PUBLIC INFORMATION IN MARKOV GAMES

ANDREW KLOOSTERMAN

UNIVERSITY OF VIRGINIA
DEPARTMENT OF ECONOMICS
P.O. BOX 400182
CHARLOTTESVILLE, VA 22904-4182
ASK5B@VIRGINIA.EDU
434-924-6753

KEYWORDS: MARKOV GAMES, VALUE OF INFORMATION, COLLUSION
JEL CODE: C73

ABSTRACT. In a Markov game, players engage in a sequence of games determined by a Markov process. In this setting, this paper investigates the impact of varying the informativeness of public information, as defined by Blackwell 1951 and 1953, pertaining to the games that will be played in future periods. In brief, when a curvature condition on payoffs is satisfied, the finding is that, for any fixed discount factor, the set of strongly symmetric subgame perfect equilibrium payoffs of a Markov game with more informative signals is contained in this set of equilibrium payoffs if the Markov game is played with any less informative signals. The second result shows that larger equilibrium payoffs are possible with less informative signals when the curvature condition fails, but only for some discount factors. The third result strengthens the curvature condition, but generalizes the first result to all subgame perfect equilibrium payoffs. Finally, a collusion application is presented to illustrate the curvature condition.

1. Introduction

A well-known fact of dynamic games is that intertemporal incentives can support equilibrium outcomes that are not possible in static games; threats of punishment in the future induce players to choose actions that are not myopically optimal. Many dynamic environments feature uncertainty about the nature of future interactions.

I am especially grateful to my advisors Ennio Stacchetti and David Pearce for their help on this project. I also wish to thank Tomasz Sadzik, Dilip Abreu, Laurent Mathevet, Jack Fanning, Elliot Lipnowski, Ilari Paasivirta, Peter Troyan and everyone else who has discussed it with me for many insightful comments.
As the future determines what punishments are possible, it is reasonable to suppose that information regarding it will impact equilibrium outcomes. However, it is unclear what this impact will be, and investigating this question forms the focus of this paper.

Dynamic environments with uncertainty about the future are ubiquitous in economics. Consider the example of collusion over the business cycle (Rotemberg and Saloner 1986, Haltiwanger and Harrington Jr. 1991, Kandori 1991, Bagwell and Staiger 1997). There is an infinite horizon over which an oligopoly of firms would like to collude, but, unlike the standard repeated oligopoly problem, demand fluctuates over time as the economy goes through booms and busts. Examples of information that firms may possess regarding future demand include industry demand reports and macroeconomic indicators. This paper speaks to whether it is high or low quality information that makes collusion easier. As shall be seen, the answer turns out to be low quality information.

Formally, the environment is a game-theoretic setting called Markov games (also commonly called stochastic games). Markov games generalize infinitely repeated games by allowing the game played each period to vary. There is a set of possible games and one is selected each period according to a Markov process. That is, the game in each period is randomly drawn from a distribution that depends on the game and potentially, although not in this paper, the players’ actions in the previous period.

In this paper, the Markov games model is modified with the addition of public information. Specifically, in every period, before they choose their actions, the players receive a public signal about the game in the next period.\footnote{The game in the following period must be known by some agent outside the Markov game (usually nature) that generates the public signal. As signals arrive before actions are taken, this is why the Markov process cannot depend on the actions of the players in the previous period.} Because the current game is known, there are no direct payoff consequences of the signals. However, as mentioned previously, information is relevant for equilibrium behavior because beliefs about future games determine which intertemporal incentives can be supported. For example, returning to the collusion game, a signal that predicts high demand in the next period would make collusion easier, as a cheating firm would forgo more expected future profit.

Employing Blackwell’s 1951 and 1953 criterion to define how one set of signals is more informative than another, this paper compares the equilibrium payoffs of two hypothetical worlds where the same Markov game is played but with more informative public signals in one world, called world M, than in the other world, called world L.
There are three results. First, under a curvature condition on payoffs, the set of strongly symmetric subgame perfect equilibrium payoffs in world M is contained in this set of equilibrium payoffs in world L. The result holds for any fixed discount factor. In the sense that larger equilibrium payoffs are possible in world L, the value of information is negative. Second, when no curvature condition is imposed, there are still some discount factors for which the best feasible payoff in the Markov game is an equilibrium payoff only in world L. Therefore, the value of information is still negative for these discount factors. Finally, the first result is generalized to the case of all equilibria. Strong symmetry weakens the curvature condition, but is not the crucial component of the argument.

Complementing the general results, an application section is provided to show that the curvature assumption holds for the collusion game mentioned above, whether the competition is price (Bertrand) or quantity (Cournot). The analysis characterizes the most collusive equilibrium to show that when the curvature assumption holds weakly, price competition, the most collusive equilibrium is only strictly larger in world L for a few discount factors, but when it holds strictly, quantity competition, the most collusive equilibrium is strictly larger in world L for every discount factor except for those where full collusion is possible in both worlds.

The theory developed in this paper has other applications in development economics, labor economics, and political economy. Farmers in villages face stochastic crop yields due to weather, but try to smooth consumption by setting up informal insurance with each other. The information in this environment is weather reports. The same risk-sharing game applies to workers who are subject to layoffs, and information could be anything about the company’s future plans. Finally, leaders choose their policies contingent on the possibility of losing power (Acemoglu, Golosov, and Tsyvinski 2011), and public information could be polling data if the government is democratic or arms-stockpiling if the government is autocratic.

The impact of information in environments with uncertainty is a very natural question, and has been studied extensively in the past. Kandori 1992 is the most related, because he also studies information in an infinite horizon context where intertemporal incentives are present. Kandori shows that more informative signals about past actions increase equilibrium payoffs in infinitely repeated games with imperfect public

---

monitoring. The findings here show that information about an uncertain future has a very different impact than information about an uncertain path.

The seminal work on this question is Blackwell 1951 and 1953. He shows that more informative signals increase payoffs in decision problems with uncertainty. His work has been extended to games in a number of ways, and the results are more mixed. Though many of these papers share the result of a negative value of information, the environment and analysis are very different; information is directly payoff-relevant rather than indirectly relevant through its impact on intertemporal incentives. It should be specially noted that a large portion of this literature examines the impact of information in one-shot oligopoly games (Ponssard 1979, Vives 1984, Sakai 1985, Einy et al 2003).

The paper proceeds as follows. Section 2 presents the model, Section 3 presents the theory, Section 4 considers the collusion application, and Section 5 concludes. Some proofs are in the appendix while others are relegated to the online appendix.

2. Model

2.1. Preliminaries. A Markov game is a tuple \( \mathcal{G} = (S, q_0, Q, \delta) \). \( S \) is a finite set of normal form stage games so each \( s \in S \) is a tuple \( s = (N, (A_i, u_i^s)_{i \in N}) \) where \( N \) is the set of players, \( A_i \) is the set of actions for player \( i \), and (with \( A = \times_{i=1}^N A_i \)) \( u_i^s : A \to \mathbb{R} \) is the utility function for player \( i \) in game \( s \). It is assumed, except where noted below, that \( A_i \) is a convex and compact subset of \( \mathbb{R}^{p_i} \) for some \( p_i \in \mathbb{N} \) for all \( i \), and that \( u_i^s \) is continuous for all \( i \) and \( s \).

There is an infinite horizon and \( q_0 \) and \( Q \) describe the Markov process that determines which game from \( S \) is played each period. \( q_0 \in \mathbb{R}^{|S|} \) is a unit length vector that is the distribution from which the period 0 game is selected. \( Q \) is an \( |S| \times |S| \) matrix whose rows are the distributions from which the period \( t + 1 \) game is selected conditional on the period \( t \) game. In particular, for each current game \( s \) and potential game next period \( s' \)

\[ Q_{ss'} = \text{Prob}(s'|s) \]

\(^3\)Kandori's results apply to a different model of collusion where there is uncertainty about current demand (Green and Porter 1984).


\(^5\)It is without loss of generality that only \( u_i^s \) depends on \( s \) (see Mailath and Samuelson 2006).

\(^6\)The only condition imposed on \( Q \) is that each \( s \) occurs with positive probability in some period.
The sequence of a period is as follows; the game is revealed, actions are taken simultaneously, all actions are observed, and finally payoffs are realized. The last object in $G$ is the common discount factor $\delta \in [0, 1)$.

Every period, after the current game is revealed but before actions are taken, the players receive a public signal about the game in the next period. Formally, public information is described by a tuple $I = (X^s, F^s)_{s \in S}$. $X^s$ is the finite set of public signals that can be received when the current game is $s$. $F^s$ is a $|S| \times |X^s|$ matrix whose rows are the distributions from which a signal is drawn conditional on the true game $s'$ in period $t+1$.\footnote{The easiest way to interpret information here is to suppose that nature has already selected the game for next period and draws a signal about it to show to the players.} In particular, for the true game $s'$ next period and signal $x$

$$F_{s'x}^s = \text{Prob}(x|s, s')$$

To reiterate, $s$ is known while $s'$ is unknown. The dependence on $s$ of $X^s$ and $F^s$ is only to allow different signals and distributions in different current games.

The states in the Markov game with public information are the game and signal pairs, and the state space is denoted by $Z$:

$$Z = \bigcup_{s \in S} (\{s\} \times X^s)$$

Some results impose symmetry for which it is assumed that every game is symmetric. This means $A_i = A_j$ and $u^s_i(a) = u^s_j(\bar{a})$ if $a_i = \bar{a}_j$ and $a_{-i}$ is any permutation of $\bar{a}_{-j}$ for every pair of players $i$ and $j$. This unique action set is denoted $\bar{A}$.

