A Structural Model of Health Plan Choice and Health Care Demand in the Medicare Managed Care Program

Keith Brand†
Department of Economics
University of Virginia

October 2005

Abstract

I examine the joint determination of health plan choice and subsequent health care utilization in the Medicare managed care and supplemental insurance markets. The objectives are to evaluate the welfare impact of the Medicare managed care program taking into account the effect of endogenous selection, and to advance the applied literature on health plan choice and health care demand.

Standard models of insurance markets suggest that Medicare Advantage will have ambiguous welfare implications because of endogenous selection. While previous research has found evidence that Medicare Advantage plans experience favorable selection, this research is, to my knowledge, the first to consider the welfare impact.

In the model, beneficiaries choose health plans by taking expectations of indirect utilities over a known distribution of health states. Health care utilization is simultaneously determined in five dimensions: inpatient care, outpatient care, doctor visits, prescription drugs, and dental care. The plan choice and utilization models are stochastically dependent. The demand model accounts for uncertainty about the efficacy of treatment, substitutability across types of treatment, diminishing marginal product both intensively and extensively, and limits on the efficacy of treatment.

To estimate the model, I use health plan choice data at both aggregate and individual levels as well as individual level data on subsequent utilization. The aggregate health plan choice data is observed at the county/age/gender level. I also observe counties of operation and health plan characteristics of all Medicare managed care organizations and data on county level Medigap premiums.

JEL Classification: C34, H51, I10, I18
Keywords: Medicare, Health Insurance, Health Care Demand, Demand Systems

† I thank my advisors Steven Stern, Leora Friedberg, and John Pepper for their invaluable suggestions and support. Financial support from the Bankard Fund for Political Economy and the University of Virginia Graduate School of Arts and Sciences is gratefully acknowledged. I also thank the CMS, Resource Data Assistance Center, Robert Town, and Su Liu for their assistance with data. Correspondence: 2015 Ivy Road, PO Box 400182, Charlottesville, VA 22904-4182. Email: kb6r@virginia.edu.
1 Introduction

Much of the past and current debate on Medicare reform centers on the role of private health plans. Although Medicare was designed as a single-payer health plan at its inception in 1965, it has contracted with private health plans to provide health care services and insurance since 1982. More recently, Congress has sought to address the long-term solvency issues confronting Medicare in part by increasing or restructuring the role of private health plans. The Medicare Modernization Act of 2003 includes a significant role for private health plans in the provision of the outpatient prescription drug benefit. A series of Medicare reform proposals in the late 1990s suggested placing Medicare in direct competition with private health plans through a competitive bidding system commonly referred to as premium support.¹

Opposition to such proposals is based in part on the issue of endogenous selection. It is argued that if traditional Medicare competes directly with private health plans and if healthier beneficiaries systematically choose private health plans, then Medicare would have average costs exceeding those of private plans. This could lead to higher premiums for traditional Medicare resulting in a welfare loss for some. As described in Rothschild and Stiglitz (1976), endogenous selection implies that the welfare impact of moving from a single-payer system to a more competitive system will vary by risk type.

Currently, Medicare offers a managed care option to its beneficiaries through the Medicare Advantage program. Under this program, Medicare contracts with private managed care organizations (MCOs) on a county-by-county basis to bear risk and provide health care services to any beneficiary in that county who wishes to enroll. Endogenous selection of Medicare beneficiaries into Medicare Advantage has been a concern for policy-makers and researchers alike, and while most studies have found evidence that Medicare MCOs generally experience favorable selection, some find evidence that Medicare MCOs experience adverse selection in particular types of care.

Despite the continued policy debate, there has been little research on the welfare effect

¹See McClellan (2000) for a review of Medicare reform proposals.
of Medicare Advantage. Town and Liu (2003) estimate aggregate consumer and producer surplus due to Medicare Advantage and conclude that the total surplus generated net of the associated tax burden is significant.\(^2\) However, their study does not consider the issue of selection and therefore assumes that consumer surplus changed uniformly across the population.

To assess the welfare impact of the Medicare Advantage, I estimate distributions of the change in consumer surplus within age/gender groups brought about by Medicare Advantage. To estimate these distributions, I examine Medicare beneficiary decisions in the Medicare managed care market and the closely related, nongroup, partially regulated market for supplemental health insurance known as Medigap. Previous researchers (for example, Baker and Corts (1996) and Baker (1997)) have suggested that, because of selection, Medigap premiums may be higher in markets with large Medicare MCO market penetration than they would be in the absence of Medicare Advantage. To my knowledge, there has been no research that attempts to quantify the welfare implications of this issue.

I construct and estimate a structural discrete/continuous choice model of health plan choice and subsequent health care utilization. I model health plan choice as a discrete choice in which the consumer’s value of a health plan is based on an expected indirect utility. Health plans are differentiated by premiums, cost sharing arrangements, and a set of unobservable characteristics that affect the productivity of health care. I model health care demand as jointly determined in five types of care. Given a realized pretreatment health state, consumers maximize expected utility subject to the budget and technological constraints implied by the health plan choice where expectations are taken over a known distribution of outcomes for a given level of treatment. The health plan choice and utilization decisions are stochastically dependent.

To estimate the model, I use health plan choice data at both aggregate and individual levels as well as individual level data on subsequent utilization. The aggregate health plan choice data is observed at the county/age/gender level. This allows me to observed selection based on observable characteristics at both individual and market levels.

\(^2\)Under their conservative estimate, net welfare over the period 1993-2000 was $24.8 billion.
A second focus of this research is to contribute to the applied literature on health plan choice and health care demand. Since Grossman (1972), the theoretical literature has treated health care utilization as a derived demand, i.e., there is some technological relationship between the commodity demanded, health care, and the utility generating commodity, health. My principal contribution to this literature is to directly apply the notion of health care as a derived demand by modeling the technological relationship between health care inputs and health, and imposing the constraints implied by this relationship on the consumer choice problem. I model several aspects of health production commonly not treated in the applied literature. These include limits on the efficacy of treatment, uncertainty about the efficacy of treatment, diminishing marginal product on both intensive and extensive margins\(^3\), and substitutability and complementarity across types of treatment. Although I model the relationship between health care inputs and health, my goal is only to model consumer choice in a market in which budget and technological constraints apply, not to predict treatment outcomes per se. No data on health outcomes are used in this analysis.

To facilitate the incorporation of these technological properties into the consumer choice problem, I depart in a fundamental way from the previous literature on health care demand. To my knowledge, all of the previous literature has modeled health care demand by expressing some outcome variable as an explicit function of data, parameters, and unobservables. In this research, I simply specify the consumer choice problem and then define demand implicitly by the first order conditions of the consumer. This allows me to capture the effect of the technological constraint in a straightforward manner.

I also depart from the previous literature by simultaneously estimating a multi-dimensional demand system. I use data on five types of health care: inpatient hospital care, outpatient hospital care, doctor visits, prescription drugs, and dental care. Inpatient care, outpatient care, and doctor visits are the types of health care that traditional, or fee-for-service (FFS), Medicare was originally designed to cover and comprise a large share of expenditures. Pre-

\(^3\)Intensive diminishing marginal product is the usual notion that marginal product declines with inputs. Extensive diminishing marginal product states that the marginal product at any given quantity demanded is lower if the consumer is in a healthier pretreatment state.
scription drug and dental expenditures are included because they are not covered by FFS Medicare and are considered to be the most valued types of extra benefits that a participating MCO may provide. To my knowledge, only one study, Hubbard-Rennhoff (2005), has estimated the joint determination of different types of health care demand, in that case, mental and physical health care. The rest of the literature has either examined one type of care or estimated demand for more than one type of health care independently.

While modeling demand in five dimensions simultaneously creates significant computational burden, there are three important benefits. First, it captures the intuition that consumers often purchase health care in bundles for a given pretreatment state. This implies that demand models for more than one type of care are stochastically dependent. Second, since the model captures a large share of expenditures, it gives a relatively complete picture of consumer demand and, as will be discussed below, consumer valuation of health plans. Third, it allows the model to capture risk, or propensity of illness, in multiple dimensions.

The model of health care demand based on utility maximization provides an intuitive basis for health plan choice. In the model, the value of a health plan for a risk averse consumer is an expected indirect utility where expectations are taken over a known distribution of health states. This distribution of health states depends on the risk type of the consumer, which is a function of observed and unobserved characteristics. As Cameron et al (1988) point out, this approach, while intuitive, is computationally burdensome in the absence of simplifying assumptions on preferences or health production that permit analytical solutions to the consumer choice problem on utilization. These assumptions may be used in conjunction with distributional assumptions that allow closed form integration. In order to model the intuitive properties of health care demand and health plan choice described above, I avoid such assumptions in this research. Instead, I rely on numerical and simulation methods to estimate the model.

The relationship between health plan choice and expected utilization is the basis for the

4 A common assumption in both the theoretical and applied literatures is a deterministic, linear relationship between health care and health. See Dardanomni and Wagstaff (1991), Zabinski (1994), and Blomqvist (1997) for examples.

5 For example, Zabinski (1994).
endogeneity problem commonly cited in the literature on the effect of health plan choice on subsequent utilization. Health plan characteristics in a utilization model are endogenous if unobservables in the two decisions are correlated. Most researchers use an instrumental variables approach to correct for this endogeneity. To my knowledge, only three studies, Dowd et al. (1991), Zabinski (1994), and Hubbard-Rennhoff (2005), estimate health plan choice and subsequent utilization jointly. Like these studies, I estimate a model of joint determination that allows for stochastic dependence.

The principal data source for this study is the Centers for Medicare and Medicaid Services (CMS). Individual level data on health plan choice and subsequent utilization is contained in the 2000 Medicare Current Beneficiary Survey Cost and Use file. Aggregate data on Medicare beneficiary health plan enrollment and eligibility from the year 2000 has been made available at the county/age/gender group level. Data on MCO plan characteristics and counties of operation from the year 2000 was provided by the Resource Data Assistance Center. Data on Medigap premiums was provided by the American Association of Retired Persons (AARP).

The remainder of the paper is organized as follows. Section 2 provides background on Medicare, Medicare Advantage, and Medigap. Section 3 reviews the literature on the welfare impact of and on endogenous selection into Medicare Advantage. Section 4 describes the data. Sections 5 and 6 provide the structural model and empirical strategy, respectively. Section 7 gives reduced form results and a discussion, and section 8 concludes.

2 Background

2.1 Medicare

Medicare provides public health insurance to over 40 million Americans who are either over 65 years of age, disabled, or have end stage renal disease (ESRD). It is the single largest payer of health care services in the United States, covering more than 16.8% of national health expenditures in 2003. In fiscal 2002, Medicare outlays were $253.7 billion, making

---

\(^{6}\)Centers for Medicare and Medicaid Services, Office of the Actuary
it the third largest federal program, behind only Social Security and Defense. Spending on the Medicare program comprised 12.6% of all federal outlays and 2.6% of GDP in 2002.

Medicare was set up in two main parts. Part A, which covers hospitalization and skilled nursing care, is compulsory and is financed by a 2.9% payroll tax. Most people are automatically enrolled in Part A when they turn 65. Part B, which covers physician services and most outpatient care, is voluntary and is financed through general revenues and a monthly premium ($45.50 in 2000) paid by the beneficiary. The Part B premium is approximately 25% of the average cost of services under Part B. With this degree of subsidization, participation in Part B is very high (95% in 2000).

Both parts of Medicare leave the beneficiary exposed to significant risk. Part A coverage has a $776 per event deductible and runs out after 150 hospital days. Part B has a $100 annual deductible, a 20% coinsurance rate for covered medical care services, a 50% coinsurance for outpatient mental health services, and no maximum annual out-of-pocket expenditure limit. In addition, it excludes dental care, eye care, many types of preventive care (until recently routine mammography), long-term care, and outpatient prescription drugs.

Because of this exposure to risk, over 90% of beneficiaries have some type of supplemental coverage. Common sources are Medicaid, employer provided retiree coverage, and nongroup Medigap policies. Another option for Medicare beneficiaries to obtain extra coverage is to enroll in a Medicare MCO or private FFS plan through the Medicare Advantage program.\footnote{The distribution of sources of supplementary coverage in 2000 was: Employers provided coverage (36%), Medigap (27%), Medicare Advantage (17%), Medicaid (11%), No supplementary coverage (9%). (Newhouse, 2001)}

2.2 Medicare Advantage

Since the passage of the Tax Equity and Fiscal Responsibility Act (TEFRA) in 1982, the CMS has allowed federally qualified MCOs to enter into risk contracts under Section 1853 of the Social Security Act. Under this arrangement, the MCO agrees to provide all necessary care to any beneficiary who wishes to enroll in a specified service area (usually a county)
in exchange for a capitated albeit risk adjusted monthly payment per enrollee.\textsuperscript{8} This risk adjustment is based on age, gender, Medicaid eligibility, and residence in a long-term care facility. Until 1998, the base reimbursement rate, known as the Adjusted Area Per Capita Cost (AAPCC), was based on a five year moving average of 95\% of the average expenditure in that county by beneficiaries enrolled in FFS Medicare.

The TEFRA program’s reimbursement methodology was criticized for, among other issues\textsuperscript{9}, not taking into account the possibility of selection. Since the reimbursement is based on the average expenditures of those remaining in FFS Medicare, MCOs would earn supernormal profits if they experienced favorable selection. The general consensus in the literature is that this is indeed the case.\textsuperscript{10}

Medicare beneficiaries enrolled in Medicare Advantage must also enroll in Part B and pay the Part B premium. Participating MCOs must offer benefits at least equal to those offered by Parts A and B of FFS Medicare. MCOs have the option to offer more benefits such as prescription drug coverage or a dental plan and may charge an extra premium for this coverage subject to approval by the CMS. As of 2000, MCOs were precluded from offering

\textsuperscript{8}MCOs were also allowed to enter into non-risk contracts with the CMS under which the MCO provides care and is reimbursed on a fee-for-service basis. This type of contract, generally known as a cost contract, is no longer offered by the CMS and all such existing contracts were scheduled to sunset as of December 31, 2004.

\textsuperscript{9}Another prominent issue was payment rate volatility which may discourage MCO participation, particularly in rural areas. The Balanced Budget Act (BBA,1997) addressed this by setting payment rates as the maximum of a weighted average of the local AAPCC and the national mean rate (the weight on the national mean was scheduled to increase 10\% each year until a 50-50 blend was reached in 2003), a minimum of $367 per month, and a 2\% increase from the previous year’s rate. The Balanced Budget Refinement Act of 1999 established special payment increases for plans in underserved counties as a response to plan withdrawals following the implementation of the BBA. The Benefits Improvement and Protection Act (BIPA) of 2001, among other provisions, raised the minimum increase to 3\% and established different payment floors for rural ($475) and urban ($525) counties. The BIPA modifications made the payment floor binding for 40\% of beneficiaries (Newhouse, 2001).

\textsuperscript{10}See Hellinger and Wong (2000) for a review. Commonly cited evidence is that mortality rates are typically 25\% lower for Medicare Advantage enrollees, suggesting that, on average, Medicare Advantage MCOs have fewer expensive, end-of-life cases (Newhouse, 2001).
a negative premium, i.e., returning all or part of the Plan B premium to its enrollees.\textsuperscript{11} In 2000, 32\% of Medicare Advantage plans were offered at a zero premium and these plans comprised 59\% of the Medicare Advantage market in terms of enrollment.

One of the attractions of the Medicare Advantage program is that premiums are typically much lower than Medigap premiums while still providing comparable supplemental coverage. Beneficiaries sacrifice open provider choice and subject themselves to more extensive utilization review by taking the Medicare Advantage option.

\section*{2.3 Medigap}

Medigap policies generally cover the consumer cost sharing associated with Medicare Parts A and B and are sold in the nongroup market. Since 1992, Medigap policies have been regulated in that benefits are standardized by Congressional statute. Except in Massachusetts, Minnesota, and Wisconsin, insurers may offer Medigap supplements in only ten standardized plans, referred to as Medigap Plans A through J. All of these plans cover Medicare Parts A and B coinsurance and most cover the Part A deductible as well. The two most commonly purchased, Plans C and F, cover the deductible and coinsurance in both Parts A and B, reducing the first-dollar and marginal prices for any Medicare covered service to zero. Medigap premiums are not regulated. Table 4 provides some descriptive statistics on premiums and benefits for the four most popular (A, B, C, F) Medigap options and Medicare Advantage health plans.

There are significant dynamic issues involved in choosing a Medigap supplement that the static model presented below will not be able to capture. From the time of Part B enrollment, the Medicare beneficiary has a six month period in which to choose from any available Medigap plan with an enrollment guarantee. If a Medicare beneficiary discontinues his/her Medigap plan to join a Medicare MCO, the beneficiary has a twelve month window in which the beneficiary can leave the MCO and be guaranteed renewal of that Medigap plan.

\textsuperscript{11}Reforms under BIPA (2001), implemented in January 2003, allowed MCOs to charge a negative premium of up to 125\% of the Plan B premium with 80\% of the reduction going to enrollees and the remainder going to Medicare.
if it is still offered. If the Medigap plan is no longer offered, the beneficiary receives another open enrollment guarantee on any available Medigap plan A, B, C, or F. This twelve month window is available only if the MCO enrollment was the first Medicare MCO enrollment for that beneficiary. If the beneficiary enrolled in an MCO when the beneficiary first became eligible for Medicare, the beneficiary may leave the MCO within the first twelve months and be guaranteed enrollment in any available Medigap plan.

