unit quantity variation has on his market price, that he regards the ‘marginal revenue’ of an extra unit sold as always less than the price he gets for that incremental unit. I can see how the impact of, say, a one-unit output variation on price would be small, but what I can’t understand is why it would be any more perceptible when a monopolist varies quantity than when an output variation by one of many producers creates precisely the same small movement along the market demand curve.”

It is, of course, quite true that equal quantity variations in aggregate quantity supplied would have the same effect on price no matter whether they were accomplished by small producers or large ones. Understanding why large and small producers act differently requires that one adopt a slightly different perspective on the problem, one that can be appreciated by using the simple market demand curve model embodied in Exhibit 6.5. Assume that the quantity supplied would be 10. At that quantity, the demand curve tells us that the market-clearing price would be $20. If one more unit were to be supplied, however, price would have to fall to $19 in order to induce consumers voluntarily to purchase the larger quantity. At the original quantity, the total revenues paid by consumers to producers will be 10 X $20 = $200 while at the larger output the payments from consumers are 11 x $19 = $209. For producers in the aggregate, therefore, the marginal revenue of the 11th unit is only $9 even though it sells at a price of $19.
Exhibit 6.5: Expansion versus Market Spoilage

It is instructive to observe the price and revenue change in terms of Exhibit 6.5 where the components of the change can be analyzed graphically. The total revenue corresponding to each situation is equivalent to the area of the rectangle formed by each price-and-quantity combination. These large rectangles have in common the large, shaded rectangle labeled B, so that the net change in revenues can be seen as the additional tall vertical segment C minus the loss of the horizontal segment A. Segment A represents what we can call a “market spoilage effect” whereby the sale of the 11th unit requires that the price on the other 10 units be cut by $1 apiece. The area of segment A is \((10 \times 1) = 10\). In turn, segment C is the revenue increase attributable to selling the incremental product at its market price and will be referred to as the “expansion effect.” Since C’s area is \((1 \times 19) = 19\), the net of these two areas, $19 - $10, is exactly the $9 change in producers’ receipts derived in the previous paragraph.

But we have been analyzing the impact of the output variation in terms of producers in the aggregate rather than a single producer. The dawning of insight into the problem should come when one realizes that the seller of the additional unit gets the full benefit of the revenue increment that has just been called the expansion effect, but normally suffers only part of the loss flowing from the market spoilage effect. For instance, suppose that the original quantity supplied came from three suppliers, I, II, and III, who would have marketed 1, 3, and 6 units, respectively. If producer I markets an extra unit, his individual market spoilage effect only extends to the single unit he would have sold at original $20 price. Thus, his marginal revenue from the one-unit volume expansion is the $19 expansion effect minus only a $1 market spoilage effect, for a net of $18. Similar computations for II and III, yield figures of $16 and $13, respectively.

You should now recognize the rather direct application of the externality and Prisoner’s Dilemma models in the opening section of these materials to this context. “Market spoilage” due to supply expansion is an external cost borne by producers in proportion to their market share. “Expansion” is an internal benefit whose magnitude bears no relation to market share. A small producer bears so little of the market spoilage effect that his incentives are dominated by the expansion effect. As his market share
increases, the market spoilage effect is less and less exported to others and becomes a significant internal cost of the producer’s own output variations. The limiting case is, of course, the single seller (monopolist) for whom the market spoilage effect is felt undiluted.

A bit of a generalization in symbolic form will recap the argument thus far and also pave the way for the next step. Let $Q$ represent the total quantity supplied before a one-unit variation, $q$ the original quantity supplied by some particular producer, $\Delta P$ the price diminution required to accommodate the one-unit change, and $P$ the market price after the change. Then the individual producer’s own marginal revenue will be:

$$\text{MR} = P - \left(\frac{q}{Q}\right) [Q \Delta P]$$

where the term in parentheses is the producer’s market share, the term in square brackets is the market spoilage effect, and $P$ is of course the expansion effect. For the monopolist—or any producer with a large market share—the value of $(q/Q)$ approaches one and the market spoilage effect becomes a significant output-restraining cost. For such producers, marginal revenue is always significantly less than market price $P$. By contrast, a small producer will find that $\text{MR} \approx P$ as market share $(q/Q)$ approaches zero.

“But wait,” our critical student objects. “The textbooks say that for the competitive firm $\text{MR} = P$ not that $\text{MR} \approx P$. Markets are represented as competitive even when single firms have 5% or more of the market.” Fair enough. All we have done thus far is to show why competitive firms would have little incentive to forego any profitable expansion effect (by failing to produce a unit when $MC < \text{MR} \leq P$). An additional step remains. Even if a single producer were willing to produce one unit less, our argument can be completed by showing that this forebearance may have little or no effect on aggregate quantity supplied by the industry as a whole. Unless there is some effect on aggregate quantity supplied, the desired price maintenance impact will obviously not be achieved.

Rather than analyze the effect of a single-unit forebearance from producing, let us first over-dramatize by assuming a very large such forebearance. The principles derived from this more emphatic case can then easily be analogized to the single-unit situation. Specifically, suppose that, in the Henry Simons model, 10 firms unaccountably decide to do everyone else a favor by withholding their normal production, $q$, from the market and, instead, produce nothing. Will the reduction in actual quantity brought to market be equal to $10 \times q$? No. As soon as price begins to rise slightly in response to the initial supply constriction, other producers will be motivated to respond by expanding to fill the “hole” left by the forebearing producers.

In the short run, this “hole filling” effect will not completely counterbalance the reduction by the missing 10 producers. The other producers must, after all, expand out along their own short run supply curves, i.e., the upward-sloping marginal cost curves of the individual firms. You can, however, use the Exhibits developed above to tell exactly what the impact will be. Recall that the situation is now, in effect, one wherein ten too many firms have left the industry. What is the relevant supply curve for that situation? What is its intersection with the demand curve? And, ultimately, how much difference did the complete forebearance of ten producers make?

The previous paragraph addresses only the short run hole-filling effect. If the ten firms were to persist in their behavior, how long would it take before the productive “hole” will be filled without a remaining trace? All that is necessary for this to happen is that somewhere out there are ten other investors who are willing to replace the ten who are attempting to create a reduction in the aggregate quantity supplied.

One of the opening questions in this section was: “As a competitive firm contracts or expands its own production, does not the aggregate quantity supplied on the market change, moving the price up or down the demand curve within the range of the producer’s own quantity variations?” Despite its superficial plausibility, that question should be answered negatively.

In sum, a competitive firm tends to ignore the effects of its own output on price for two conceptually distinct reasons. One is that, even if such output variations affected industry output, what we have called the market spoilage effect is very dilute. The second is that a firm may have good reason to believe that
aggregate output, and therefore price, will be about the same whatever its own output decisions. In combination, these factors justify the MR = P assumption of the competitive model.

Implicitly, our discussion above also suggests how anticompetitive supply restrictions may come about when the factual underpinnings of the competitive model are violated by factors such as small numbers of (“oligopolistic”) producers, firms with large market shares, explicit collusive arrangements, and barriers to entry of new firms.

**QUESTIONS**

1. Review the Prisoner’s Dilemma and externality concepts developed in the first part of these materials. How do these concepts apply to competitive behavior?
2. The standard price-and-output determination model for a monopolist is not provided in this text. (Many people are already familiar with it. If you are not, you may now know enough to guess at how it works. If not, look it up.) Draw that standard diagram and indicate the monopoly price and output. If the monopolist could be made to act like a competitive firm, i.e., ignore his own effect on price, what would price and output be? Who gains and loses as between the two outcomes? Show that consumers lose more by the monopolistic output reduction than the monopolist gains.
3. The text of this section deals with competitive selling markets. To what extent do the same considerations carry over to buying markets, i.e., competitive purchases rather than sales?

**SEGREGATED MARKETS: PROVING “EXPLOITATION”**

The opinions that follow are from a Chicago housing discrimination case that involves both a novel legal theory and interesting facts. On the one hand, you should reflect on what the future implications for the housing market would be if plaintiffs’ liability theory were

Need first page //415 in casebook//
tiffs’ case in chief, Judge Perry granted defendants’ motion for directed verdict, holding in opposition to Judge Will’s theory of the case that:

* * * [C]ounsel for the plaintiffs have not painted a pretty picture of the defendants, but that picture is a picture of exploitation for profit, and not racial discrimination.

* * *

Nowhere in the six weeks trial is there one scintilla of evidence that the defendants or any of them or their agents ever refused to sell to a white person or a black person or a non-white person any house or refused to sell one or the other at a higher or lower price, absolutely no positive evidence of discrimination in this record

Accordingly, for want of any evidence in support of the complaint, the motion for a directed verdict by all of the defendants now on trial is hereby granted, and the complaint of all of the plaintiff’s is hereby dismissed as to all of the defendants.

Under Judge Perry’s theory of the case, absent evidence of defendants’ sales of the same or similar housing to whites on more favorable terms and prices, namely, the traditional theory of racial
Plaintiffs raise numerous issues in this appeal the most important of which of course is the correctness of the grant of a directed verdict for the defendants. The other issues can broadly be categorized as challenges to the correctness of certain of the trial judge’s evidentiary rulings and other procedural rulings.

In judging the propriety of the grant of the directed verdict, we are confronted with two issues. We must first resolve the conflict as to the scope of section 1982. That is, we must determine whether section 1982 covers only the so-called traditional type of discriminatory conduct, or whether a claim may be stated under section 1982 by proof of exploitation of a discriminatory situation already existing and created in the first instance by the action of persons other than defendants. If we determine that section 1982 is violated under the latter theory we then must determine whether the evidence, both the admitted evidence and erroneously excluded evidence, when viewed in the light most favorable to plaintiffs, together with all inferences that may be reasonably drawn therefrom, is such that it can be found that plaintiffs have made out a prima facie case.

Section 1982 of the Civil Rights Act of 1866 provides:

All citizens of the United States shall have the same right, in every State and Territory, as is enjoyed by white citizens thereof to inherit, purchase, lease, sell, hold, and convey real and personal property.

Plaintiffs’ “exploitation” theory of liability under this section can briefly be restated as follows: As a result of racial discrimination there existed two housing markets in Chicago, one for whites and another for blacks, with the supply of housing available in the black market far less than the demand. Defendants entered the black market selling homes for prices far in excess of their fair market value and far in excess of prices which whites pay for comparable homes in the white market and on more onerous terms than whites similarly situated would encounter. Plaintiffs contend that by so acting defendants seized upon and took advantage of the opportunity created by racial residential segregation to exploit blacks in violation of section 1982.

It is asserted that to countenance such actions by the defendant would be in direct contravention of the express language of section 1982 which provides by clear implication that black citizens “shall have the same right * * * as is enjoyed by white citizens * * * to * * * purchase * * * real * * * property.” Moreover, plaintiffs claim that a ruling which would hold defendants’ asserted acts of exploitation to be outside of the coverage of section 1982 would be contrary to the Supreme Court’s declaration in Jones v. Mayer Co., 392 U.S. 409, 443, 88 S.Ct. 2186, 2205, 20 L.Ed.2d 1189 (1968), and that section 1982 was meant to be a vehicle through which “to assure that a dollar in the hands of a Negro will purchase the same thing as a dollar in the hands of a white man.”

Defendants contend that absent a showing of the traditional form of discrimination, namely, that defendants refused to sell to blacks because of their race, or offered to sell the same houses to whites at lower prices or on more favorable terms than they offered to sell to blacks, plaintiffs have not made out a case under section 1982. Defendants assert that the houses they sold to plaintiffs in the black market were available to whites and would have been sold on the same terms and for the same prices as sales to the black plaintiffs. Therefore, defendants argue, plaintiffs enjoyed “the same right” as enjoyed by white citizens to purchase houses in the black market. Moreover, it is urged that other sellers and not defendants discriminated against plaintiffs in the first instance by refusing to sell to plaintiffs housing in other urban areas, thereby excluding them from the white market. The defendants claim that an extension of section 1982 so as to proscribe their alleged acts of exploitation would be tantamount to holding defendants liable for the discrimination of others without a showing of any discrimination by defendants. With respect to the Supreme Court’s decision in Jones v. Mayer Co., the defendants suggest that it lends no support to the plaintiffs’ desired interpretation of section 1982. Rather, defendants view Jones v. Mayer Co. as authority for their contention that section 1982 prohibits only the so-called traditional type of discrimination and does not encompass plaintiffs’ theory of exploitation liability. Also, defendants urge that plaintiffs’ interpretation
of section 1982 would render that section unconstitutionally vague and would expose defendants to risk or detriment without fair warning of the nature of the proscribed conduct. It is on the basis of these arguments that defendants claim that no case can be made under section 1982 for exploitation absent a showing that defendants offered to sell the same or similar homes to whites at lower prices and on more favorable terms than they made available to blacks. We do not agree.

At the outset we note that section 1982 is framed in broad yet clear language. It provides that:

All citizens of the United States shall have the same right * * * as is enjoyed by white citizens thereof to * * * purchase * * * real * * * property.

Facially, therefore, the scope of section 1982 would appear to be rather far reaching; indeed such a reading of the statute is supported by the Supreme Court’s interpretation of section 1982 in Jones v. Mayer Co., 392 U.S. 409, 88 S.Ct. 2186, 20 L.Ed.2d 1189 (1968). In that case the Court was confronted with questions as to the scope and constitutionality of section 1982. The plaintiff in that case, a black person, brought an action pursuant to section 1982 claiming that the defendants refused to sell him a house on the basis of his race. The district court dismissed the plaintiff’s complaint, and the court of appeals affirmed, concluding that section 1982 applied only to state action and not private action. The Supreme Court rejected the argument for a narrow construction of section 1982, holding in broad language:

That section 1982 bars all racial discrimination, private as well as public, in the sale or rental of property, and that the statute, thus construed, is a valid exercise of the power of Congress to enforce the Thirteenth Amendment. 392 U.S. 413, 88 S.Ct. 2189 [Emphasis in original].