2.2. Strategies and Equilibrium. Let $H$ denote the set of all histories; past games, past public signals, and past actions. A pure strategy $\sigma = (\sigma_1, \ldots, \sigma_N)$ consists of a mapping $\sigma_i : H \times Z \rightarrow A_i$ that specifies an action for every history and current state for every player $i$. Except where noted, only pure strategies will be considered in this paper, a common assumption with continuum action spaces.\footnote{Pure strategy equilibria are not guaranteed to exist without additionally assuming quasiconcavity on the utility functions, but in this paper the results are vacuously true when there is no equilibrium.}

For a fixed strategy $\sigma$, let $v(\sigma) \in \mathbb{R}^{|H|Z}$ be the expected discounted sum of payoffs if the players follow $\sigma$. Adopting common notation, this object will be called the value associated with $\sigma$. Each component $v_{i,sx}(\sigma)\footnote{Subscripted states omit the extra comma throughout the paper.}^{9}$ is

$$v_{i,sx}(\sigma) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t E_{h^t, s_t, x_t} \left[ u^{s_t}_i(\sigma(h^t, s_t, x_t)) \right] (s_0, x_0) = (s, x)$$
For any strategy $\sigma$ and history $h \in \mathcal{H}$, the continuation strategy $\sigma|h$ is defined, for each history $\hat{h} \in \mathcal{H}$ and state $(s, x)$, by

$$\sigma|h(\hat{h}, s, x) = \sigma((h, \hat{h}), s, x)$$

The equilibrium concept is subgame perfect equilibrium, hereafter called SPE. A strategy $\sigma$ is a SPE if for every history $h \in \mathcal{H}$, state $(s, x) \in \mathcal{Z}$, player $i$, and strategy $\hat{\sigma}_i$ for player $i$

$$v_{i,sx}(\sigma|h) \geq v_{i,sx}(\hat{\sigma}_i, \sigma|_{i \neq h})$$

If $\sigma$ is a SPE, then $v(\sigma)$ is called a SPE value and, for a given discount factor $\delta$,

$$V(\delta) \subseteq \mathbb{R}^{N|Z|}$$

The recursive methods developed by Abreu, Pearce, and Stacchetti 1986 and 1990 can be adapted to characterize $V(\delta)$ by defining decomposability and self-generation in Markov games with information. This is possible, because Markov games with information are themselves Markov games and Markov games have a recursive structure like infinitely repeated games.\(^\text{10}\) Adding information maintains the Markov structure where the states define the games and the Markov process is given by the $|Z| \times |Z|$ matrix $P$ with

$$P_{sx,s'x'} = \frac{Q_{ss'}F^{s}_{s'} \sum_{\hat{s} \in \hat{S}} Q_{s\hat{s}}F^{\hat{s}}_{s\hat{s}'}}{\sum_{\hat{s} \in \hat{S}} Q_{s\hat{s}}F^{\hat{s}}_{s\hat{s}'} F^{s}_{s'}_{sx}}$$

The characterization is only used for the analysis of strong symmetry, the restriction that $\sigma_i(h, s, x) = \sigma_j(h, s, x)$ for all $(h, s, x)$ and for all players $i$ and $j$, so it is introduced in that environment. To simplify notation, define the functions\(^\text{11}\) $\tilde{u}^s : \tilde{A} \rightarrow \mathbb{R}$ and $\hat{u}^s : \hat{A} \rightarrow \mathbb{R}$

$$\tilde{u}^s(a) = u^s_i(a, \ldots, a) \text{ and } \hat{u}^s(a) = \max_{a' \in \hat{A}} u^s_i(a', a, \ldots, a)$$

Let $\hat{a}$ denote an action that maximizes the latter function.

Finally, with strong symmetry all players are identical, so the player dimension for payoff sets can be dropped. In the following definition, for the set $W \subseteq \mathbb{R}^{|Z|}$, $W_{sx}$ refers to the projection of $W$ onto the state $(s, x)$.

**Definition 1.** Let $W \subseteq \mathbb{R}^{|Z|}$. $v_{sx} \in \mathbb{R}$ is decomposable on $W$ at state $(s, x)$ if there exists an action $a_{sx} \in \tilde{A}$ and a function $\gamma_{sx} : A \times Z \rightarrow \mathbb{R}$ such that $\gamma_{sx}(a, s', x') \in W_{sx}$.

\(^\text{10}\)This tractability advantage is the reason for the assumptions that the underlying stochastic process is Markov and that signals are about games in the next period only. In principle, both of these assumptions can be relaxed, but the characterization of equilibrium payoffs would be greatly encumbered.

\(^\text{11}\)Using player 1’s utility is an arbitrary choice as each game is symmetric.
for all \((a, s', x') \in A \times Z\), and
\[
v_{sx} = (1 - \delta)\hat{u}(a_{sx}) + \delta E_{s'x'}[\gamma_{sx}((a_{sx}, \ldots, a_{sx}), s', x')|s, x]
\geq (1 - \delta)\hat{u}(a_{sx}) + \delta E_{s'x'}[\gamma_{sx}((\hat{a}_{sx}, a_{sx}, \ldots, a_{sx}), s', x')|s, x]
\]

Adopting some other common notation, the action \(a_{sx}\) and function \(\gamma_{sx}\) that decompose \(v_{sx}\) are called the enforcing action and enforcing continuation function respectively. Also, the number \(\gamma_{sx}((a_{sx}, \ldots, a_{sx}), s', x')\) is called the enforcing continuation value for next period state \((s', x')\). Furthermore, the equality is called the value recursion while the inequality is called the incentive constraint. The restriction that \(\gamma_{sx}(a, s', x') \in W_{s'x'}\) is called the feasibility of the enforcing continuation function.

**Definition 2.** \(W\) is self-generating if for every \(v \in W\) and state \((s, x)\), \(v_{sx}\) is decomposable on \(W\) at state \((s, x)\).

The following proposition characterizes \(V(\delta)\).\(^{12}\)

**Proposition 1.** Let \(\delta \in (0, 1)\) and \(W\) be a bounded subset of \(\mathbb{R}^{|Z|}\).

1. If \(W\) is self-generating, then \(W \subseteq V(\delta)\).
2. \(V(\delta)\) is the largest bounded self-generating set.

2.3. Information. In order to determine the impact of information, the equilibrium payoffs of a Markov game modified with more informative signals, called world M (for more), are compared to the equilibrium payoffs of the same Markov game modified with less informative signals, called world L (for less). World M adopts the notation for information introduced above, \(\mathcal{I}^M = (X^s, F^s)_{s \in S}\) while world L information is denoted by \(\mathcal{I}^L = (Y^s, G^s)_{s \in S}\).\(^{13}\) More informative signals are defined formally following the classic definition of Blackwell. Recall that a stochastic matrix is a matrix with rows summing to 1.

**Information:** (Blackwell 1951 and 1953) World M has more informative signal structures than world L. That is, for each \(s \in S\), there is a stochastic matrix \(R^s\) of size \(|X^s| \times |Y^s|\) such that
\[
G^s = F^s R^s
\]

The matrix \(R^s\) has a standard interpretation. Both \(x\) and \(y\) signals are informative about next period’s game \(s'\). However, the less informative signals \(y\) do not directly

\(^{12}\)The proof follows similarly to Abreu, Pearce, and Stacchetti 1990 and is therefore omitted.

\(^{13}\)The superscripts \(M\) and \(L\) are used extensively to denote the world in which objects lie.
inform players about $s'$ but rather take the more informative signals and garble them. The garbling motivates an “information identity” that is crucial to the one-sided nature of the results in the next section:

$$\text{Prob}(y|s, x, s') = \text{Prob}(y|s, x)$$

A classic example is noisy signals. The signal tells players the true game in the following period but is a mistake with some probability $\epsilon \in (0, 1/2)$. In this example, $X^s = S$ for every game $s$ and the diagonal components of $F^s$ are $1 - \epsilon$ while the off-diagonal components are $\epsilon/(|S| - 1)$. The same is true for world L, except the noise parameter is $\epsilon'$ where $\epsilon' > \epsilon$. This is a special case of Blackwell’s informativeness. For instance, when $S = \{s_1, s_2\}$, then (for both $s = s_1$ and $s = s_2$) $F^s$, $G^s$, and $R^s$ are

$$F^s = \begin{pmatrix} 1 - \epsilon & \epsilon \\ \epsilon & 1 - \epsilon \end{pmatrix}, \quad G^s = \begin{pmatrix} 1 - \epsilon' & \epsilon' \\ \epsilon' & 1 - \epsilon' \end{pmatrix}, \quad R^s = \begin{pmatrix} 1 - \epsilon - \epsilon' & \epsilon - \epsilon' \\ \epsilon' - \epsilon & 1 - 2\epsilon' \end{pmatrix}$$

2.4. Comparison. Blackwell imposes no restrictions on $X^s$ and $Y^s$ so $V^M(\delta)$ and $V^L(\delta)$ are not necessarily subsets of the same space and are therefore not directly comparable. The natural comparison is to consider payoffs at the interim stage within a period; after the game for the current period is revealed but before the signal is observed.

For a fixed strategy $\sigma$, this is called the interim value associated with $\sigma$ and denoted $v^M_{i,s}(\sigma) \in \mathbb{R}^{N|S|}$ and $v^L_{i,s}(\sigma) \in \mathbb{R}^{N|S|}$ in the respective worlds. For a given player $i$ and game $s$, $v^M_{i,s}$ and $v^L_{i,s}$ are

$$v^M_{i,s}(\sigma) = E_x[v^M_{i,sx}(\sigma)|s] \quad \text{and} \quad v^L_{i,s}(\sigma) = E_y[v^L_{i,sy}(\sigma)|s]$$

Interim values that are the expectation of SPE values are called interim SPE values and, for a given discount factor $\delta$, the sets of all SPE interim values are denoted by $V^M(\delta)$ and $V^L(\delta)$ in worlds M and L respectively.

An important remark is that just because values were introduced first and interim values constructed from them, interim values should not be thought of as any less canonical than values. An equivalent exposition would have defined payoffs from the interim standpoint in Section 2.2 and then construct ex-post values for the purpose of equilibrium incentive constraints.

---

14Note that the conditional expectation is taken over a vector rather than a function. It should be interpreted in the obvious way: the components of the vector correspond to the domain of a function; i.e. in the case of world M, the function is $f_i : Z^M \to \mathbb{R}$ where $f_i(s, x) = v^M_{i,sx}(\sigma)$.
The public signals affect equilibria though the role they play in supporting intertemporal incentives. Signals that predict future games where large losses are possible (harsh punishments) support many actions in the current period, whereas signals that predict future games where only small losses are possible support few actions in the current period. More informative signals of the first type support more actions than less informative signals, but more informative signals of the second type support fewer actions so the net effect is ambiguous. The results in this section resolve this ambiguity.

There are many complicating issues, two of which figure prominently. First, the scope for punishment is partly exogenous as it depends on the utility functions in each game, but also partly endogenous as it depends on the strategy employed. Second, it is not necessarily easy to classify signals as predicting games where large losses are possible or predicting games where only small losses are possible, because signals predict a distribution over games rather than a specific game.

3. Strong Symmetry. The main assumption for the first theorem relates to the symmetric utility function $\bar{u}$ and the maximal deviation function $\hat{u}$ defined in Section 2.2.

Assumption CURV 1: For every finite integer $n$, points $a, a_1, \ldots, a_n \in \bar{A}$ and weights $\lambda_1 \ldots \lambda_n \in [0, 1]$ summing to 1,

$$\text{if } \bar{u}(a) = \sum_{k=1}^{n} \lambda_k \bar{u}(a_k), \text{ then } \hat{u}(a) \leq \sum_{k=1}^{n} \lambda_k \hat{u}(a_k).$$

Intuitively, Assumption CURV 1 says that actions that deliver intermediate payoffs have a smaller gain from deviation than the average gain from actions that deliver more extreme payoffs so intermediate actions may be easier to support as equilibria. One illustrative sufficient condition for Assumption CURV 1 is also given in Section 4.