These considerations complicate a model of health plan choice for Medicare beneficiaries in that choices made in a given year affect the choice set in other years. However, accounting for these consideration in a dynamic framework would add excessive complication to the model. I follow the previous literature by assuming that health plan choice in a given year is independent of choices made in other years.

3 Existing Literature

In this section, I briefly review some of the relevant literature on the welfare impact of the Medicare Advantage and endogenous selection in health insurance markets. Town and Liu (2003) analyze the welfare effect of Medicare Advantage (then known as Medicare + Choice) using the methods of Berry (1994). They use Medicare MCO county level market shares data from 1993-2000 and premium data provided by the AARP to estimate the model. Explanatory variables are the premium, a drug benefit indicator, and an interaction of the drug benefit indicator and an indicator for year if later than 1996, to control for the impact of the BBA. As noted earlier, they estimate a significant increase in net welfare with a consumer surplus of $15.6 billion and producer surplus of $52 billion. Annual per beneficiary surplus in the year 2000 is estimated to be $113. The two principal sources of consumer surplus are the availability of prescription drug benefits and price competition. The drawback of this study is that it does not incorporate the effects of selection on welfare. Beneficiaries are treated uniformly and so the estimated increased consumer surplus masks the variation in the change in consumer surplus across risk groups that is the focus of this work.

Atherly, Dowd, and Feldman (2003) use individual level data from the Medicare Cur-
rent Beneficiary Survey 1998 to estimate the effect of Medicare MCO plan characteristics on health plan enrollment. They estimate a model of health plan choice between Medicare Advantage plans and a Medigap option using a nested logit model. They find that age, poor self-assessed health status, and high income are all negatively and significantly associated with Medicare Advantage enrollment. However, they also find that indicators of chronic illnesses such as hypertension and diabetes are positively associated with Medicare Advantage when interacted with a prescription drug or eye benefit indicator. This provides evidence that although MCOs may generally experience favorable selection, the presence of extra benefits may lead to adverse selection.

Call et al (1999) examine endogenous selection into Medicare MCOs using 1993 expenditures to predict 1994 enrollment decisions. They find that Medicare MCOs experienced favorable selection and that this effect declined as the total, county-level market share of Medicare MCOs increased. However, they find evidence of adverse selection into Medicare MCOs among short-term enrollees. Their data does not allow them to estimate the relationship between enrollment and specific plan characteristics such as a prescription drug benefit.

Feldman et al (2003) use data provided by the CMS to construct a plan level measure of risk based on inpatient claims data. They explain this measure of risk with plan characteristics and find that while risk is positively associated with offering a drug benefit, it is negatively associated with offering a dental plan.

Baker and Corts (1996) discuss the trade-off between increased competition and adverse selection in health insurance markets because of the rise of managed care. They propose that, because of selection and increased price competition, the effect of managed care on traditional, fee-for-service health insurance premiums is ambiguous. Baker (1997) looks at the additional possibility that practice style spillovers resulting from the rise of managed care could result in lower health care expenditures in traditional insurance as health care providers in a given market begin to adopt more conservative practice styles. His instrumental variables estimates indicate that health care expenditures in traditional insurance are concave in managed care market share with a peak market share of 15-18%. He acknowledges, however, that his study is unable to separate a spillover effect from a selection effect.
Cutler and Reber (1998) examine the welfare effect of adverse selection resulting from a policy change on health insurance subsidization at Harvard University. They look at employee health plan choices after Harvard moved from a proportional subsidy rule to a voucher system. The policy change raised the effective premium of a more generous health plan relative to basic plans and the response to the change was a nonrandom sorting in which the generous plan experienced adverse selection. In an example of what is commonly referred to as a “death spiral”, this led to further premium increases for the generous plan, and, subsequently, the generous plan was withdrawn. The authors also emphasize the importance of considering endogenous selection in evaluating public policy on social insurance programs. This study is the most similar to what I present here in that both studies examine the welfare consequences of endogenous selection in health insurance markets. The research I present here contributes to this literature by examining both health plan choice and utilization data and by using a structural model capable of predicting equilibrium outcomes in counterfactual policy scenarios.

4 Data

The data for this research will come from multiple sources with three primary data sets provided by the CMS. First, the Medicare Current Beneficiary Survey (MCBS) 2000 is a nationally representative sample of 12,305 Medicare beneficiaries containing data on health care utilization and health plan choice. The Cost and Use file of the MCBS provides demographic data such as age, gender, and zip code, and is linked to claims data from CMS records. The file provides information on beneficiary health care expenditures over the course of the year, the type of supplemental coverage in which the beneficiary was enrolled, and the identity of the firm providing this coverage if the beneficiary was enrolled in a Medicare MCO.

Survey respondents are asked if their former (or current) employer has provided health insurance or has paid the cost of a Medigap supplement or Medicare MCO. The MCBS also provides data on Medicaid eligibility and on the source of Medicare eligibility (age, disability,
or ESRD). Beneficiaries were removed from the sample if they reported having an employer sponsored health plan\textsuperscript{12} (n=3838) or were eligible for Medicaid\textsuperscript{13} (n=2407). Beneficiaries were also dropped if the source of Medicare eligibility is disability or ESRD (n=1944). With 1,135 beneficiaries in more than one of these categories, this leaves a sample of 5,251 beneficiaries.

Second, the Medicare Managed Care Quarterly/State/County/Plan database provides enrollment and eligibility counts from January 2000 at the county level for each Medicare MCO. This data has been made available by the CMS disaggregated into twelve gender/age groups. The age groups are: less than 65, 65-69, 70-74, 75-79, 80-84, greater than 84. I use all age groups except the less than 65 group, since, by definition, the source of Medicare eligibility is either disability or ESRD. To my knowledge, this research is the first use of the disaggregated data in economic research.

Third, the Medicare Health Plan Compare database provides detailed information on product characteristics of all Medicare Advantage plans in every county such as the premium, cost sharing arrangements for each category of health care, and the scope of services provided. The current version of this data is available for download at any time from the CMS web site. The Resource Data Assistance Center compiled a year 2000 version of this database for this research. This data allows me to construct the complete choice set of Medicare Advantage MCOs, with the relevant product characteristics, for all Medicare beneficiaries.

\textsuperscript{12}I drop those with employer sponsored coverage because no data on these health plans is observed so the choice set cannot be well-specified. The assumption is that anyone who is offered such coverage takes it up. In addition, dropping beneficiaries with employer sponsored coverage avoids complications associated with Medicare as secondary payer legislation. See Glied and Stabile (2001) for a discussion.

\textsuperscript{13}I drop all individuals who are dual eligible or who are a Qualified Medicare Beneficiary (QMB). QMBs are low income beneficiaries who do not qualify for full Medicaid benefits but are eligible for Medicaid coverage of the Medicare Part B premium and all cost sharing of Medicare Parts A and B. Since this is roughly equivalent to a Medigap supplement at a zero premium, I assume that all QMB eligibles take up this coverage. I retain all Specified Low-Income Beneficiaries (SLMB). These are low income beneficiaries who are eligible for a Medicaid subsidy of the Medicare Part B premium but still face the cost sharing associated with Medicare Parts A and B.
4.1 Data Limitations

There are three important data limitations with respect to choice of supplemental insurance coverage that affect the aggregate data. First, counties typically have many firms offering multiple Medigap plans, but premiums and market shares for these plans are not available. Second, the share of Medicare beneficiaries enrolled in Medicaid is observed at the state level but not the county level. Third, market shares of non-Medicare Advantage private plans, such as those provided as part of a retiree benefits package, and their product characteristics, are not observed.

I take the following steps to account for these limitations. I follow Town and Liu (2003) and Atherly et al (2004) by reducing the choice set among Medigap options to a single plan. Following Town and Liu (2003), I have obtained premium data from the AARP on its Plan F Medigap policy. Plan F is the most popular Medigap option nationally and the AARP, with over 2 million enrollees, is one of the largest providers of Medigap insurance (Town and Liu, 2003). I also use data from the CMS on county level per beneficiary FFS Medicare Parts A and B reimbursements. This measure should be highly correlated with Medigap premiums. I use a principal components analysis on the AARP and CMS data to generate a county level Medigap Plan F premium. The resulting data is scaled to have the same mean and variance as the AARP premium data. I use this premium, and the benefits defining the Plan F Medigap supplement, as a proxy for the entire Medigap choice set. So for each beneficiary in the model, the choice set includes the proxy Medigap supplement, all of the Medicare Advantage plans offered in that county, and no supplemental coverage.

To account for Medicaid eligibility, I again follow Town and Liu (2003) by using information on state level Medicaid enrollments to deflate the size of the Medicare market of each county within that state. I assume that the state level enrollment rates are invariant across the age/gender groups. To account for employer sponsorship, I use the MCBS data

---

14 Medigap Plan F covers the deductibles and coinsurance for Plan A and Plan B, skilled nursing coinsurance, two additional pints of blood each year, and foreign travel emergency expenditures. It does not include a prescription drug benefit or a dental plan.

15 The correlation between per beneficiary Medicare payments and the AARP premium data is .62.
to adjust the market sizes by using the within age/gender group percentage of those who reported having an employer sponsored health plan. Since the MCBS data set is small, I use the same rate to adjust the market size of each age/gender group across all counties in the country.

Finally, many Medicare Advantage MCOs sell multiple products in the same county. However, both the individual and aggregate plan choice data identify only the MCO and not the specific plan. In the aggregate data, I follow previous research by assuming the plan selected was the least expensive plan offered by that company. In the MCBS, respondents were asked about the premium and drug coverage of their supplemental coverage. Where possible, I matched this information to the data in the Medicare Health Plan Compare database. In cases where a unique matched occurred, I assigned the beneficiary to that specific plan. Otherwise, I again followed precedent by assuming the plan selected was the least expensive offering.

### 4.2 Descriptive Statistics

In this section, I describe the basic characteristics of the MCBS data and the Medicare Managed Care Quarterly/State/County/Plan database and provide descriptive statistics on health care utilization and health plan choice. Figure 1 plots the distribution of log-expenditures of the sum of the five type of health care examined in this study for the Medicare Advantage and FFS groups in the full MCBS. The left intercept is the share of beneficiaries in each group that consumed no health care. This figure gives a clear picture of first order stochastic domination by the FFS group. Over 8% of the Medicare Advantage group consumed no health care compared to only 4% of the FFS group.

Figure 2 gives the log-expenditure distributions for the subsample of the MCBS used in this study. The figure gives the log-expenditure distributions for the Medigap, Medicare Advantage, and FFS-only groups. First order stochastic domination of the Medigap group is evident, but now the Medicare Advantage group dominates the FFS-only group through the 70th percentile. At that point, the FFS-only group dominates the Medicare Advantage group. These patterns are not surprising since Medigap enrollees generally face no cost
sharing for inpatient care, outpatient care, or doctor visits, and while Medicare Advantage enrollees often face very little cost sharing, they are also subject to supply side controls that may hold down expenditures. Of course, part of the difference in the expenditure distributions could be a result of endogenous selection.

Table 1 gives mean expenditures for the five types of health care used in this study by insurance group. On average, the market value of health care service consumed by the Medigap subsample was $7777, follow by the FFS-only group with $5485, and the Medicare Advantage group with $4860. Note the shift toward prescription drugs and dental care in the Medicare Advantage group. 17.8% and 6.0% of expenditures went towards prescription drugs and dental care, respectively, while the corresponding percentages for the Medigap and FFS-only groups are 12.2% and 3.3%, and 9.6% and 1.8%.

Table 2 gives the demographic characteristics of the Medigap, Medicare Advantage, and FFS-only groups. For comparison, I’ve included the same data for the two groups excluded from the sample: those with employer provided health insurance and those eligible for full Medicaid benefits or QMB status. The table excludes those whose Medicare eligibility is based on either disability or ESRD. Of the three included groups, Medigap enrollees are, on average, oldest, have the highest income, and most commonly female.

Table 3 gives average total health care spending by age/gender group. In four of the five age groups, men spend more than women, on average, and for both men and women, average expenditures peaks in the 75-79 age group. Chart 1 gives the share of men and women enrolled in Medicare Advantage by age group. Aside from the less than 65 age group, which is excluded in this study, the pattern shows enrollment shares falling in age for both men and women. Overall, 16.7% of women were enrolled in a Medicare Advantage health plan as were 16.5% of men.

There are 3782 county/MCO/plan combinations in 910 different counties in the Medicare Advantage program. Figure 4 plots the proxy Medigap premium on Medicare Advantage market share for the 910 counties with at least one Medicare MCO. The dashed line is the average Medigap premium for those counties with no Medicare MCOs. The figure indicates that MCOs enter the Medicare Advantage program in counties with generally higher Medi-
gap premiums. Since MCOs generally cluster in urban areas, where establishing provider networks is less costly but the general price level is higher, this is not surprising. However, this could also reflect endogenous entry of MCOs into Medicare Advantage in markets with larger mark-ups among Medigap insurers. This research is not able to examine this possibility because of the data restrictions mentioned above. In the model, the entire Medigap choice set is reduced to one plan so the level of competition among Medigap insurers cannot be assessed. The solid line is a sixth order polynomial trend. The upward trend is consistent with the hypothesized effect of endogenous selection on Medigap premiums. Of course, the figure cannot be taken as evidence of this causal relationship because Medigap premiums and Medicare Advantage market shares are jointly determined.

Both the individual and aggregate data are broadly consistent with previous findings that Medicare Advantage plans experience favorable selection. Generally, Medicare Advantage plans are more likely to attract younger Medicare beneficiaries and, on average, health care expenditures are lower for Medicare Advantage enrollees. There is also evidence that Medigap premiums are higher in counties with greater Medicare Advantage market share, a result that is consistent with the effect of endogenous selection. As noted earlier, lower expenses in the Medicare Advantage group could result from lower risk beneficiaries choosing Medicare Advantage, differences in consumer cost sharing arrangements, or unobserved MCO characteristics such as stringent utilization review practices. Separating these effects is a focus of the model presented next.

5 Economic Model

The consumer’s problem is modeled in two stages involving discrete and continuous choices. In the first stage, the consumer chooses from available health plans given an income and health endowment, knowledge of the distribution of health states, and complete information on the product characteristics of each health plan available in the market. In the second stage, the consumer receives a draw from the distribution of health states and then optimally chooses a vector of health care inputs subject to the budget and productivity constraints
implied by the choice of health plan. The consumer has imperfect information about the
efficacy of treatment and so chooses health care inputs to maximize expected utility where
expectations are taken over a known distribution of outcomes for a given level of treatment.

Health care utilization is decomposed into five different types that are observed in the
MCBS data: inpatient hospital care, outpatient hospital care, doctor visits, prescription
drugs, and dental care. I index these categories with the set $T = \{IP, OP, DV, PM, DC\}$. Together these five types of care comprise 67.4% of all health care spending documented in the MCBS data. Most of the remaining expenses (80.9%), take place in long-term care facilities. So the types of care examined in this study comprise 91.6% of all nonfacility based health events.

The first issue is selecting an appropriate choice variable for the consumer. Many studies
on health care utilization, for example Dowd, et al (1991), Cameron, et al (1988), and Deb
and Trivedi (1997), use event counts such as the number of doctor visits or the number
of inpatient days as the dependent variable. Hunt-McCool, et al (1994) estimate models
explaining the number of doctor visits or hospital admissions as well as budget shares of
expenditures on doctor visits or inpatient care. In presenting the results of the RAND Health

In the MCBS Cost and Use file, utilization is documented at the event level with a dollar
amount associated with each event. I use total expenditures as opposed to event counts
because there is considerable variation in spending within a given event for each category of
utilization. This variation seems to be important in estimating the distribution of unobserved
components of consumer health characteristics. Since the quantities and intensity of specific
treatments at each event are not observed, I use total expenditures for the year as a reasonable
proxy for both the quantity and intensity of treatment. I account for price variation across
regions by deflating total expenditures by a county level price index. The price index I use
is the hospital wage index provided by the CMS.

As noted earlier, no objective measures of posttreatment health are used in this analysis.
The model assumes that there is some unobserved yet quantifiable value of pretreatment
health that is consistent with defining the observed data on health care utilization as the
solution to a consumer constrained optimization problem. As will be discussed later, the estimation algorithm is based on fitting the model to the data by solving for the set of the unobserved components of pretreatment health that is consistent with the data and the model.

5.1 Utilization Choice

5.1.1 Basic Model

In this section, I develop the consumer choice model based on some intuitive properties about health production and basic consumer theory. The model is developed to capture the following properties about health production:

i) Limits on the efficacy of treatment.

ii) Uncertainty about the efficacy of treatment for given quantities of health care.

iii) Diminishing marginal product in the production of health both intensively and extensively.

iv) Substitutability and complementarity across types of care in the production of health.

To simplify the exposition, I first consider the simple one-dimensional consumer choice problem

\[
\max_m U(C, H) \quad \text{s.t.} \quad C = y - pm \quad \text{and} \quad H = H(m, \theta).
\]

Here, \(\theta\) denotes pretreatment health state with larger values indicating worse states of health (so \(-\theta\) is a health endowment), \(m\) denotes a scalar quantity of health care sold at unit price \(p\), \(y\) denotes income, \(H(\cdot)\) is the health production function, \(H\) is posttreatment health, and \(C\) is a numeraire.