The Court went on to note that section 1982 was an attempt by Congress to provide “that the right to purchase and lease property was to be enjoyed equally throughout the United States by Negro and white citizens alike” and that Congress “plainly meant to secure that right against interference from any source whatever, whether governmental or private.” 392 U.S. at 423-424, 88 S.Ct. at 2195. The Court concluded its analysis by stating that section 1982 must be accorded “a sweep as broad as its language.” 392 U.S. at 437, 88 S.Ct. at 2202. Observing that “when racial discrimination herds men into ghettos and makes their ability to buy property turn on the color of their skin, then it too is a relic of slavery,” the Court proceeded to uphold the constitutionality of section 1982 stating:

Negro citizens, North and South, who saw in the Thirteenth Amendment a promise of freedom-freedom to “go and come at pleasure” and to “buy and sell when they please”-would be left with “a mere paper guarantee” if Congress were powerless to assure that a dollar in the hands of a Negro will purchase the same thing as a dollar in the hands of a white man. At the very least, the freedom that Congress is empowered to secure under the Thirteenth Amendment includes the freedom to buy whatever a white man can buy, the right to live wherever a white man can live. If Congress cannot say that being a free man means at least this much, then the Thirteenth Amendment made a promise the Nation cannot keep. 392 U.S. 443, 88 S.Ct. 2205.

Clearly the Court’s decision in Jones v. Mayer Co. does not, contrary to defendants’ assertions, detract from plaintiffs’ contention as to the scope of section 1982. Rather, the decision is support for plaintiffs’ theory for in Jones the Court viewed section 1982 as a broad based instrument to be utilized in eliminating all discrimination and the effects thereof in the ownership of property. Accordingly, Jones v. Mayer Co. does not stand as an obstacle to plaintiffs’ case, but supports it. Defendants insist that section 1982 cannot be construed to encompass other than the traditional type of discrimination, that is, that defendants offered to sell to whites on more favorable terms and prices than to plaintiffs. Keeping in mind the Supreme Court’s admonition in Lane v. Wilson, 307 U.S. 268, 275, 59 S.Ct. 872, 876, 83 L.Ed. 1281 (1939), that the Constitution and statutes promulgated in its enforcement nullify “sophisticated as well as simple-minded modes of discrimination,” we reject defendants’ notion of adherence to a strict, rigid, and traditional type of discrimination. We need not resort to a labeling exercise in categorizing certain activity as discriminatory and others as not of such character for section 1982 is violated if the facts demonstrate that defendants exploited a situation created by socioeconomic forces tainted by racial discrimination. Indeed, there is no difference in results between the traditional type of discrimination and defendants’ exploitation of a discriminatory situation. Under the former situation blacks either pay excessive prices or are refused altogether from purchasing housing, while
under the latter situation they encounter oppressive terms and exorbitant prices relative to the terms and prices available to white citizens for comparable housing.

To avoid this conclusion, defendants contend that even though the results obtained under both the traditional and exploitation theories are similar, they come about through significantly different means. Under the traditional theory a black man is denied the “same right” as a white man in that the seller offers to sell the same house to each but at different prices and terms due to the differences in race of the prospective buyer. It is defendants’ position that they offered the plaintiffs the same terms and prices they would offer whites. Therefore it is asserted that plaintiffs had the same right as white citizens. This argument ignores current realities of racial psychology and economic practicalities. Defendants can find no justification for their actions in a claim that they would have sold on the same terms to those whites who elected to enter the black market and to purchase housing in the ghetto and segregated inner-city neighborhoods at exorbitant prices, far in excess of prices for comparable homes in the white market. It is no answer that defendants would have exploited whites as well as blacks. To accept defendants’ contention would be tantamount to perpetuating a subterfuge behind which every slumlord and exploiter of those banished to the ghetto could hide by a simple rubric: The same property would have been sold to whites on the same terms.

Defendants urge that other sellers and not they were the active agents of discrimination. That is, blacks were excluded from the white market by other sellers who refused to sell to plaintiffs and that accordingly plaintiffs’ action lies solely against those other owners and real estate operators and not the defendants. But, we repeat, defendants cannot escape the reach of section 1982 by proclaiming that they merely took advantage of a discriminatory situation created by others. We find repugnant to the clear language and spirit of the Civil Rights Act the claim that he who exploits and preys on the discriminatory hardship of a black man occupies a more protected status than he who created the hardship in the first instance. Moreover, defendants’ actions prolong and perpetuate a system of racial residential segregation, defeating the assimilation of black citizens into full and equal participation in a heretofore all white society. Through the medium of exorbitant prices and severe, long-term land contract terms blacks are tied to housing in the ghetto and segregated inner-city neighborhoods from which they can only hope to escape someday without severe financial loss. By demanding prices in excess of the fair market value of a house and in excess of what whites pay for comparable housing, defendants extract from blacks resources much needed for other necessities of life, thereby reducing their standard of living and lessening their chances of escaping the vestiges of a system of slavery and oppression. Indeed, defendants’ activity encourages overt discrimination by offering the long oppressed black an unattractive yet alternative choice to that of a confrontation for equal buyers’ rights in a white neighborhood.

Defendants in effect contend that this is solely a matter of economics and not of discrimination. We cannot accept this contention for although the laws of supply and demand may function so as to establish a market level for the buyer in the black housing areas, it is clear that these laws are affected by a contrived market condition which is grounded in and fed upon by racial discrimination—that is, the available supply of housing is determined by the buyer’s race. In other contexts the law has prevented sellers from charging whatever the market will bear when special circumstances have occasioned market shortages or superior bargaining positions. In such instances sellers were denied the opportunity to exploit others merely because the opportunity existed.

Contrary to the trial court’s stance, the shortage of housing here was triggered not by an economic phenomenon but by a pattern of discrimination that has no place in our society. Accordingly, neither prices nor profits—whether derived through well-intentioned, good faith efforts or predatory and unethical practices—may reflect or perpetuate discrimination against black citizens. We agree with Judge Will’s statement that “there cannot in this country be markets or profits based on the color of a man’s skin.” Contract Buyers League v. F & F Investment, 300 F.Supp. 210, 216 (N.D.Ill.1969). Price and profit differentials between individual buyers may be justified on a multitude of grounds; for example, the prospective purchaser’s reputation or his financial position and potential earning power. But price or profit may not turn on whether the prospective buyer has dark or light pigmentation.
Defendants urge that acceptance of plaintiffs’ asserted interpretation of section 1982 would render the statute unconstitutional and that the standards of liability utilized to deploy such an expanded interpretation would be unconstitutionally vague and indefinite. They argue that a narrow interpretation is compelled so as to sustain the constitutionality of the statute and to prevent the exposure of “a potential actor to some risk or detriment without giving him fair warning of the nature of the proscribed conduct.” Rowan v. Post Office Dept., 397 U.S. 728, 740, 90 S.Ct. 1484, 1492, 25 L.Ed.2d 736 (1970). Defendants’ argument misses the mark for we have not here a penal statute to be strictly construed. Rather, we have a civil statute which is remedial in nature and, as viewed by the Supreme Court, is to be accorded “* * * a sweep as broad as its language.”

dollar in the hands of a white man. Such practices render plaintiffs’ dollars less valuable than those of white citizens—a situation that was spawned by a discarded system of slavery and is nurtured by vestiges of that system. Courts in applying section 1982 must be vigilant in preventing toleration of this deplorable circumstance.

We hold accordingly that plaintiffs state a claim under section 1982 since they allege that (1) as a result of racial residential segregation dual housing markets exist and (2) defendant sellers took advantage of this situation by demanding prices and terms unreasonably in excess of prices and terms available to white citizens for comparable housing. If the plaintiffs sustain the burden of proof on these elements they make out a prima facie case, whereupon, as recently made clear by the Supreme Court, the burden of proof shifts to the defendants “to articulate some legitimate, nondiscriminatory reason” for the price and term differential. McDonnell Douglas Corp. v. Green, 411 U.S. 792, 802, 93 S.Ct. 1817, 1824, 36 L.Ed.2d 668 (1973).

II

Having determined the substantive framework upon which an action may be brought pursuant to section 1982 we address ourselves to the correctness of the directed verdict entered in favor of the defendants at the close of plaintiffs’ case in chief. Keeping in mind that we must view all the evidence, plus the reasonable inferences to be drawn therefrom, in the light most favorable to the plaintiffs, Hannigan v. Sears, Roebuck & Co., 410 F.2d 285, 287 (7th Cir.), cert. denied, 396 U.S. 902, 90 S.Ct. 214, 24 L.Ed.2d 178 (1969), we hold that the admitted evidence was sufficient to establish a prima facie case under section 1982 pursuant to the exploitation theory of liability and, accordingly, warranted submission of the issues to the jury for resolution. Moreover, we rule that on the basis of the evidence erroneously excluded by the trial judge the plaintiffs have made out a case for section 1982 liability under the so-called traditional theory of discrimination and that, therefore, a directed verdict for defendants on this separate ground was likewise improperly granted. To demonstrate the error committed we proceed to review the evidence adduced at trial.

There was sufficient evidence to establish, prima facie, the existence of dual housing markets in the Chicago metropolitan area as a result of racial residential segregation. Dr. Karl E. Taeuber, a professor of sociology, testified as an expert witness about the results of his extensive research on the dispersion of population in the city of Chicago. His statistical analysis indicated that Chicago was a highly segregated city and that there was a very high degree of residential segregation between whites and blacks. Moreover, despite the decrease of white population in the city accompanied by a rapid increase of the black population, the supply of new housing available to whites was much greater than that available to blacks. Also, during the pertinent time period the expanding suburban housing market was limited almost completely to whites. Dr. Taeuber testified that the main obstacle to the movement of blacks into the white areas of Chicago and suburban residential areas was the high degree of discrimination against blacks in the white market. As a result the supply of housing available to whites was far greater in both absolute and relative terms to the supply of new housing available to blacks.

Plaintiffs produced additional proof concerning the existence of dual housing markets through the testimony of another expert witness, Scott Tyler, a real estate broker and appraiser with many years of experience in the real estate business in Chicago. Significantly, defendants seemingly concede the existence of a dual housing market in Chicago. Nor do we think it beyond the strictures of judicial notice to observe that there exists in Chicago and its environs a high degree of racial residential segregation.
We turn to the second element of the case, whether there was a sufficient *prima facie* showing of an unreasonable differential in price and sale terms between the housing sold or offered by defendants to plaintiffs and comparable housing available to whites. With respect to this phase plaintiffs offered the appraisal testimony of five expert witnesses; the testimony of two, however, was excluded from the jury by the trial judge. Concerning the testimony of Scott Tyler and John Hank which the judge did allow, both witnesses utilized the market data method of appraisal in arriving at a fair market value of a sampling of plaintiffs’ homes. The fair market value appraisals were based on sales of comparable homes in all-white neighborhoods which were located in close geographical proximity to plaintiffs’ homes and had similar communal amenities such as transportation, schools, churches, and quality of neighborhood. Both witnesses testified that the comparable white housing was sold at prices substantially below the prices commanded by defendants. Expert witness Tyler’s appraisals demonstrated that on the average the contract prices charged by defendants exceeded the fair market value of the homes by $6,508, or 34.5 percent. Expert witness Hank was of the opinion that on the average defendants’ prices exceeded fair market value by $4,209, or 20.6 percent. Plaintiffs adduced additional appraisal testimony from Paul Underwood who had been an appraiser for a savings and loan association which had loaned money to one of the defendant land companies for construction of houses sold to plaintiffs. Pursuant to this financing arrangement Underwood appraised thirty of defendants’ houses for which he testified that the sales prices charged by defendants, on the average, exceeded fair market value by $4,296, or 20.9 percent.” Based on the foregoing we think a jury could reasonably reach the conclusion that defendants’ price differential was unreasonably in excess of fair market value and prices available to white citizens for comparable housing. Accordingly, there was sufficient evidence on the price differential to send the case to the jury.

Turning to the issue of the reasonableness of the sale terms differential the evidence at trial indicated that defendants refused to sell other than on land contract to plaintiffs. There was testimony to the effect that defendants refused to participate in any sales through a deed and mortgage arrangement despite the prospective buyer’s ability to obtain mortgage financing. The evidence indicates that plaintiffs were of the equivalent economic status as many whites who routinely obtained mortgages to finance the purchase of houses and that a competing construction company in the black market sold the vast majority of its homes on deed and mortgage to blacks similarly situated economically to plaintiffs. Also, the evidence demonstrates that some plaintiffs made down payments of up to forty-five percent of the contract price—well above the amount needed to qualify for mortgages—and yet defendants refused to deal on terms other than contract. On the basis of this evidence it could reasonably be inferred that defendants utilized the contract method of sales to facilitate their exorbitant pricing practices and not because of significant differences between plaintiffs’ economic status and that of whites similarly situated who were able to utilize mortgage financing. Hence, a jury could find that the different treatment accorded plaintiffs by defendants’ sales terms was discriminatory.

Furthermore, plaintiffs offered evidence, which was erroneously excluded by the trial judge, sufficient to establish a *prima facie* case pursuant to the traditional theory of discrimination. Under that theory there must be a showing of “treating, in similar circumstances, a member or members of one race different from the manner in which members of another race are treated.” Love v. DeCarlo Homes, Inc., 482 F.2d 613, 615 (5th Cir. 1973). That is, a black prospective buyer of a dwelling demonstrates discriminatory conduct if he proves that an owner utilizes different pricing policies with respect to blacks and whites similarly situated.

The proffered evidence which was rejected at trial involved the sales of new houses to white buyers in Deerfield, Illinois and Park Forest South, Illinois, both being suburban residential developments. The sellers of the houses in Deerfield were Universal Construction Company and a joint venture comprised of the Deerfield Home Development Company and Universal Builders, Inc., while the houses in Park

---

39 The trial court improperly excluded appraisal evidence of Walter Tomlinson and Francis Parker. See, e.g., Illinois Power & Light Co. v. Talbott, 321 Ill. 538, 543, 152 N.E. 486 (1926). Tomlinson and Parker were employed by savings and loan institutions and as a consequence thereof they had appraised certain of defendants' constructions, for the purpose of extending construction mortgage funds. Tomlinson was of the opinion that defendants' sales prices were, on the average, $4,099 or 20.4 percent in excess of appraised value while Parker's appraisals reflected an average excess of $3,729 or 16.6 percent over the fair cash market value of the respective houses.
Forest South were sold by the P.F.S. Development Company. The plaintiffs contended that these corporations were owned and managed by the same persons that owned and managed defendants and that these persons were engaged in discriminatory conduct through the use of different pricing practices in Deerfield and Park Forest South from those used in pricing defendants’ houses. The trial judge excluded the evidence on two grounds: (1) the plaintiffs were not allowed to disregard the separate identities of the three corporations and (2) the lack of a basis of comparability for homes in Deerfield and Park Forest South with those of plaintiffs in Chicago. We rule that the trial court erred in excluding plaintiffs’ evidence on the Deerfield and Park Forest South pricing policies.