Theorem 1. Assume CURV 1 holds. Then $V^M(\delta) \subseteq V^L(\delta)$ for every discount factor $\delta \in [0, 1)$.

The proof constructs a self-generating set $W^L$ in world $L$ from $V^M(\delta)$ in world $M$. For each state $(s, y)$ in world $L$, an enforcing action and enforcing continuation function are constructed by averaging out the information in the world $M$ signals.
conditional on the information in \((s, y)\). The main difficulty is that continuation payoffs condition on the current state. However, the law of iterated expectation ensures that the signals in the current state in world M have no effect, because the continuation payoff is further conditioned on the less informative signal in world L. The “information identity” is crucial for applying the law of iterated expectation. Additionally, the role of CURV 1 is to ensure that gain from deviating from the enforcing action, an average action by construction, is not too large.

The theorem says that the players can do better and do worse in world L. However, focusing on the largest equilibrium payoff as the criterion for comparison, the value of information is negative. Section 4 illustrates this point by characterizing the most collusive equilibrium in the application of collusion over the business cycle.

3.2. No Curvature Assumption. How crucial is Assumption CURV 1? The second theorem shows that for some discount factors it is not crucial at all as the largest feasible payoff in the Markov game is a SPE value for a smaller discount factor in world L.

For each \(s \in S\), consider \(a_s^* \in \text{argmax}_{a \in \tilde{A}} \tilde{u}(a)\) and a Nash equilibrium action \(a^{NE}_s\).\(^{15}\) This section analyzes the following grim trigger strategy. The players all choose \(a_s^*\) for the current game \(s\) unless any player deviates after which the players all choose \(a^{NE}_s\) for the current game \(s\). Denote the values associated with the grim trigger strategy by \(v^{M,*}\) and \(v^{L,*}\) in worlds M and L respectively. Similarly, denote the values associated with perpetual play of the Nash equilibria by \(v^{M,NE}\) and \(v^{L,NE}\).\(^{16}\)

Theorem 2 is a comparison of the cutoff discount factors for which the grim trigger strategy is a SPE. The result is usually strict, and the interest in a strict result will be evident momentarily, but there are a few cases where it is not. To rule out the two fairly obvious cases where the result is not strict, a Markov game in world L is called Non-Trivial if there exists an \(s\) such that \(a_s^*\) is not a Nash Equilibria for game \(s\), and for any \((s, y)\) such that \(a_s^*\) is not a Nash Equilibria, there exists an \(\hat{s}\) that occurs with positive probability in some future period given \((s, y)\) and \(\tilde{u}(a_s^*) > \tilde{u}(a^{NE}_s)\).\(^{17}\) The first piece of this condition rules out the case where the cutoff discount factors are both zero as static and dynamic incentives are aligned. The second piece rules out the case where both cutoff discount factors are one as the grim trigger strategy is never a SPE.

\(^{15}\)If the only Nash equilibria are in mixed strategies, then the restriction to pure strategies can be relaxed to allow the players to choose a mixed strategy Nash equilibrium.

\(^{16}\)It is intuitive and straightforward to show that \(\pi^{M,*} = \pi^{L,*}\) and \(\pi^{M,NE} = \pi^{L,NE}\).

\(^{17}\)The latter condition is only imposed in world L as this implies it also holds in world M.
The third case is more interesting. It is possible that every signal predicts that the same expected loss would be inflicted after a deviation.\textsuperscript{18} A Markov game in world M is called \textit{Future Informative with respect to discount factor} $\delta$, if for all $(s, x) \in Z^M$, if
\[
(1 - \delta)(\hat{u}^s(a^*_s) - \bar{u}^s(a^*_s)) = \delta E_{s'x'}[v^{M*,s}_{s'x'} - v^{M,NE}_{s'x'} | s, x]
\]
then, for all $y \in Y^s$, there exists $\hat{x} \in X^s$ such that $R^s_{\hat{x}y} > 0$ and
\[
E_{s'x'}[v^{M*,s}_{s'x'} - v^{M,NE}_{s'x'} | s, \hat{x}] \neq E_{s'x'}[v^{M*,s}_{s'x'} - v^{M,NE}_{s'x'} | s, x]
\]
In words, Future Informativeness says that if, in state $(s, x)$, the incentive constraint binds (the first equality), then there exists a signal $\hat{x}$ that occurs with positive probability ($R^s_{\hat{x}y} > 0$) where the expected future loss from deviating ($E_{s'x'}[v^{M*,s}_{s'x'} - v^{M,NE}_{s'x'} | s, \hat{x}]$) is different than after the signal $x$.

**Theorem 2.** There exists $\delta^{M,*}$ (and $\delta^{L,*}$) such that the grim trigger strategy is a SPE iff $\delta \geq \delta^{M,*}$ (and $\delta \geq \delta^{L,*}$) in world $M$ (and world $L$).

- $\delta^{M,*} \geq \delta^{L,*}$, and
- $\delta^{M,*} > \delta^{L,*}$ iff the Markov game is also Non-Trivial in world $L$ and Future Informative with respect to discount factor $\delta^{M,*}$ in world $M$.

The existence of cutoff discount factors is standard. The interesting result is the comparison of these cutoffs and the intuition for this result is straightforward. The cutoff discount factor in each world is determined by the state where the gains from deviating are the largest and future expected losses are the smallest. The gains from deviating are the same in both worlds since the same action $a^*_s$ is taken regardless of the signal, but the smallest expected future loss is smaller in world $M$ because there is a more accurate signal predicting that games in future periods will only permit these small punishments. Therefore, more patience is required in world $M$.

By construction, $v^{M,*}$ and $v^{L,*}$ are the maximum feasible payoffs in the respective Markov games with information. To obtain a result more comparable to Theorem 1, the Nash equilibrium payoff also must deliver the minmax payoff. Then, perpetual play of the Nash equilibria is an optimal punishment and $\delta^{M,*}$ and $\delta^{M,*}$ are the cutoffs for $v^{M,*}$ and $v^{L,*}$ to be SPE values. Set containment also requires assuming the sets of SPE values are convex.

\textsuperscript{18}One situation that needs to be ruled out is deterministic Markov processes (such as the special case of repeated games), but there are also cases where the signals are informative, just not about losses. See the online appendix for such an example.
Corollary 1. If for every game s there is Nash equilibrium delivering the minmax payoff to all players, then the largest feasible interim payoff in the Markov game is a SPE interim value in world L but not in world M for every $\delta \in [\delta^L, \delta^M]$. Furthermore, if $V^M(\delta)$ is convex, then $\nabla^M(\delta) \subset \nabla^L(\delta)$ for every $\delta \in [\delta^L, \delta^M]$.

The smallest interim SPE value comes from playing the Nash equilibria when the equilibria all minmax the players so the players cannot do strictly worse in world L. They do strictly better for some discount factors, so in this sense, the value of information is strictly negative for these discount factors.

The corollary is interesting when $\delta^L$ is strictly less than $\delta^M$. Future Informative-ness is a very weak condition. However, the assumption of minmaxing Nash Equilibria is inherently stronger. Still, Theorem 2 has one property that makes it a stronger assertion than Theorem 1. The theorem applies to games with finite action spaces, such as the prisoner’s dilemma, in addition to games with continuous action spaces as assumed in the model.

One final remark is that there is nothing special about considering $a^*_s$ in Theorem 2. The result holds for any set of actions, one for each game s. However, playing an action that is not conditioned on the information in the public signal does not make sense for attaining other payoffs. The best feasible payoff is attainable only by choosing $a^*_s$, which is why this section is interesting. Other payoffs are more easily achieved by choosing different actions after different signals in a way that uses the information available in the signals.

3.3. No Symmetry. The analysis now turns to showing that the main idea in Theorem 1 is robust to the general case of all SPE. Denote the (entire) maximal deviation function by $\hat{u}_s^i : A_{-i} \to \mathbb{R}$ where $\hat{u}_s^i(a_{-i}) = \max_{a'_i \in A_i} u_s^i(a'_i, a_{-i})$.

Assumption CURV 2: $\hat{u}_s^i$ is convex and $u_s^i$ is concave for each $s \in S$ and $i \in N$.

Theorem 3. Assume CURV 2 holds and for every game s there is a Nash Equilibrium delivering the minmax payoff to all players. For each discount factor $\delta \in (0, 1)$ and for every $v^M \in V^M(\delta)$, there exists a $v^L \in V^L(\delta)$ such that $\bar{v}^L \geq \bar{v}^M$.

The central issue in extending Theorem 1 is that the Intermediate Value Theorem cannot be used to obtain the existence of an action $a^L_{xy}$ such that $u^i_s(a^L_{xy}) = E_x[u^i_s(a^M_{xy})|s,y]$ for every player $i$. The Intermediate Value Theorem would yield an action $a$ for player $i$ to get the desired equality and an action $a'$ for player $j$. With strong symmetry, $a = a'$ as player $i$ and player $j$ are equivalent but this is no longer
the case. Instead, $a_{sy}^L$ is constructed as the expectation across actions in world $M$: $a_{sy}^L = E_x[a_{sx}^M | sy]$.

Assumption CURV 2 is used like CURV 1, but is adapted to the new construction of the enforcing action. Because $u^*_i$ is concave, the construction creates a larger payoff in world $L$ than in world $M$. Because $\hat{u}^*_i$ is convex, the incentive to deviate is lower. Also, as larger values are constructed in world $L$, the set containment result is lost.

A technical issue makes the proof of Theorem 3 seem more complicated than it really is. From the standpoint of decomposition, it is difficult to obtain feasible enforcing continuation functions because the constructed values are larger in world $L$ than world $M$ rather than equal. The same property causes issues for constructing punishments so, for simplicity, minmaxing equilibria are also assumed in the assertion. Still, the proof follows the same logic of Theorem 1, but applies it to the strategies directly rather than using self-generation techniques. Strategies are far more cumbersome, but the ideas are all the same.

Finally, one might ask about an analogous result to Theorem 2. A similar result holds for maximizing feasible welfare, though there are some additional issues.\[19\]

4. Collusion

In this section, the model is applied to the game of collusion over the business cycle in the setting of both price (Bertrand) and quantity (Cournot) competition. For simplicity, the marginal cost of all firms is set to 0, although similar results are obtained when cost is the variable that changes over time.