For simplicity, I assume that \(H(\cdot)\) is a second order polynomial in \(m\). Higher order polynomials introduce the possibility of nonconvex upper contour sets, which would greatly
complicate estimation. I also assume the initial condition

\[ H(0, \theta) = -\theta. \]

This simply says that if the consumer chooses no health care, then the posttreatment health state equals the health endowment. Finally, I assume \( U_C > 0, \ U_H > 0, \ U_{CC} < 0, \ U_{HH} < 0 \) for all values of \( C \) and \( H \).

The first property I consider is limits on the efficacy of treatment. This implies that there is some finite value of health care, denoted \( \bar{m} \) such that \( H_m > 0 \) for \( m < \bar{m} \) and \( H_m < 0 \) for \( m > \bar{m} \). If we consider consumer preferences on \( C \) and \( m \) (as opposed to \( C \) and \( H \)), the value \( \bar{m} \) can thought of as a satiation point in \( m \) because the indifference curves slope up for \( m > \bar{m} \), i.e., health care becomes a “bad”. This is illustrated in Figure 5.

The assumption that \( H(\cdot) \) is a second order polynomial in \( m \) implies

\[ H_m(m, \theta) = \lambda (\bar{m} - m), \text{ for some } \lambda > 0. \]

Integrating and applying the initial condition gives

\[ H(m, \theta) = \lambda \left( \bar{m}m - \frac{1}{2}m^2 \right) - \theta. \]

Diminishing marginal product on the extensive margin can be incorporated by assuming that \( \bar{m} \) is an increasing function of \( \theta \), i.e., the marginal product at any given level of treatment is higher if the consumer is in a worse pretreatment state. Hence,

\[ H(m, \theta) = \lambda \left( \bar{m}(\theta)m - \frac{1}{2}m^2 \right) - \theta, \text{ with } \bar{m}'(\theta) > 0. \] (1)

The function (1) is the basis for the detailed model of health production described below.

5.1.2 Endowments and Preferences

Consumer \( i \) is exogenously endowed with an income \( y_i \) and a \( 5 \times 1 \) ex ante health state vector \( \xi_i \). I assume that \( y_i \) does not depend on \( \xi_i \). Since most of the individuals in the data are no longer working, this assumption seems reasonable. Each component of \( \xi_i \) is assumed be the sum of a linear-in-parameters index of observable health characteristics \( X_i \) and heterogeneity
that is unobserved to health plans and the researcher but known to the consumer. Define the vector of these unobservables as $\eta_i$ and each component of $\xi_i$ as

$$\xi_{it} = X_i \beta_t + \eta_{it}, \forall t \in T,$$

where $\beta_t$ is a vector of parameters to be estimated. The vector $X_i$ is composed of age and its square and a gender indicator. The vector $\xi_i$ can be thought of as the risk type of the consumer.

After the choice of health plan is made, the consumer receives a vector draw $\epsilon_i$. Define the pretreatment health state of the consumer as the vector

$$\theta_{it} = X_i \beta_t + \eta_{it} + \epsilon_{it}, \forall t \in T.$$

As in the basic model, the consumer has preferences over posttreatment health $H_i$, consumption of a numeraire $C_i$, and a $5 \times 1$ vector of health care purchased from health plan $k$ offered by MCO $j$ (hereafter referred to as plan $jk$), denoted $m_{ijk} = \{m_{itjk}\}_{t \in T}$. Health care enters utility both indirectly, through $C_i$ and $H_i$, and directly. The direct effect captures the disutility associated with the consumption of health care as explained below. I assume that preferences can be represented by the utility function

$$U(C_i, H_i, m_{ijk}) = \frac{C_i^{1-\gamma_1} - 1}{1 - \gamma_1} + \delta \frac{H_i^{1-\gamma_2} - 1}{1 - \gamma_2} - \kappa_1 m_{ijk} + \kappa_2 m_{ijk}^2 - \kappa_{IP}^F 1[m_{iIPjk} > 0]$$

where $\delta, \gamma_1, \gamma_2, \kappa_1, \kappa_2, \kappa_{IP}^F$ are parameters to be estimated. The parameters $\gamma_1$ and $\gamma_2$ measure the degree of risk aversion in the consumption of the numeraire $C$ and the level of health $H$, respectively. The parameter $\delta$ measures the rate at which the consumer is willing to give up utility derived from $C$ for utility derived from $H$.

The parameters $\kappa_1$ and $\kappa_2$ are $5 \times 1$ vectors that capture continuous direct effects of health care on utility. The parameter $\kappa_{IP}^F$ denotes a fixed disutility that is incurred only if some inpatient hospital care is consumed. In addition to being intuitive, the disutility parameters are important if the model is to fit the data well. It is not uncommon in the data for a consumer to choose a zero amount of some type of health care despite a zero first dollar price. As long as the first unit of health care is productive, some nonpecuniary cost is
required if the model is to explain the data. The coefficients on \( m_{ijk} \) and \( m_{ijk}^2 \) capture this nonpecuniary cost. Intuition suggests that each element of \( \kappa_1 \) is positive but each element of \( \kappa_2 \) could be positive or negative. The fixed cost associated with inpatient care \( \kappa_{IP} \) is included because the lowest nonzero inpatient care expenditure exceeds $300 while each of the other four types of care have positive observations below $5. The fixed cost is included to explain this discontinuity.

5.1.3 The Production of Health

In this section, I extend the simple model of health production in (1). To begin, I assume the satiation point is stochastic and define \( \nu_{it} \) as a productivity shock associated with health care type \( t \). In the model, the value of \( \nu_{it} \) is unknown at the time treatment decisions are made. To introduce plan specific heterogeneity in the production of health, I assume the satiation point and the parameter \( \lambda \) depend on plan \( jk \). The nature of this heterogeneity is discussed in the next section. Let \( \chi_{1tj} \) and \( \chi_{2tj} \) denote plan specific unobserved characteristics. Note that \( \chi_{1tj} \) and \( \chi_{2tj} \) are not indexed by \( k \), so the unobserved product characteristics are assumed to be constant across plans within an MCO. Since I assume \( \lambda > 0 \) and \( \bar{m}(\theta) > 0 \), I define \( h_{itjk} \), the posttreatment health for consumer \( i \) in dimension \( t \) from consuming health care purchased from plan \( jk \), as

\[
h_{itjk} = \exp \{ \chi_{1tj} \} \left( \bar{m}_{itj} m_{itjk} - \frac{1}{2} m_{itjk}^2 \right) - \theta_{it}, \tag{2}
\]

where \( \bar{m}_{itj} = \exp \{ \chi_{2tj} + \nu_{it} + \zeta_t g(\theta_{it}) \} \).

where \( g(\theta_{it}) \) is an increasing function the specification of which is described below and \( \zeta_t \) is a parameter to be estimated. The MCO unobservable \( \chi_{2tj} \) shifts the expected satiation point and the MCO unobservable \( \chi_{1tj} \) determines the total product for a given satiation point and \( m_{itjk} \).

To model substitutability and complementarity across types of care in the production of health, I redefine (2) as

\[
h_{itjk} = \exp \{ \chi_{1tj} \} \left( \bar{m}_{itj} \sum_{r \in T} \Psi_{tr} m_{irjk} - \frac{1}{2} m_{itjk}^2 \right) - \theta_{it}. \tag{3}
\]
where $\Psi$ is a $5 \times 5$ matrix the diagonal elements of which are fixed at one. The off-diagonal elements of $\Psi$ allow health care in one dimension to shift the satiation point in another. Specifically, the total and marginal products of care from treatment $t$ increases in $m_{irjk}$ if $\Psi_{tr} > 0$. In this case, health care types $t$ and $r$ are complements. If $\Psi_{tr} < 0$, they are substitutes. To keep the specification as parsimonious as possible while still capturing basic intuition, I assume that dental care is not a substitute or complement for the other types of health care, so $\Psi_{DCt} = \Psi_{tDC} = 0$, $\forall t \in T \backslash DC$.

To map the vector $h_{ijk}$ to the scalar argument $H_i$, I use the constant elasticity of substitution function

$$H_i = \left( \sum_{t \in T} \alpha_t h_{itjk}^\rho \right)^{\frac{1}{\rho}},$$

where $1/(1 - \rho)$ is the elasticity of substitution and the parameters $\{\alpha_t\}_{t \in T}$ denote weights associated with the elements of $\{h_{itj}\}_{t \in T}$.

**Unobserved Product Characteristics** As described above, each health plan has two unobserved (to the researcher) product characteristics that affect the productivity of health care in each dimension. These product characteristics are assumed to be observed by the Medicare beneficiaries at the time the enrollment decision is made. The intuition behind these product characteristics is based on work by Jin (2005) and Cutler et al (2005). In both of these studies, MCOs are differentiated in terms of the quality of health care services provided. Since many MCOs have a restricted provider network or may limit access to specific technologies, it seems natural to differentiate health plans in dimensions other than the premium, benefits provided, and cost sharing arrangements and to suppose that this differentiation will affect consumer choice. The role of the vectors $\{\chi_{1tj}, \chi_{2tj}\}_{t \in T}$ is to capture such differentiation.

This model of product differentiation is inadequate in the following sense. Some health plans may be characterized as providing a low quality product and not very restrictive in the provision of services. Others may exercise considerable control on access to services, but these services may be of a high quality. In the latter case, these controls would be a supply side constraint imposed on the consumer in the utilization choice problem. Unfortunately, there is
nothing in the data that will allow me to separate these two effects. In this model, consumers are assumed to solve an optimization problem free of constraints imposed by the health plan that may lead to inequalities in the first order conditions at interior solutions. Hence, this model of product differentiation confounds a unrestrictive low quality product and a restrictive high quality product. Nonetheless, I believe the model is adequate to the extent that I am only interested in accounting for health plan specific unobserved heterogeneity in the health plan choice problem.

Finally, indexing the Medigap option with $M$ and FFS-only with $FFS$, I assume that $\chi_{1M} = \chi_{1FFS}$ and $\chi_{2M} = \chi_{2FFS}$ in each county. This is intuitive because both Medigap enrollees and FFS only beneficiaries have open provider choice, so these choices should be differentiated by cost sharing arrangements only.

**Specification of $g(\theta_{it})$** In capturing the property of diminishing marginal product on the extensive margin, the specification of $g(\theta_{it})$ requires some care if the model is to retain some intuitive comparative statics. Specifically, if health production grows too quickly in $\theta_{it}$, then a marginal increase in $\theta_{it}$ could lead to lower optimal quantity of health care. Intuitively, this occurs because if the marginal product is higher at some level of $m_{itjk}$, then the marginal product is higher at the inframarginal units as well. In the extreme, it could be that health care becomes so productive that the model may imply that a consumer is better off in more adverse pretreatment states. Therefore, if the model is to maintain the comparative statics that a more adverse state of pretreatment health leads to greater optimal quantities of health care and lower indirect utility, then the expression

$$\frac{\partial \bar{m}_{itj}}{\partial \theta_{it}} = \bar{m}_{itj} \zeta_t g'(\theta_{it})$$

will need to be carefully specified.

Early experiments indicated that the identity function $g(\theta_{it}) = \theta_{it}$ would be a poor choice for the reasons described above. Simulations of optimal health care choices under counterfactual health states often resulted in a positive relationship between indirect utility and $\theta_i$ for large values of $\theta_i$. The problem is that, in this case, $\bar{m}_{itj}$ explodes in $\theta_{it}$ leading to large values of $H_i$ even for very small values of $\zeta_t$. 

23
After a number of other experiments, I concluded that the best choice is

\[ g(\theta_{it}) = \left(1 + \exp\left(\frac{\theta_{it} - a_t}{b_t}\right)\right)^{-1} \]

where \(a_t\) and \(b_t\) are parameters to be estimated. Under this specification, the function \(\bar{m}_{itj}\) is bounded above by \(\exp\{\chi_{2jt} + \nu_{it} + \zeta_t\}\), and (4) becomes

\[ \frac{\partial \bar{m}_{itj}}{\partial \theta_{it}} = \bar{m}_{itj}\zeta_t g(\theta_{it})(1 - g(\theta_{it}))/b_t \]

which is strictly positive, bounded above, and approaches zero as \(\theta_{it}\) becomes arbitrarily large or small. Early tests using this specification were successful in that the optimal quantities are increasing and indirect utilities are decreasing in \(\theta_{it}\). The property of diminishing marginal product on the extensive margin is maintained because \(\frac{\partial \bar{m}_{itj}}{\partial \theta_{it}}\) is strictly positive. A more detailed discussion is presented in Appendix 9.1.

5.1.4 The Budget Constraint

As noted earlier, I use deflated expenditures, which proxies for quantity and intensity of treatment, as the choice variable for the consumer. For many choices of health plans, this will present complication in defining out-of-pocket expenses and marginal prices. If the consumer is facing a simple coinsurance rate such as 20% of expenses after some annual deductible, then defining out-of-pocket expenses and marginal prices is straightforward. If, however, the consumer pays a flat fee per event, then defining total out-of-pockets and marginal prices is not obvious. Some relationship must be established between deflated expenditures and number of events if the model is to conform to the data. To address this, I use the MCBS data to pre-estimate a reduced form relationship between event counts and a polynomial in expenditures, types of insurance, and exogenous county characteristics. I then use the parameter estimates to evaluate the conditional expectation of event counts given the data. Total out-of-pockets for each type of care is the copayment times the predicted number of events and the marginal price is the copayment times the derivative of the conditional expectation of events with respect to expenditures. I present the details in Appendix 9.2.

Let \(p_{tjk}\) denote the appropriately defined marginal price for health care type \(t\) in health
plan $jk$ face by the consumer. Let $oop_{it}$ denote total out-of-pockets incurred by consumer $i$ from health care expenses of type $t$ and
\[ oop_i = \sum_{t \in T} oop_{it} \]
denote total out-of-pockets. Finally, let $\pi_{jk}$ denote the premium for plan $jk$ and $\pi_B$ denote the Medicare Part B premium. The budget constraint is then
\[ C_i = y_i - \pi_{jk} - \pi_B - oop_i. \]

### 5.1.5 Optimization

Given the realization of $\epsilon_i$, the beneficiary chooses $m_{ijk}$ so as to maximize $U$ given the constraints imposed by the choice of health plan $jk$ and income $y_i$. The vector $m_{ijk}^*$ therefore satisfies the vector of first order conditions
\[
U_{m_{ijk}}(m_{ij}, \epsilon_i, j)|_{m_{ijk}^*} = -\frac{p_{tjk}}{[y_i - \pi_{jk} - \pi_B - oop_i]^\gamma_1} - \kappa_{it} + 2\kappa_{it}m_{itjk} \\
+ \delta \int \cdots \int \sum_{r \in T} \alpha_r \Psi_r h^{-1}_{irjk} m_{irj} - \alpha_t h^{-1}_{itjk} m_{itjk} \frac{dF_\nu(\nu_i)}{H_2^{\gamma_2 + \rho - 1}} \leq 0, \forall t \in T, \tag{5}
\]
where $F_\nu$ denotes the distribution function of the vector of productivity shocks $\nu_i$. The components of system $U_{m_{ijk}}^* \leq 0$ corresponding to types of care for which $m_{itjk}^*$ is positive implicitly define demand functions in a neighborhood of $m_{ijk}^*$.\(^{16}\)

### 5.1.6 Corner Solutions

As in most survey data on health care utilization, corner solutions are common in the MCBS utilization data. Only 731 out of 12,305 beneficiaries surveyed and 288 of 5,251 beneficiaries included consume a positive amount of all five types of medical care modeled in this study. The nonlinearities and simultaneity in the demand model described above present significant complication in the treatment of censored data. Many studies have used a two-part approach

\(^{16}\)This is true if the elements of the matrix $\frac{\partial^2 U(m_{ijk}, \theta_i, j)}{\partial \theta_i \partial m_{ijk}}$ corresponding to interior solutions is full rank at $m_{ijk}^*$.\)
to censoring, i.e., first estimating the probability of any care and then estimating a linear model with the appropriate error correction. (For example, see Dowd, et al. (1991).) Other studies that use a discrete dependent variable employ a hurdle model which can be thought of as the product of a probit model and a count model left censored at one.\footnote{The probit component of the model captures the probability of any care and the count model captures the probability of a given number of events given at least one event. The hurdle model attempts to capture the idea that the decisionmaking process for the first event is driven solely by consumer preferences, but the decisionmaking process for subsequent events are driven by both consumer and physician preferences. See Pohlmeier and Ulrich (1995) for an example.} The relationship between the unobservables and observed quantities implied by the structure of this model suggests that a different method will be required.

The approach is analogous to a simple Tobit model in which the analyst integrates the density of the unobserved component of a linear model over the region of the support that is consistent with some observed, censored outcome. The task here is to integrate the joint density of $\epsilon_i$ over the region of $\mathbb{R}^5$ that is consistent with an observed vector of quantity choices that may be a mix of interior and corner solutions. If the beneficiary consumes a positive amount of all five types of care, there is one vector $\epsilon_i \in \mathbb{R}^5$ that is consistent with the observed utilization data. This vector is the value of $\epsilon_i$ that solves the system (5) with equality. However, if at least one component of $m^{*}_{ijk}$ is zero, there will be a region in $\mathbb{R}^5$ that will be consistent with the observed outcome. To establish some notation, let $\Theta$ denote the set of model parameters, $\epsilon_{iC}$ denote the elements of $\epsilon_i$ for which a corner solution is observed, $F_{\epsilon C}$ denote the marginal distribution of $\epsilon_{iC}$, and $F_{\epsilon I}(\cdot | \epsilon_{iC})$ denote the conditional distribution of the elements of $\epsilon_i$ for which an interior solution is observed.