Defendants’ argument that the corporate formalities should not be disregarded in the context of the issues is without worth. As this court has stated, corporate formalities “may be disregarded in exceptional situations where it otherwise would present an obstacle to the due protection or enforcement of public or private rights.” Ohio Tank Car Co. v. Keith Ry. Equip. Co., 148 F.2d 4, 6 (7th Cir. 1945), cert. denied, 326 U.S. 730, 66 S.Ct. 38, 90 L.Ed. 434 (1945). In situations such as here, where common ownership and management exists, corporate formalities must not be rigidly adhered to when inquiry is made of civil rights violations. Accordingly, the objection raised by defendants presents no obstacles to the comparison of defendants’ pricing policies implemented in Deerfield and Park Forest South by the other selling organizations.

Turning to the second reason given for excluding the Deerfield and Park Forest South evidence we find that there was sufficient comparability between those operations and defendants’ sales in the relevant black market so as to render the exclusion an abuse of discretion. Factors such as geographical proximity, the date of the sale, type of construction, and materials used are factors going only to the weight to be accorded the evidence by the jury and do not go to its admissibility. See, e.g., Winston v. United States, 342 F.2d 715, 721 (9th Cir. 1965); United States v. 124.84 Acres of Land, Warrick County, Ind., 387 F.2d 912 (7th Cir. 1968). Plaintiffs met the requirement of sufficient comparability through the use of a comparative statistical analysis of accounting data reflecting defendants’ sales operations and pricing policies in Chicago with those of Deerfield and Chicago and with those of Deerfield and Park Forest South. We find plaintiffs’ statistical evidence to have been sufficiently competent; see United States v. Certain Interests in Property, etc., 326 F.2d 109 (2d Cir. 1964), cert. denied, 377 U.S. 978, 84 S.Ct. 1884, 12 L.Ed.2d 747 (1964), and probative of plaintiffs’ claim that defendants sold houses to blacks on price terms different from those they sold to white buyers similarly situated. Indeed, as the Eight Circuit has stated, “statistical evidence can make a prima facie case of discrimination.” Carter v. Gallagher, 452 F.2d 315, 323 (8th Cir. 1972), cert. denied, 406 U.S. 950, 92 S.Ct. 2045, 32 L.Ed.2d 338 (1972).

Plaintiffs’ expert witnesses testified that in the housing industry prices are established on the basis of direct costs, consisting generally of the investment in the land and the cost of construction, including materials. Once the direct costs are calculated, an allowance for overhead and profit is added to the direct costs to attain the sales price. The allowance for overhead and profit is added to the direct costs to attain the sales price. The allowance for overhead and profit is the gross profit on sales which plaintiffs’ expert witness testimony indicated was generally found in the real estate industry to be from fifteen to nineteen percent of the sales price.40 Defendants testified to the utilization of this method of pricing in their sales to plaintiffs; however, the statistical evidence presented by plaintiffs tends to refute that assertion.41

---

40 To the extent that differences in land location and building materials affect sales price, those differences are reflected in the absolute amount of direct costs. The differences do not affect the pricing methods employed. That is, the expert testimony and defendants were agreed that the type of house or area in which it is built does not affect the method of setting prices. The selling price is arrived at by establishing direct costs and adding an amount for overhead and profit. The amount added for overhead and profit represents the gross profit and is an important comparative figure in analyzing the substance of the pricing methods employed by defendants.

41 Plaintiffs’ statistical evidence was prepared by John Royer, an independent certified public accountant. To assure a substantial basis for statistical comparability between the Chicago and suburban operations, Royer made appropriate adjustments for differences in accounting methods and financial presentations between the various operations. Moreover, with respect to percentage comparisons of gross profit and mark-up, Royer testified that before calculating such percentages he determined that the components of direct costs were the same for the suburban companies as for the defendant companies.
Viewing the evidence, first in absolute terms, it shows that defendants’ pricing policy in Chicago produced an average gross profit substantially in excess of that produced by the Deerfield and Park Forest South operations. In relative percentage terms defendants reaped a gross profit of 27.6 percent of sales, well above the industry figure of 14 to 19 percent and considerably in excess of the gross profit percentages of the Deerfield and Park Forest South operations. Second, analyzing the same data in terms of the mark-up of the sales price over the direct costs, it is clear that the mark-up-as a percentage of direct costs-was much higher in the sales to plaintiffs than in the sales to white buyers in Deerfield and Park Forest South. The statistical evidence substantiates the claim that defendants priced plaintiffs’ houses much higher relative to direct costs than the houses sold in Deerfield and Park Forest South and belies any defendants utilized comparable pricing policies in, their sales to plaintiffs. Plaintiffs presented Dr. Richard Freeman, a Professor of Economics, who analyzed the statistical data pertaining to the difference in gross profits between defendants’ sales to plaintiffs and suburban operations. His analysis demonstrated that the difference in pricing practices was due to the race of the buyer and not economic factors.

In summary, it is difficult as a prima facie matter to infer that the substantial disparity between the pricing practices of defendants in the black real estate market and the pricing practices in the white market was attributable to some factor other than the race of the buyers. Whether defendants afforded plaintiffs the “same right” to purchase housing as offered white buyers was, based on the foregoing evidence, an issue to be properly submitted to the jury.

**QUESTIONS**

1. Assume that all of the factual contentions made by plaintiffs are true. Looking to the future, would application of the legal theory advocated by plaintiffs be more likely to help or to hurt potential home purchasers who are blacks?

---

42 In those instances where direct costs were comparable in absolute amounts, the gross profit garnered by defendants was substantially in excess of that obtained in the suburban operations. For example, the following chart compares the operations of Universal Construction Co. in Deerfield with defendants’ efforts in Chicago:

<table>
<thead>
<tr>
<th></th>
<th>(1) Universal Construction Company</th>
<th>(2) Chicago Operation</th>
<th>(3) Chicago Difference (Column 2 Minus Column 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Sales Price</td>
<td>$22,644.60</td>
<td>$25,172.53</td>
<td>$2,527.93</td>
</tr>
<tr>
<td>Average Direct Costs</td>
<td>$18,779.45</td>
<td>$18,246.14</td>
<td>$-553.31</td>
</tr>
<tr>
<td>(Land Plus Building)</td>
<td>$18,779.45</td>
<td>$18,246.14</td>
<td>$-553.31</td>
</tr>
<tr>
<td>Average Gross Profit</td>
<td>$3,865.15</td>
<td>$6,926.39</td>
<td>$3,061.24</td>
</tr>
</tbody>
</table>

Although the average direct costs were slightly lower in the Chicago operations, the average sales price and average gross profit were substantially higher in those operations than in the Deerfield operations.

43 In Deerfield, Universal Construction realized a gross profit of 17.1 percent while the gross profit of the Deerfield Home-Universal Builders joint venture was 17.6 percent. The operations in Park Forest South turned an average gross profit of 15.5 percent of sales price.

44 The average mark-up or ratio of gross profit to direct costs was 20.58 percent for Universal Construction while the Deerfield Home joint venture operations reflected a mark-up of 21.42 percent; the mark-up in Park Forest South was 18.35 percent. In contrast, defendants’ markup on their sales to plaintiffs averaged 37.96 percent.
2. A “market-clearing” price is one which equates quantity demanded at that price with the quantity supplied. Does the court’s holding require a builder such as Universal to charge less than a market-clearing price? If so, how does the quantity available get rationed out among the potential purchasers? Would resold houses presumably be subject to the same rule?

3. One reading of this case is that the seller must sell the houses at the price they would have sold for “but for” the segregated condition of the housing market? What difficulties would there be in knowing exactly what this price would be? Could the rule be attacked as unconstitutionally vague?

4. The assumption in this case is that racial bigotry and segregated housing markets will result in higher housing prices for blacks. In fact, the same premise is consistent with a result in which the price of housing to the discriminated-against group is lower rather than higher. What are the theoretical effects that work in opposite directions on the price of housing for blacks? How would one know what the true empirical result is? See Lapham, “Do Blacks Pay More for Housing,” 79 J.Polit.Econ. 1244 (1971).

5. Is entry into the house building industry in downtown Chicago limited by any substantial obstacles? What implications can be drawn from your answer?

6. Did the defendant have a “monopoly” or substantial price-fixing market power over any product? If so, what? Of what relevance is cross elasticity of demand in answering this question?

7. Which of the factual assertions of plaintiffs’ case seem most dubious? What kind of empirical evidence or theoretical arguments might be adduced to rebut them?

8. Pursuant to the remand ordered by the Seventh Circuit, the Clark case was subjected to a full trial on the merits several years later. The opinion in that case is excerpted below. Before reading the opinion, however, try to guess what evidence would have been presented by each party and how persuasive the evidence would have been.

---

**CLARK V. UNIVERSAL BUILDERS, INC.**

*United States District Court, N.D.Illinois, 1983.*

No. 69 C 115.

BUA, Judge.

---

**III. The Traditional Theory of Discrimination**

The traditional theory of discrimination requires proof that the defendants utilized different pricing policies in their sales of homes to the plaintiffs than they did in selling comparable homes in Deerfield to similarly situated white homebuyers.

A. Plaintiffs failed to show either that the Deerfield and South side areas were comparable community environments or that the houses built in those communities were similar in terms of lot size, building materials, design and layout.

The plaintiffs attempted to prove that homes sold to whites by Universal Construction Company and Deerfield Home Development Company (DHDC) were comparable to those sold to the members of the plaintiff class by the South side land companies. This they failed to do.

The evidence shows that Universal Construction Company and DHDC sold approximately 50 homes in Deerfield between 1953 and 1958. During these years, Deerfield was a developing suburb that still retained a somewhat rural character. Ample evidence of this fact was presented at trial. For example, it was not until 1958 that the Tri-State connector was opened linking Deerfield to the Tri-State highway system, and the Kennedy-Evans expressway system was not opened to link Deerfield to downtown Chicago until late 1960. Additionally, Deerfield did not have a high school during this period, making it necessary for local high school students to attend school in Highland Park. Finally, Deerfield residents were dependent on local roads and a single railroad line for commuter service to Chicago. In contrast, the South side neighborhood in which the defendants operated was a mature urban environment with extensive mass transit and railroad service available to its residents, and with shopping, schools, churches, parks and other facilities in the area.
In addition to the dissimilarity between Deerfield in the years 1953 to 1958 and Chicago in the years 1958 to 1969, the evidence indicates that the homes built in Deerfield were not comparable to the homes sold to plaintiffs. The plaintiffs offered the testimony of Edward Noonan, an architect and real estate developer, in their attempt to establish the comparability of the Deerfield and South side homes.

Mr. Noonan’s testimony that the homes in Deerfield were comparable in size and superior in quality to the South side homes was based on his inspection of two homes in Deerfield constructed by DHDC and 12-20 homes on the South side. All of the inspected homes were selected by plaintiffs’ counsel. Moreover, Mr. Noonan did not identify which models he inspected, and did not state whether he inspected one model or 20. It is to be noted that defendants constructed approximately 30 different home models on the South side. Mr. Noonan conceded that houses with the same direct costs of construction cannot be said to be comparable from the point of view of “living or location,” or “value to the purchaser.” Noonan, p. 678. In fact, he admitted that sales prices are not necessarily related to construction costs.

Additionally, the ranch houses built in Deerfield in 1956 and 1957 were of frame construction. Many did not have basements. In contrast, the South side homes were brick and were typically built with basements. In addition, many of the South side homes were bi-level, split level, or two story homes. In sum, this court concludes that Mr. Noonan’s testimony is unpersuasive on the issue of comparability, and that the evidence as a whole shows a complete dissimilarity between the Deerfield and South side homes in terms of community environment, lot size, building materials, design, and layout.

B. Evidence of the Appraised Value of the Deerfield and South side homes does not support the inference that defendants discriminated against plaintiffs in setting the sales price.

In order to sustain their burden under the so-called traditional theory of discrimination, the plaintiffs submitted in evidence appraisals of the value of homes built in Deerfield and on the South side. These appraisals were made by a number of savings and loan associations that provided construction and long term mortgage financing to the defendants. At the time these appraisals were made, Illinois and federal law required savings and loan associations to prepare written appraisals of the value of any real estate given as security for a mortgage loan. The appraisal reports were offered as evidence of the fair market value of the homes appraised in an effort to show that the average sales price approximated the appraised value for the Deerfield homes but greatly exceeded appraised value for the South side homes. Plaintiffs argue that this evidence demonstrates that the defendants’ pricing policies on the South side were discriminatory. This court is not persuaded.

The testimony establishes that an appraisal is a subjective estimate of property value that may vary considerably with regard to a single parcel of real estate depending on the individual responsible for the appraisal. Plaintiffs presented no persuasive evidence that the appraisal reports offered in evidence are reliable indicators of the fair market value of the homes appraised. In fact, the testimony of defendants’ expert Mr. Shlaes indicates that there is no necessary relationship between an appraised value and fair market value or actual sales price.

Additionally, plaintiffs’ own witnesses in several instances refused to assert that the appraisal reports reflected fair market value. For example, Mr. Langworthy of Rogers Park Savings & Loan Association testified that appraisals made for that association were used for the purposes of determining the size of the loan that could be approved for a particular home. Likewise, Mr. Del Campo, a field appraiser for three years, testified that the appraisal reports to Fairfield Savings & Loan Association were made to enable Fairfield to render an opinion of the mortgage dollar amount it might grant to a borrower. Testimony also indicated that the appraisal values for homes financed by Chicago Federal Savings & Loan Association were based solely on an estimate of the reproduction cost of a home derived from cost manuals containing data six months to a year old. The Chicago Federal appraisals were not intended to reflect the fair market value of the property appraised.

