Federal Reserve Bank of Minneapolis
Research Department

October 19, 2010

Beliefs and Private Monitoring*

Christopher Phelan

University of Minnesota
Federal Reserve Bank of Minneapolis

Andrzej Skrzypacz

Graduate School of Business
Stanford University

ABSTRACT

This paper develops new recursive, *set based* methods for studying repeated games with private monitoring. For any *finite state* strategy profile, we find necessary and sufficient conditions for whether there exists a distribution over initial states such that the strategy, together with this distribution, form a correlated sequential equilibrium. Also, for any given correlation device for determining initial states (including degenerate cases where players' initial states are common knowledge), we provide necessary and sufficient conditions for the correlation device and strategy to be a correlated sequential equilibrium, or in the case of a degenerate correlation device, for the strategy to be a sequential equilibrium. We also consider several applications. In these, we show that the methods are computationally feasible, show how to construct and verify equilibria in a secret price-setting game and show that with private monitoring, stick-and-carrot strategies cannot support equilibrium in a classic repeated duopoly game.

*The authors thank V. Bhaskar, Peter DeMarzo, Glenn Ellison, Larry Jones, Narayana Kocherlakota, David Levine, George Mailath, Stephen Morris, Ichiro Obara, Larry Samuelson, Itai Sher, Ofer Zeitouni, seminar participants at the Federal Reserve Bank of Minneapolis, the Harvard/MIT joint theory seminar, Stanford University, Iowa State University, Princeton University, the University of Chicago, the University of Minnesota, University College London, the London School of Economics and the 2006 meetings of the Society for Economic Dynamics and three anonymous referees for helpful comments as well as the excellent research assistance of Songzi Du, Kenichi Fukushima and Roozbeh Hosseini. Financial assistance from National Science Foundation Grant # 0721090 is gratefully acknowledged. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1. Introduction

This paper develops new methods for studying repeated games with private monitoring. In particular, we develop tools that allow us to answer when a particular strategy is consistent with equilibrium. For an important subclass of strategies - those which can be represented as finite automata - we provide readily checkable and computable necessary and sufficient conditions for equilibrium.

The importance of these methods is as follows: While checking the equilibrium conditions in *public* monitoring games and perfect public equilibria is relatively simple, for games with *private* monitoring, for almost all strategies, checking the equilibrium conditions has previously been considered difficult if not impossible. For instance, consider the following repeated game with private monitoring taken from Malaith and Morris (2002): Two partners, privately, either cooperate or defect, and in each period each, privately, has either a good or bad outcome. While each player can neither observe his partner's action, nor his partner's outcome, outcomes are correlated: the vector of joint outcomes is a probabilistic function of the vector of joint actions. (A player cooperating makes it more likely that both players have a good outcome.)

At issue is that even for the simplest games, such as the one presented above, and even the simplest strategies, such as tit-for-tat, there are an infinite number of possible histories where incentives must be checked, and to check incentives one must calculate beliefs for all of them. (This difficulty is not confined to the example above. See for example the work of Kandori (2002) and Mailath and Samuelson (2006), Chap. 12.) In this paper, for a very large class of strategies, we resolve this issue by showing the necessity and sufficiency of checking incentives only for "extreme beliefs" (as opposed to checking incentives for all possible histories).

The focus of our analysis is strategies which can be represented by a finite automaton (*finite state strategies*). A key point (first made by Mailath and Morris (2002)) is that if all players' strategies are finite automata, a particular player's private history is relevant only to the extent that it gives him information regarding the private states of his opponents. This lets us summarize a player's history as a *belief* over a finite state space, a much smaller object than the belief over the private histories of opponents (a point also made by Mailath and Morris (2002)). Moreover, unlike the set of possible private *histories*, the set of possible private *states* for one's opponents does not grow over time.

While many private histories may put a player in the same state of his automaton, they will, in general, induce different beliefs regarding the state of his opponents. Given this, there are two advantages to working with *sets* of beliefs representing all possible beliefs a player can have in a given private state. One is that it is necessary and sufficient to check incentives only for extreme points of those sets instead of looking at beliefs after all histories. The other advantage is that these sets can be readily calculated using recursive methods (operators from sets to sets) that we describe and demonstrate computationally.