4.1. Assumption CURV 1. Suppose first that $N$ identical firms compete in price. For current demand $s$, the firm that quotes the lowest price $p \in \mathbb{R}_+$ makes sales determined by the continuous demand function $Q^*(p)$ and all other firms make zero sales; if multiple firms quote the same lowest price, then the sales are split evenly among them. Strong symmetry in this setting considers the case where each of the $N$ firms sets the same price each period and sales are split evenly so

$$\tilde{u}^*(p) = 1/NQ^*(p)p$$

Any firm can undercut the agreed price by a negligible amount and capture the entire market at essentially the agreed price so the payoff of the best deviation is

$$\hat{u}^*(p) = Q^*(p)p$$

\[19\] See Kloosterman 2014 for a rigorous analysis of this case.
The maximal deviation function is a linear transformation of the utility function so it is easy to see that CURV 1 is satisfied. However, it is only weak in the sense that the inequality for \( \hat{u} \) in CURV 1 is an equality.

Alternatively, suppose that the N identical firms compete in quantity. For current demand \( s \), the price is determined by the linear inverse demand function \( P^s(q) \), parameterized by \( a^s \) and \( b^s \), where \( q \in \mathbb{R}_+^N \) is the vector of firms’ quantities:\(^{20}\)

\[
P^s(q) = a^s - b^s \left( \sum_{i=1}^{N} q_i \right)
\]

Strong symmetry in this setting considers equilibria in which each of the \( N \) firms sets the same quantity \( q \in \mathbb{R}_+ \) in every period and profits are split evenly so

\[
\bar{u}^s(q) = (a^s - b^s Nq)q
\]

Algebra yields the maximal deviation function

\[
\hat{u}^s(q) = \max \left\{ \frac{1}{4b^s} (a^s - (N - 1)b^s q)^2, 0 \right\}
\]

\( \bar{u} \) is concave while \( \hat{u} \) is convex. To show that CURV 1 is satisfied, the following proposition can be invoked.\(^{21}\)

**Proposition 2.** Suppose \( \tilde{A} \subseteq \mathbb{R} \), \( \bar{u} \) and \( \hat{u} \) are both increasing or both decreasing, \( \bar{u} \) is concave, and \( \hat{u} \) is convex. Then Assumption CURV 1 is satisfied.

Unlike for the case of price competition, here CURV 1 is strictly satisfied.\(^{22}\)

4.2. **The Most Collusive Equilibrium.** The collusion applications satisfy the conditions for Theorem 1 to apply. In order to illustrate this result and clarify the impact of information, two examples are provided here that solve for the most collusive SPE (the equilibrium in which the firms make the most profit) as a function of the discount factor. For both examples, suppose there are two games, \( H \) (a high demand game)

---

\(^{20}\)The price can be negative. This allows negative profits (an important part of the equilibria in the upcoming Proposition 4) even though marginal cost is normalized to 0. The proposition also holds for the more realistic case where the price is restricted to be positive as long as the marginal cost is also positive so negative profits are possible.

\(^{21}\)\( \bar{u} \) is actually increasing for quantities less than than the monopoly quantity, but it can be assumed that such quantities are never chosen as, for every quantity less than monopoly, there is a quantity larger than monopoly that delivers the same payoff with a smaller gain from deviation.

\(^{22}\)A few other elements of the model need to be tweaked for these applications. First, for the price case, only the symmetric portion of the utility function is continuous, but this is no problem as only continuity of the symmetric portion is necessary. Second, the action sets are unbounded but this is acceptable because there is the upper bound of monopoly on profits.
and $L$ (a low demand game). The information considered is the case of noisy signals introduced in Section 2.3. For clarity, let $X^s = Y^s = \{h, l\}$ so lowercase letters refer to signals and uppercase letters refer to the game. The set of signals in worlds $M$ and $L$ are the same and the analysis will refer to arbitrary signals in either world as $x$’s.

The point made in this section is that for each fixed $\delta$, as $\epsilon$ increases to $\epsilon'$, the interim value associated with the most collusive equilibrium at least weakly increases as well. While the information considered does not encompass the full generality that Blackwell permits, it covers a one-dimensional continuum of possibilities and is quite practical for many real world situations.

Price Competition:
Let $\Pi^s$ denote the monopoly profit per firm for demand $s$ and assume $\Pi^H > \Pi^L$. The most collusive equilibrium is described by the following grim trigger strategy.

**Proposition 3.** The most collusive equilibrium can be characterized by a grim trigger strategy consisting of a vector of four agreed profits per firm $\Pi = (\Pi_{HH}, \Pi_{HL}, \Pi_{LH}, \Pi_{LL})$. The firms all choose a price that delivers profit $\Pi_{sx}$ corresponding to the current state $(s, x)$ every period unless any firm deviates after which the firms all choose price zero in all future periods. The vector $\Pi$ is the (unique) largest vector in $[0, \Pi^H]^2 \times [0, \Pi^L]^2$ such that, for each $(s, x)$,

$$(I - \delta P)_{sx}^{-1} \cdot \Pi \geq N\Pi_{sx} \quad (=) \text{ if } \Pi_{sx} < \Pi^s$$

Figures 1 and 2 illustrate the vector of agreed profits and the interim values for the most collusive SPE for the parameters $\Pi^H = 25$, $\Pi^L = 5$, $N = 2$, $\epsilon = .05$, $\epsilon' = .25$, and

$$Q = \begin{pmatrix} .75 & .25 \\ .25 & .75 \end{pmatrix}$$

For $\delta < .5$ no collusion is possible, and for $\delta > .65$ full collusion is possible so information only matters for $\delta \in [.5, .65]$. In this range, agreed profits are more widespread in world $M$ than in world $L$ for game $H$. More profit is possible in state $(H, h)$ in world $M$ because the signal $h$ predicts high demand is very likely in the future. However, for the same reason, less profit is possible in state $(H, l)$. As can be seen in Figure 2, the effects exactly cancel out in the interim payoffs until $\Pi^H$ is possible in state $(H, h)$ in world $M$. Theorem 2 still holds so players can achieve full collusion with strictly less patience in world $L$. 
A strict difference in interim values only occurs for a small range of discount factors. This is a consequence of CURV 1 holding weakly. Still, due entirely to Theorem 2, some discount factors exist for which the result is strict so the value of information is negative.

**Quantity Competition:**

Let $\rho^s$ denote the monopoly quantity in game $s$. The most collusive equilibrium can be characterized using the standard tool for this game in the repeated game setting; a stick-carrot scheme with a single period stick (Abreu 1986).

**Proposition 4.** The most collusive equilibrium can be characterized by a stick-carrot scheme with a single period stick consisting of two vectors of quantities, carrots...
Figure 3. Quantity Competition: Quantities for \( \epsilon = .05 \).

\[ q^c = (q^c_{HH}, q^c_{HL}, q^c_{LH}, q^c_{LL}) \] and sticks \( q^{st} = (q^{st}_{HH}, q^{st}_{HL}, q^{st}_{LH}, q^{st}_{LL}) \). The firms choose the component of \( q^c \) corresponding to the current state every period unless any firm deviates. Following any deviation, firms choose the component of \( q^{st} \) corresponding to the realized state next period and then the corresponding components in \( q^c \) every period after that. For each state \((s, x)\), the carrots and sticks are uniquely determined by the two equations

\[
\hat{u}^s(q^{st}_{sx}) = \tilde{u}^s(q^{st}_{sx}) + \delta E_{s'x'}[(\tilde{u}^{s'}(q^{c}_{s'x'}) - \tilde{u}^{s'}(q^{st}_{s'x'}))]|s, x| \\
\hat{u}^s(q^{c}_{sx}) \leq \tilde{u}^s(q^{c}_{sx}) + \delta E_{s'x'}[(\tilde{u}^{s'}(q^{c}_{s'x'}) - \tilde{u}^{s'}(q^{st}_{s'x'}))]|s, x| \quad (=) \text{ if } q^{c}_{sx} > \rho^s
\]

Figures 3, 4, and 5 illustrate the quantities and interim values in the most collusive SPE for \( b^H = b^L = 1, a^H = 5, \) and \( a^L = 1 \) and the same \( N, Q, \epsilon, \) and \( \epsilon' \) from the price collusion example.\(^{23}\)

When partial collusion is the most collusive outcome, just as for price competition, larger profits are possible after the signal \( h \) in world M and only smaller profits are possible after the signal \( l \) (the carrots are more widespread in Figure 3(a) than in 4(a)). Unlike price competition, the effects do not cancel out. For all \( \delta \) such that full collusion is not possible in both worlds, although the difference is small so Figure 5(b) is significantly zoomed, the interim values are strictly larger in world M. This is a consequence of CURV 1 holding strictly.

\(^{23}\)These parameter values for \( a^s \) and \( b^s \) make it clear that game \( H \) displays higher demand. Generally, the analysis works for any parameters although varying the slope coefficients has an ambiguous effect on demand.
Figure 4. Quantity Competition: Quantities for $\epsilon' = .25$.

Figure 5. Quantity Competition: Interim Values for $\epsilon = .05$ and $\epsilon' = .25$.

5. Conclusion

This paper examines the impact of public information in Markov games pertaining to the game in the next period. The results present a number of situations where less informative signals allow the players to obtain larger equilibrium payoffs.

There are some limiting assumptions made in the model, a few of which are worth mentioning. Most notably, players’ actions do not affect the Markov process. This assumption rules out the case where agents could use their actions to change the evolution of the Markov game when signals tell them it is evolving in an undesirable direction. However, a model where players’ actions do affect the Markov process must include an independent, exogenous process and signals can only be about this exogenous process so it becomes quite tricky. Second, two of the assumptions that
limit the kinds of static games that can be analyzed are that the sets of actions
are convex and compact and that there are Nash equilibrium actions which minmax
players. Though both these assumptions are only imposed for some of the results,
they definitely limit the analysis to specific types of Markov games, and would be
interesting to relax in future work.

One more limiting assumption is that the information in this paper is only about the
game in the following period. In a more general model of information, a classic puzzle
of repeated games can be addressed as a story of information. It is well-known that
more actions and payoffs are possible in indefinitely repeated games (an isomorphic
transformation of infinitely repeated games) than in finitely repeated games. This
result can be understood by modeling the indefinitely repeated game as a Markov
game with two stage games, namely the game that is being repeated and a null game
where no game is played, with the transition matrix

\[
Q = \begin{pmatrix}
\delta & 1 - \delta \\
0 & 1
\end{pmatrix}
\]

The indefinitely repeated game is this Markov game with no information whereas the
finitely repeated game is this Markov game with perfect information about the period
where the game transits to the null game.