This structure is similar to the demand system analyzed in Wales and Woodland (1983). However, in their model, the unobservables enter the demand equations linearly and the $t^{th}$ unobservable enters the $t^{th}$ demand equation only. These restrictions do not hold in this demand model, causing some complication in integrating over the region of interest. First, the limits of integration are defined by the first order conditions and so are functions of the model parameters and the data. Second, the optimality conditions $U^*_{m_{ijk}} \leq 0$ are nonlinear in the elements of $\epsilon_i$, implying that the region of integration is not Euclidean, but rather a
nonlinear manifold. This manifold is defined by the system $U_{m_{ijk}}^* = 0$ for all $t \in T$ such that $m_{ijk}^* > 0$ and its dimension equals the number of corner solutions.

For example, suppose $m_{iIPjk}^* = m_{iOPjk}^* = m_{iDCjk}^* = 0$, while $m_{iDVjk}^* > 0$, and $m_{iPMjk}^* > 0$ and consider the vector $\tilde{\epsilon}_i$ that solves the full system $U_{m_{ijk}}^* = 0$. Then the components of $\tilde{\epsilon}_i$ corresponding to $IP$, $OP$, and $DC$ represent what I refer to as reservation states in that they are the realizations of $\epsilon_i$ at which the consumer is just indifferent between purchasing the first increment of health care and not in the dimensions corresponding to the corner solutions, given the observed positive quantities purchased. So the values of $\tilde{\epsilon}_i$ corresponding to $IP$, $OP$, and $DC$, which I denote $(\epsilon_{iIPjk}(\Theta), \epsilon_{iOPjk}(\Theta), \epsilon_{iDCjk}(\Theta))$, are the appropriate censoring points and, therefore, the limits of integration in the expression for the joint density of $\epsilon_i$ that is consistent with the data.

The limits of integration are indexed by $jk$ since they are functions of plan characteristics via the system $U_{m_{ijk}}^* = 0$. Different product characteristics result in different censoring points for the same observed $m_{ijk}^*$. In this way, the MCO characteristics influence the probability (loosely speaking) of any care as well as the quantity of care given some care.

For any combination of $(\epsilon_{iIP} < \epsilon_{iIPjk}(\Theta), \epsilon_{iOP} < \epsilon_{iOPjk}(\Theta), \epsilon_{iDC} < \epsilon_{iDCjk}(\Theta))$, there exists some pair $(\epsilon_{iDV}, \epsilon_{iPM})$ that solves the system $U_{m_{ijk}}^*|_{\forall t \in T:m_{i}^*} = 0$. The set of these pairs comprise of the region of interest. So the expression for the joint density of $\epsilon_i$ that is consistent with the data is

$$\int_{-\infty}^{\epsilon_{iDV}(\Theta)} \int_{-\infty}^{\epsilon_{iOP}(\Theta)} \int_{-\infty}^{\epsilon_{iDC}(\Theta)} f_{\epsilon_i}(\epsilon_{iDV}(\Theta), \epsilon_{iPM}(\Theta)|\epsilon_{iC}) \|\mathbb{J}_{ij}\| dF_{\epsilon_i}(\epsilon_{iC}) \quad (6)$$

where $(\epsilon_{iDV}(\Theta), \epsilon_{iPM}(\Theta))$ denote implicit functions relating the elements of $\epsilon_i$ along the intersection of the manifolds defined by the first order conditions corresponding to the types of health care for which there is positive consumption, $DV$ and $PM$. The Jacobian matrix is denoted by $\mathbb{J}_{ij}$ (in this case a $2 \times 2$ matrix), and $\|\mathbb{J}_{ij}\|$ is the Jacobian of the transformation from $\epsilon_i$ to $m_{ijk}$ when $m_{iIPjk}^* = m_{iOPjk}^* = m_{iDCjk}^* = 0$. 

27
5.1.7 Value of a Health Plan

Finally, I define the value function of plan $jk$ to beneficiary $i$ as the expected indirect utility where expectations are taken over the known distribution of health states.

$$V_{ijk}(\eta_i) = \int \cdots \int [U \left(m_{ij}^*, \theta_i, y_i \right) \mid \epsilon_i] \, dF(\epsilon_i)$$  \hspace{1cm} (7)

5.2 Health Plan Choice

In the first stage of the model, beneficiaries choose a health plan from the available options. The choice set will vary by county since each will have a different set of MCOs participating in Medicare Advantage.

In addition to the health plan valuation described in the previous section, the beneficiary’s valuation of a health plan may vary for idiosyncratic reasons. Let $e_{ijk}$ denote this idiosyncratic variation in the value of beneficiary $i$ for plan $jk$. I assume $e_{ijk}$ is distributed Generalized Extreme Value (GEV). Define $e_i$ as a $2 + \sum_{j=1}^{J} K_j$ vector of idiosyncratic terms where $J$ is the number of Medicare MCOs in the county and $K_j$ is the number of plan offered by MCO $j$. This idiosyncratic component represents valuation of the plan to the beneficiary that is independent of valuation based on expectations of health state and the consumption of other goods. Examples of such variation are the amount of experience the beneficiary has with receiving care in a managed care setting, a convenient location of a clinic, or the presence of a familiar physician in the provider network of MCO $j$. I assume the value of plan $jk$ to beneficiary $i$ conditional on $\eta_i$ is

$$\tilde{V}_{ijk}(\eta_i) = V_{ijk}(\eta_i) + e_{ijk}.$$

Since a component of this idiosyncratic variation reflects the beneficiary’s general attitude towards managed care, it is not appropriate to assume that the draws $e_{ijk}$ are mutually independent across MCOs or across plans within MCOs. Therefore, I assume a nested multinomial logit choice probability conditional on the vector $\eta_i$. In this nesting structure, the Medicare beneficiary is modeled as first choosing between the Medigap plan, no supplemental coverage, and a Medicare Advantage option. If Medicare Advantage is chosen, the beneficiary then chooses among MCOs and then among plans if the selected MCO offers multiple plans.
With these assumptions, and under the assumption that the beneficiary chooses the health plan with the largest $V_{ijk} (\eta_i)$, the probability that beneficiary $i$ chooses plan $jk$ conditional on $\eta_i$ is

$$P_{ijk}(\eta_i) = \frac{\exp \left\{ \frac{V_{ijk}(\eta_i)}{(1-\sigma_a)(1-\sigma_w)} \right\}}{I^\sigma_a_j J^\sigma_a \sum_{j=1}^J \exp \left\{ \frac{V_{ijk}(\eta_i)}{(1-\sigma_w)(1-\sigma_a)} \right\} + \sum_{j=1}^J J^{1-\sigma_w}}$$

(8)

where $I_j = \sum_{k=1}^{K_j} \exp \left\{ \frac{V_{ijk}(\eta_i)}{(1-\sigma_w)(1-\sigma_a)} \right\}$ and $J = \sum_{j=1}^J J^{1-\sigma_w}$.

The parameter $\sigma_a$ is approximately the correlation of the errors across MCOs and $\sigma_w$ is approximately the correlation of the GEV errors across plans within an MCO.

5.3 Aggregate Data

The definition of the choice probability at the individual level is also used to model the aggregate health plan choice data. I take the mixed logit approach similar to Berry, Levinsohn, and Pakes (BLP, 1995), Brownstone and Train (1999), and related literature.

As noted earlier, the aggregate health plan choice data is observed at the level of the MCO, not health plan. Beneficiary income is not observed as well. I estimate a parametric income distribution using data from the MCBS and the Census Bureau for each county/gender/age group. Let $F_{yX_i}$ denote this estimated distribution for age/gender group $X_i$ in a given county. Given (8), we can express the predicted $X_i$ specific market share of MCO $j$ as

$$P_{X_i,j} = \int \left[ \int \cdots \int P_{X_i,j}(\eta_i, y_i) dF_{\eta}(\eta_i) \right] dF_{yX_i}(y_i).$$

(9)

where $P_{X_i,j}(\eta_i, y_i) = \frac{\exp \left\{ \frac{V_{X_i,j}(\eta_i, y_i)}{(1-\sigma_a)} \right\}}{J^\sigma_a \sum_{j=1}^J \exp \left\{ \frac{V_{X_i,FFS}(\eta_i, y_i)}{(1-\sigma_w)} \right\} + \sum_{j=1}^J J^{1-\sigma_w}}$

and now

$$J = \sum_{j=1}^J \exp \left\{ \frac{V_{X_i,j}(\eta_i, y_i)}{(1-\sigma_a)} \right\}.$$

Note that both the nesting structure and the unobserved heterogeneity due to $\eta_i$ and $y_i$ remove the unreasonable substitution patterns in the market shares across health plan.
options that are associated with simple multinomial logit models\textsuperscript{18}.

Finally, aggregate health plan choice data for Medigap supplements is not available at the county level. I impose the restriction that for each county/age/gender group, those not enrolling into a Medicare Advantage plan are apportioned into the Medigap supplement at the rate observed nationally. Since 36\% of Medicare beneficiaries chose a Medigap supplement in 2000 and 9\% chose no supplementary coverage, I set

\[ P_{X_iM} = \frac{4}{5} \left( 1 - \sum_{j=1}^{J} P_{X_{ij}} \right). \]

6 Estimation

The model parameters are estimated using the Method of Simulated Moments using two sets of moments following the approach of Imbens and Lancaster (1994). One set of moments is derived from the derivatives of a simulated log likelihood function constructed using the MCBS data, and a second set of moments is derived from the aggregate health plan choice data. Below I discuss distributional assumptions, the construction of the likelihood function and the moments conditions and the respective simulators, the objective function and estimation algorithm, and identification.

6.1 Distributional Assumptions

As noted above, I assume the vector defining the idiosyncratic component of beneficiary health plan valuation, \( e_i \), is distributed GEV,

\[ e_i \sim F_e(e_i) = \exp[-f_e(\exp{-e_iFFS}, \exp{-e_iM}, ...., \exp{-e_iJK})] \]

where the function \( f_e \) will be selected on a county-by-county basis to reflect the size and structure of the choice set faced by the Medicare beneficiaries in that county. The general

form of \( f_e \) is

\[
f_e(e_i) = \left[ \exp \{-e_{iF}S\} + \exp \{-e_{iM}\} + \left( \sum_{j=1}^{J} \sum_{k=1}^{K_j} \exp \{-e_{ijk}\} \right) \right]^{1-\sigma_w}\left(1-\sigma_a\right).
\]

I assume that the vectors \( \epsilon_i, \eta_i, \) and \( \nu_i \) are each distributed normally but the vectors are mutually independent. The assumption that \( \nu_i \) is independent of \( \epsilon_i \) and \( \eta_i \) is made for simplicity. The assumption \( \eta_i \) and \( \epsilon_i \) are independent can be made without loss of generality because the sum of two normal variables can always be defined as the sum of two orthogonal components. While variances of \( \epsilon_i \) and \( \eta_i \) are separately identified, the means are not. In addition, the variance of \( \nu_i \) is identified but the mean is not separately identified from the unobserved vector \( \chi_{2j} \). Therefore, I assume

\[
\epsilon_i \sim N(\mu, \Omega_\epsilon) \\
\eta_i \sim N(0, \Omega_\eta) \\
\nu_i \sim N(0, \Omega_\nu).
\]

For simplicity, I assume that \( \Omega_\nu \) is a diagonal matrix. The vectors \( \chi_{1j} \) and \( \chi_{2j} \) are not treated as random variable but rather as fixed effects that are numerically solved for in the estimation algorithm.

### 6.2 Individual Data

The likelihood contribution of an individual is the probability of choosing the observed health plan multiplied by the joint density of \( \epsilon_i \) that is consistent with the observed utilization data given the health plan choice, integrated over the joint density of \( \eta_i \). Consider the example in section 5.1.6. The likelihood contribution of an individual with that configuration of interior and corner utilization solutions and having selected health plan \( jk \) is

\[
L_i(\Theta) = \frac{\int \cdots \int [P_{ijk}(\eta_i) \int_{-\infty}^{\epsilon_{iDC}[\Theta]} \int_{-\infty}^{\epsilon_{iDV}[\Theta]} \int_{-\infty}^{\epsilon_{iPM}[\Theta]} f_{\epsilon_i}(\epsilon_{iDV}(\Theta), \epsilon_{iPM}(\Theta)) | \epsilon_{iC}) | \| \eta_i \| | dF_{\epsilon_i}(\epsilon_{iC})] dF_{\eta}(\eta_i) \] \[
\int \cdots \int P_{ijk}(\eta_i) dF_{\eta}(\eta_i)
\]

(10)
I approximate (10) using simulation and numerical methods. Numerical methods are used to solve for the set of unobservables that are consistent with the observed utilization data as optimal choices. Numerical methods are also used to solve for optimal quantities of health care for counterfactual pretreatment health states. Simulation methods are used to approximate integrals and to draw from marginal distributions of the unobservables in the cases in which corner solutions are observed. I present the details of simulation and numerical methods in Appendix 9.3. Table 5 provides some Monte Carlo results on the approximation error associated within simulating the integral (6).

6.3 Aggregate Data

I use the aggregate health plan choice data to uncover the health plan specific vector \( \chi_j = (\chi_{1j}, \chi_{2j}) \) and form moment conditions in a manner analogous to BLP. The algorithm involves first solving for the vector \( \chi_j \) that makes the observed market shares fit the theoretical market shares for a given guess of the model parameters and then forming moment conditions by assuming the vector \( \chi_j \) is mean independent of a set of instruments.

For the Medigap plan and each MCO in a county, there are ten elements of \( \chi_j \) and ten age/gender market shares. (Recall that I assume \( \chi_{FFS} = \chi_M \) and I drop the less than 65 age group.) Let \( \chi \) denote the \( 10(J + 1) \) vector of unobserved health characteristics in the market , \( P(\chi) \) denote the \( 10(J + 1) \) vector of market shares with typical element (9), and \( P \) denote the corresponding vector of observed market shares. The simulator of (9) is

\[
\tilde{P}_{Xij} = \frac{1}{R} \sum_{r=1}^{R} \tilde{P}_{Xij}(\eta_{ir}, y_{ir}),
\]

where

\[
\tilde{P}_{Xij}(\eta_{ir}, y_{ir}) = \frac{\exp \left\{ \tilde{V}_{Xij}(\eta_{ir}, y_{ir}) \right\}}{\mathcal{J}^{\sigma_a} \left[ \exp \left\{ \tilde{V}_{X_{FFS}}(\eta_{ir}, y_{ir}) \right\} + \exp \left\{ \tilde{V}_{X_{M}}(\eta_{ir}, y_{ir}) \right\} \right] + \mathcal{J}^{1-\sigma_a}}.
\]

Let \( \tilde{P}(\chi) \) denote the \( 10(J + 1) \) vector of simulated market shares.

For each county, I numerically solve for the vector \( \chi \) by solving the system

\[
\frac{\partial \tilde{P}(\chi)}{\partial \chi'} \left( P - \tilde{P}(\chi) \right) = 0.
\]
The standard assumption in this literature is that the unobserved product characteristics are mean independent of all observed product characteristics other than the price. I continue that assumption here. Let $Z$ denote set of all observed characteristics of all MCOs in the county other than the premium. Then by assumption

$$E [\chi_j(\Theta)|Z] = 0.$$  \hfill (13a)

### 6.4 Objective Function

To develop an objective function for the model, I follow Imbens and Lancaster (1994). They discuss the efficient application of aggregate data to the estimation of microeconometric models. In their model, score contributions based on individual data are treated as moment conditions and estimation is made more efficient by adding moments derived from aggregate data. They apply the technique using individual and aggregate data on labor market outcomes. The analogous outcome variable here is health plan choice. The distinction in this research is that I do not use the aggregate health plan choice data per se to construct additional moments, but rather I use the moment conditions defined by (13a) which are derived using the aggregate health plan choice data.

Let $\tilde{L}_i(\Theta)$ denote the simulator of the likelihood contribution in (10) as described in Appendix 9.3 and define the derivatives of the simulated log likelihood function as

$$\Lambda_1(\Theta) = \frac{1}{N} \sum_{i=1}^{N} \frac{\partial \ln \tilde{L}_i(\Theta)}{\partial \Theta}. \hfill (14)$$

By definition $E[\Lambda_1(\Theta_o)] = 0$ where $\Theta_o$ denotes the true parameter values. As in Imbens and Lancaster (1994), I treat (14) as a moment condition. I use the sample analog of (13a) as a second set of moments. Let $G_j(Z)$ denote a $L \times 10$ matrix of functions of the instruments and define

$$\Lambda_2(\Theta) = \frac{1}{910} \sum_{c=1}^{910} \frac{1}{J_c + 1} \sum_{j=1}^{J_c+1} G_j(Z) \chi_j(\Theta)$$

where $c$ indexes counties and $J_c$ is the number of Medicare MCOs in county $c$. (Recall that there are 910 counties with at least one Medicare MCO.) By (13a), $E [G_j(Z)\chi_j(\Theta_o)] = 0$, implying $E [\Lambda_2(\Theta_o)] = 0$. 

33
The MSM estimator of $\Theta$ is then

$$\hat{\Theta}_{MSM} = \arg\min_{\Theta} [\Lambda(\Theta)' W \Lambda(\Theta)]$$

where $\Lambda(\Theta)' = [\Lambda_1(\Theta)' \Lambda_2(\Theta)']$ and $W$ is a positive definite weighting matrix. Following Hansen (1982), the optimal weighting matrix is

$$W = E [\Lambda(\Theta)\Lambda(\Theta)']^{-1}.$$  

In the model, there are 102 parameters and $102 + L$ moments. I solve the model in two rounds using a block diagonal weighting matrix where the upper left block is an identity matrix and the lower right block is the inverse of the expected inner product of the instruments in the first round to obtain a consistent estimate $\hat{\Theta}$ and the sample analog of the optimal weighting matrix evaluated at $\hat{\Theta}$ in the second.