The testimony of defendants’ expert, Mr. Shlaes, casts a further shadow on the probative value of the appraisals. His testimony indicates that mortgage lenders “tend to use the appraisals they commission to support their own decisions which are based on underwriting concerns [i.e.] whether or not they want to make loans at the time or in the area to the borrower.” (Shlaes 1498). If the savings and loan association is reluctant to make the mortgage loan, it may express or imply to the appraiser a desire for a
low appraisal so that it can avoid making the loan in the amount requested. If the lender is eager to make the loan, it may encourage the appraiser to take a liberal approach in his or her estimate of value.

Shlaes’ testimony also corroborates the testimony of Mr. Tomlinson that at the time the appraisals were made cost manuals relied on in estimating reproduction costs did not contain current data. Moreover, plaintiffs’ comparison of the average appraisal and average sales price of homes sold by Universal Corporation Company and DHDC in Deerfield with 15 homes sold on the South side in the year ending January 31, 1958 fails to support an inference that the defendants were using different pricing policies on the South side than were used in Deerfield. Plaintiffs’ argument predicated on averages is a distortion. It avoids individual comparisons of price to appraisal. Individual comparisons show that on the South side some of the ratios of average sales price to average appraisal were higher and some were lower. The range of such averages relating to both Deerfield and South side homes were similar.

C. Comparisons of gross profits realized on homes sold on the South side with those sold in Deerfield do not justify a conclusion that defendants utilized discriminatory pricing policies in their sales of homes to plaintiffs.

The plaintiffs also argue that evidence of the gross profits made by defendants on the South side homes when compared to the gross profits made on homes sold by Universal Construction Company and DHDC in Deerfield justifies the inference that defendants utilized different pricing policies on the South side than did Universal Construction and DHDC in Deerfield.

The plaintiffs have presented voluminous evidence comparing the “gross profit ratios” of Deerfield and South side homes. A “gross profit ratio” is obtained by dividing gross profit by the total sales amount, and is expressed as a percentage.

The plaintiffs presented testimony in an attempt to demonstrate that during the period 1957 to 1969, there existed a standard gross profit ratio of between 15% and 20% which was generally sought by members of the home-building industry. This court finds that the preponderance of credible evidence does not support such a conclusion.

Plaintiffs’ evidence consists, in the main, of the testimony of Mr. Edward Noonan and Mr. Herbert Rosenfeld. Mr. Noonan has been identified earlier in this opinion. See p. 18, supra. Mr. Rosenfeld is an accountant and real estate developer and was a one-third owner of Surf Builders from 1963 to 1969. Surf Builders was one of the defendants’ competitors during this period, constructing and selling homes in the same area as the defendants on the South side of Chicago.

Mr. Noonan’s testimony does not establish that home builders set prices based on an estimate of cost plus a standard gross profit. Rather, his testimony establishes at most that, in his opinion, builders during the period 1957 through 1969 enjoyed on the average a 15%-20% gross profit ratio. Mr. Rosenfeld’s testimony indicates that his company sold approximately 150 homes during this period to black purchasers. In choosing plans for homes to be constructed he first viewed Universal’s models and noted defendants’ prices. While his testimony at one point indicates that he hoped to make a gross profit equal to 15% to 20% of sales price, he admitted that the gross profit figure was determined by market conditions and the amount of earnings the company wished to achieve. Like Noonan’s, Herbert Rosenfeld’s testimony does not support plaintiffs’ contention that a standard gross profit percentage existed or that such a desired percentage dictated pricing policies.

Therefore, it is difficult for this court to infer that a comparison of gross profit percentages of homes sold on the South side with homes sold in Deerfield evidences discriminatory pricing by the defendants. Moreover, the defendants presented credible evidence that there is no standard gross profit percentage in the housing industry and that no standard gross profit percentage is used as a price setting mechanism.
Defendants offered the testimony of Martin Bartling, Jr. on this issue. Mr. Bartling has been the executive vice president of the Home Builders Association of Greater Chicago since February 1, 1978 and in 1960 was the president of the National Home Builders Association. Prior to moving to the Chicago area, he had over 23 years experience as a builder of single family homes in Tennessee. Mr. Bartling credibly testified that he had never heard of a standard gross profit in all of his years in the industry or that builders set their sales prices with regard to a formula based on or related to a standard gross profit. This court finds it hard to believe that a standard gross profit or standard gross profit ratio could have existed or been recognized by real estate developers or home builders as an aid in setting sales prices in the face of this testimony. Mr. Schlaes' testimony corroborates this conclusion. Mr. Turoff's testimony indicates that prices were set after taking into account estimated costs including indirect costs, competitive factors, and differences between models with an eye toward achieving a net profit of 10%-15%. No attempt was made to determine gross profit. In setting prices, a variable dollar figure was added to estimated direct costs. This figure varied depending on overhead, intention to advertise, the number of homes it was believed would be sold and competition. Price adjustments were then made to preserve marketing differences between different model homes.

IV. The Exploitation Theory of Discrimination

The Court of Appeals' opinion in this case, Clark v. Universal Builders, Inc., 501 F.2d 324 (7th Cir. 1974), held that “plaintiffs state a claim under section 1982 since they allege that (1) as a result of racial residential segregation dual housing markets exist, and (2) defendant sellers took advantage of this situation by demanding prices and terms unreasonably in excess of prices and terms available to white citizens for comparable housing.” This court begins by noting that judicial notice has been taken of the fact that during all time periods relevant to this case Chicago residential patterns have been characterized by a high degree of racial segregation.

Dual Housing Markets

In order to sustain their burden of proof on the existence of “dual housing markets” plaintiffs relied primarily on the testimony of Karl E. Taeuber, Ph.D.. Mr. Taeuber is a Professor of Sociology and the Assistant Director of the Institute for Research on Poverty at the University of Wisconsin. His area of specialization within sociology is demography. The McGraw Hill Dictionary of Scientific and Technical Terms, Second Edition, defines demography as, “The statistical study of populations with reference to nationality, mortality, migratory movements, age, and sex, among other social, ethnic, and economic factors.”

Prof. Taeuber has studied and published on the demographics of the black population, residential and school segregation, and population redistribution trends in the United States. His educational and professional background do not, so far as the evidence shows, include training, experience, or expertise in the field of economics, more specifically in the area of price theory and the effects of racism or racial segregation on prices.

Prof. Taeuber testified that there was a “dual housing market” in the Chicago metropolitan area during the 1950’s and 1960’s. The segments of the dual housing market cannot be identified on a geographical basis. As used by Prof. Taeuber, the term “dual housing market” refers to the operation of two separate markets: in one market housing is made available almost exclusively to whites, and in the other market, housing is made available primarily to blacks. The so-called white market in housing presents a number of barriers to the entry of blacks as a result of several factors including race discrimination on the part of white sellers. The so-called black market presents few impediments to the entry of white buyers other than the lack of white demand for housing in areas that are commonly thought of as populated by black residents.

Prof. Taeuber testified that the ability too afford housing does not explain the pattern of residential segregation because both blacks and whites are represented at all income levels and that most members of both races are in middle income brackets. There was no testimony as to the distribution of income by race, however, and defendants’ expert, Pierre De Vise, testified that there were both black and white residents in virtually every census tract in the metropolitan area. Thus, Taeuber’s testimony does not exclude the ability to afford housing as a factor affecting racial residential patterns.
Taeuber also testified that most areas of Chicago contain a mix of high and low-priced housing and thus, even considering the ability to afford housing as a factor affecting residential patterns in the city, this factor would not serve to explain the exclusion of blacks from those areas of the city containing both high and low-priced housing. However, as pointed out above, exclusion of blacks from identifiable areas is a relative thing and the existence of racial discrimination cannot alone serve to explain racial residential patterns.

Taeuber concludes that blacks suffer from a form of discrimination different from that encountered by white ethnic minorities. However, such a conclusion is not relevant to the issue of how that discrimination affects the price that blacks must pay for housing. Taeuber’s testimony indicates that when various immigrant groups came to Chicago from Europe, their initial areas of settlement were concentrated in core residential areas. As they gained experience in the city and secured a sound economic base, they quickly adopted the American preference for living in single family dwellings removed from poor areas and this resulted in their dispersal throughout the metropolitan area. However, black “migration” to the city has not resembled this pattern. Although black “immigrants” from the South, after the first World War, had the advantages of speaking English, citizenship, and a preference to live in single family residences removed from poor areas, the same degree of dispersal did not take place. According to Taeuber, these observations support the conclusion that race discrimination caused residential segregation.

While race discrimination has clearly been a substantial contributing factor in creating patterns of past and present residential segregation, the evidence offered by plaintiffs through Taeuber is inadequate to persuade this court that these conditions presented the defendants with the ability to charge a price unreasonably in excess of that which white homebuyers would be charged for comparable housing, or a price in excess of what plaintiffs would have paid for their homes in a perfectly integrated market. Thus, try as it might, this court is unable to find that plaintiffs have met their burden of showing that the patterns of residential segregation existing in Chicago and contributed to, in significant degree, by race discrimination, presented the defendants with the opportunity to charge exorbitant prices amounting to a tax on plaintiffs’ race.

In attempting to show that market conditions tainted by race discrimination made it possible for defendants to charge excessive prices to black homebuyers, Taeuber testified that population trends are related to the need for additional housing. Taeuber stated that an increase in population, or an increase in the number of households in an area, causes a need for additional housing. However, such testimony bears little, if any, relation to the defendants’ ability to extract exorbitant prices for their homes. A desire for single family homes cannot be translated into a statement of economic demand for housing that affects its price.

A simple and common example suffices to illustrate the leap of logic made by Taeuber in offering his opinion to the contrary. It is clear at the present time that there are many people who desire to purchase single family homes who do not constitute part of the demand for those homes available due to economic reasons. On the other hand, everyone needs housing but many prefer to rent, or are unable to afford the down payment or make the periodic payments necessary to purchase. In short, this court has been presented with absolutely no persuasive evidence from which to infer that market conditions in the so-called black segment of the “dual housing market” presented the defendants with the opportunity to create a monopoly or oligopoly for single family homes in their market area. In other words, there is no persuasive evidence from which this court can conclude that the demand for new homes among blacks exceeded the supply to a proportionately greater extent than among whites, giving defendants the opportunity to exploit the “dual housing market” by charging plaintiffs unreasonably high prices. This court simply does not believe that Taeuber’s testimony regarding the “demand” for new single family housing among blacks is credible, and the court is not persuaded that Taeuber has the expertise to draw such a conclusion. Furthermore, given the plaintiffs’ failure to show by a preponderance of the evidence that a standard gross profit existed in the housing industry, this court is unwilling to draw the conclusion that the difference in gross profits earned on the South side homes and the Deerfield homes supports the inference that defendants were able to exploit plaintiffs by setting prices at an unreasonable level.
In addition, the testimony of defendants’ experts, Brian Barry, Pierre De Vise and Yale Brozen, support the conclusion that the defendants did not, and were unable to, extract higher prices from plaintiffs because of the operation of a “dual housing market” and that the prices charged would not have been lower in the absence of the pattern of racial residential segregation existing in Chicago during the years defendants operated on the South side.

The fact that defendants sold homes to plaintiffs on installment contracts, rather than through a mortgage financing arrangement, does not support the inference that the homes were sold on terms more onerous than those available to white citizens. The record in the present case demonstrates that white citizens do purchase homes on installment contracts and that white citizens did purchase homes on contract from Universal Construction Company and DHDC. In addition, none of the witnesses stated that they actually applied for a mortgage in connection with the purchase of these homes, nor did the plaintiffs present any evidence from which this court could determine whether they would have qualified for a mortgage if in fact they had applied. Plaintiffs’ evidence with respect to the size of the down payments made is inconclusive on this question. Also, many homes were sold with down payments substantially below the then current 10% standard down payment.

More significant, however, is the fact that the installment contracts contained a full prepayment privilege without penalty which allowed the buyer to obtain mortgage financing at any time. Additional evidence that the terms of the sale were not onerous is the fact that 275 of the purchasers were represented by attorneys and they advised their clients to proceed with the transaction. Furthermore, 18 of the homes were sold with mortgage financing. On the basis of the above factors, the court is unwilling to conclude that the terms of the sale were onerous or oppressive. In addition, the court is not persuaded that defendants engaged in deceptive practices related to the negotiation of the contracts of sales.

V. Conclusion

The above findings of fact, analyzed in relation to the applicable legal standards, results in the following legal conclusions. In order to establish a traditional violation of 42 U.S.C. Sec. 1982, plaintiffs must prove by a preponderance of the evidence that a seller of realty has sold comparable homes in comparable areas to black persons at prices which were higher, or on terms which were more onerous than those sold to white persons. Jones v. Alfred J. Mayer Co., 392 U.S. 409, 88 S.Ct. 2186, 20 L.Ed.2d 1189 (1968). In addition, plaintiffs must demonstrate that the difference in prices or terms was due to the race of the purchaser, and that the seller’s actions were motivated by a discriminatory purpose or intent. Mescall v. Burrus, 603 F.2d 1266 (7th Cir. 1979); Resident Advisory Board v. Rizzo, 564 F.2d 126, (3d Cir. 1977), cert. denied, 435 U.S. 908, 98 S.Ct. 1458, 55 L.Ed.2d 499 (1978); Chicano Police Officer’s Association v. Stover, 552 F.2d 918 (10th Cir. 1977); Croker v. Boeing Co., 437 F.Supp. 1138 (E.D.Penn. 1977); Lewis v. Bethlehem Steel Corp., 440 F.Supp. 949 (D.Md.1977); cf. Washington v. Davis, 426 U.S. 229, 96 S.Ct. 2040, 48 L.Ed.2d (1976); Village of Arlington Heights v. Metropolitan Housing Development Corporation, 429 U.S. 252, 97 S.Ct. 555, 50 L.Ed.2d 450 (1977); Mt. Healthy City School District v. Doyle, 429 U.S. 274, 97 S.Ct. 568, 50 L.Ed.2d 471 (1977). The evidence in the present case failed to demonstrate by a preponderance of credible evidence that the Deerfield and South side community areas were comparable, or that the homes sold were similar in design, layout, lot size, or building material. In addition, the evidence presented by plaintiffs, specifically the evidence related to appraised values, comparison of gross profits, and the method of financing purchases, does not persuade this court that the seller’s actions were motivated by a discriminatory purpose or intent. For these reasons, plaintiffs have failed to establish a violation of 42 U.S.C. Sec. 1982 under the traditional theory of discrimination.