Even though the results depend on many limiting assumptions, there are still many
examples to which they can be applied. For instance, agencies which collect and
disseminate information about the future of the economy may produce a positive externality
on anti-trust by providing high quality information. In other examples,
cooperative behavior is good for society, so agencies that control information may
consider withholding it.

**References**


---

24 Thanks to an anonymous referee for making this interesting observation.


Appendix

Law of Iterated Expectation:
The main role of information in the proof of Theorem 1 (and Theorem 3) is through the law of iterated expectation.

Proposition 5. Let \( a \) be an event, \( B \) and \( C \) be sets of events, and \( f(a, b, c) \) be a function. If \( \text{Prob}(c|b) = \text{Prob}(c|a, b) \) for all \( b \in B \) and \( c \in C \), then

\[
E_b[E_c[f(a, b, c)|b]|a] = E_{bc}[f(a, b, c)|a]
\]

Proof:

\[
E_b[E_c[f(a, b, c)|b]|a] = \sum_{b \in B} \text{Prob}(b|a) \sum_{c \in C} \text{Prob}(c|b) f(a, b, c)
\]

\[
= \sum_{b \in B} \text{Prob}(b|a) \sum_{c \in C} \text{Prob}(c|a, b) f(a, b, c)
\]

\[
= \sum_{b \in B, c \in C} \text{Prob}(b, c|a) f(a, b, c)
\]

\[
= E_{bc}[f(a, b, c)|a]
\]

Proposition 5 is used for 3 equations in the proof of Theorem 1.

For Equation (4): \( a = (s, y) \), \( B = (s, y) \times (Z^L)' \), and \( C = X^s \times X^{s'} \). Then, because \( a \) is a component of each \( b \), \( \text{Prob}(c|b) = \text{Prob}(c|a, b) \).\(^{25}\)

For Equation (5): \( a = (s, y) \), \( B = s \times X^s \), and \( C = (Z^M)' \). The necessary equality for \( \text{Prob}(c|b) = \text{Prob}(c|a, b) \) is that \( \text{Prob}(s', x'|s, x, y) = \text{Prob}(s', x'|s, x) \). \( x' \) depends only on \( s' \) so \( \text{Prob}(s', x'|s, x, y) = \text{Prob}(x'|s') \text{Prob}(s'|s, x, y) \) and \( \text{Prob}(s', x'|s, x) = \text{Prob}(x'|s') \text{Prob}(s'|s, x) \). Hence, it is sufficient to show that \( \text{Prob}(s'|s, x, y) = \text{Prob}(s'|s, x) \).

\(^{25}\)The notation is slightly different in Equation (4) (and (5) and (9)) because notation for summing over a singleton is omitted and \( f \) does not depend on \( a, b, \) and \( c \) so some objects sum out to 1.
This follows from Bayes rule and the “information identity:"

\[
\Pr(s'|s, x, y) = \frac{\Pr(x, y|s, s') \Pr(s'|s)}{\Pr(x, y|s)} = \frac{\Pr(y|s, x, s') \Pr(x|s, s') \Pr(s'|s)}{\Pr(y|s, x) \Pr(x|s)} = \frac{\Pr(x|s, s') \Pr(s'|s)}{\Pr(x|s)} = \Pr(s'|s, x)
\]

For Equation (9): \(a = s, B = s \times Y^s, \) and \(C = X^s.\) Again, because \(a\) is a component of each \(b, \) \(\Pr(c|b) = \Pr(c|a, b).\)

**Convex Equilibrium Value Sets:**

The curvature assumption guarantees that the sets of equilibrium values are convex which is used to show the feasibility of the enforcing continuation function in the proof of Theorem 1.

**Proposition 6.** Assume CURV 1 holds. Then \(V^M(\delta)\) and \(V^L(\delta)\) are convex sets for every discount factor \(\delta \in (0, 1).\)

**Proof:** See Online Appendix.

**Proof of Theorem 1:**

Fix any arbitrary \(\delta.\) The proof proceeds by constructing a set \(W^L\) that is self-generating in world \(L\) so, by Proposition 1, it follows that \(W^L \subseteq V^L(\delta)\) and therefore \(\overline{W^L} \subseteq \overline{V^L}(\delta).\) It is then shown that the set \(W^L\) also has the property that \(\overline{V^M}(\delta) = \overline{W^L}.\) Combining the latter two set equations completes the proof.

**Construction of \(W^L\)**

Let \(T : \mathbb{R}^{|Z^M|} \to \mathbb{R}^{|Z^L|}\) be the operator where, for each component \((s, y)\) and \(v \in \mathbb{R}^{Z^M},\)

\[
T_{sy}(v) = \mathbb{E}_x[v_{sx}|s, y]
\]
Note that such a conditional expectation is defined. Explicitly, it is
\[
\text{Prob}(x|s, y) = \frac{\text{Prob}(x|s) \text{Prob}(y|s, x)}{\sum_{s'} F_{s'x}^s Q_{s's'}^s P_{xy}^s}
\]

The set \( W^L \) is constructed by applying \( T \) to each SPE value in world \( M \):
\[
W^L = \{ T(v^M) : v^M \in V^M(\delta) \}
\]

There are three properties needed to show that \( W^L \) is self-generating; the value recursion, the incentive constraint, and the feasibility of the enforcing continuation functions.

The value recursion

Consider some arbitrary \( w^L \in W^L \) and arbitrary component \( w_{s'y}^L \). By construction, there exists some \( v^M \in V^M(\delta) \) such that \( w_{s'y}^L = T_{s'y}(v^M) = E_x[v_{sx}^M|s, y] \). Therefore, the value recursion amounts to constructing an enforcing action \( a_{s'y}^L \) and enforcing continuation values \( \gamma_{s'y}^L((a_{s'y}^L, \ldots, a_{s'y}^L), s', y') \), one for each \((s', y')\), such that
\[
(1) \quad (1 - \delta)\tilde{u}_s(a_{s'y}^L) + \delta E_{s'y}[\gamma_{s'y}^L((a_{s'y}^L, \ldots, a_{s'y}^L), s', y')|s, y] = E_x[v_{sx}^M|s, y]
\]

To save on notation, denote \( \gamma_{s'y}^L((a_{s'y}^L, \ldots, a_{s'y}^L), s', y') \) by simply \( \gamma_{s'y}^L(a_{s'y}^L, s', y') \) (and similarly in world \( M \)) for the rest of this proof. As \( V^M(\delta) \) is self-generating (by Proposition 1 again), each component \( v_{sx}^M \) of \( v^M \) can be decomposed by an enforcing action \( a_{sx}^M \) and enforcing continuation function \( \gamma_{sx}^M \) on \( V^M(\delta) \) at state \((s, x)\). Hence, by plugging in the value recursions that hold in world \( M \) (and simplifying some from the linearity of the expectation operator), the right hand side of Equation (1) is
\[
(2) \quad E_x[v_{sx}^M|s, y] = (1 - \delta)E_x[\tilde{u}_s(a_{sx}^M)|s, y] + \delta E_x[E_{s'x'}[\gamma_{sx}^M(a_{sx}^M, s', x')|s, x]|s, y]
\]

The construction of \( a_{sx}^L \) and \( \gamma_{sx}^L(a_{sx}^L, s', x') \) proceeds by equating each piece of the sum in Equations (1) and (2). So \( a_{sy}^L \) is chosen to satisfy the equation
\[
(3) \quad \tilde{u}_s(a_{sy}^L) = E_x[\tilde{u}_s(a_{sx}^M)|s, y]
\]

As the expectation is simply a convex combination and \( u^s \) is continuous (and therefore \( \tilde{u}_s \) is continuous), the Intermediate Value Theorem guarantees existence of such an action. The enforcing continuation value for state \((s', y')\) is constructed as
\[
\gamma_{sy}^L(a_{sy}^L, s', y') = E_{xx'}[\gamma_{sx}^M(a_{sx}^M, s', x')|s, x, y, s', y']
\]

\(^{26}\)Any signal \( y \) that does not occur with positive probability can be ignored.
Plugging the enforcing continuation values into Equation (1) and applying Proposition 5 yields
\[
E_{s'y'}[E_{xx'}[\gamma_{sx}^M(a_{sx}^M, s', x')|s, y, s', y']|s, y] = E_{xx'}[\gamma_{sx}^M(a_{sx}^M, s', x')|s, y]
\]

Alternatively, applying Proposition 5 to the right hand side of Equation (2) yields
\[
E_x[E_{s'x'}[\gamma_{sx}^M(a_{sx}^M, s', x')|s, x]|s, y] = E_{xs'}[\gamma_{sx}^M(a_{sx}^M, s', x')|s, y]
\]

As the right hand sides of Equations (4) and (5) are identical, Equation (1) holds.

The incentive constraint

For each \(a \in A\) with \(a \neq (a_{sy}^L, a_{sy}^L, \ldots, a_{sy}^L)\) and each state \((s', y')\), the rest of the enforcing continuation function \(\gamma_{sy}^L\) is constructed as
\[
\gamma_{sy}^L(a, s', y') \equiv E_{xx'}[\gamma_{sx}^M((a_{sx}^M, a_{sx}^M, \ldots, a_{sx}^M), s', x')|s, y, s', y']
\]

To save on notation, denote \(\gamma_{sy}^L((\hat{a}_{sy}^L, a_{sy}^L, \ldots, a_{sy}^L), s', y')\) by simply \(\gamma_{sy}^L(\hat{a}_{sy}^L, s', y')\) (and similarly in world \(M\)) for the rest of this proof. With the same argument as for Equations (4) and (5)
\[
E_{s'y'}[\gamma_{sy}^L(\hat{a}_{sy}^L, s', y')|s, y] = E_x[E_{s'x'}[\gamma_{sx}^M(\hat{a}_{sx}^M, s', x')|s, x]|s, y]
\]

By the construction of \(a_{sy}^L\) in Equation (3) and Assumption CURV 1 it follows that
\[
\hat{u}^s(a_{sy}^L) \leq E_x[\hat{u}^s(a_{sx}^M)|s, y]
\]

In summary, the incentive constraint follows as
\[
w_{sy}^L = E_x[v_{sx}^M|s, y]
\]
\[
\geq E_x[(1 - \delta)\hat{u}^s(a_{sx}^M) + \delta E_{s'x'}[\gamma_{sx}^M(a_{sx}^M, s', x')|s, x]|s, y]
\]
\[
= E_x[(1 - \delta)\hat{u}^s(a_{sx}^M)|s, y] + \delta E_{s'y'}[\gamma_{sy}^L(\hat{a}_{sy}^L, s', y')|s, y]
\]
\[
\geq (1 - \delta)\hat{u}^s(a_{sy}^L) + \delta E_{s'y'}[\gamma_{sy}^L(\hat{a}_{sy}^L, s', y')|s, y]
\]

The second line is the incentive constraints, one for each signal \(x\), that hold in world \(M\). The third line is Equation (6) and the fourth line is Equation (7).