The estimation involves iteratively evaluating the moment conditions for a given guess of the model parameters and then updating the guess of the model parameters using standard Newton based methods. Each iteration requires the evaluation of the likelihood contribution of each beneficiary in the MCBS data as described in section 6.2 and solving for the vector of MCO unobserved characteristics in each county as described in section 6.3.

### 6.5 Identification

Identification of the model parameters comes from both variation in the data and functional form assumptions. The necessary condition for identification in this context is that the matrix $\frac{\partial \Lambda(\Theta)}{\partial \Theta}$ must be full rank. As in Imbens and Lancaster (1994), all of the model parameters are identified using the individual level data only. The aggregate data allows me to solve for the unobserved MCO characteristics and increases the efficiency of the estimation. Generally, all of the parameters associated with the health production constraint and the beneficiary’s health endowment and preferences ($\gamma_1, \gamma_2, \delta, \kappa, \beta, \alpha, \rho, \zeta, a, b, \Psi, \Omega, c, \mu$) are identified through covariation in observable beneficiary characteristics ($y_i, X_i$), observed characteristics of health plans, and the observed health care expenditures only. The parameters $\Omega_\eta$ and $\Omega_c$ require the individual health plan choice data and variation in expected...
indirect utilities across health plan choices for them to be separately identified. Without this variation, only $\Omega_\epsilon + \Omega_\eta$ is identified.

The linear independence of the columns of $\frac{\partial \Lambda(\Theta)}{\partial \Theta}$ comes from the covariation in the data and the functional form assumptions. The economic interpretation of the parameters, of course, relies on the structure of the model. Although space considerations preclude a presentation of the linear independence of the columns of $\frac{\partial \Lambda(\Theta)}{\partial \Theta}$ for all parameters, I discuss one identification issue commonly cited in the literature on health plan choice and health care utilization.

As noted earlier, health plan characteristics in a model of health care demand are endogenous if the unobservables in the health plan choice and health care demand equations are correlated. A model of joint determination captures the stochastic dependence but still leaves a question of identifying the effect of a covariate on the two components of the model. For example, the effect of an exogenous variable such as age or income on health plan choice may differ from its effect on health care demand. An instrumental variables approach could use exclusion restrictions to derive instruments, but often plausible exclusion restrictions may not exist. In this case, it would be difficult to assume, for example, that age or gender affects health care demand but does not affect health plan choice. Another method requiring stronger distributional assumptions is the error correction method as described in Heckman (1979). Cameron et al (1988) employ a method less dependent on distributional assumptions by estimating a plan choice probability and then using this probability as an instrument in the health care demand equation.

In this model, the specification of expected indirect utility as the basis for health plan valuation implies parameter restrictions that enable the model to identify these two marginal effects in the absence of exclusion restrictions. This is because a covariate, say age, affects both the utilization and health plan choice decisions only through the vector $\theta_i$ and the parameter vector giving the marginal effect of age on $\theta_i$, $\beta_{age}$, is the same in both decisions. For example, the marginal effect of the covariate age on health care demand given health
plan choice is given by the vector

\[
\frac{\partial m_{ijk}^*}{\partial X_{i, age}} = -\left[ U_{m_{ijk}m_{ijk}}^* \right]^{-1} \left[ U_{m_{ijk}\theta_i}^* \frac{\partial \theta_i}{\partial X_{i, age}} \right]
\]

and the marginal effect of age on the probability of choosing health plan \( jk \), conditional on \( \eta_i \), is

\[
\frac{\partial P_{ijk}(\eta_i)}{\partial X_{i, age}} = \frac{\partial P_{ijk}(\eta_i)}{\partial V_{ijk}(\eta_i)} \sum_{t \in T} \frac{\partial V_{ijk}(\eta_i)}{\partial \theta_{it}} \frac{\partial \theta_{it}}{\partial X_{i, age}}
\]

\[
= \frac{\partial P_{ijk}(\eta_i)}{\partial V_{ijk}(\eta_i)} \sum_{t \in T} \frac{\partial V_{ijk}(\eta_i)}{\partial \theta_{it}} \beta_{age,t}.
\]

The parameter vector of interest \( \beta_{age} \) is same in both of these marginal effects, so parameter restrictions, as opposed to exclusion restrictions, are used to identify the parameters.

### 6.6 Policy Simulation

With the estimated parameters, I evaluate consumer surplus for each county/gender/age group using methods described in Small and Rosen (1981). The consumer surplus for age/gender \( X_i \) in county \( c \) is

\[
CS_{X_i c} = \int \cdots \int \frac{1}{U_y} \ln \left[ \left( \sum_{j=1}^{J} \left[ \sum_{k=1}^{K} \exp \left\{ \frac{V_{X_{ij}}(\eta_i, y_i)}{(1 - \sigma_a)(1 - \sigma_w)} \right\} \right]^{1 - \sigma_w} \right)^{1 - \sigma_a} + \exp \{V_{X_i FF}(\eta_i, y_i)\} + \{V_{X_i M}(\eta_i, y_i)\} \right] dF_{\eta}(\eta_i) dF_{yX_i}(y_i)
\]

where \( U_y \) denotes the marginal utility of income.

I then remove all Medicare MCOs from the choice set and solve for how Medicare beneficiaries would sort into Medigap or no supplemental coverage if they were faced with these two options only. This requires finding a premium for the Medigap option that is consistent with this sorting. If endogenous selection has the hypothesized effect, this counterfactual Medigap premium should be lower than the observed Medigap premium in most counties. I again evaluate consumer surplus under for each county/gender/age group. The difference will provide an estimate of the distribution of the change in consumer surplus within age/gender.
groups across counties due to Medicare Advantage. It will reveal the cost of endogenous selection into Medicare Advantage, which drives up costs for those who remain outside the program.

7 Reduced Form Analysis

In this section, I present the results of a reduced form analysis on the distributions of the change in consumer surplus within age/gender groups due to Medicare Advantage. I use a standard regression equation using market shares data based on a nested logit model of health plan choice. The unit of observation is age/gender/county group. I predict Medigap premiums in the absence of Medicare Advantage using a regression model in which the Medigap premium is expressed as a function of Medicare Advantage market share and an index that captures the distribution of risk characteristics of the Medicare Advantage enrollment given the market share.

The primary data sources are the Medicare Managed Care Quarterly/State/County/Plan database, the Medicare Health Plan Compare database, the AARP Medigap premium data, and the Area Resource File (ARF). The ARF provides county level data on health care provider counts, health care facilities, the number of HMOs in the commercial sector, and the market share of the commercial HMOs.

7.1 Reduced Form Model

As in section 6, define the probability that an individual in age/gender group \( i \) chooses MCO \( j \) in county \( c \) as

\[
P_{icj} = \frac{\exp \left\{ \frac{V_{icj}}{1-\sigma} \right\}}{\mathcal{J}_c^\sigma \left[ \exp \left\{ V_{icM} \right\} + \mathcal{J}_c^{1-\sigma} \right]},
\]

where

\[
\mathcal{J}_c = \sum_{j=1}^{J_c} \exp \left\{ \frac{V_{icj}}{1-\sigma} \right\}.
\]

Similarly, define the analogous probability for the Medigap option as

\[
P_{icM} = \frac{\exp \left\{ V_{icM} \right\}}{\exp \left\{ V_{icM} \right\} + \mathcal{J}_c^{1-\sigma}}.
\]
and the probability that an individual in age/gender group \( i \) chooses MCO \( j \) in county \( c \) conditional on having chosen a Medicare MCO as

\[
P_{icj|MA} = \frac{\exp \left\{ \frac{V_{icj}}{1 - \sigma} \right\}}{J_c}.
\]

Then it can be shown that

\[
\ln P_{icj} - \ln P_{icM} = V_{icj} - (1 - \sigma)V_{icM} + \sigma \ln P_{icj|MA}.
\]

(15)

I define

\[V_{icj} = X_{cj}\beta + Z_{icj}\gamma - \alpha\pi_{cj} + u_{icj},\]

where \( X_{cj} \) denotes a vector of observed product characteristics of MCO \( j \) in county \( c \). This vector contains the inpatient and outpatient admission fees, the physician copay, a drug benefit indicator, the generic drug copay, a dental benefit indicator, and the dental copay. Since an MCO may offer several plans in a given county, I use the average of these terms in cases where more than one plan is offered. The vector \( Z_{icj} \) denotes a set of interactions of the terms in \( X_{cj} \) and the consumer characteristics age, gender, and income. I use the midpoint of the age range of each age/gender group (87 for the 85 and over group) and I use the county level median household income for the age group.\(^{19}\) The premium is denoted \( \pi_{cj} \), and \( u_{icj} \) denotes unobserved valuation of MCO \( j \) for age/gender group \( i \). I also define the value of the Medigap option as

\[V_{icM} = -\alpha\pi_{cM}.
\]

Then, equation (15) can be written as the regression equation

\[
\ln P_{icj} - \ln P_{icM} = X_{cj}\beta + Z_{icj}\gamma - \alpha\pi_{cj} + (1 - \sigma)\alpha\pi_{cM} + \sigma \ln P_{icj|MA} + u_{icj}.
\]

(16)

I follow Berry (1994) and BLP by assuming that the MCO premium and \( P_{icj|MA} \) are endogenous but the disturbance \( u_{icj} \) is mean independent of \( X_{cj} \) and \( Z_{icj} \). To correct for this endogeneity, and to account for the nonlinear parameter restriction on \( \pi_{cM} \), I estimate (16) using nonlinear instrumental variables. I use the county level AAPCC payment and the

\(^{19}\)The median household income data was obtained from the Census Bureau Summary File 3.
set of AAPCC payments in all counties of operation of each Medicare MCO operating in county $c$ as instruments. I generate variation at the age/gender group level by interacting the instruments with the age, gender, and income of group $i$. The identifying assumptions are the premium responds to the payment received by the MCO from Medicare, but the payment does not directly affect the odds ratio of a beneficiary joining the MCO. Similarly, the AAPCC payments in other counties of operation of each Medicare MCO operating in county $c$ are reflected in their respective premiums, and therefore are related to $P_{icj|MA}$, but do not directly affect the odds ratio of a beneficiary joining MCO $j$.

To estimate the change in consumer surplus, I will need to predict the Medigap premium in the absence of Medicare Advantage for each county. I estimate a regression model that explains the Medigap premium with a set of exogenous covariates and a measure of the Medicare Advantage market share. Ideally, the Medigap premium would be expressed as function of all ten age/gender group market shares. However, since the Medigap premium and Medicare Advantage market shares are jointly determined, this would require finding instruments for at least ten endogenous variables. Instead, I express the Medigap premium as a second order polynomial in the total Medicare Advantage market share in county $c$ and develop an index that captures the distribution of risk characteristics given the overall market share. To develop this index, I use the group level market shares and the average group level health care spending from the MCBS data. Let $M_i$ denote the average health care spending in the MCBS survey for age/gender group $i$. I define the Medicare Advantage risk index for county $c$ as

$$MA risk_c = \frac{\sum_i M_i \sum_j P_{icj}}{\sum_i \sum_j P_{icj}}.$$  \hspace{1cm} (17)

So the index is just the weighted average of the average group level MCBS expenses where the weights are the group level Medicare Advantage market shares conditional on the overall Medicare Advantage market share in that county.

I estimate the regression equation

$$\ln \pi_{cM} = \beta_1 M share_c + \beta_2 M share_c^2 + \beta_3 M risk_c + Z_c \gamma + u_c$$  \hspace{1cm} (18)

where $M share_c$ denotes the overall market share of Medicare Advantage MCOs in county $c$. 

39
and \( Z_c \) denotes a set of exogenous covariates. The vector \( Z_c \) contains the data from the ARF, county level data from the Census Bureau on per capita income, population density, and an urban indicator, the CMS hospital wage index, a county level risk index provided by the CMS that is based on FFS Medicare expenses from previous years, and the number of Medicare eligibles in each age/gender group. I include the number of HMOs and HMO market share in commercial sector to control for the possibility of spillover effects as described in Baker (1997).

I estimate (18) using instrumental variables in which the Medicare Advantage market share and Medicare Advantage risk index are treated as endogenous. I use polynomials of the AAPCC as instruments. The identifying assumption is that Medicare MCO premiums respond to the AAPCC and so affect Medicare Advantage market shares, but the AAPCC should not affect the Medigap premium conditional on the Medicare Advantage market share. With the estimated parameters, I derive the predicted Medigap premium in the absence of Medicare Advantage as

\[
\hat{\pi}_{cM} = \exp \left\{ \hat{\beta}_3 MArisk_{c0} + Z_c \hat{\gamma} + \hat{u}_c \right\}
\]

where \( MArisk_{c0} \) is (17) evaluated at \( P_{icj} = 0 \) for all \( i \) and \( j \).\(^{20}\) Finally, following Sherman and Rosen (1981), the change in consumer surplus due to Medicare Advantage for age/gender group \( i \) in county \( c \) can be written as

\[
\Delta CS_{ic} = \frac{1}{\alpha} \ln \left[ \exp \{ V_{icM} \} + \left( \sum_{j=1}^{J_c} \exp \left\{ V_{icj} \left( \frac{1}{1-\sigma} \right) \right\} \right)^{1-\sigma} \right] - \frac{V_{icM}}{\alpha} \left. \hat{\pi}_{cM} \right|_{\hat{\pi}_{cM}}.
\]

### 7.2 Results

The nonlinear IV estimation results of equation (16) are given in Table 6. All dollar denominated variables are scaled by 1,000. The estimate of the coefficient on MCO premium \( \alpha \) is 0.6073 and is precisely estimated (\( p < 0.01 \)). This is an order of magnitude lower than the result in Town and Liu (2003). (They estimate \( \alpha \) to be 0.0053 using nonscaled premium data.) This may be because they assume that the GEV error associated with Medigap is

\(^{20}\)This expression is not zero and can be easily derived using L’Hopital’s rule.
correlated with the GEV errors associated with Medicare MCOs, and so the model they estimate does not include the nonlinear parameter restriction on the Medigap premium. The correlation between the Medicare MCO GEV errors is estimated to be 0.6868 and is also precisely estimated \((p < 0.001)\). Town and Liu (2003) estimate this parameter to be 0.58. The presence of a drug benefit is positively associated with Medicare Advantage enrollment as is the presence of a dental benefit. Generic drug and dental copays, conditional on the presence of a benefit, are negatively associated, although the coefficient on dental copay is not statistically significant \((p = 0.173)\). Medicare Advantage enrollment is negatively associated with the inpatient admission fee, but positively associated with the outpatient admission fee and the physician copay. The coefficient on the physician copay is not statistically significant \((p = .124)\), but the coefficient on outpatient admission fee is significant \((p = 0.001)\). These two counterintuitive results may be a result of specification error that would arise if the cost sharing terms are correlated with the unobservable, e.g., MCOs with higher copays also offer a higher unobserved quality of service.

The IV estimation results of equation (18) are given in Table 7. The results indicate a concave relationship between Medicare Advantage market share and the log Medigap premium. The marginal effect peaks at a Medicare Advantage market share of .442. Since only 18 of the 910 counties have a Medicare Advantage market share that exceeds .45, the estimated relationship between the variables is almost always positive in the data. The coefficient on the Medicare Advantage risk index is precisely estimated \((p = 0.001)\) and has the negative sign that is consistent with endogenous selection. Conditional on Medicare Advantage market share, a higher risk index, indicating a distribution of risk characteristics less favorable to Medicare MCOs, is associated with lower Medigap premiums.

The predicted Medigap premiums in the absence of Medicare Advantage are on average $26 lower than the observed Medigap premiums. In 13 counties, the predicted Medigap premium is actually higher than the observed premium. As expected, this occurs in counties with very high values of the of the Medicare risk index.\(^{21}\)

\(^{21}\)In the 897 counties in which the predicted Medigap premium is lower than the observed Medigap premium, the average of \(MA_{risk}\) is $63,035 while in the 13 counties in which it is higher, the average of \(MA_{risk}\)
Tables 8 through 10 give the main welfare results of the reduced form analysis. Table 8 gives the weighted average of the change in consumer surplus due to Medicare Advantage for each age/gender group. The weights are the number of Medicare eligibles in each county. The distributions within each group is highly skewed so the mean, standard deviation, and median are provided. The average of the change in consumer is positive for all groups. However, as expected, the increase in surplus diminishes with age (except for the male 70-74 group, which has a higher change in surplus than the male 65-69 group).

In each age/gender group, there are some counties for which the change in consumer surplus is actually negative. Aggregating the Medicare eligible population by age/gender group over these counties and dividing by the total eligible population gives the percentage of each age/gender group for which the change in consumer surplus is negative. The results are given in table 9. The percentages range from 0.5% to 3.4%. Consistent with the intuition behind the Rothschild-Stiglitz model, older, i.e., higher risk Medicare beneficiaries are more likely to be less well off when multiple health plan choices are available and beneficiaries systematically sort into health plans based on risk type.