Plaintiffs have also offered evidence to establish a violation of 42 U.S.C. Sec. 1982 under a theory of exploitation of a dual housing market. To establish a violation of 42 U.S.C. Sec. 1982 under the “exploitation” theory, plaintiffs must establish by a preponderance of credible evidence that as a result of racial residential segregation dual housing markets exist and that defendants took advantage of this situation by demanding prices and terms unreasonably in excess of those prices and terms available to white citizens for comparable property. Clark v. Universal Builders, Inc., 501 F.2d 324 (7th Cir. 1974); Mescall v. Burrus, 603 F.2d 1266 (7th Cir. 1979). Plaintiffs have demonstrated by a preponderance of credible evidence the exis-
tence of a dual housing market that exists by virtue of the historical patterns of residential segregation in the City of Chicago. Plaintiffs have failed to demonstrate, however, that defendants were presented with the opportunity to extract unreasonable prices or terms which amounted to a tax on plaintiffs’ race, or that the prices and terms charged were motivated by a discriminatory purpose or intent.

Finally, plaintiffs have alleged that defendants’ activities were conducted in a willful and wanton manner and that plaintiffs are therefore entitled to punitive damages. As previously stated, the court has concluded that plaintiffs have not established by a preponderance of the evidence the claimed statutory violations. Even assuming that these violations had been established, the court concludes that punitive damages could not be imposed as a matter of law. The alleged statutory violations occurred prior to Jones v. Mayer, 392 U.S. 409, 88 S.Ct. 2186, 20 L.Ed.2d 1189 (1968) which held for the first time that Sec. 1982 was applicable to private action until the decision in Griffin v. Breckenridge, 403 U.S. 88, 91 S.Ct. 1790, 29 L.Ed.2d 338 (1971). As a matter of law, one cannot be held to have willfully violated a statute which one would have legitimate reason to conclude was inapplicable. See generally Jones v. Mayer, 392 U.S. 409, 88 S.Ct. 2186, 20 L.Ed.2d 1189 (1968); Lee v. Southern Home Sites Corporation, 429 F.2d 290 (5th Cir. 1970).

All findings of fact deemed to be conclusions of law, and all conclusions of law deemed to be findings of fact are incorporated herein by reference.

For the foregoing reasons, the court concludes that the plaintiffs did not meet their burden of proof and judgment is therefore entered in favor of defendants.

QUESTIONS

1. The opinion relates that “the testimony of defendants’ expert Mr. Shlaes indicates that there is no necessary relationship between an appraised value and fair market value or actual sales price.” Does this testimony surprise you? Is there a sense in which it could be true and not be damaging to the plaintiffs’ evidence (consisting of a group of appraisals)?
2. What about the testimony that mortgage lenders “tend to use the appraisals they commission to support their own decisions which are based on underwriting concerns [i.e.] whether or not they want to make loans at the time or in the area to the borrower”? Is this plausible? Is it damaging?
3. The court regarded as the “Keystone” of plaintiffs’ argument their assertion that builders and developers set prices in relation to the direct costs of construction and seek to recover a certain gross profit. Gross profit is the difference between sales price and the direct costs of constructing the building and purchasing the site. Other “indirect” costs were acknowledged as including items such as wages for clerical help, salary, and other expenses associated with a business but which cannot be directly allocated as part of the cost of any particular house. Can you think of any indirect costs which might systematically be greater (or smaller) in downtown Chicago than in the suburbs?
4. In comparing two somewhat different products, would you expect the profit margins (or “markups”) over all costs to be very similar if computed over some reasonable length of time? Why? Would the same hold true for subcategories of costs such as variable costs, direct costs, etc.?
5. In deciding that Deerfield and downtown houses are not comparable, the trial court alludes to significant differences in local facilities and infrastructure: mass transit, parks, shopping, schools, etc. Is the trial court incorrect in thinking (fn. 16) that the values of such amenities would be reflected in the cost of land acquisition and, hence, in the direct cost comparisons?
6. Assume that the defendants in fact had no exploitative power. Why would they have been willing to sell on contract to some buyers who presumably could not qualify for a mortgage loan? (After all, doesn’t the mortgagee bank in effect have title to the house much as the contract seller does?) Would it provide one hypothesis if you found out that there was a state usury law? If there were no usury law, is there an alternative explanation: that the developer is a better risk-bearer than an ordinary lending institution? Why?
7. Plaintiffs offered evidence that at least one of Universal’s competitors, Surf Builders, sold what were alleged to be comparable homes on substantially better terms (lower prices, mortgages instead of
contracts, etc.). Are we to understand that Surf sold below market price or that Universal sold above? Surf sold homes in the same neighborhood, and therefore presumably the same market, as Universal. What could account for substantially the same product selling at two different prices in the same market?

8. If Universal were indeed attempting to be exploitative, why might it prefer the contract sale to sale via mortgages?
9. On what basis could defendants’ experts have testified that “defendants did not, and were unable to, extract higher prices” due to what was admittedly a “dual housing market.” Is that conclusion a logical proposition derived from economic theory or is it an empirical question?
10. If you could decide the issues of law and fact in Clark, what would your holdings be? If there are additional factual inquiries that would affect your opinion, indicate what they are, why they are relevant, and what sort of evidence might resolve them.

THE THEORY OF THE SECOND BEST

Many of us learned at our mother’s knee the old adage that “Two wrongs don’t make a right.” Nonetheless, an economic axiom known as the “Second Best Theory” cautions us that, the time-honored cliché may sometimes be incorrect, that two wrongs may indeed be better than only one.a

Reduced to simple terms, the Theory of the Second Best points out that what appears to be an improvement in one part of an interdependent process or system may have undesirable feedback effects in another. (In the field of systems analysis, a similar problem is captured in the warnings against “suboptimizing,” i.e., optimizing the solution to one component part of a large problem while disregarding the effects on other aspects of the overall problem.) In the economic context, the Second Best Theory teaches that we can be sure that perfecting one market will also improve overall resource allocation only when all markets except the one being tinkered with are already working perfectly. In this case, the solution reached is a “First Best.” If, however, some of the other markets are functioning incorrectly, optimization of a related market may make things worse, rather than better, from an overall perspective.

Most of the applications of the Second Best theorem in the economic and legal literature have been drawn from antitrust and regulatory economics. b A typical contrived pedagogical example will illustrate. An earlier section in this chapter suggested why “too little” of a monopolized product is sold as judged by the product’s production cost and its value in use: price is set higher than marginal cost. Suppose, then, that potatoes were monopolized and pasta (spaghetti, macaroni, noodles, etc.) were not. Too few potatoes would be consumed, but some of the demand for starchy food would be shifted to a substitute, pasta. Relative to the First Best quantities of potatoes and pasta that would be purchased if both markets were functioning ideally, we would now find too few potatoes and too much pasta being consumed. A Second Best theorist would then suggest that, rather than allow the competitive pasta market to produce out to where P = MC for pasta, pasta should be taxed with the purpose of artificially raising the pasta price above its (net of tax) marginal cost and causing consumers both to cut back on pasta and also shift some of their demand back to potatoes. This Second Best policy involves, in effect, the interposition of a second distortion in order partially to counterbalance a preexisting distortion. The First Best solution would, of course, be to get rid of the monopoly in the potato market in the first place. Although the monopoly application is technically correct, students of the law frequently find the application an neficicious one. One reason is of course, that the distributive consequences of the Second Best policy are in this case seen as improving the position of the monopolist, even though the efficiency of resource use is arguably improved. Once one understands the principle of offsetting distortions, however, more attractive applications of the Second Best principle can easily be supplied.

---

a Although the notion involved is a Lancaster, "The General Theory of the common-sense one, the theorem was first Second Best," 24 Rev. Econ. Stud. 11 formalized and named by Lipsey and (1956).

One rather persuasive example is supplied by the problem of pricing urban transportation. Ordinarily, economists would argue that the cost of a ride on an urban mass transportation system should be set equal to its marginal cost. Since consumers will buy rides up to the point where the marginal benefits of a ride equals its price, a price below marginal cost will result in inputs being used that produce less output value than their resource cost: \( MB = P \) and \( P < MC \), and therefore \( MB < MC \). In this case, however, little or no charge at all is made for the use of public highways, one of the substitutes for mass transit rides. By subsidizing transit rides, travelers may be pulled away from the highways, where their resource use is even more wasteful. First Best would be an appropriate user-charge for highways rides, but if this is impractical the Second Best may require “too low” prices for mass transit rides.

In the legal literature, a thought-provoking example of the application of Second Best theory is James Henderson’s analysis of the extension of strict liability to segmented markets. See Henderson, “Extending the Boundaries of Strict Products Liability: Implications of the Theory of the Second Best,” 128 U.Pa.L.Rev. 1036 (1980). Imposition of strict liability is often premised upon some defect in the affected product market. Even assuming \textit{arguendo} that underlying premise, strict liability may make matters worse if the relevant market is divided into segments—such as commercial and non-commercial sellers—and one (or more) of these is not reachable by the strict liability. For instance, one of Henderson’s discussions compares the sale or lease of used products through commercial and non-commercial channels. If political or other exigencies require that non-commercial channels will not be included within the strict liability rule, then imposition of the rule on commercial markets may motivate a shift into non-commercial outlets. Why might this “half a loaf” be worse than an environment in which a “needed” strict liability is imposed on neither sub-market?

The basic logic of the Second Best principle is easily extended to other interdependent systems or processes, including adjudication. If a bias in one direction can be identified and it is impractical to correct that bias directly (the First Best solution), then the toleration or even creation of an offsetting bias may be desirable.

Taken to its logical extreme, the Second Best theorem is a counsel of agnosticism. Since it is seldom the case that one can have confidence that the rest of the system is optimally fine-tuned, it is technically true that apparent improvements of subsectors of the system may make things worse. Some might prefer, however, to couch the issue in terms of rebuttable presumption. An apparent improvement might then be regarded as a \textit{prima facie} case in favor of a policy and the burden shifted to those who oppose it. That this burden can sometimes be met is demonstrated by some of the examples given above.

Generally speaking, the Second Best principle will have a practical relevance in the economic environment when there is a defective but “uncorrectable” market that has a “closely related” market of either complements or substitutes. How does the policy reaction depend on whether the related market involves substitutes or complements?

\textbf{QUESTIONS}

1. Could a Second Best analysis be applied to the treatment of damages for future lost income? (Recall the knotty problems discussed in Chapter IV above.)

2. Are some kinds of punitive damages a reflection of the Second Best principle?

3. Outline in some detail how a parallel might be drawn between (a) Henderson’s example of the commercial and non-commercial markets and (b) the urban transport pricing problem. What is the loss if the Second Best principal is not applied?

4. Following on the example in the text above, suppose that the pasta market were monopolized. What implications does this have for the market in spaghetti sauce?

5. “In applying the Second Best theorem to economic systems, it is better to look at the way in which quantities are distorted than the way in which prices are distorted.” What does this mean? Return to this question again after you have reflected on the application in the following section.

STATEMENT OF U.S. DEPT. OF TRANSPORTATION BEFORE THE INTERSTATE COMMERCE COMMISSION, WESTERN COAL INVESTIGATION-GUIDELINES FOR RAILROAD RATE STRUCTURE

[The ICC was investigating appropriate rate guidelines for large-volume western coal movements originating in the Rocky Mountain-Upper Missouri River Region. These routes represent areas where the railroads occupy a strong competitive position with respect to alternative transport modes such as trucks or barges. Hence, the railroads would, with ICC approval, potentially be able to recover not only the variable costs of shipping on these routes but also to partially amortize fixed costs. DOT’s statement is sympathetic to the railroads’ desire to extract differentially higher prices on western coal shipments since the competitive situation makes it possible to do so and higher general rates on all shippers would otherwise be required to cover fully allocated costs.]

* * *

IV. Setting a Ceiling on Rates for Western Coal Rail Movements A. Considerations of National Energy Policy

* * DOT notes that this nation’s energy policy is based in large part on conservation of energy resources as well as on a shift, wherever possible, to coal, a plentiful energy resource, and away from scarce petroleum products. Much of our dependence on petroleum fuels is due to an artificial reduction in the price of those fuels through a history of price controls. DOT is anxious to assure that in shifting to coal, market tests of resource allocation efficiency are retained in order to prevent this sort of misallocation in the future. As long as the delivered price of coal does not rise above the delivered price of alternative energy resources, our nation’s fuel resources can be said to be efficiently and desirably allocated.

To assure that energy is appropriately priced and resources are properly allocated, DOT believes that the Commission should establish the principle that, in general, the maximum rate on coal is one that will equate the delivered price of coal per BTU to the delivered price per BTU of any other available fuel. Such a ceiling on transportation rates for coal would encourage a shift by utilities toward the use of coal and away from less desirable sources of energy.

Specifically, DOT believes that where a utility does have the opportunity to select from a number of fuels, including western coal, the Commission should not (except in a few circumstances described below) permit the imposition of a rate for large-volume coal movements which would result in the delivered price of coal being higher than the delivered price of other competing fuels. In a properly functioning market, this is the maximum a railroad could charge and still retain the traffic. However, where legal or institutional constraints require a utility to burn coal, DOT contends that the Commission should permit railroads to set rates up to a level that equates the delivered price of coal per unit of electricity with the delivered price of alternative fuels as though there were no price controls on those fuels. Such rate levels will contribute to the health of the railroads, allow allocational efficiency, and encourages conservation of energy.