Feasibility of the enforcing continuation functions
The proof is given for $\gamma_{Ls}^{Ls}(a_{Ls}^{Ls}, s', y')$ though the proof for the deviant actions is identical. The key is to split the conditional expectation in the construction of $\gamma_{Ls}^{Ls}(a_{Ls}^{Ls}, s', y')$ into its two arguments:

$$E_{xx'}[\gamma_{sx}^{Msx}(a_{sx}^{Msx}, s', x')|s, y] = E_{x'}[E_{x}[\gamma_{sx}^{Msx}(a_{sx}^{Msx}, s', x')|s, y]|x']$$

Equation (8) holds because $x'$ is a signal about two periods ahead so the current period state $(s, y)$ does not affect it (given that tomorrow’s game $s'$ is also being conditioned on).

Note that $\gamma_{sx}^{Msx}(a_{sx}^{Msx}, s', x') \in V_{sx}^{Msx}(\delta)$ for every signal $x$ because of the feasibility of the enforcing continuation functions in world M. Also, $V^{M}(\delta)$ is convex by Proposition 6 so $E_{x}[\gamma_{sx}^{Msx}(a_{sx}^{Msx}, s', x')|s, y] \in V_{sx}^{Msx}(\delta)$. Therefore, there exists $w^{M} \in V^{M}(\delta)$ (choose any SPE value such that $w^{M}_{s'y} = E_{x'}[\gamma_{sx}^{Msx}(a_{sx}^{Msx}, s', x')|s, y]$) with

$\gamma_{Ls}^{Ls}(a_{Ls}^{Ls}, s', y') = E_{x'}[w_{s'y}^{M}|s', y']$

As this is how $W_{s'y}^{Ls}$ is constructed, it follows that $\gamma_{Ls}^{Ls}(a_{Ls}^{Ls}, s', y') \in W_{s'y}^{Ls}$ as desired.

Equality of interim values

The final step is to show that $\overline{V}^{M}(\delta) = \overline{W}^{L}$. Note that, for any arbitray $w^{L} \in W^{L}$, the $v^{M} \in V^{M}(\delta)$ from which it was constructed, and $s$

$$\overline{w}^{L}_{s} = E_{y}[w^{L}_{s'y}|s] = E_{y}[E_{x}[v^{Msx}_{sx}|s, y]|s] = E_{x}[v^{Msx}_{sx}|s] = \overline{v}^{M}_{s}$$

The only part of this set of equalities that is not by construction or definition is the second to the last equality. It holds by application of the law of iterated expectation one more time. Therefore $\overline{w}^{L}_{s} = \overline{v}^{M}_{s}$ and so the proof is complete. \qed
Proof of Proposition 6:

It is shown here that $V^M(\delta)$ is convex though the proof for $V^L(\delta)$ is identical except that world $L$ notation is used. For notational simplicity, the superscript $M$ is dropped. The proof proceeds by showing that $V(\delta) = \text{co}(V(\delta))$ where $\text{co}$ stands for convex hull. Obviously, $V(\delta) \subseteq \text{co}(V(\delta))$ so the proof shows the other direction. By Proposition 1, it suffices to show that the $\text{co}(V(\delta))$ is a self-generating set.

Let $w \in \text{co}(V(\delta))$ and consider an arbitrary state $(s, x)$. By definition of the convex hull, there exists an integer $K$, values $v^k \in V(\delta)$, and weights $\alpha^k \in (0,1)$ summing to 1 such that

$$w_{sx} = \sum_{k=1}^{K} \alpha^k v^k_{sx}$$

The goal is to show that $w_{sx}$ is decomposable on $\text{co}(V(\delta))$ at state $(s, x)$. As $V(\delta)$ is self-generating (by Proposition 1 again) for each $v^k_{sx}$ there exists enforcing actions $a^k_{sx}$ and enforcing continuation values $\gamma^k_{sx}((a^k_{sx}, \ldots, a^k_{sx}), s', x')$, one for each state $(s', x')$, that decompose $v^k_{sx}$ on $V(\delta)$ at state $(s, x)$. Define a candidate $a_{sx}$ to decompose $w_{sx}$ by the equation

$$\tilde{u}^s(a_{sx}) = \sum_{k=1}^{K} \alpha^k \tilde{u}^s(a^k_{sx})$$

Such an action exists by the Intermediate Value Theorem which holds because $A$ is convex and $u^s$ (and therefore $\tilde{u}^s$) is continuous. For notational simplicity, denote $\gamma_{sx}((a_{sx}, \ldots, a_{sx}), s', x')$ by $\gamma_{sx}(a_{sx}, s', x')$ (and similarly for the $\gamma^k_{sx}$). Define candidates for the enforcing continuation values $\gamma_{sx}(a_{sx}, s', x')$, one for each $(s', x')$, by

$$\gamma_{sx}(a_{sx}, s', x') = \sum_{k=1}^{K} \alpha^k \gamma^k_{sx}(a^k_{sx}, s', x')$$

Finally, for all actions other than $a_{sx}$ by every player, when the next period state is $(s', x')$, the enforcing continuation function specifies the lowest SPE value in state $(s', x')$ which is denoted $w^s_{s'x'}$ (the lowest value is achievable due to well-known arguments that $V(\delta)$ is closed).

All that needs to be shown is that the candidates do indeed decompose $w_{sx}$. There are three parts to show; the value recursion, incentive constraint, and that the enforcing continuation function specifies feasible values. The value recursion follows immediately from the value recursions that hold for each $v^k_{sx}$ (and the linearity of the
\[ w_{sx} = (1 - \delta)\tilde{u}^s(a_{sx}) + \delta E_{s'x'}[\gamma_{sx}(a_{sx}, s', x')|s, x] \]
\[ = (1 - \delta)\sum_{k=1}^{K} \alpha^k \tilde{u}^s(a^k_{sx}) + \delta E_{s'x'} \left[ \sum_{k=1}^{K} \alpha^k \gamma^k_{sx}(a^k_{sx}, s', x') | s, x \right] \]
\[ = \sum_{k=1}^{K} \alpha^k v^k_{sx} \]

The incentive constraint follows from the incentive constraints that hold for each \( v^k_{sx} \) (the first inequality) and CURV 1 (the second inequality).

\[ w_{sx} = \sum_{k=1}^{K} \alpha^k v^k_{sx} \]
\[ \geq \sum_{k=1}^{K} (1 - \delta)\alpha^k \tilde{u}^s(a^k_{sx}) + \delta E_{s'x'}[v_{s'x'}|s, x] \]
\[ \geq (1 - \delta)\tilde{u}^s(a_{sx}) + \delta E_{s'x'}[v_{s'x'}|s, x] \]

By feasibility of the enforcing continuation functions for each \( v^k \), it follows that \( \gamma^k_{sx}(a^k_{sx}, s', x') \in V_{s'x'}(\delta) \) for every \( k \) and \( (s', x') \) so the convex combination \( \gamma_{sx}(a_{sx}, s', x') \) is an element of \( \text{co}V_{s'x'}(\delta) \). Since \( \text{co}(V_{s'x'}(\delta)) = \text{co}(V(\delta))_{s'x'} \) it follows that \( \gamma_{sx}(a_{sx}, s', x') \) is a feasible continuation value. For all \( a \) other than \( a_{sx} \) by every player, the enforcing continuation function specifies \( v_{s'x'} \) which is feasible by definition. \( \square \)

For the rest of the online appendix, two more pieces of commonly used notation are employed. First, a history is called an on-path history if it occurs with positive probability. Second, a superscript is used to denote finite sequences of signals across time. For example, \( x^t = (x_0, \ldots, x_t) \).

**Proposition 7.** Fix a player \( i \) and \( K \) strategies \( \{\sigma^{M,k}\}_{k=1}^{K} \) for world \( M \) and one strategy \( \sigma^L \) for world \( L \). Let \( \alpha^k \in (0, 1) \) for each \( k \) be numbers that sum to 1 in total. If for all periods \( t \), on-path histories \( h^{M,t} \in \mathcal{H}^M \) and \( h^{L,t} \in \mathcal{H}^L \), and states

---

\[ ^{27} \text{The linearity of the expectation operator is used frequently for simplifications in the rest of the appendix. For succinctness, this is the last time it will be explicitly mentioned.} \]
$(s_t, x_t) \in Z^M$ and $(s_t, y_t) \in Z^L$

$$u^s_t(\sigma^L(h^{L,t}, s_t, y_t)) \geq \sum_{k=1}^K \alpha^k E_{x_t}[u^s_t(\sigma^{M,k}(h^{M,t}, s_t, x_t))|h^{L,t}, s_t, y_t]$$

Then, for every (period 0) state $(s, y)$,

$$v^L_{i,sy}(\sigma^L) \geq \sum_{k=1}^K \alpha^k E_x[v^M_{i,sx}(\sigma^{M,k})|s, y]$$

Equality of first relationship for all $h^{L,t}, s_t, y_t$, implies equality of second.