In another measure of the welfare effect of Medicare Advantage, table 10 provides the predicted share of the 2000 Medicare Advantage enrollment that would have selected Medigap if it were priced as it would have been in the absence of Medicare Advantage. The results show that 3.89%-5.25% would have done so with women and older beneficiaries more likely to have done so. The overall percentage is 4.5%. So the prediction is that a significant share of Medicare Advantage enrollees were not enrolled in Medicare Advantage because of extra benefits or lower premiums, but because their preferred health plan was too expensive.

8 Conclusion

This paper studies how Medicare beneficiaries respond to the range of health plan choices they face, including traditional fee-for-service Medicare, Medigap policies, and managed care options through Medicare Advantage. A major goal of the paper is to understand is $99,508.
the welfare trade-off between providing more health plan options through the Medicare Advantage program and the possibility that endogenous selection into Medicare Advantage raises premiums for Medigap supplementary coverage. Theory predicts that the welfare effect will vary across risk types with higher risk types being more likely to be adversely affected.

To address this issue, this paper has presented a structural model of health plan and health care demand that contributes to the existing literature by incorporating the notion of health care as a derived demand into the consumer choice problem. The model captures the technological relationship between health care and health by including many of the aspects thought to be important in the production of health but commonly not treated in the applied literature. These include limits on the efficacy of treatment, uncertainty about the efficacy of treatment, diminishing marginal product both intensively and extensively, and substitutability and complementarity across types of treatment.

This paper has also presented reduced form evidence that the diversity of health plan choices brought about by Medicare Advantage has resulted in significant increases in consumer surplus across all age/gender groups. However, the reduced form evidence also suggests that the change in consumer surplus is not uniform within or across age/gender groups and that higher risk beneficiaries are more likely to be adversely affected by Medicare Advantage. Finally, the reduced form evidence also suggests that 4.5% of all Medicare Advantage enrollees would have preferred to be enrolled in traditional Medicare with a Medigap supplement if the Medigap supplement were priced as it would be in the absence of Medicare Advantage.

These results are, of course, sensitive to the method of deriving the Medigap premium in the absence of Medicare Advantage. The primary benefit of the structural model in this context is that it will provide a more credible estimate of this counterfactual premium because it will be based on predicted optimal behavior in the health plan choice and health care utilization decisions of Medicare beneficiaries. The result will offer some evidence on the welfare impact of the current role of private health plans in Medicare as policy makers continue to debate the future structure of the Medicare program.
In this appendix, I describe the issues considered and some results of early experiments that led to my choice of \( g(\theta) \). For ease of exposition, I examine a simple one dimensional, deterministic case. The issue is a potential conflict between incorporating the notion of diminishing marginal product on the extensive margin and maintaining the intuitive comparative static that indirect utilities should fall with the pretreatment state \( \theta \). Consider posttreatment health in the simple model described in section 5.1.1,

\[
H = \lambda \left( \bar{m}(\theta)m - \frac{1}{2}m^2 \right) - \theta
\]

where

\[
\bar{m}(\theta) = \exp \{ \chi_2 + \zeta g(\theta) \}.
\]

At an optimum, we have

\[
\frac{\partial H^*}{\partial \theta} = \lambda \exp \{ \chi_2 + \zeta g(\theta) \} \zeta g'(\theta)m^* - 1
\]

and

\[
\frac{\partial U^*}{\partial \theta} = \frac{\partial U}{\partial H} \frac{\partial H}{\partial \theta} \bigg|_{m^*} = \frac{\partial U^*}{\partial H} \left[ \lambda \exp \{ \chi_2 + \zeta g(\theta) \} \zeta g'(\theta)m^* - 1 \right].
\]

So \( \frac{\partial U^*}{\partial \theta} < 0 \) if and only if

\[
\lambda \exp \{ \chi_2 + \zeta g(\theta) \} \zeta g'(\theta)m^* < 1.
\]

The intuition is that the consumer should not be healthier after treatment resulting from an increase in \( \theta \) than if the consumer hadn’t experienced the increase in \( \theta \) to begin with. This can be guaranteed by removing the property of diminishing marginal product on the extensive margin, i.e., \( \zeta g'(\theta) = 0 \). If this property is maintain, it is possible that \( \frac{\partial U^*}{\partial \theta} > 0 \) would be true at some data points for some values of the model parameters. My object is to choose \( g'(\theta) \) so as to minimize this possibility.

My first choice was \( g(\theta) = \theta \) with a small starting guess of \( \zeta = 0.001 \). I tested this by plotting the path of \( \theta \) that makes an observed value of \( m \) utility maximizing and the
corresponding indirect utility. I use the values $\gamma_1 = 0.4, \gamma_2 = 0.6, \delta = 2, \lambda = 1, \chi_2 = 4,$ and $y = 24$. I input medical expenses $m$ from 1 to 100 assuming the consumer pays a constant coinsurance rate of 0.2. For each value of $m$, I solve for the $\theta$ that satisfies the consumer’s first order condition, and the resulting indirect utility, and the value of $\frac{\partial H^*}{\partial \theta}$. Figure 5 plots the path of $m^*$ and $U^*$ on $\theta$. The path of $m^*$ has the intuitive upward slope. The path of $U^*$ is initially downward sloping but later rises. Figure 6 illustrates the problem. The marginal effect of pretreatment health on posttreatment health is not only positive, it eventually explodes. So even though the consumer spend more of his income on health care, the care is becoming so productive that the consumer actually is better off being very sick and requiring a large amount of health care.

Figures 7 and 8 give the same results for $g(\theta) = \left(1 + \exp\left(\frac{\theta - a}{b}\right)\right)^{-1}$ and $a = 0, b = 10,000,$ and $\zeta = 1$. As before, the path of $m^*$ has the intuitive upward slope. In this case, the path of $U^*$ also is intuitive in that it slopes down and is monotone. This is because the path of $\frac{\partial H^*}{\partial \theta}$ is always negative and goes to -1 as $\theta$ becomes very large or small. To my knowledge, it cannot be guaranteed that $\frac{\partial H^*}{\partial \theta}$ will always be negative for all data points and for all values of the parameters of the model. I test for this in the search algorithm by evaluating $\frac{\partial U^*}{\partial \theta}$ (in the model, a $5 \times 1$ vector) at 500 randomly selected observations at each iteration of the likelihood function evaluation. To date, no element of $\frac{\partial U^*}{\partial \theta}$ has been positive under this specification.

9.2 Out-of-Pocket Expenses

In this appendix, I discuss the reduced form models used to estimate the relationship between health care expenditures and out-of-pocket expenses and marginal prices for Medicare Advantage enrollees. This is required because I use deflated health care expenditures as the choice variable for the consumer and health insurance cost sharing is usually defined as a fixed copayment per event. Therefore, some relationship must established between health care expenditures and the variable relevant for consumer cost sharing in order to define an appropriate total out-of-pocket and marginal price for the consumer choice problem. This is unnecessary for Medigap and FFS-only enrollees because, for all types of care other than
inpatient hospital care, either consumer cost sharing is defined in terms of coinsurance rates or the consumer faces the full marginal price.

For doctor visits and outpatient events, I use ordinary least squares as opposed to a count data model since there are a large number of trips and the distribution is relatively continuous. In addition, Poisson and negative binomial regressions resulted in implausibly large predicted values for the number of events in some cases. For example, the maximum number of doctor visit events in the MCBS data is 444. The maximum predicted value is 314 using least squares but 56,820 using a negative binomial model and 750 using a Poisson regression. Since there is high degree of correlation between doctor visit events and outpatient expenditures and vice versa, I use both types of expenditures to predict both types of event counts. I use the all of the MCBS data and county level data from the Area Resource File. The results are given in Table 11.

For prescription drug events, there is further complication in that copays are different for branded and generic drugs and much of the total expenditure is on off-formulary drugs. Classifying each event as either generic, branded, or off-formulary is impractical because the data required is not available. I make the simplifying assumption that the consumer’s share of the total expenditure is a function of consumer and area characteristics and the particular cost sharing terms of the health plan. These cost sharing terms include an indicator of branded drug coverage, and indicator for the type of cost sharing for generic and branded drugs (coinsurance or copay) and the amount of the cost sharing given the type. I also include an indicator of the presence of an annual maximum benefit and the level of the maximum if one exists. Defining $y$ as the consumer’s share of total expenditures, I regress $\ln y - \ln(1 - y)$ on the above covariates. The results are given in Table 12. Note that this assumption implies that, in the model, the beneficiary faces a constant marginal price for prescription drugs given health plan choice.

Dental care events are complicated by the fact that, in all cases, Medicare MCOs only cover preventive dental events and charge a copay per event. Fortunately, the data allows me to separate preventive from nonpreventive events. The MCBS records if specific services were performed at each event. I classify an event as nonpreventive if any of the following services
were performed: extractions, fillings, root canals, crowns, bridge work, orthodontics. I then use total expenditures and total preventive expenditures to predict the number of preventive events using a Poisson regression. I use Poisson rather than negative binomial because a test that the dispersion parameter is zero is not rejected at the 0.50 level. The results are given in Table 13. I assume the beneficiary pays the full cost of the nonpreventive events plus the predicted number of preventive events times the preventive event copay.

Out-of-pocket expenses for inpatient events are significantly massed at zero. 90% of all Medicare Advantage enrollees who consumed some inpatient care paid nothing. This percentage is not significantly different if the MCO charges an admission fee (89.5%) or not (90.3%). 89.9% of Medicare Advantage enrollees were in a plan that did not charge an inpatient admission fee. 94.6% of all Medigap enrollees and 66.0% of all FFS-only beneficiaries who consumed some inpatient care paid nothing. Of the FFS-only beneficiaries who consumed some inpatient care, 19.4% paid the inpatient event deductible of $776. Count model regressions of inpatient expenditures on inpatient events yielded nothing significant. Because of this, I make the simplifying assumption that the inpatient admission fee, FFS-only or Medicare Advantage, is an annual fixed cost. If a beneficiary consumed any inpatient care, I assume he/she paid the admission fee, but the marginal price is zero.

### 9.3 Simulating the Likelihood Contribution

In this appendix, I discuss simulating the individual contributions to the likelihood function. Recall from section 5.1.6 that in the case of a mix of corner and interior solutions, the region support of the unobservables that is consistent with the observed utilization data is a non-Euclidean subspace of $\mathbb{R}^5$. I approximate the integral in (6) using simulation methods in a two-step process. First, conditional on the observed data for beneficiary $i$, and a guess of the model parameters (including $\chi_j$), I numerically solve the full system $U_{mijk}^* = 0$ for $\tilde{\epsilon}_i$, that is, I find the realization of $\epsilon_i$ that makes the observed quality choices, including the zeros, utility maximizing. This provides values for the limits of integration.
\((\epsilon^R_{iIPjk}(\Theta), \epsilon^R_{iOPjk}(\Theta), \epsilon^R_{iDCjk}(\Theta))\). The simulator of the first order conditions (5) is

\[
\tilde{U}_{mitj}(m_{ijk}, \epsilon_i, j) = -\frac{p_{itj}}{y_{i} - \pi_{jk} - \pi_B - oop_i} + \kappa_{it} + 2\kappa_2 m_{itjk} + \delta \frac{1}{R_v} \sum_{r=1}^{R_v} \frac{\sum_{s \in T} \alpha_s h_{isjk}(\nu_{sr})^{\rho - 1} \partial h_{isjk}(\nu_{sr})}{H_i(\nu_r)^{\gamma_2 + \rho - 1}},
\]

where \(\frac{\partial h_{isjk}(\nu_{sr})}{\partial m_{isjk}} = \begin{cases} \exp\left\{\chi_{1tj}\right\} \left(\bar{m}_{itj}(\nu_{tr}) - m_{itjk}\right), & \text{if } s = t \\ \exp\left\{\chi_{1sj}\right\} \Psi_{ts}\bar{m}_{isj}(\nu_{sr}), & \text{else} \end{cases}, \forall t \in T. \tag{19}\)

I then simulate (6) by randomly drawing from the marginal density \(F_{\epsilon_C}\) using the GHK\(^{22}\) algorithm and then numerically solving the 2 × 1 system \(\tilde{U}^*_m{|}_{it \in T; m^{*}_{itjk} > 0} = 0\) for the combination \((\epsilon_{idV}, \epsilon_{iPM})\) that makes \(m^{*}_{itjk}\) utility maximizing for that draw. This numerical search is done for each replication of \((\epsilon_{iIP}, \epsilon_{iOP}, \epsilon_{iDV})\). Finally, I evaluate the matrix \(\mathbb{J}_{ijk}\) using the simulated values of \((\epsilon_{iIP}, \epsilon_{iOP}, \epsilon_{iDV})\) and the pair \((\epsilon_{idV}, \epsilon_{iPM})\) that solve the simulated first order conditions. In this way, I trace out the manifold in \(\mathbb{R}^5\) that is defined by the intersection of the implicit functions given by \((\epsilon_{idV}(\Theta), \epsilon_{iPM}(\Theta))\).

To set some notation, let \(\mu^C\) denote the vector of elements of \(\mu\) that correspond to the types of health care for which a corner solution is observed and define \(\Omega^C\) analogously to \(\Omega\). So the marginal distribution of \(\epsilon_{iC}\), denoted in section 5.1.6 as \(F_{\epsilon_C}\), is \(N(\mu^C, \Omega^C)\). I order the elements of \(\epsilon_{iC}\) such that the diagonal elements of \(\Omega^C\) are increasing.\(^{23}\) The simulator of (6) is then

\[
\frac{1}{R_v} \sum_{r=1}^{R_v} f_{\epsilon_{idV}(\Theta, \epsilon_{iC}), \epsilon_{iPM}(\Theta, \epsilon_{iC})} \|\mathbb{J}_{ijk}\| \prod_{k=1}^3 Q_{kr}
\]

where \(Q_{kr} = \Phi\left(\frac{\epsilon^R_{\epsilon_C}(\Theta) - \mathbf{m}_k}{\mathbf{v}_k}\right)\)

and \(\mathbf{m}_k = \mu^C_k + [\Omega^C_{k,1:k-1}][\Omega^C_{1:k-1,1:k-1}]^{-1}(\epsilon_{1:k-1,r} - \mu^C_{1:k-1})\),

\[
\mathbf{v}_k = \sqrt{\Omega^C_{k,k} - [\Omega^C_{k,1:k-1}][\Omega^C_{1:k-1,1:k-1}]^{-1}[\Omega^C_{1:k-1,1:k-1}]}. \tag{20}
\]

In this notation, \(\Omega^C_{k,1:k-1}\) denotes the elements of the \(k^{th}\) row, columns 1 through \(k - 1\) of \(\Omega^C\).

\(^{22}\)See Stern(1997) for a description of the GHK algorithm and related simulation methods.

\(^{23}\)This is important because simulations using the GHK algorithm are not invariant to the ordering of the draws. See Stern (1997) for a discussion.
The simulator of $P_{ijk}(\eta_i)$ is

$$
\bar{P}_{ijk}(\eta_i) = \frac{\exp\left\{ \frac{\tilde{V}_{ijk}(\eta_i)}{(1-\sigma_a)(1-\sigma_w)} \right\}}{\mathcal{J}_j} = \left[ \exp\left\{ \tilde{V}_{ijFS}(\eta_i) \right\} + \exp\left\{ \tilde{V}_{iM}(\eta_i) \right\} + \mathcal{J}^{1-\sigma_a} \right]^{\mathcal{J}^{1-\sigma_a}}
$$

where

$$
\mathcal{J}_j = \sum_{k=1}^{K_j} \exp\left\{ \frac{\tilde{V}_{ijk}(\eta_i)}{(1-\sigma_w)(1-\sigma_a)} \right\}, \quad \mathcal{J} = \sum_{j=1}^{J} \mathcal{J}_j^{1-\sigma_w},
$$

and

$$
\tilde{V}_{ijk}(\eta_i) = \frac{1}{R_e} \sum_{r=1}^{R_e} U \left( m_{ijk}^*, X_i \beta + \eta_i + \epsilon_{ir}, y_i \right).
$$

Evaluating (22) requires a numerical search for $m_{ijk}^*$ for each replication of $\epsilon_{ir}$.