B. The Market Structure For Fuels

In an efficiently functioning market, electric utilities would determine what fuel, as delivered to the plant, was least costly and would contract to purchase both the fuel and any transportation services needed to secure that fuel. Today’s markets for energy and transportation do not match this ideal. The institutions and information necessary to simultaneously get price quotes on all alternatives, and to make long-term contracts for the lowest cost fuel and transportation combination, are simply not
currently available. In the absence of contract rail rates, those utilities which choose (or are required) to employ coal for generation of electric power must enter into contracts to buy coal without knowing what their rail transportation costs will be over time. Another evident market defect is created by the price controls on oil and natural gas which make those fuels artificially attractive relative to coal and incidentally work to frustrate national energy policy. Finally, a host of laws, directly or indirectly, work to influence the decision as to what fuel a particular utility will burn.

D Establishment of a Rate Ceiling Where the Utility is Required to Utilize Coal

Where a particular utility is required to use coal to generate electricity, energy policy considerations are already achieved. Under those circumstances, economic efficiency and the policies of the 4R Act require only that the maximum rate for any new coal movements should leave the delivered price of coal per BTU no higher than the “economic cost” of using any other energy source. Such an analysis permits the setting of railroad rates at levels no higher than they could be set were there economically efficient markets for both energy and transportation.

In light of the serious market imperfections described in Section IV B above, DOT contends that in order to ascertain the maximum rail rate for coal, it is necessary to calculate a “shadow price” associated with the delivery and purchase costs of other fuels.

For a variety of historical reasons, price controls have been established on domestic oil and natural gas which encourage excessive use of those fuels relative to coal. This has artificially forced down coal prices and rail coal rates. Limiting coal rates to a level that left the delivered price of coal no higher than the artificially depressed delivered prices of oil would unreasonably reduce railroad revenues without improving allocation of fuel or transportation resources in the economy. The major reason why there has been pressure to shift from other fuels to coal has been to conserve dwindling supplies of natural gas and to reduce dependence on oil. It would be inconsistent to require use of coal to conserve use of other fuels while at the same time treating the price of those fuels as though they were lower than the free market level. Using the free market price of the alternative fuels, rather than the artificially low regulated price results in rate ceilings set on the basis of economic values of alternative resources. It also allows the railroads to obtain necessary revenues to maintain operations and to attract and retain capital.

QUESTIONS

1. In the text at footnotes 27 and 28, the DOT staff uses “serious market imperfections” to advocate higher delivered prices for coal. In which markets are the imperfections? How does this affect the coal market? How is the concept of cross elasticity of demand (see Ch. V) relevant here?
2. By the DOT argument, what would be the “First Best” solution to the alleged imperfections in energy markets? Why doesn’t DOT argue for that solution?
3. Is it fair to characterize DOT’s position as based on Second Best considerations?
4. Suppose that you represented western utilities who are obligated to burn coal. How could you argue that, on its own premises, the DOT recommendation is exactly 180 degrees around from the correct conclusion?
5. If the best argument were made in the previous question, is there any way in which the DOT position might be reformulated so that it at least arguably follows from a line of Second Best logic?

26 The economic cost of any resource is the price determined in the market in the absence of price controls or other constraints on the market.
27 A shadow price is the price which would exist had the market been operating properly and efficiently and had all alternatives been available to the utility at the time the decision to utilize coal was made.
28 The current government energy policy encouraging a shift from petroleum fuels to coal has, to some degree, corrected this artificial imbalance.
NELSON, D.J.

This is a libel to recover the value of a fin-back whale. The libellant lives in Provincetown and the respondent in Wellfleet. The facts, as they appeared at the hearing, are as follows:

In the early spring months the easterly part of Massachusetts Bay is frequented by the species of whale known as the fin-back whale. Fishermen from Provincetown pursue them in open boats from the shore and shoot them with bomb-lances fired from guns made expressly for the purpose. When killed they sink at once to the bottom, but in the course of from one to three days they rise and float on the surface. Some of them are picked up by vessels and towed into Provincetown. Some float ashore at high water and are left stranded on the beach as the tide recedes. Others float out to sea and are never recovered. The person who happens to find them on the beach usually sends word to Provincetown, and the owner comes to the spot and removes the blubber. The finder usually receives a small salvage for his services. Tryworks are established in Provincetown for trying out the oil. The business is of considerable extent, but, since it requires skill and experience, as well as some outlay of capital, and is attended with great exposure and hardship, few persons engage in it. The average yield of oil is about twenty barrels to a whale. It swims with great swiftness, and for that reason cannot be taken by the harpoon and line. Each boat’s crew engaged in the business has its peculiar mark or device on its lances, and in this way it is known by whom the whale is killed.

The usage on Cape Cod, for many years, has been that the person who kills a whale in the manner and under the circumstances described, owns it, and this right has never been disputed until this case. The libellant has been engaged in this business for ten years past. On the morning of April 9, 1880, in Massachusetts Bay, near the end of Cape Cod, he shot and instantly killed with a bomb-lance the whale in question. It sunk immediately, and on the morning of the 12th was found stranded on the beach in Brewster, within the ebb and flow of the tide, by one Ellis, seventeen miles from the spot where it was killed. Instead of sending word to Provincetown, as is customary, Ellis advertised the whale for sale at auction, and sold it to the respondent, who stripped off the blubber and tried out the oil. The libellant heard of the finding of the whale on the morning of the 15th, and immediately sent one of his boat’s crew to the place and claimed it. Neither the respondent nor Ellis knew the whale had been killed by the libellant, but they knew or might have known, if they had wished, that it had been shot and killed with a bomb-lance, by some person engaged in this species of business.

The libellant claims title to the whale under this usage. The respondent insists that this usage is invalid. It was decided by Judge Sprague, in Taber v. Jenny, 1 Sprague, 315, that when a whale has been killed, and is anchored and left with marks of appropriation, it is the property of the captors; and if it is afterwards found, still anchored, by another ship, there is no usage or principle of law by which the property of the original captors is diverted, even though the whale may have dragged from its anchorage. The learned judge says:

“When the whale had been killed and taken possession of by the boat of Hillman (the first taker) it became the property of the owners of that ship, and all was done which was then practicable in order to secure it. They left it anchored, with unequivocal marks of appropriation * * *”

In Swift v. Gifford, 2 Low. 110, Judge Lowell decided that a custom among whalemen in the Arctic seas, that the iron holds the whale, was reasonable and valid. In that case a boat’s crew from the respondent’s ship pursued and struck a whale in the Arctic Ocean, and the harpoon and the line attached to it remained in the whale, but did not remain fast to the boat. A boat’s crew from the libellant’s ship continued pursuit and captured the whale, and the master of the respondent’s ship claimed it on the spot. It was held by the learned judge that the whale belonged to the respondents. It was said by Judge Sprague, in Bourne v. Ashley, an unprinted case referred to by Judge Lowell in Swift
v. Gifford, that the usage for the first iron, whether attached to the boat or not, to hold the whale was fully established; and he added that, although local usages of a particular port ought not to be allowed to set aside the general maritime law, this objection did not apply to a custom which embraced an entire business, and had been concurred in for a long time by every one engaged in the trade.

I see no reason why the usage proved in this case is not as reasonable as that sustained in the case cited. Its application must necessarily be extremely limited, and can affect but a few persons. It has been recognized and acquiesced in for many years. It requires in the first take the only act of appropriation that is possible in the nature of the case. Unless it is sustained, this branch of industry must necessarily cease, for no person would engage in it if the fruits of his labor could be appropriated by any chance finder. It gives reasonable salvage for securing or reporting the property. That the rule works well in practice is shown by the extent of the industry which has grown up under it, and the general acquiescence of a whole community interested to dispute it. It is by no means clear that without regard to usage the common law would not reach the same result. That seems to be the effect of the decisions in Taber v. Jenny and Bartlett v. Budd. If the fisherman does all that it is possible to do to make the animal his own, that would seem to be sufficient. Such a rule might well be applied in the interest of trade, there being no usage or custom to the contrary. Holmes Com. Law, 217. But be that as it may, I hold the usage to be valid, and that the property in the whale was in the libellant.

The rule of damages is the market value of the oil obtained from the whale, less the cost of trying it out and preparing it for the market, with interest on the amount so ascertained from the date of conversion. As the question is new and important, and the suit is contested on both sides, more for the purpose of having it settled than for the amount involved, I shall give no costs.

Decree for the libellant for $71.05, without costs.

QUESTIONS

1. For expository simplicity, consider the whale industry to be composed of two stages of production, harpooning and finding. Harpooners actually produce two distinguishable “outputs”; (1) the harpooned whales that they are immediately able to retrieve; and (2) the “found” whales that initially got away. A case such as Ghen v. Rich determines the extent of the harpooners’ property rights in the second type of output. Let $/3$ be the fraction of a found whale’s value that a harpooner can appropriate (implicitly, the finder’s share is $1 - /3$) and $V$ the market value of a whale. If the fraction of whales that get away initially is $f$ and the fraction of those ultimately found is $x$, then the marginal revenue curve that a harpooner sees is $[ (1 - f) V + (x f) /3 V]$ for each whale harpooned. Why? Would a harpooner want the magnitude of $/3$ to be as high as possible?

2. A finder produces only one kind of output: found whales. The marginal revenue curve that a finder sees is $[ (1 - f) V + (x f) /3 V]$ since his share of the whale’s value is the remainder not kept by the harpooner. As the value of $/3$ moves toward zero, the finder’s cost curves are likely to shift upward. Why? From the finder’s perspective, what determines the optimal value of $/3$?

3. Do you think that the finders may have wanted to lose this case? Is there anything in the opinion pertinent to this question?

4. What is the “best” value for $/3$ from the standpoint of society: the harpooners, the finders, and the users of whale products? Why might this value change over time? Should a court readjust it? How else might it be done?

5. Show that any value set for $/3$ falls short of a “First Best” result. [Hint: Do both the harpooners and the finders underproduce? What would be the joint maximization outcome and why do they not achieve it?]

6. What problems in earlier chapters are related to that of the harpooners and the finders? Does this current context present any difficulties that were not involved in the earlier problems?

NONMARKET SYSTEMS: INTERLOCAL MIGRATION

Although markets are perhaps the most important and well analyzed type of interactive economic system, other important systemic processes are amenable to economic analysis. For instance, attempts to regulate local development and, implicitly, interlocal migration patterns have raised some provocative legal issues. Although spatial location models can be among the most complex of all economic models,
the flavor of the problem can be conveyed in the form of a very simplistic model.

Assume a very simple world with two possible living sites, City A and City B. City A is for various reasons—climate, topography, access to other resources, etc.—a more “favorable” site to develop. At each site, however, the net benefits of residence depend upon the number of other people sharing the location. Exhibit 6.6 has been drawn to reflect the net benefits per capita obtainable at each site as a function of the number of residents. Over an early range of population growth, there are “agglomeration economies” as the ability to share the cost of facilities and to interact beneficially increases. After the “optimal” population size is reached, the effects of congestion and other unfavorable interactions become dominant and increasing population causes benefits to decline.

**EXHIBIT 6.6**

Net Benefits per capita Population

![Diagram](image-url)
Be forewarned that Exhibit 6.6 has been drawn in a particular way in order to make a number of pedagogical points. The really important underlying assumption, however, is only that there are both economies and diseconomies of population size change over certain ranges. For simplicity, we shall also assume that the world is full of more or less homogeneous people who evaluate the two sites in similar ways.

If the world starts with one person, Exhibit 6.6 suggests that the person will choose to locate in City A. As population size increases up to the optimal size N, everyone is perfectly happy with the influx. Now let population increase to 2N. The “quality of life” in City A will deteriorate, but the B2 benefit level in A is still as good as the best one can do at the site of the potential City B. Is there any sense, then, in which a misallocation of population might be said to exist? At what point will people begin to leave A and move to B?

Although the Exhibit 6.6 model is a very crude one, it will suffice to pose some interesting questions. Relaxation of the model’s simplifying assumptions raises even more fascinating issues, but those intricacies are beyond the scope of the present treatment.¹

QUESTIONS

1. Assume the 2N level of total population. Show that average benefits of the population would be maximized if there were N people in each city. How could that population distribution be achieved, or maintained if it were achieved?
2. At 2.4N, all of the people would still be at site A. Why? Wouldn’t the last 0.4N be better off if they moved to A? Who bears the “entrepreneurial” costs of developing B?
3. What “externalities” are involved in this interlocal migration process? When the second N people move into City A, do the external costs that they impose exceed the internal benefits they receive from being in A rather than B?
4. Does the law in any way seek to “internalize” this type of externality? Should it?

CONSTRUCTION INDUSTRY ASSOCIATION OF SONOMA COUNTY V. CITY OF PETALUMA

522 F.2d 897 (1975), certiorari denied 424 U.S. 934, 96 S.Ct. 1148,
47 L.Ed.2d 342 (1976).

CHOY, Circuit Judge:
The City of Petaluma (the City) appeal from a district court decision voiding as unconstitutional certain aspects of its five-year housing and zoning plan. We reverse.

Statement of Facts

The City is located in southern Sonoma County, about 40 miles north of San Francisco. In the 1950’s and 1960’s, Petaluma was a relatively self-sufficient town. It experienced a steady population growth from 10,315 in 1950 to 24,870 in 1970. Eventually, the City was drawn into the Bay Area metropolitan housing market as people working in San Francisco and San Rafael became willing to commute longer distances to secure relatively inexpensive housing available there. By November 1972, according to unofficial figures, Petaluma’s population was at 30,500, a dramatic increase of almost 25 per cent in little over two years.

Alarmed by the accelerated rate of growth in 1970 and 1971, the demand for even more housing, and the sprawl of the City eastward, the City adopted a temporary freeze on development in early 1971. The construction and zoning change moratorium was intended to give the City Council and the City planners an opportunity to study the housing and zoning situation and to develop short and long range plans. The Council made specific findings with respect to housing patterns and availability in Petaluma, including

¹ Exhibit 6.6 is taken from Goetz “Fiscal Incentives for the Cities,” Committee on Urban Public Economics Seminar Series, Nov., 1969. The additional “intricacies” and their policy implications were in fact developed in that paper.
the following: That from 1960-1970, 88 per cent of housing permits issued were for single-family detached homes; that in 1970, 83 per cent of Petaluma’s housing was single-family dwellings; that the bulk of recent development (largely single-family homes) occurred in the eastern portion of the City, causing a large deficiency in moderately priced multi-family and apartment units on the east side.