**Proof of Proposition 7:**

By noting how values are calculated (see Section 2.2) it is sufficient to show that for every period $t$,

$$E_{h^{L,t}, s_t, y_t}[u^s_t(\sigma^L(h^{L,t}, s_t, y_t))|s, y] \geq \sum_{k=1}^K \alpha^k E_x[E_{h^{M,t}, s_t, x_t}[u^s_t(\sigma^{M,k}(h^{M,t}, s_t, x_t))|s, x]|s, y]$$

For the left hand side of this inequality, plug in the asserted inequality in the proposition and simplify using the law of iterated expectations (which follows here similarly to Equation (4) above).

$$E_{h^{L,t}, s_t, y_t}[u^s_t(\sigma^L(h^{L,t}, s_t, y_t))|s, y] \geq E_{h^{L,t}, s_t, y_t} \left[ \sum_{k=1}^K \alpha^k E_x[u^s_t(\sigma^{M,k}(h^{M,t}, s_t, x_t))|h^{L,t}, s_t, y_t]|s, y \right]$$

$$= \sum_{k=1}^K \alpha^k E_{h^{M,t}, s_t, x_t}[u^s_t(\sigma^{M,k}(h^{M,t}, s_t, x_t))|s, y]$$

For the right hand side, the law of iterated expectation (which follows here similarly to Equation (5) above by noting that all objects in periods greater than 0 cancel here just as $x'$ did in that case) yields the desired equality for each $k$.

$$E_x[E_{h^{M,t}, s_t, x_t}[u^s_t(\sigma^{M,k}(h^{M,t}, s_t, x_t))|s, x]|s, y] = E_{h^{M,t}, s_t, x_t}[u^s_t(\sigma^{M,k}(h^{M,t}, s_t, x_t))|s, y]$$

The only inequality is the asserted inequality so if it holds at equality, then the values are equal as well. \qed
Proof of Theorem 2:
First, it is shown that $\delta^{M,*}$ exists (and the proof of $\delta^{L,*}$ is identical). Consider any $(s, x) \in Z^M$. Due to the stationarity of the grim trigger strategy, the value $v^{M,*}_{sx}$ can be written recursively as

$$v^{M,*}_{sx} = (1 - \delta)\tilde{u}^s(a^*_s) + \delta E_{s'x'}[v^{M,*}_{s'x'}|s, x]$$

The incentive constraints to support $v^{M,*}$ as a SPE value are, for each $(s, x)$,

$$v^{M,*}_{sx} \geq (1 - \delta)\hat{u}^s(a^*_s) + \delta E_{s'x'}[v^{M,NE}_{s'x'}|s, x]$$

For the following, it is easier to consider the rearranged incentive constraint

$$(11) \quad (1 - \delta)(\hat{u}^s(a^*_s) - \tilde{u}^s(a^*_s)) \leq \delta E_{s'x'}[v^{M,*}_{s'x'} - v^{M,NE}_{s'x'}|s, x]$$

Now, because deviations from the Nash equilibrium every period are never optimal, the grim trigger strategy is a SPE iff Equation (11) holds for every state $(s, x)$.

Note that both sides of Equation (11) are weakly positive. The left hand side is obviously so and $\tilde{u}^s(a^*_s) \geq \hat{u}^s(a^{NE}_s)$ for every $s$ so the right hand side is an infinite polynomial in $\delta$ with (weakly) positive coefficients.

There are three cases to consider. First, if $\hat{u}^s(a^*_s) = \tilde{u}^s(a^*_s)$ the left hand side is zero for every $\delta$ so Equation (11) holds for every $\delta$. Second, if $\hat{u}^s(a^*_s) > \tilde{u}^s(a^*_s)$ and $\hat{u}^s(a^*_s) = \tilde{u}^s(a^{NE}_s)$ for every $s$ that occurs in some future period with positive probability (so the right hand side is 0 for all $\delta$), then Equation (11) never holds.

The third case is where both sides are strictly positive. Note that then the left hand side is continuous, strictly decreasing, and converges to 0 as $\delta \to 1$ while the right hand side is continuous and strictly increasing (the polynomial has at least one strictly positive coefficient) so there exists $\delta^{M,*}_{sx} < 1$ such that Equation (11) holds iff $\delta \geq \delta^{M,*}_{sx}$.

As Equation (11) must hold for every state, letting $\delta^{M,*}_{sx} = 0$ if the first case holds and $\delta^{M,*}_{sx} = 1$ if the second case holds, $\delta^{M,*}$ is simply the maximum of the $\delta^{M,*}_{sx}$.

For the weak inequality, by Proposition 7 with $K = 1$, because $\sigma^L(h^{L,t}, s_t, y_t) = \sigma^M(h^{M,t}, s_t, x_t) = a^{*}_{s_t}$ for on-path histories and $a^{NE}_{s_t}$ for on-path histories when the Nash equilibria played, two equations hold.

$$(12) \quad E_{s'y'}[v^{L,*}_{s'y'}|s, y] = E_x[E_{s'x'}[v^{M,*}_{s'x'}|s, x]|s, y]$$
\[ E_{s'x'}[v_{s'y'}^{L,NE}|s,y] = E_x[E_{s'x'}[v_{s'x'}^{M,NE}|s,x]|s,y] \]

If \( \delta^{M,*} = 1 \), the result is immediate so suppose that \( \delta^{M,*} < 1 \) and let \((s,y)\) be any state. Consider a signal \( \hat{x} \in X^s \) such that
\[
\hat{x} \in \arg\min_{\hat{x} \in X^s} E_{s'x'}[v_{s'y'}^{M,*} - v_{s'y'}^{M,NE}|s,\hat{x}] \]

As \( \delta^{M,*} < 1 \), Equation (11) evaluated at \( \delta^{M,*} \) holds for state \((s,x)\). Also, as \( x \) minimizes the difference in continuation values it follows that it is less than the conditional expectation of differences:
\[
E_{s'x'}[v_{s'y'}^{M,*} - v_{s'y'}^{M,NE}|s,x] \leq E_x[E_{s'x'}[v_{s'y'}^{M,*} - v_{s'y'}^{M,NE}|s,\hat{x}]|s,y] \]

Combining Equations (11) for state \((s,x)\) and (15) yields
\[
(1 - \delta^{M,*})(\hat{u}^*(a^*_s) - \bar{u}^*(a^*_s)) \leq \delta^{M,*} E_x[E_{s'x'}[v_{s'y'}^{M,*} - v_{s'y'}^{M,NE}|s,x]|s,y] \]

Combing Equations (12), (13), and (16) yields
\[
(1 - \delta^{M,*})(\hat{u}^*(a^*_s) - \bar{u}^*(a^*_s)) \leq \delta^{L,*} E_{s'y'}[v_{s'y'}^{L,*} - v_{s'y'}^{L,NE}|s,y] \]

Of course, Equation (17) is Equation (11) for state \((s,y)\) so \( \delta^{L,*} \leq \delta^{M,*} \). As \((s,y)\) was arbitrary, \( \delta^{L,*} \leq \delta^{M,*} \).

For the strict inequality, first assume that the Markov game is Non-Trivial in world L and Future Informative with respect to \( \delta^{M,*} \) in world M. By Non-Triviality, Cases 1 and 3 from the first part of the proof are not true so \( \delta^{M,*} \in (0,1) \). Therefore, as Equation (11) is continuous in \( \delta \) it suffices to show that Equation (17) holds strictly since it would therefore hold for a slightly smaller \( \delta \) than \( \delta^{M,*} \). It therefore suffices to show that if Equation (11) is an equality for state \((s,x)\), then (15) is a strict inequality.

Note that the \( \text{Prob}(x|s,y) > 0 \) iff \( R_{xy}^s > 0 \) (see the part of the proof of Theorem 1 about the construction of \( W^L \)). Therefore, if Equation (11) is an equality for state \((s,\hat{x})\) at \( \delta^{M,*} \), then Future Informativeness implies that Equation (15) is strict as there exists \( \hat{x} \) with \( \text{Prob}(\hat{x}|s,y) > 0 \) and
\[
E_{s'x'}[v_{s'y'}^{M,*} - v_{s'y'}^{M,NE}|s,\hat{x}] > E_{s'x'}[v_{s'y'}^{M,*} - v_{s'y'}^{M,NE}|s,\hat{x}] \]

The proof in the other direction is by contrapositive. Assume that the Markov game is either not Non-Trivial or not Future Information with respect to \( \delta^{M,*} \). If it
is not Non-Trivial in regards to the first part, then \( \delta^{M,*} = \delta^{L,*} = 0 \) because Case 1 from above is true for every state. If it is not Non-Trivial in regards to the second part, then \( \delta^{L,*} = 1 \), because Case 2 from above is true for some state in world L. Also, \( 1 = \delta^{L,*} \leq \delta^{M,*} \leq 1 \) implies that \( \delta^{M,*} = 1 \).

If it is not Future Informativeness, it suffices to show that Equation (17) is an equality for some state \((s,y)\) because then Equation (11) will not hold for any smaller \( \delta \) than \( \delta^{M,*} \).

Not Future Informativeness means that there exists some state \((s,x)\) where Equation (11) holds with equality for \( \delta^{M,*} \) and there exists some \( y \in Y^s \) such that for every \( \hat{x} \in X^s \) with \( R^{s,y}_{\hat{x}} > 0 \)
\[
(18) \quad E_{s'x'}[v^{M,s}_{s'x'} - v^{M,NE}_{s'x'} | s, \hat{x}] = E_{s'x'}[v^{M,s}_{s'x'} - v^{M,NE}_{s'x'} | s, x]
\]

Because Equation (11) holds with equality at \( \delta^{M,*} \), it must be that \( x \) can be taken as \( x \) from Equation (14). This is because the left hand side of Equation (11) is constant in the signal, so a state where \( \delta^{M,*} = \delta^{M,*} \) must be one with the smallest right hand side. Equation (18) implies that Equation (15) holds with equality.

\( \square \)

**Proof of Theorem 3:**

Fix \( \delta \in (0,1) \). Let \( \sigma^M \) be a SPE in world M. The proof constructs a SPE strategy \( \sigma^L \) with \( v^L \geq v^M \).

For every on-path history \( h^{L,t} \in H^{L,t} \) (the construction builds up from period 0 to determine the on-path histories for each subsequent period), player \( i \), and period \( t \) state \((s_t, y_t)\) construct \( \sigma^L_t \) as follows.
\[
(19) \quad \sigma_i^L(h^{L,t}, s_t, y_t) = E_{x_t}[\sigma_i^M(h^{M,t}, s_t, x_t)|h^{L,t}, s_t, y_t]
\]

For all off-path histories, let \( \sigma_i^L(h^{L,t}, s_t, y_t) = a_i^{NE} \) (the Nash action specified in the theorem). It is clearly not profitable to deviate from the Nash actions forever, so only on-path histories must be checked for incentives. For the rest of the proof, only on-path histories are considered and the term on-path is dropped to keep the exposition concise.
Letting $v^{L,NE}$ be the value of the Nash actions forever, by the one-shot deviation principle, subgame perfection requires that for every history $h$ and state $(s, y)$:

$$v^L_{i,sy}(\sigma^L|_h) \geq (1 - \delta)h^s(\sigma^L_{i,t}(h,s,y)) + \delta E_{s'y'}[v^{L,NE}_{i,s'y'}|s,y]$$

(20)

Let $h^{L,t}$ be any arbitrary history in world $L$. By Equation (19), for any other history $h^{L,\tau}$ and state $(s_\tau, y_\tau)$,

$$\sigma^L_{i}|_{h^{L,t}}(h^{L,\tau}, s_\tau, y_\tau) = E_{x^{t+\tau}}[\sigma^M_{i}|_{h^{M,t}}(h^{M,\tau}, s_\tau, x_\tau) | (h^{L,t}, h^{L,\tau}), s_\tau, y_\tau]$$

(21)