The simulator for $L_i(\Theta)$ for the case described in 5.1.6 is

$$
\tilde{L}_i(\Theta) = \frac{\sum_{s=1}^{R_q} P_{ijk}(\eta_{is}) \left[ \frac{1}{R_q} \sum_{r=1}^{R_e} f_{\epsilon} (\epsilon_iDVR(\Theta, \epsilon_{icr}), \epsilon_iPMr(\Theta, \epsilon_{icr}), \eta_{is}) \| \mathcal{J}_{ijk} \| \prod_{k=1}^{3} Q_{kr} \right]}{\sum_{s=1}^{R_q} P_{ijk}(\eta_{is})}.
$$

In the case of five interior solutions, I solve the system $\tilde{U}_{m_{ijk}} = 0$ for the vector $\epsilon_i$ that is consistent with the data. The simulated likelihood contribution is

$$
\tilde{L}_i(\Theta) = \frac{\sum_{s=1}^{R_q} P_{ijk}(\eta_{is}) \left[ f_{\epsilon} (\epsilon_i(\Theta)) \| \mathcal{J}_{ijk} \| \right]}{\sum_{s=1}^{R_q} P_{ijk}(\eta_{is})}
$$

where $\mathcal{J}_{ijk}$ is a $5 \times 5$ matrix. In the case of five corner solutions, the simulated likelihood contribution is

$$
\tilde{L}_i(\Theta) = \frac{\sum_{s=1}^{R_q} P_{ijk}(\eta_{is}) \left[ \frac{1}{R_q} \sum_{r=1}^{R_e} \prod_{k=1}^{5} Q_{kr} \right]}{\sum_{s=1}^{R_q} P_{ijk}(\eta_{is})}.
$$
References


### Table 1: Health Care Expenses and Budget Share
**MCBS Sample (N=5251)**

<table>
<thead>
<tr>
<th></th>
<th>Medigap Mean (Stan Dev)</th>
<th>Medigap Share</th>
<th>Medicare Advantage Mean (Stan Dev)</th>
<th>Medicare Advantage Share</th>
<th>FFS Only Mean (Stan Dev)</th>
<th>FFS Only Share</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inpatient Care</strong></td>
<td>3086.21 (9372.51)</td>
<td>39.7%</td>
<td>2164.41 (7901.79)</td>
<td>44.5%</td>
<td>2778.17 (10260.50)</td>
<td>50.1%</td>
</tr>
<tr>
<td><strong>Outpatient Care</strong></td>
<td>790.29 (2023.57)</td>
<td>10.2%</td>
<td>581.90 (4552.42)</td>
<td>12.0%</td>
<td>529.04 (4552.42)</td>
<td>9.6%</td>
</tr>
<tr>
<td><strong>Doctor Visits</strong></td>
<td>2696.72 (3844.04)</td>
<td>34.7%</td>
<td>952.42 (1778.12)</td>
<td>19.6%</td>
<td>1551.57 (2729.00)</td>
<td>28.3%</td>
</tr>
<tr>
<td><strong>Prescription Drugs</strong></td>
<td>944.96 (1090.95)</td>
<td>12.2%</td>
<td>867.50 (1490.41)</td>
<td>17.8%</td>
<td>527.53 (864.59)</td>
<td>9.6%</td>
</tr>
<tr>
<td><strong>Dental Care</strong></td>
<td>258.82 (1123.94)</td>
<td>3.3%</td>
<td>293.95 (814.16)</td>
<td>6.0%</td>
<td>99.02 (737.68)</td>
<td>1.8%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>7777.00 (12648.88)</td>
<td>100%</td>
<td>4860.17 (10473.85)</td>
<td>100%</td>
<td>5485.319 (12587.74)</td>
<td>100%</td>
</tr>
</tbody>
</table>

### Table 2: Age, Gender, and Income by Insurance Type
**MCBS, Age Entitlement Only**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Share of Sample</th>
<th>Age Mean (SD)</th>
<th>Female Mean (SD)</th>
<th>Income Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medigap</td>
<td>3103</td>
<td>30.4%</td>
<td>78.72 (7.58)</td>
<td>.61 (.49)</td>
<td>28,087.65 (49,066.55)</td>
</tr>
<tr>
<td>Medicare Advantage</td>
<td>1439</td>
<td>14.1%</td>
<td>76.49 (7.07)</td>
<td>.59 (.49)</td>
<td>24,650.54 (24,009.07)</td>
</tr>
<tr>
<td>FFS Only</td>
<td>709</td>
<td>7.0%</td>
<td>77.91 (8.40)</td>
<td>.50 (.50)</td>
<td>17,852.31 (15,543.33)</td>
</tr>
<tr>
<td>Employer Provided</td>
<td>3519</td>
<td>34.5%</td>
<td>76.37 (6.79)</td>
<td>.54 (.50)</td>
<td>34,999.07 (40,964.77)</td>
</tr>
<tr>
<td>Medicaid or QMB</td>
<td>1430</td>
<td>14.0%</td>
<td>79.80 (8.78)</td>
<td>.73 (.44)</td>
<td>11,811.97 (41,201.77)</td>
</tr>
<tr>
<td>Full MCBS</td>
<td>10,361</td>
<td>100%</td>
<td>77.70 (7.61)</td>
<td>.59 (.49)</td>
<td>26,964.91 (41,125.48)</td>
</tr>
</tbody>
</table>
Table 3: Mean Total Health Care Expenditures by Age/Gender
MCBS Sample (N=5251)
Mean/(Stan Dev)/ N

<table>
<thead>
<tr>
<th></th>
<th>65-69</th>
<th>70-74</th>
<th>75-79</th>
<th>80-84</th>
<th>85 and over</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5634.49</td>
<td>6248.629</td>
<td>6855.212</td>
<td>6386.13</td>
<td>6337.73</td>
</tr>
<tr>
<td></td>
<td>(10547.75)</td>
<td>(12,241.93)</td>
<td>(11,381.55)</td>
<td>(8,913.13)</td>
<td>(10,096.61)</td>
</tr>
<tr>
<td></td>
<td>416</td>
<td>649</td>
<td>589</td>
<td>656</td>
<td>800</td>
</tr>
<tr>
<td>Male</td>
<td>6210.49</td>
<td>5939.152</td>
<td>8465.763</td>
<td>7475.57</td>
<td>8103.799</td>
</tr>
<tr>
<td></td>
<td>(16,665.41)</td>
<td>(11,345.19)</td>
<td>(15,738.76)</td>
<td>(11,430.83)</td>
<td>(14,884.14)</td>
</tr>
<tr>
<td></td>
<td>385</td>
<td>567</td>
<td>463</td>
<td>387</td>
<td>339</td>
</tr>
</tbody>
</table>

Table 4: Descriptive Statistics of Medigap and Medicare Advantage Plans

Medigap Basic Benefit:
   i) Part A Coinsurance. Coverage is extended for an additional 365 days.
   ii) Part B Coinsurance.
   iii) First three pints of blood.

Additional Benefits

<table>
<thead>
<tr>
<th>BENEFIT</th>
<th>PLAN A</th>
<th>PLAN B</th>
<th>PLAN C</th>
<th>PLAN F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skilled Nursing Coinsurance</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part A Deductible</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Part B Deductible</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part B Excess</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foreign Travel Emergency</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Plan F Premiums (weight=Medicare eligibles):
National Weighted Average: $131.24
Weighted Average in Medicare Advantage Counties: $136.27

Medicare Advantage:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Premium, National Weighted Average</td>
<td>$37.17</td>
</tr>
<tr>
<td>Percentage of plans offered at a zero premium</td>
<td>32.5%</td>
</tr>
<tr>
<td>Percentage of plans that offer a drug benefit</td>
<td>73.8%</td>
</tr>
<tr>
<td>Average generic copay</td>
<td>$7.45</td>
</tr>
<tr>
<td>Branded drug benefit given a drug benefit</td>
<td>82.9%</td>
</tr>
<tr>
<td>Average branded copay</td>
<td>$17.70</td>
</tr>
<tr>
<td>Percentage of plans that offer a dental benefit</td>
<td>20.9%</td>
</tr>
<tr>
<td>Average preventive dental copay</td>
<td>$7.33</td>
</tr>
<tr>
<td>Average inpatient admission fee</td>
<td>$28.83</td>
</tr>
<tr>
<td>Percentage of plans with no inpatient admission fee</td>
<td>86.6%</td>
</tr>
<tr>
<td>Average outpatient admission fee</td>
<td>$26.86</td>
</tr>
<tr>
<td>Percentage of plans with no outpatient admission fee</td>
<td>87.44%</td>
</tr>
<tr>
<td>Average physician copay</td>
<td>$7.89</td>
</tr>
<tr>
<td>Percentage of plans with no physician copay</td>
<td>7.51%</td>
</tr>
</tbody>
</table>
I randomly selected eight observations from the MCBS data (two each having an order of integration one through four) and evaluated the simulator (12) using 10, 20, and 30 draws. The values of the parameters of the distribution of $\varepsilon$ are given in the first table. In the second table, the first five columns give the limits of integration for the types of health care for which a corner solution is observed. The sixth column gives the value of the integral taken to be the truth (based on 20,000 evaluations). Columns seven through twelve give the mean squared error and root mean squared error divided by the simulated true value for the given number of draws. The approximation errors are based on 500 trials. All trials use antithetic acceleration.

**Table 5: Monte Carlo Results on Approximation Error Associated with Simulating the Integral (6).**

<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$E(\varepsilon)$</th>
<th>VAR(\varepsilon)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC 10,611.051</td>
<td>11,723,023.246</td>
<td></td>
</tr>
<tr>
<td>PM 4,869.303</td>
<td>5,214,651.680</td>
<td>16,058,068.019</td>
</tr>
<tr>
<td>DV -51,107.120</td>
<td>-4,513,472.936</td>
<td>1,228,286.571</td>
</tr>
<tr>
<td>OP -118,454.835</td>
<td>-3,398,527.615</td>
<td>9,726,062.339</td>
</tr>
<tr>
<td>IP -479,910.973</td>
<td>-600,217.984</td>
<td>8,536,139.766</td>
</tr>
</tbody>
</table>

**LIMITS OF INTEGRATION**

<table>
<thead>
<tr>
<th>LIMITS OF INTEGRATION</th>
<th>R=10,000</th>
<th>R = 10</th>
<th>R = 20</th>
<th>R = 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC 12417 10734 13206 7622 12603 8067</td>
<td>7.450E-09 5.859E-22 0.0032</td>
<td>2.899E-22 0.0023</td>
<td>1.770E-22 0.0018</td>
<td>1.411E-23 0.0059</td>
</tr>
<tr>
<td>PM -259405 -464141 -120286 6235</td>
<td>6.371E-10 3.657E-23 0.0095</td>
<td>1.989E-23 0.0070</td>
<td>1.411E-23 0.0059</td>
<td>1.612E-21 0.0069</td>
</tr>
<tr>
<td>DV -259405 -464141 -120286 6235</td>
<td>4.627E-06 7.494E-15 0.0187</td>
<td>3.731E-15 0.0132</td>
<td>2.877E-15 0.0116</td>
<td>2.877E-15 0.0116</td>
</tr>
<tr>
<td>OP -259405 -464141 -120286 6235</td>
<td>5.844E-09 4.386E-21 0.0133</td>
<td>2.685E-21 0.0089</td>
<td>1.612E-21 0.0069</td>
<td>1.612E-21 0.0069</td>
</tr>
<tr>
<td>IP 259405 464141 120286 6235</td>
<td>2.197E-04 1.575E-11 0.0057</td>
<td>8.036E-13 0.0041</td>
<td>5.495E-13 0.0034</td>
<td>5.495E-13 0.0034</td>
</tr>
<tr>
<td>-- 74034 224429</td>
<td>3.572E-06 8.459E-15 0.0257</td>
<td>4.187E-15 0.0181</td>
<td>2.880E-15 0.0152</td>
<td>2.880E-15 0.0152</td>
</tr>
<tr>
<td>-- 121736 459866</td>
<td>6.991E-03 2.057E-08 0.0205</td>
<td>9.969E-09 0.0143</td>
<td>6.625E-09 0.0116</td>
<td>6.625E-09 0.0116</td>
</tr>
<tr>
<td>-- 121386 470643</td>
<td>8.981E-03 3.350E-08 0.0204</td>
<td>1.604E-08 0.0141</td>
<td>1.175E-08 0.0121</td>
<td>1.175E-08 0.0121</td>
</tr>
</tbody>
</table>
Table 6: Nonlinear IV Results for Equation (20)

Dependent Variable: ln(MCO / Market Share/Medigap Market Share)
N = 19,404
F(30, 19373) = 2,077.43
Pr[F > 2077.43] = 0.000

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Stan Error</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inpatient Admission Fee</td>
<td>-0.0217</td>
<td>0.0079</td>
<td>0.006</td>
</tr>
<tr>
<td>Outpatient Admission Fee</td>
<td>0.0250</td>
<td>0.0079</td>
<td>0.001</td>
</tr>
<tr>
<td>Physician Copay</td>
<td>0.0362</td>
<td>0.0235</td>
<td>0.124</td>
</tr>
<tr>
<td>Drug Benefit</td>
<td>1.2376</td>
<td>0.2885</td>
<td>0.000</td>
</tr>
<tr>
<td>Generic Copay</td>
<td>-0.0969</td>
<td>0.0325</td>
<td>0.003</td>
</tr>
<tr>
<td>Dental Benefit</td>
<td>1.2170</td>
<td>0.2839</td>
<td>0.000</td>
</tr>
<tr>
<td>Dental Copay</td>
<td>-0.0249</td>
<td>0.0183</td>
<td>0.173</td>
</tr>
<tr>
<td>Age*Inpatient Adm Fee</td>
<td>0.0002</td>
<td>0.0001</td>
<td>0.080</td>
</tr>
<tr>
<td>Age*Outpatient Adm Fee</td>
<td>-0.0002</td>
<td>0.0001</td>
<td>0.038</td>
</tr>
<tr>
<td>Age*Physician Copay</td>
<td>-0.0012</td>
<td>0.0003</td>
<td>0.000</td>
</tr>
<tr>
<td>Age*Drug Benefit</td>
<td>-0.0106</td>
<td>0.0031</td>
<td>0.001</td>
</tr>
<tr>
<td>Age*Generic Copay</td>
<td>0.0005</td>
<td>0.0004</td>
<td>0.157</td>
</tr>
<tr>
<td>Age*Dental Benefit</td>
<td>-0.0119</td>
<td>0.0031</td>
<td>0.000</td>
</tr>
<tr>
<td>Age*Dental Copay</td>
<td>-0.0003</td>
<td>0.0002</td>
<td>0.197</td>
</tr>
<tr>
<td>Female*Inpatient Adm Fee</td>
<td>-0.0004</td>
<td>0.0010</td>
<td>0.706</td>
</tr>
<tr>
<td>Female*Outpatient Adm Fee</td>
<td>0.0004</td>
<td>0.0010</td>
<td>0.713</td>
</tr>
<tr>
<td>Female*Physician Copay</td>
<td>-0.0026</td>
<td>0.0028</td>
<td>0.362</td>
</tr>
<tr>
<td>Female*Drug Benefit</td>
<td>-0.0176</td>
<td>0.0340</td>
<td>0.605</td>
</tr>
<tr>
<td>Female*Generic Copay</td>
<td>0.0043</td>
<td>0.0040</td>
<td>0.285</td>
</tr>
<tr>
<td>Female*Dental Benefit</td>
<td>-0.0186</td>
<td>0.0338</td>
<td>0.582</td>
</tr>
<tr>
<td>Female*Dental Copay</td>
<td>0.0006</td>
<td>0.0022</td>
<td>0.787</td>
</tr>
<tr>
<td>Income*Inpatient Adm Fee</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.084</td>
</tr>
<tr>
<td>Income*Outpatient Adm Fee</td>
<td>-0.0002</td>
<td>0.0001</td>
<td>0.027</td>
</tr>
<tr>
<td>Income*Physician Copay</td>
<td>0.0007</td>
<td>0.0002</td>
<td>0.001</td>
</tr>
<tr>
<td>Income*Drug Benefit</td>
<td>0.0071</td>
<td>0.0028</td>
<td>0.010</td>
</tr>
<tr>
<td>Income*Generic Copay</td>
<td>0.0005</td>
<td>0.0003</td>
<td>0.105</td>
</tr>
<tr>
<td>Income*Dental Benefit</td>
<td>-0.0103</td>
<td>0.0027</td>
<td>0.000</td>
</tr>
<tr>
<td>Income*Dental Copay</td>
<td>-0.0003</td>
<td>0.0002</td>
<td>0.039</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-0.6073</td>
<td>0.2231</td>
<td>0.007</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.6868</td>
<td>0.0037</td>
<td>0.000</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.2497</td>
<td>0.0325</td>
<td>0.000</td>
</tr>
</tbody>
</table>
### Table 7: IV Results for Equation (22)

Dependent Variable: ln(Medigap Premium)
N = 910
F( 29,  880) =  8.24
Pr[F > 8.24] = 0.000