To correct the imbalance between single-family and multi-family dwellings, curb the sprawl of the City on the east, and retard the accelerating growth of the City, the Council in 1972 adopted several resolutions, which collectively are called the “Petaluma Plan” (the Plan).

The Plan, on its face limited to a five-year period (1972-1977), fixes a housing development growth rate not to exceed 500 dwelling units per year. Each dwelling unit represents approximately three people. The 500-unit figure is somewhat misleading, however, because it applies only to housing units (hereinafter referred to as “development-units”) that are part of projects involving five units or more. Thus, the 500-unit figure does not reflect any housing and population growth due to construction of single-family homes or even four-unit apartment buildings not part of any larger project.

The Plan also positions a 200 foot wide “greenbelt” around the City, to serve as a boundary for urban expansion for at least five years, and with respect to the east and north sides of the City, for perhaps ten to fifteen years. One of the most innovative features of the Plan is the Residential Development Control System which provides procedures and criteria for the award of the annual 500 development-unit permits. At the heart of the allocation procedure is an intricate point system, whereby a builder accumulates points for conformity by his projects with the City’s general plan and environmental design plans, for good architectural design, and for providing low and moderate income dwelling units and various recreational facilities. The Plan further directs that allocations of building permits are to be divided as evenly as feasible between the west and east sections of the City and between single-family dwellings and multiple residential units (including rental units), that the sections of the City closest to the center are to be developed first in order to cause “infilling” of vacant area, and that 8 to 12 per cent of the housing units approved be for low and moderate income persons.

In a provision of the Plan, intended to maintain the close-in rural space outside and surrounding Petaluma, the City solicited Sonoma County to establish stringent subdivision and appropriate acreage parcel controls for the areas outside the urban extension line of the City and to limit severely further residential infilling.

The purpose of the Plan is much disputed in this case. According to general statements in the Plan itself, the Plan was devised to ensure that “development in the next five years will take place in a reasonable, orderly, attractive manner, rather than in a completely haphazard and unattractive manner.” The controversial 500-unit limitation on residential development-units was adopted by the City “[i]n order to protect its small town character and surrounding open space.” The other features of the Plan were designed to encourage an east-west balance in development, to provide for variety in densities and building types and wide ranges in prices and rents, to ensure infilling of close-in vacant areas, and to prevent the sprawl of the City to the east and north. The Construction Industry Association of Sonoma County (the Association) argues and the district court found, however, that the Plan was primarily enacted “to limit Petaluma’s demographic and market growth rate in housing and in the immigration of new residents.” Construction Industry Assn. v. City of Petaluma, 375 F.Supp. 574, 576 (N.D.Cal.1974).

Substantive Due Process

1 The district court found that although the Plan is ostensibly limited to a five-year period, official attempts have been made to perpetuate the Plan before 1977. Such attempts include the urban extension line (see text infra) and the agreement to purchase from the Sonoma County Water Agency only 9.8 million gallons of water per day through the year 1990. This flow is sufficient to sup port a population of 55,000. If the City were to grow at a rate of about 500 housing units per year (approximately three persons per unit), the City would reach a population of 55,000 about the year 1990. The 55,000 figure was mentioned by City officials as the projected optimal (and maximum) size of Petaluma. See, e.g., R.T. at 135-43, 145-46.
Appellees claim that the Plan is arbitrary and unreasonable and, thus, violative of the due process clause of the Fourteenth Amendment. According to appellees, the Plan is nothing more than an exclusionary zoning device, designed solely to insulate Petaluma from the urban complex in which it finds itself. The Association and the Landowners reject, as falling outside the scope of any legitimate governmental interest, the City’s avowed purposes in implementing the Plan—the preservation of Petaluma’s small town character and the avoidance of the social and environmental problems caused by an uncontrolled growth rate.

In attacking the validity of the Plan, appellees rely heavily on the district court’s finding that the express purpose and the actual effect of the Plan is to exclude substantial numbers of people who would otherwise elect to move to the City. 375 F.Supp. at 581. The existence of an exclusionary purpose and effect reflects, however, only one side of the zoning regulation. Practically all zoning restrictions have as a purpose and effect the exclusion of some activity or type of structure or a certain density of inhabitants. And in reviewing the reasonableness of a zoning ordinance, our inquiry does not terminate with a finding that it is for an exclusionary purpose. We must determine further whether the exclusion bears any rational relationship to a legitimate state interest. If it does not, then the zoning regulation is invalid. If, on the other hand, a legitimate state interest is furthered by the zoning regulation, we must defer to the legislative act. Being neither a super legislature nor a zoning board of appeal, a federal court is without authority to weigh and reappraise the factors considered or ignored by the legislative body in passing the challenged zoning regulation. The reasonableness, not the wisdom, of the Petaluma Plan is at issue in this suit.

It is well settled that zoning regulations “must find their justification in some aspect of the police power, asserted for the public welfare.” Village of Euclid v. Ambler Realty Co., 272 U.S. 365, 387, 47 S.Ct. 114, 118, 71 L.Ed. 303 (1926). The concept of the public welfare, however, is not limited to the regulation of noxious activities or dangerous structures. As the Court stated in Berman v. Parker, 348 U.S. 26, 33, 75 S.Ct. 98, 102, 99 L.Ed. 27 (1954): “The concept of the public welfare is broad and inclusive. The values it represents are spiritual as well as physical, aesthetic as well as monetary. It is within the power of the legislature to determine that the community should be beautiful as well as healthy, spacious as well as clean, well-balanced as well as carefully patrolled.” (citations omitted). Accord, Village of Belle Terre v. Boraas, 416 U.S. 1, 6, 9, 94 S.Ct. 1536, 39 L.Ed.2d 797 (1974).

In determining whether the City’s interest in preserving its small town character and in avoiding uncontrolled and rapid growth falls within the broad concept of “public welfare,” we are considerably assisted by two recent cases. Belle Terre, supra, and Ybarra v. City of Town of Los Altos Hills, 503 F.2d 250 (9th Cir. 1974), each of which upheld as not unreasonable a zoning regulation much more restrictive than the Petaluma Plan, are dispositive of the due process issue in this case.

---

10 "Exclusionary zoning" is a phrase popularly used to describe suburban zoning regulations which have the effect, if not also the purpose, of preventing the migration of low and middle-income persons. Since a large percentage of racial minorities fall within the low and middle income brackets, exclusionary zoning regulations may also effectively wall out racial minorities. See generally Alo, Goldberg & White, Racial and Economic Segregation by Zoning: Dean Knell for Home Rule?, 1969 U.Tol.L.Rev. 65 (1969); Bigham & Bostick, Exclusionary Zoning Practices: An Examination of the Current Controversy, 25 Vand.L.Rev. 111 (1972); Davidoff & Davidoff, Opening the Suburbs: Toward Inclusionary Land Use Controls, 22 Syracuse L.Rev. 509 (1971); Note, Exclusionary Zoning and Equal Protection, 84 Harv.L.Rev. 1645 (1971). Most court challenges to and comment upon so-called exclusionary zoning focus on such traditional zoning devices as height limitations, minimum square footage and minimum lot size requirements, and the prohibition of multifamily dwellings or mobile homes. The Petaluma Plan is unique in that although it assertedly slows the growth rate it replaces the past pattern of single-family detached homes with an assortment of housing units, varying in price and design.
In Belle Terre the Supreme Court rejected numerous challenges to a village’s restricting land use to one-family dwellings excluding lodging houses, boarding houses, fraternity houses or multiple-dwelling houses. By absolutely prohibiting the construction of or conversion of a building to other than single-family dwelling, the village ensured that it would never grow, if at all, much larger than its population of 700 living in 220 residences. Nonetheless, the Court found that the prohibition of boarding houses and other multi-family dwellings present urban problems, such as the occupation of a given space by more people, the increase in traffic and parked cars and the noise that comes with increased crowds. According to the Court, “A quiet place where yards are wide, people few, and motor vehicles restricted are legitimate guidelines in a land-use project addressed to family needs. This goal is a permissible one within Berman v. Parker, supra. The police power is not confined to elimination of filth, stench, and unhealthy places. It is ample to lay out zones where family values, youth values, and the blessings of quiet seclusion, and clean air make the area a sanctuary for people” 416 U.S. at 9, 94 S.Ct. at 1541. While dissenting from the majority opinion in Belle Terre on the ground that the regulation unreasonably burdened the exercise of First Amendment associational rights, Mr. Justice Marshall concurred in the Court’s express holding that a local entity’s zoning power is extremely broad: “[L]ocal zoning authorities may properly act in furtherance of the objectives asserted to be served by the ordinance at issue here: restricting uncontrolled growth, solving traffic problems, keeping rental costs at a reasonable level, and making the community attractive to families. The police power which provides the justification for zoning is not narrowly confined. And, it is appropriate that we afford zoning authorities considerable latitude in choosing the means by which to implement such purposes.” 416 U.S. at 13-14, 94 S.Ct. at 1543 (Marshall, J., dissenting) (emphasis added) (citations omitted).

Following the Belle Terre decision, this court in Los Altos Hills had an opportunity to review a zoning ordinance providing that a housing lot shall contain not less than one acre and that no lot shall be occupied by more than one primary dwelling unit. The ordinance as a practical matter prevented poor people from living in Los Altos Hills and restricted the density, and thus the population, of the town. This court, nonetheless, found that the ordinance was rationally related to a legitimate governmental interest—the preservation of the town’s rural environment—and, thus, did not violate the equal protection clause of the Fourteenth Amendment. 503 F.2d at 254.

Both the Belle Terre ordinance and the Los Altos Hills regulation had the purpose and effect of permanently restricting growth; nonetheless, the court in each case upheld the particular law before it on the ground that the regulation served a legitimate governmental interest falling within the concept of the public welfare: the preservation of quiet family neighborhoods (Belle Terre) and the preservation of a rural environment (Los Altos Hills). Even less restrictive or exclusionary than the above zoning ordinances is the Petaluma Plan which, unlike those ordinances, does not freeze the population at present or near-present levels. Further, unlike the Los Altos Hills ordinance and the various zoning regulations struck down by state courts in recent years, the Petaluma Plan does not have the undesirable effect of walling out any particular income class nor any racial minority group.16

13 The plaintiffs in Belle Terre claimed inter alia that the ordinance interfered with a person's right to travel and right to migrate to and settle within a state. The Supreme Court held that since the ordinance was not aimed at transients, there was no infringement of anyone's right to travel. 416 U.S. at 7, 94 S.Ct. 1536. Although due to appellees' lack of standing we do not reach today the right to travel issue, we note that the Petaluma plan is not aimed at transients, nor does it penalize those who have recently exercised their right to travel. See CEEED v. California Coastal Zone Conservation Comm'n, 43 Cal.App.3d 306, 118 Cal.Rptr. 315, 332-34 (1974); cf. Dunn v. Blumstein, 405 U.S. 330, 342, 92 S.Ct. 995, 31 L.Ed.2d 274 (1972).

16 Although appellees have attempted to align their business interest in attacking the Plan with legitimate housing needs of the urban poor and racial minorities, the Association has not alleged nor can it allege, based on the record in this case, that the Plan has the purpose and effect of excluding poor persons and racial minorities. Cf. Board of County Supervisors of Fairfax County v. Carper, 200 Va. 653, 107 S.E.2d 390 (1959). Contrary to the picture painted by appellees, the Petaluma Plan is "inclusionary" to the extent that it offers new opportunities, previously unavailable, to minorities and low and moderate-income persons. Under the pre-Plan system single family, middle-income housing dominated the Petaluma market, and as a result low and moderate income persons
Although we assume that some persons desirous of living in Petaluma will be excluded under the housing permit limitation and that, thus, the Plan may frustrate some legitimate regional housing needs, the Plan is not arbitrary or unreasonable. We agree with appellants that unlike the situation in the past most municipalities today are neither isolated nor wholly independent from neighboring municipalities and that, consequently, unilateral land use decisions by one local entity affect the needs and resources of an entire region. See, e.g., Golden v. Planning Board of Town of Ramapo, 30 N.Y.2d 359, 334 N.Y.S.2d 138, 285 N.E.2d 291, appeal dismissed, 409 U.S. 1003, 93 S.Ct. 436, 34 L.Ed.2d 294 (1972); National Land & Investment Co. v. Kohn, 419 Pa. 504, 215 A.2d 597 (1965); Note, Phased Zoning: Regulation of the Tempo and Sequence of Land Development, 26 Stan.L. Rev. 585, 605 (1974). It does not necessarily follow, however, that the due process rights of builders and landowners are violated merely because a local entity exercises in its own self-interest the police power lawfully delegated to it by the state. See Belle Terre, supra; Los Altos Hills, supra. If the present system of delegated zoning power does not effectively serve the state interest in furthering the general welfare of the region or entire state, it is the state legislature’s and not the federal courts’ role to intervene and adjust the system. As stated supra, the federal court is not a super zoning board and should not be called on to mark the point at which legitimate local interests in promoting the welfare of the community are outweighed by legitimate regional interests. See Note, supra, at 608-11.

We conclude therefore that under Belle Terre and Los Altos Hills the concept of the public welfare is sufficiently broad to uphold Petaluma’s desire to preserve its small town character, its open spaces and low density of population, and to grow at an orderly and deliberate pace.

Commerce Clause

The district court found that housing in Petaluma and the surrounding areas is produced substantially through goods and services in interstate commerce and that curtailment of residential growth in Petaluma will cause serious dislocation to commerce. 375 F.Supp. at 577, 579. Our ruling today, however, that the Petaluma Plan represents a reasonable and legitimate exercise of the police power obviates the necessity of remanding the case for consideration of appellees’ claim that the Plan unreasonably burdens interstate commerce.