The $x$ signals only depend on the current period and following period games. Hence, (where $s_t$ is the period 0 game in $h^{L,\tau}$ or $s_\tau$ if $h^{L,\tau} = \emptyset$),

$$\text{Prob}(x^{t-1}|(h^{L,t}, h^{L,\tau}), s_\tau, y_\tau, x_\tau) = \text{Prob}(x^{t-1}|h^{L,t}, s_t)$$

and

$$\text{Prob}(x_\tau| (h^{L,t}, h^{L,\tau}), s_\tau, y_\tau) = \text{Prob}(x_\tau | h^{L,\tau}, s_\tau, y_\tau)$$

Putting these together with Equation (21)

$$\sigma^L_{i}|_{h^{L,t}}(h^{L,\tau}, s_\tau, y_\tau) = E_{x^{t-1}}[E_{x^t}[\sigma^M_{i}|_{h^{M,t}}(h^{M,\tau}, s_\tau, x_\tau) | h^{L,\tau}, s_\tau, y_\tau] | h^{L,t}, s_t]$$

(22)

Because $u^s_{i,s}$ is concave due to CURV 2,

$$u^s_{i,s}(\sigma^L|_{h^{L,t}}(h^{L,\tau}, s_\tau, y_\tau)) = u^s_{i,s}(E_{x^{t-1}}[E_{x^t}[\sigma^M_{i}|_{h^{M,t}}(h^{M,\tau}, s_\tau, x_\tau) | h^{L,\tau}, s_\tau, y_\tau] | h^{L,t}, s])$$

$$\geq E_{x^{t-1}}[E_{x^t}[u^s_{i,s}(\sigma^M_{i}|_{h^{M,t}}(h^{M,\tau}, s_\tau, x_\tau)) | h^{L,\tau}, s_\tau, y_\tau] | h^{L,t}, s]$$

This means that Proposition 7 with the strategies $\sigma^{M,k} = \sigma^M|_{h^{M,t}}$ for each $h^{M,t}$ (there are a finite number of on-path histories) with corresponding weights $\alpha^k = \text{Prob}(x^{t-1}|h^{L,t}, s)$ applies. That is, for every state $(s, y)$,

$$v^L_{i,sy}(\sigma^L|_{h^{L,t}}) \geq E_{x^{t-1}}[E_{x}[v^M_{i,x,s}(\sigma^M|_{h^{M,t}})|s, y] | h^{L,t}, s]$$

(23)

The bulk of the rest of the argument follows from the incentives of $\sigma^M$ in world $M$ (the second inequality below). The convexity portion of CURV 2 yields the third inequality (note the relationship between $\sigma^L|_{h^{L,t}}$ and the strategies $\sigma^M|_{h^{M,t}}$ to see that it applies). The first equality holds because $v^M_{s'y'}$ does not depend on $x^{t-1}$ while the

\[28\] As a caution, note that vertical lines are used for both continuation strategies and conditional probability.
second has already been shown in the proof of Theorem 2.

\[ v^L_{i,sy}(\sigma^L|_{h^{L,t}}) \geq E_{x^{t-1}}[E_x[v^M_{i,sx}(\sigma^M|_{h^{M,t}})|s,y]|h^{L,t},s] \]

\[ \geq (1-\delta)\hat{u}^L(\sigma^L_{-1}(h^{L,t},s,y)) + \delta E_{x^{t-1}}[E_x[v^M_{i,sx'}(\sigma^M,NE)|s,x'|s,y]|h^{L,t},s] \]

\[ = (1-\delta)\hat{u}^L(\sigma^L_{-1}(h^{L,t},s,y)) + \delta E_{x}[E_{s'y'}[v^M_{i,sx'}(\sigma^M,NE)|s,x'|s,y]|h^{L,t},s] \]

As Equation (20) holds, \( \sigma^L \) is a SPE strategy. Finally note that, letting \( h^{L,t} = \emptyset \), it has been shown above that for any period 0 state \((s,y)\) and player \( i \)

\[ v^L_{i,sy} \geq E_x[v^M_{i,sx}|s,y] \]

Finally, following the same logic behind Equation (9) in the proof of Theorem 1,

\[ v^L_{i,s} \geq v^M_{i,s} \]

**Proof of Proposition 2:**

Let \( a, a_1, \ldots, a_n \in \tilde{A} \) and weights \( \lambda_1 \ldots \lambda_n \in [0,1] \) summing to 1 satisfy

\[ \tilde{u}(a) = \sum_{k=1}^{n} \lambda_k \tilde{u}(a_k) \]

By concavity of \( \tilde{u} \),

\[ \tilde{u}(a) = \sum_{k=1}^{n} \lambda_k \tilde{u}(a_k) \leq \tilde{u} \left( \sum_{k=1}^{n} \lambda_k a_k \right) \]

If both functions are increasing then

\[ a < \sum_{k=1}^{n} \lambda_k a_k \quad \text{and} \quad \hat{u}(a) \leq \hat{u} \left( \sum_{k=1}^{n} \lambda_k a_k \right) \]

If both functions are decreasing then

\[ a > \sum_{k=1}^{n} \lambda_k a_k \quad \text{and} \quad \hat{u}(a) \leq \hat{u} \left( \sum_{k=1}^{n} \lambda_k a_k \right) \]
In either case, the same conclusion follows. By convexity of \( \hat{u} \),
\[
\hat{u}(a) \leq \hat{u} \left( \sum_{k=1}^{n} \lambda_k a_k \right) \leq \sum_{k=1}^{n} \lambda_k \hat{u}(a_k)
\]

Proof of Proposition 3:
The Nash Equilibrium of each stage game is \( p = 0 \) which is also the minmax payoff (any firm can guarantee a profit of 0 by choosing a price of 0) so it is sufficient to assume that all deviations are followed by \( p = 0 \) forever by all firms.

Moving to on-path behavior, suppose the current state is \((s, x)\) and the strategy specifies profit \( \Pi_{sx} \) per firm today and some expected continuation value, denoted \( CV_{sx} \). A firm that deviates captures the entire market today and gets 0 profit in all future periods, so the incentive constraint is

\[
(1 - \delta)\Pi_{sx} + \delta CV_{sx} \geq (1 - \delta) N \Pi_{sx}
\]

The incentive constraint illustrates why the profit in the most collusive equilibrium is stationary; \( \Pi = (\Pi_{Hh}, \Pi_{Hl}, \Pi_{Lh}, \Pi_{Ll}) \), a profit for each state to be agreed on given the current state. Suppose for contradiction that it were not. Then consider the stationary strategy that, for each state, prescribes the profit that maximizes profit over all periods in that state. The maximum profit was incentive compatible in the period in which it was prescribed. The only difference between Equation (24) in the new and original strategy is that \( CV_{sx} \) is larger in the stationary profile, because all future profits are at least as large. So the stationary profile is incentive compatible and delivers a higher payoff, and thus the contradiction is reached.\(^{29}\)

Given stationarity, the value \( v \) can be written recursively as the vector equation
\[
v = (1 - \delta)\Pi + \delta P v
\]

In closed form, where the inverse matrix is known to exist due to standard arguments for Markov games,
\[
v = (1 - \delta)(I - \delta P)^{-1}\Pi
\]

The incentive constraint for the grim trigger strategy to be a SPE is therefore
\[
I - \delta P)^{-1}\Pi \geq N \Pi
\]

\(^{29}\)If the maximum does not exist, the supremum can be used. As incentives are weak inequalities, this poses no problem.
While the maximization problem for the most collusive equilibrium consists of maximizing four distinct profits, the solution is unique, because increasing profits in any state increases the continuation value in all states and therefore only helps with incentives in other states.

Example: A Markov game that is not Future Informative

The Markov game has two players and three stage games, labeled A, B, and C, and the payoff matrices for the three games are

<table>
<thead>
<tr>
<th></th>
<th>Game A</th>
<th>Game B</th>
<th>Game C</th>
</tr>
</thead>
<tbody>
<tr>
<td>a₁</td>
<td>1,1</td>
<td>2,2</td>
<td>3,3</td>
</tr>
<tr>
<td>a₂</td>
<td>-1,2</td>
<td>-1,3</td>
<td>-1,4</td>
</tr>
<tr>
<td>b₁</td>
<td>2,2</td>
<td>3,1</td>
<td>4,1</td>
</tr>
<tr>
<td>b₂</td>
<td>0,0</td>
<td>0,0</td>
<td>0,0</td>
</tr>
</tbody>
</table>

The Markov game is equally likely to be in any of the stage games initially, \( q₀(A) = (1/3, 1/3, 1/3) \), and the states are completely iid so the transition matrix is

\[
Q = \begin{pmatrix}
    1/3 & 1/3 & 1/3 \\
    1/3 & 1/3 & 1/3 \\
    1/3 & 1/3 & 1/3
\end{pmatrix}
\]

There are two public signals in world M, \( x₁ \) and \( x₂ \). \( x₁ \) suggests that the game in the following period is equally likely to be A or C while \( x₂ \) suggests that the game in the following period is B. There is just one public signal in world L, \( y \), which therefore says nothing. Formally, for each game \( s \)

\[
F^s = \begin{pmatrix}
    1 & 0 \\
    0 & 1 \\
    1 & 0
\end{pmatrix}, \quad G^s = \begin{pmatrix}
    1 \\
    1 \\
    1
\end{pmatrix}, \quad R^s = \begin{pmatrix}
    1 \\
    1
\end{pmatrix}
\]

The maximum feasible payoff is obtained by choosing \((a₁, a₁)\) when the current period game is A, \((b₁, b₁)\) when the current period game is B, and \((c₁, c₁)\) when the current period game is C.

The main point of the example is that the signals in world M are informative, but they both predict exactly the same loss from a deviation. For the first signal, players miss out on the gain in A which is 1 or C which is 3, which on average is 2. For the second signal, players miss out on the gain from B which is also 2.

To formalize, consider A and signal \( x₁ \). Following the strategy yields 1 today, is equally likely to get 1 or 3 tomorrow, and is equally likely to get 1, 2, or 3 in all
future periods. The value of deviating is 2 today and 0 in the future so the incentive constraint is

\[(1 - \delta) + \delta[(1 - \delta)(1/2 + 3(1/2)) + \delta(1/3 + 2(1/3) + 3(1/3))] \geq (1 - \delta)^2\]

The deviation payoff is obviously the same for signals \(x_2\) and \(y\). By construction, the values are the same too. For signals \(x_2\) and \(y\) respectively they are

\[(1 - \delta) + \delta[(1 - \delta)2 + \delta(1/3 + 2(1/3) + 3(1/3))]\]

\[(1 - \delta) + \delta[(1 - \delta)(1/3 + 2(1/3) + 3(1/3)) + \delta(1/3 + 2(1/3) + 3(1/3))]\]

All three incentives constraints are the same. Solving for \(\delta\), \(\delta^{M,*} = \delta^{L,*} = 1/3\) so strictness fails.