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Stan Error</th>
<th>p  value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicare Advantage Market Share</td>
<td>3.60064</td>
<td>0.95417</td>
<td>0.000</td>
</tr>
<tr>
<td>Medicare Advantage Market Share²</td>
<td>-4.07250</td>
<td>2.58338</td>
<td>0.115</td>
</tr>
<tr>
<td>Medicare Advantage Risk Index/1000</td>
<td>-0.00301</td>
<td>0.00115</td>
<td>0.009</td>
</tr>
<tr>
<td>County CMS Wage Index</td>
<td>-0.01849</td>
<td>0.08924</td>
<td>0.836</td>
</tr>
<tr>
<td>Urban indicator</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.010</td>
</tr>
<tr>
<td>Population Density</td>
<td>0.00001</td>
<td>0.00001</td>
<td>0.258</td>
</tr>
<tr>
<td>Per Capita Income</td>
<td>-6.55E-6</td>
<td>1.04E-5</td>
<td>0.528</td>
</tr>
<tr>
<td>Per Capita Income²</td>
<td>2.21E-10</td>
<td>2.05E-10</td>
<td>0.283</td>
</tr>
<tr>
<td>Medicare Risk Index</td>
<td>0.90712</td>
<td>0.24683</td>
<td>0.000</td>
</tr>
<tr>
<td>Medical Doctors</td>
<td>-0.00001</td>
<td>0.00002</td>
<td>0.744</td>
</tr>
<tr>
<td>Surgeons</td>
<td>-0.00199</td>
<td>0.00546</td>
<td>0.715</td>
</tr>
<tr>
<td>Hospitals</td>
<td>0.00164</td>
<td>0.00412</td>
<td>0.692</td>
</tr>
<tr>
<td>Operating Rooms</td>
<td>-0.00228</td>
<td>0.00072</td>
<td>0.002</td>
</tr>
<tr>
<td>Register Nurse FTE</td>
<td>0.00006</td>
<td>0.00003</td>
<td>0.021</td>
</tr>
<tr>
<td>Licensed Practical Nurse FTE</td>
<td>0.00012</td>
<td>0.00010</td>
<td>0.243</td>
</tr>
<tr>
<td>Skilled Nursing Facilities</td>
<td>0.00108</td>
<td>0.00183</td>
<td>0.554</td>
</tr>
<tr>
<td>Commercial HMOs</td>
<td>-0.00345</td>
<td>0.00292</td>
<td>0.239</td>
</tr>
<tr>
<td>Commercial HMO Market Share</td>
<td>-0.62347</td>
<td>0.13797</td>
<td>0.000</td>
</tr>
<tr>
<td>Female 65-69 eligibles</td>
<td>0.00010</td>
<td>0.00009</td>
<td>0.223</td>
</tr>
<tr>
<td>Female 70-74 eligibles</td>
<td>-0.00010</td>
<td>0.00012</td>
<td>0.384</td>
</tr>
<tr>
<td>Female 75-79 eligibles</td>
<td>-0.00002</td>
<td>0.00012</td>
<td>0.872</td>
</tr>
<tr>
<td>Female 80-84 eligibles</td>
<td>0.00004</td>
<td>0.00008</td>
<td>0.611</td>
</tr>
<tr>
<td>Female 85 and over eligibles</td>
<td>0.00004</td>
<td>0.00007</td>
<td>0.593</td>
</tr>
<tr>
<td>Male 65-69 eligibles</td>
<td>-0.00023</td>
<td>0.00007</td>
<td>0.001</td>
</tr>
<tr>
<td>Male 70-74 eligibles</td>
<td>0.00025</td>
<td>0.00008</td>
<td>0.003</td>
</tr>
<tr>
<td>Male 75-79 eligibles</td>
<td>-0.00013</td>
<td>0.00009</td>
<td>0.139</td>
</tr>
<tr>
<td>Male 80-84 eligibles</td>
<td>0.00007</td>
<td>0.00010</td>
<td>0.488</td>
</tr>
<tr>
<td>Male 85 and over eligibles</td>
<td>-0.00010</td>
<td>0.00012</td>
<td>0.410</td>
</tr>
<tr>
<td>Constant</td>
<td>5.84240</td>
<td>1.36659</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Table 8: Estimated change in Consumer Surplus due to Medicare Advantage
Mean (Standard Deviation) Median

<table>
<thead>
<tr>
<th>AGE GROUP</th>
<th>MEN</th>
<th>WOMEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>65-69</td>
<td>576.22 (388.50)</td>
<td>516.39 (437.86)</td>
</tr>
<tr>
<td>70-74</td>
<td>636.23 (456.80)</td>
<td>562.88 (464.62)</td>
</tr>
<tr>
<td>75-79</td>
<td>564.70 (418.29)</td>
<td>501.83 (442.92)</td>
</tr>
<tr>
<td>80-84</td>
<td>529.24 (407.53)</td>
<td>452.93 (398.34)</td>
</tr>
<tr>
<td>85 and over</td>
<td>447.65 (374.80)</td>
<td>337.73 (367.04)</td>
</tr>
</tbody>
</table>

Table 9: Percentage of Medicare Population for whom the estimated change in Consumer Surplus is negative

<table>
<thead>
<tr>
<th>AGE GROUP</th>
<th>MEN</th>
<th>WOMEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>65-69</td>
<td>0.49%</td>
<td>0.64%</td>
</tr>
<tr>
<td>70-74</td>
<td>0.75%</td>
<td>1.00%</td>
</tr>
<tr>
<td>75-79</td>
<td>1.19%</td>
<td>1.14%</td>
</tr>
<tr>
<td>80-84</td>
<td>1.58%</td>
<td>1.87%</td>
</tr>
<tr>
<td>85 and over</td>
<td>2.61%</td>
<td>3.43%</td>
</tr>
</tbody>
</table>

Table 10: Estimated percentage of Medicare Advantage enrollees that would switch to Medigap if Medigap were priced as it would be in the absence of Medicare Advantage

<table>
<thead>
<tr>
<th>AGE GROUP</th>
<th>MEN</th>
<th>WOMEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>65-69</td>
<td>4.16%</td>
<td>4.07%</td>
</tr>
<tr>
<td>70-74</td>
<td>3.89%</td>
<td>3.95%</td>
</tr>
<tr>
<td>75-79</td>
<td>4.07%</td>
<td>4.22%</td>
</tr>
<tr>
<td>80-84</td>
<td>4.26%</td>
<td>4.48%</td>
</tr>
<tr>
<td>85 and over</td>
<td>4.73%</td>
<td>5.25%</td>
</tr>
</tbody>
</table>
Table 11: Doctor Visit and Outpatient Events

Least Squares Regressions

<table>
<thead>
<tr>
<th>Dependent Variable: DV Events</th>
<th>Dependent Variable: OP Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>given DV events &gt; 0</td>
<td>given OP events &gt; 0</td>
</tr>
<tr>
<td>Mean of Dependent Variable: 26.7</td>
<td>Mean of Dependent Variable: 6.3</td>
</tr>
<tr>
<td>N = 11,795</td>
<td>N = 8,937</td>
</tr>
<tr>
<td>F(28,11766) = 869.78</td>
<td>F(28,8908) = 242.26</td>
</tr>
<tr>
<td>Pr[f&gt;869.78] = 0.000</td>
<td>Pr[f&gt;242.26] = 0.000</td>
</tr>
<tr>
<td>$R^2 = .6743$</td>
<td>$R^2 = .4323$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DOCTOR VISIT EVENTS</th>
<th>OUTPATIENT EVENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>VARIABLE</td>
<td>COEFF</td>
</tr>
<tr>
<td>DV Expenses</td>
<td>0.00670</td>
</tr>
<tr>
<td>DV Expenses^2</td>
<td>-5.97E-08</td>
</tr>
<tr>
<td>MDs*DV Expenses</td>
<td>1.42E-07</td>
</tr>
<tr>
<td>OP Expenses</td>
<td>0.00146</td>
</tr>
<tr>
<td>OP Expenses^2</td>
<td>-1.03E-08</td>
</tr>
<tr>
<td>MDs*OP Expenses</td>
<td>-1.55E-08</td>
</tr>
<tr>
<td>CMS Wage Index</td>
<td>16.35628</td>
</tr>
<tr>
<td>Total Population</td>
<td>1.83E-07</td>
</tr>
<tr>
<td>Urban Indicator</td>
<td>0.45410</td>
</tr>
<tr>
<td>MDs</td>
<td>-0.00108</td>
</tr>
<tr>
<td>Surgeons</td>
<td>0.07058</td>
</tr>
<tr>
<td>Hospitals</td>
<td>-0.01057</td>
</tr>
<tr>
<td>Hospital Beds</td>
<td>0.00101</td>
</tr>
<tr>
<td>Registered Nurse FTEs</td>
<td>-0.00042</td>
</tr>
<tr>
<td>Licensed Prac. Nurse FTEs</td>
<td>-0.00097</td>
</tr>
<tr>
<td>Skilled Nursing Facilities</td>
<td>0.01119</td>
</tr>
<tr>
<td>Ambulatory Surgical Centers</td>
<td>0.03668</td>
</tr>
<tr>
<td>Commercial HMOs</td>
<td>0.03366</td>
</tr>
<tr>
<td>Comm. HMO Market Share</td>
<td>-3.06440</td>
</tr>
<tr>
<td>Oper Rms</td>
<td>-1.6E-05</td>
</tr>
<tr>
<td>Census Reg 2</td>
<td>3.88534</td>
</tr>
<tr>
<td>Census Reg 3</td>
<td>2.35060</td>
</tr>
<tr>
<td>Census Reg 4</td>
<td>3.02181</td>
</tr>
<tr>
<td>Census Reg 5</td>
<td>1.57534</td>
</tr>
<tr>
<td>Census Reg 6</td>
<td>0.09898</td>
</tr>
<tr>
<td>Census Reg 7</td>
<td>1.49314</td>
</tr>
<tr>
<td>Census Reg 8</td>
<td>0.16997</td>
</tr>
<tr>
<td>Census Reg 9</td>
<td>-0.46460</td>
</tr>
<tr>
<td>Constant</td>
<td>-9.69654</td>
</tr>
</tbody>
</table>
Table 12: Prescription Drug Out-of-Pocket Share

Least Squares Regression
Dependent Variable: $\ln(y) - \ln(1-y)$, where $y =$ total out-of-pocket/ total expenditures
Mean of $y$: 0.69
N = 1115
$F(31,1083) = 5.22$
$Pr[f > 5.22] = 0.00$
$R^2 = .130$

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>COEFFICIENT</th>
<th>STAN ERR</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic Cost Sharing Type (Coinsurance=1, Copay=0)</td>
<td>0.60069</td>
<td>0.61947</td>
<td>0.332</td>
</tr>
<tr>
<td>Generic Cost Sharing Amount</td>
<td>0.04748</td>
<td>0.02078</td>
<td>0.023</td>
</tr>
<tr>
<td>Brand Coverage Indicator</td>
<td>-1.38206</td>
<td>0.23120</td>
<td>0.000</td>
</tr>
<tr>
<td>Brand Cost Sharing Type (Coinsurance=1, Copay=0)</td>
<td>3.83710</td>
<td>0.83396</td>
<td>0.000</td>
</tr>
<tr>
<td>Brand Cost Sharing Amount</td>
<td>0.04842</td>
<td>0.00882</td>
<td>0.000</td>
</tr>
<tr>
<td>Annual Limit Indicator</td>
<td>-0.33073</td>
<td>0.18074</td>
<td>0.068</td>
</tr>
<tr>
<td>Annual Limit Amount</td>
<td>5.38E-05</td>
<td>7.84E-05</td>
<td>0.493</td>
</tr>
<tr>
<td>POS Option</td>
<td>-0.41147</td>
<td>0.24957</td>
<td>0.100</td>
</tr>
<tr>
<td>Health Plan Type (Risk=1, Cost=0)</td>
<td>1.27173</td>
<td>1.06888</td>
<td>0.234</td>
</tr>
<tr>
<td>Age</td>
<td>-0.13047</td>
<td>0.04556</td>
<td>0.004</td>
</tr>
<tr>
<td>Age²</td>
<td>0.00094</td>
<td>0.00033</td>
<td>0.005</td>
</tr>
<tr>
<td>Female</td>
<td>-0.39813</td>
<td>0.93271</td>
<td>0.670</td>
</tr>
<tr>
<td>Female*Age</td>
<td>0.01001</td>
<td>0.01244</td>
<td>0.421</td>
</tr>
<tr>
<td>Income</td>
<td>3.69E-07</td>
<td>4.08E-07</td>
<td>0.367</td>
</tr>
<tr>
<td>CMS Wage Index</td>
<td>-1.32279</td>
<td>0.60937</td>
<td>0.030</td>
</tr>
<tr>
<td>Total Population</td>
<td>-4.21E-07</td>
<td>1.57E-07</td>
<td>0.007</td>
</tr>
<tr>
<td>Medical Doctors</td>
<td>0.00014</td>
<td>0.00010</td>
<td>0.155</td>
</tr>
<tr>
<td>Surgeons</td>
<td>-0.02694</td>
<td>0.01585</td>
<td>0.089</td>
</tr>
<tr>
<td>Dentists</td>
<td>-0.00063</td>
<td>0.00043</td>
<td>0.144</td>
</tr>
<tr>
<td>Commercial HMOs</td>
<td>0.03291</td>
<td>0.01591</td>
<td>0.039</td>
</tr>
<tr>
<td>Commercial HMO Market Share</td>
<td>-2.01383</td>
<td>0.49607</td>
<td>0.000</td>
</tr>
<tr>
<td>Registered Nurse FTEs</td>
<td>-0.00020</td>
<td>0.00008</td>
<td>0.013</td>
</tr>
<tr>
<td>Hospital Beds</td>
<td>0.00012</td>
<td>0.00009</td>
<td>0.173</td>
</tr>
<tr>
<td>Ambulatory Surgical Centers</td>
<td>0.00726</td>
<td>0.00615</td>
<td>0.238</td>
</tr>
<tr>
<td>Census Region 2</td>
<td>-0.42439</td>
<td>0.32611</td>
<td>0.193</td>
</tr>
<tr>
<td>Census Region 3</td>
<td>-0.23668</td>
<td>0.38644</td>
<td>0.540</td>
</tr>
<tr>
<td>Census Region 4</td>
<td>-0.48891</td>
<td>0.49186</td>
<td>0.320</td>
</tr>
<tr>
<td>Census Region 5</td>
<td>-0.83851</td>
<td>0.36841</td>
<td>0.023</td>
</tr>
<tr>
<td>Census Region 6</td>
<td>1.30740</td>
<td>1.09398</td>
<td>0.232</td>
</tr>
<tr>
<td>Census Region 7</td>
<td>-0.98983</td>
<td>0.39309</td>
<td>0.012</td>
</tr>
<tr>
<td>Census Region 8</td>
<td>-0.47290</td>
<td>0.37771</td>
<td>0.211</td>
</tr>
<tr>
<td>Census Region 9</td>
<td>-0.27261</td>
<td>0.36519</td>
<td>0.456</td>
</tr>
<tr>
<td>Constant</td>
<td>5.51780</td>
<td>2.02999</td>
<td>0.007</td>
</tr>
</tbody>
</table>
Table 13: Dental Care Out-of-Pockets

Poisson Regression given dependent variable is greater than zero.
Dependent Variable: Number of Preventive Dental Events
given Preventive Dental Events >0.

Mean of Dependent Variable: 1.79
N = 4075
Pseudo $R^2 = 0.1129$

<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>COEFFICIENT</th>
<th>STAN ERROR</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preventive Dental Exp</td>
<td>0.00198</td>
<td>5.46E-05</td>
<td>0.000</td>
</tr>
<tr>
<td>Preventive Dental Exp$^2$</td>
<td>-3.34E-07</td>
<td>1.61E-08</td>
<td>0.000</td>
</tr>
<tr>
<td>Total Dental Exp</td>
<td>-0.00014</td>
<td>1.69E-05</td>
<td>0.000</td>
</tr>
<tr>
<td>Total Dental Exp$^2$</td>
<td>4.32E-09</td>
<td>5.62E-10</td>
<td>0.000</td>
</tr>
<tr>
<td>Metro Indicator</td>
<td>0.08336</td>
<td>0.04016</td>
<td>0.038</td>
</tr>
<tr>
<td>CMS Wage Index</td>
<td>-0.18208</td>
<td>0.10964</td>
<td>0.097</td>
</tr>
<tr>
<td>Total Population</td>
<td>-7.10E-08</td>
<td>3.67E-08</td>
<td>0.035</td>
</tr>
<tr>
<td>Dentists</td>
<td>0.00011</td>
<td>0.00006</td>
<td>0.100</td>
</tr>
<tr>
<td>Comm. HMO market share</td>
<td>0.09292</td>
<td>0.09847</td>
<td>0.345</td>
</tr>
<tr>
<td>Census Region 2</td>
<td>-0.05337</td>
<td>0.06986</td>
<td>0.445</td>
</tr>
<tr>
<td>Census Region 3</td>
<td>0.01784</td>
<td>0.06844</td>
<td>0.794</td>
</tr>
<tr>
<td>Census Region 4</td>
<td>-0.02939</td>
<td>0.07947</td>
<td>0.711</td>
</tr>
<tr>
<td>Census Region 5</td>
<td>0.01120</td>
<td>0.06977</td>
<td>0.872</td>
</tr>
<tr>
<td>Census Region 6</td>
<td>-0.10439</td>
<td>0.09092</td>
<td>0.251</td>
</tr>
<tr>
<td>Census Region 7</td>
<td>-0.04638</td>
<td>0.07837</td>
<td>0.554</td>
</tr>
<tr>
<td>Census Region 8</td>
<td>-0.06605</td>
<td>0.07671</td>
<td>0.389</td>
</tr>
<tr>
<td>Census Region 9</td>
<td>-0.03833</td>
<td>0.06735</td>
<td>0.569</td>
</tr>
<tr>
<td>Constant</td>
<td>0.14210</td>
<td>0.12066</td>
<td>0.239</td>
</tr>
</tbody>
</table>
Figure 1: Distributions of ln(Health Care Expenditures), Medicare Advantage and FFS Medicare (Full MCBS)

Figure 2: Distributions of ln(Health Care Expenditures), Medicare Advantage, Medigap, and FFS Medicare (MCBS Sample)
Figure 3: Medicare Advantage Enrollment Shares by Age/Gender
Medicare Managed Care Quarterly/State/County/Plan Database

Figure 4: Proxy Medigap Premium on Medicare Advantage Market Share
Figure 5: Satiation in Preferences because of Limits on the Efficacy of Treatment

\[ MRS = \frac{U_{ii}H_m}{U_C} \]
Figure 6: Indirect Utility and Optimal Quantity on Pretreatment State for $g(\theta) = \theta$

![Graph showing Indirect Utility and Optimal Quantity](image)

Figure 7: 

$\partial H^*/\partial \theta$ for $g(\theta) = \theta$

![Graph showing $\partial H^*/\partial \theta$](image)
Figure 8: Indirect Utility and Optimal Quantity on Pretreatment State for

\[ g(\theta) = \left(1 + \exp \left(\frac{\theta - a}{b}\right)\right)^{-1} \]

Figure 9:

\[ \frac{\partial H^*}{\partial \theta} \text{ for } g(\theta) = \left(1 + \exp \left(\frac{\theta - a}{b}\right)\right)^{-1} \]