It is well settled that a state regulation validly based on the police power does not impermissibly burden interstate commerce where the regulation neither discriminates against interstate commerce nor operates to disrupt its required uniformity. Huron Cement Co. v. Detroit, 362 U.S. 440, 448, 80 S.Ct. 813, 4 L.Ed.2d 852 (1960). As stated by the Supreme Court almost 25 years ago: “When there is a reasonable basis for legislation to protect the social, as distinguished from the economic, welfare of a community, it is not for this Court because of the Commerce Clause to deny the exercise locally of the sovereign power of the [state].” Breard v. Alexandria, 341 U.S. 622, 640, 71 S.Ct. 920, 931, 95 L.Ed. 1233 (1951). It is wholly beyond a court’s limited authority under the Commerce Clause to review state legislation by

Consequently, since the local regulation here is rationally related to the social and environmental welfare of the community and does not discriminate against interstate commerce or operate to disrupt its required uniformity, appellees’ claim that the Plan unreasonably burdens commerce must fail.\textsuperscript{18}

Reversed.

**QUESTIONS**

1. Allegedly, the Petaluma Plan was not directed at size per se but, rather, “haphazard” development. Assuming that this is true, the model in Exhibit 6.6 does not directly apply. To what extent can an analogy with the model be drawn?

2. The development cases such as Petaluma may be viewed as a question of who “owns” development rights, earlycomers or latecomers. What is your sense of how the courts seem to be answering this question?

3. Suppose that the latecomers are held to own a certain type of development right, and that right is protected from uncompensated expropriation by the earlycomers. Is there any way in which the earlycomers can keep that right from being exercised?

\textsuperscript{18} Our decision today conforms with others which have upheld reasonable state environmental legislation despite some burden incidentally placed on interstate commerce. See, e.g., Huron Cement Co. v. Detroit, supra (air pollution statute); Proctor & Gamble Co. v. City of Chicago, 509 F.2d 69 (7th Cir.), cert. denied, 421 U.S. 978, 95 S.Ct. 1980, 44 L.Ed. 2d 470 (1975) (ban on phosphate detergents); American Can Co. v. Oregon Liquor Control Commission, 15 Or.App. 618, 517 P.2d 691 (1973) (ban on non-returnable beverage containers).
CH. 1: ANALYZING CHOICE UNDER ALTERNATIVE RULES.. ERROR! BOOKMARK NOT DEFINED.

A. INTRODUCTORY CONSIDERATIONS...... ERROR! BOOKMARK NOT DEFINED.

B. ELEMENTARY GAME THEORY ............ ERROR! BOOKMARK NOT DEFINED.

1. THE PRISONER'S DILEMMA........................................................................ Error! Bookmark not defined.
   a. THE STUDENT'S DILEMMA: UNIVERSITY v. EAGER...................................................... Error! Bookmark not defined.
   b. THE PRISONER’S DILEMMA AS A MATRIX MODEL................................ Error! Bookmark not defined.
       Exhibit 1.1: PD Matrix Model........................................................................ Error! Bookmark not defined.
   c. THE “PRISONER’S DILEMMA” DEFINED......................................................... Error! Bookmark not defined.
       Exhibit 1.1: PD Matrix Model........................................................................ Error! Bookmark not defined.

2. ANOTHER IMPORTANT, CLOSELY RELATED GAME: “CHICKEN”... Error! Bookmark not defined.

3. OTHER USEFUL JARGON: “EXTERNALITIES”................................................. Error! Bookmark not defined.

4. ETHICAL AND OTHER EXTRALEGAL BEHAVIORAL REGULATIONS.......... Error! Bookmark not defined.

5. EXTERNAL BENEFITS, EXCLUDABILITY AND “FREE RIDERS”.............. Error! Bookmark not defined.

C. THE LAW AND ECONOMICS OF PROPERTY RIGHTS.... ERROR! BOOKMARK NOT DEFINED.

1. ROAD MAINTENANCE PROBLEM...................................................................... Error! Bookmark not defined.
   Exhibit 1.3 Benefits of Road Repair Levels ......................................................... Error! Bookmark not defined.

2. JOINTNESS AND EXCLUDABILITY AS ELEMENTS OF PROPERTY.... Error! Bookmark not defined.

3. CHICKEN VS. PRISONER’S DILEMMA............................................................. Error! Bookmark not defined.
   Exhibit 1.4 Road Problem As A “Chicken” Game .............................................. Error! Bookmark not defined.
   Exhibit 1.5 Road Problem As A Prisoner’s Dilemma ........................................ Error! Bookmark not defined.

4. GAMING ABOUT INFORMATION ..................................................................... Error! Bookmark not defined.
   UNTRUTHS CAN IMPROVE BUSINESS-SCHOOL GRADE; PEER PRESSURE CLASHES WITH
   ETHICS ................................................................................................................ Error! Bookmark not defined.

5. “CHICKEN” IN ACTION: LAND ASSEMBLY................................................. Error! Bookmark not defined.
   HOMEOWNER BICKERING ENDS $100,000 OFFER FOR HOUSES ..... Error! Bookmark not defined.

6. WHAT GENERALIZATIONS CAN BE MADE?..................................................... Error! Bookmark not defined.
D. TRANSFERABILITY AS AN ELEMENT OF PROPERTY ...

1. ASSIGNABILITY OF CAUSES OF ACTION ...................................
   THE CASE OF NIKE v. LOBEL ...........................................

2. OTHER KINDS OF PROPERTY RIGHTS IN LEGAL PROCESS ...
   d. Mutuality of Estoppel ...........................................
   e. Discovery of Computerized Litigation Files .................

3. COASE AND COSTS, “OPPORTUNITY” AND OTHERWISE ...
   a. Allocation of Rights in the Small-Number Context ........

CH 2: RIVALROUS AND RISKY DECISIONS

A. BEHAVIOR INVOLVING UNCERTAINTY

1. PROSPECTS AND CERTAINTY EQUIVALENTS ...........

2. EXPECTED VALUES ...........................................

3. RISK PREFERENCE ...........................................

B. APPLICATIONS ...........................................

1. LAW OF THE LEANING TREE ...................................

2. THE CASE OF MRS. CRISPY’S CHICKEN ..................

3. EX ANTE ANALYSIS OF PUNISHMENTS ..................

4. THE HUNTING OF A TAKEOVER ..........................

5. AGENT-PRINCIPAL CONFLICTS OF INTEREST .......
   a. Using Agents To Sell Real Estate ..........................
Table of Contents

b. Agent-Principal Conflicts in Legal Contracts........................................ Error! Bookmark not defined.
   HEINZMAN v. FINE, FINE, LEGUM & FINE ........................................ Error! Bookmark not defined.

   Exhibit 2.4 – Marginal Effect As Slope of Total Curve ............................... Error! Bookmark not defined.
   Exhibit 2.5 - Area Under A Marginal Curve ........................................... Error! Bookmark not defined.

7. ATTORNEY- CLIENT RELATIONSHIPS....................................................... Error! Bookmark not defined.

8. SHOULD SUBROGEEs KEEP “EXCESS” RECOVERIES? .................................... Error! Bookmark not defined.

C. LEAST-COST RISK BEARING................................................................. Error! Bookmark not defined.

1. WHAT IS “LEAST-COST” RISK BEARING? ................................................. Error! Bookmark not defined.
   Exhibit 2.6 – Binomial Distribution ............................................................. Error! Bookmark not defined.

2. SOURCES OF LEAST COST RISK-BEARING.............................................. Error! Bookmark not defined.
   Exhibit 2.7 – Pool of 5 Shared $1 Risks with 40% Probability .................... Error! Bookmark not defined.
   Exhibit 2.8 – Pool of 10 Shared $1 Risks with 40% Probability .................. Error! Bookmark not defined.
   Exhibit 2.9 – Pool of 20 Shared $1 Risks with 40% Probability .................. Error! Bookmark not defined.


D. OPTIMAL INSURANCE AND RISK ADJUSTMENT............................................. Error! Bookmark not defined.

1. “INSURANCE” AGAINST CONTRACT BREACH ............................................ Error! Bookmark not defined.
   Exhibit 2.10 (top panel) and Exhibit 2.11 (bottom panel) ............................ Error! Bookmark not defined.


3. “RISK DISTRIBUTION” INSURANCE AGAINST HARMFUL EVENTS ............... Error! Bookmark not defined.
   NATIONAL STEEL SERVICE CENTER v. GIBBONS ..................................... Error! Bookmark not defined.

E. JUSTICE AND FAIRNESS ................................................................. Error! Bookmark not defined.

   Exhibit 2.12 – Social Choice Matrix ......................................................... Error! Bookmark not defined.

2. FAIR DIVISION SCHEMES: DIVIDE AND CHOOSE .................................... Error! Bookmark not defined.

3. EQUITABLE DISTRIBUTION OF VOTING POWER ..................................... Error! Bookmark not defined.
   a. Pivotal-Voter Measure of Political Power .............................................. Error! Bookmark not defined.
      Exhibit 2.14: Power of a Voter Via Representation .................................. Error! Bookmark not defined.
   b. An Equilibrium-Displacement Measure of Voter Power ............................ Error! Bookmark not defined.
      Exhibit 2.15: Potential Displacement Via Single Voter ............................ Error! Bookmark not defined.
      Exhibit 2.16: Electoral College Power Under Alternative Theories ............ Error! Bookmark not defined.

4. EFFICIENCY ASPECTS OF FAIRNESS .................................................... Error! Bookmark not defined.
STATE OF IOWA v. HENRY PARRISH ................................................................................................................................. Error! Bookmark not defined.

5. “PATERNALISM” IN THE LAW ................................................................................................................................. Error! Bookmark not defined.

CH. 3: INTERTEMPORAL COSTS AND BENEFITS ......................................................... Error! BOOKMARK NOT DEFINED.

A. INTEREST RATES AND “PRESENT VALUES” ................................................................. Error! BOOKMARK NOT DEFINED.

   Exhibit 3.1: (1+r)n Compound Interest Growth Factors ........................................................................................................... Error! Bookmark not defined.
   Exhibit 3.2: 1/(1+r)n Discount Factors ................................................................................................................................. Error! Bookmark not defined.

B. VALUING THE BENEFITS OF INTERESTS IN LAND ............................................. Error! BOOKMARK NOT DEFINED.

   1. DETERMINING THE VALUE OF A LEASEHOLD ................................................. Error! Bookmark not defined.
   2. VALUING A PROPERTY ENCUMBERED BY A LEASE ........................................ Error! Bookmark not defined.

C. INTEREST ON DAMAGES AND PENALTIES ....................................................... Error! BOOKMARK NOT DEFINED.

   1. PREJUDGMENT INTEREST ON DAMAGES ........................................................ Error! Bookmark not defined.
      BUSIK v. LEVINE ................................................................................................................................. Error! Bookmark not defined.

      UNITED TELECOMMUNICATIONS, INC. v. AMERICAN TELEVISION AND COMMUNICATIONS CORP. ................................................................................................................................. Error! Bookmark not defined.

   3. PREJUDGMENT INTEREST ON PENALTIES ....................................................... Error! Bookmark not defined.

D. EVALUATING LOST FUTURE INCOME AS DAMAGES ....................................... Error! BOOKMARK NOT DEFINED.

   Exhibit 3.3: Lost Income Damages, Assumption Set #1 .............................................................................................................. Error! Bookmark not defined.
   Exhibit 3.4: Lost Income Damages, Assumption Set #2 .............................................................................................................. Error! Bookmark not defined.
   BEAULIEU v. ELLIOTT ................................................................................................................................. Error! Bookmark not defined.
   DOCA v. MARINA MERCANTE NICARAGUENSE, S.A .............................................................................................................. Error! Bookmark not defined.

E. COMBINED UNCERTAINTY AND DISCOUNTING ................................................. Error! BOOKMARK NOT DEFINED.

   1. Introductory Considerations ................................................................................. Error! Bookmark not defined.
      a. Tintd Hypothetical ......................................................................................................................... Error! Bookmark not defined.
      Exhibit 3.5: Tintds’ Probability of Survival to Various Birthdays .............................................................................................................. Error! Bookmark not defined.
      b. Annuity Hypothetical ......................................................................................................................... Error! Bookmark not defined.
      c. Life Tenancy and Remainder Rights .............................................................................................................. Error! Bookmark not defined.

   2. AVERAGE vs. actuarily expected LIFE earnings ..................................................... Error! Bookmark not defined.
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
</table>
| 1. | **Precautions That Minimize Net Social Cost of Accidents**

Exhibit 5.5 – Multiparty Social Cost Function #1

| 2. | **Comparative Negligence**

Exhibit 5.6 – Multiparty Social Cost Function #2

| 3. | **The “Level” or Quantitative Problem**

| 4. | **Taking Cost-Minimization Seriously (?)**

| 5. | **Promise-Breaking Analyzed as a Tort**

| 6. | **Economically Optimal Damages for Broken Promises**

| 7. | **Sec. 90, Restatement of Contracts**

| 8. | **Production Costs and Contract Breach**

| 9. | **Adjustment of a Competitive Industry**


2. Bankruptcy Laws And Fixed Costs

C. INCIDENCE ANALYSIS: SHIFTING AND PASSING ON.
   HANOVER SHOE, INC. v. UNITED SHOE MACHINERY CORP.
   ILLINOIS BRICK CO. v. ILLINOIS

D. ANTICOMPETITIVE SUPPLY RESTRICTION
   CLARK v. UNIVERSAL BUILDERS, INC.
   STATEMENT OF U.S. DEPT. OF TRANSPORTATION BEFORE THE INTERSTATE COMMERCE
   COMMISSION, WESTERN COAL INVESTIGATION-GUIDELINES FOR RAILROAD RATE
   STRUCTURE
   GHEN v. RICH
   CONSTRUCTION INDUSTRY ASSOCIATION OF SONOMA COUNTY v. CITY OF PETALUMA

Exhibit 6.2: Graph of Firm’s Cost Curves
Exhibit 6.3: Short-run Equilibrium of Supply and Demand
b. Partial Displacement And Short-Run Equilibrium Adjustment
c. Progress Toward New Equilibrium
d. Additional Implications
e. A Simple Mathematical Model [Optional]
Exhibit 6.4: Costs to Firm Paying ASD Royalty
Exhibit 6.4b: Solutions of Market Model
Exhibit 6.5: Expansion versus Market Spoilage
Exhibit 6.2a: Correctly Graphed Cost Curves
Exhibit 6.3a: Completed Graph of Demand and Shifting Supply Curves