Fixed points of our main set based operator represent the beliefs a player can have regarding his opponents' states "in the long run." We show that if incentives hold for extreme points of these sets, one can always use an initial correlation device to, in effect, start the game off as if it had been already running for a long time.[1] This technique alleviates a fundamental difficulty associated with games with private monitoring: the continuation of (sequential)

[1] An earlier version of this project entitled "Private Monitoring with Infinite Histories" focused on this point.

equilibrium play in a game with private monitoring is not a sequential equilibrium, but rather a correlated equilibrium in which private histories function as the correlation device. But as Kandori (2002) notes, the correlation device becomes increasingly more complex over time. Using randomization or exogenous correlation in period 0 of the game to make it easier to satisfy incentives and hence support an equilibrium, has been suggested by Sekiguchi (1997), Compte (2002), and Ely (2002). We present a robust way of applying this method to construct a family of Correlated Sequential Equilibria.

Our main results are presented as follows. In Section 2, we present our model, a standard repeated game with private monitoring with finiteness and full support (all signals seen with positive probability) as its only restrictive assumptions. We also present the subclass of strategies we study — *finite state strategies*, or strategies which can be represented as finite automata.

In Section 3, we show a necessary and sufficient condition for a given correlation device (choosing initial states of players) and a profile of finite automata to form a *Correlated Sequential Equilibrium (CSE)* (Theorem 1). That condition involves checking incentive constraints on a fixed point of our set operator (based on Bayes' rule) which we describe how to compute. Next, we show necessary and sufficient conditions for the *existence of a starting correlation device* such that if coupled with a given automaton they form a CSE (Theorem 2) - they involve checking incentives at extreme beliefs of fixed point of a related operator. The result implies that the best hope for incentives to hold is to start the players *as if* the game has been played for a long time (without telling them what the outcomes were, but only in which state they should be now). We also show how to verify which starting conditions can support a CSE and which not. Since we can apply these results to arbitrary correlation devices, and in particular, to degenerate ones, we can answer if a particular strategy profile is a *sequential* equilibrium — a correlated equilibrium with a degenerate correlation device.

In Section 4 we present a few applications of our methods. We start with the partnership game described above and demonstrate that the methods are easy to apply computationally and that allows us to gain new intuitions regarding how private monitoring affects incentives. In the second application we use the methods analytically to show non-robustness of stick-and-carrot strategies in *any generic* private monitoring version of the classic repeated Cournot game of Abreu, Pearce and Stacchetti (1986). In the third application we consider tacit collusion in a duopoly with competition in prices (with private prices and quantities) and show that one-period price wars are more robust to private monitoring than two-period price wars. The last application focuses on strategies which are not action-free (a class of strategies rarely studied in the literature).

In Section 5. we address the question if particular data are consistent with the model and equilibrium play by showing how to extend the methods to study correlated Nash equilibria. Unfortunately, even once we have the set of relevant beliefs for which it is necessary and sufficient to check incentives, verifying that there are no profitable deviations is not trivial. We show that one can achieve partial answers using a class of deviations that follow the recommended strategy until some history and deviate for one period only. By an example we demonstrate that this method in some cases offers definite answers. In Section 6 we conclude.

Our results complement the existing literature on the construction of belief-free equilibria (for example, the work of Ely and Välimäki (2002), Piccione (2002), Ely, Hörner and Olszewski (2005), and Kandori and Obara (2006)), in which players use mixed strategies and their best responses are independent of their beliefs about the private histories of their opponents. In contrast to belief-free equilibria, the equilibria we construct are belief-dependent; players' best responses do depend on their beliefs. (For earlier work on constructing belief-dependent sequential equilibria, see Bhaskar and Obara (2002) and Mailath and Morris (2002). The first paper constructs a particular equilibrium for an almost-perfect monitoring prisoner's dilemma game. The second describes a class of finite-monitoring equilibria in almost-public monitoring games).

In terms of the focus on strategies instead of payoffs, our work is closest to Mailath and Morris (2002) and (2006). They consider robustness of particular classes of strategies - those that are equilibria in a public monitoring game - to a perturbation of the game from public to private, yet almost-public monitoring. They show that strict equilibria in strategies which look back only a finite number of periods (a subclass of the strategies we study), are robust to such perturbations. They also show when infinite-history dependent strategies (partly covered by our analysis) are not robust. Our methods allow one to extend their analysis beyond almost-public monitoring games - as the application in Section 4B suggests.

Finally, in a recent paper Kandori (2008) studies equilibria he calls "Weakly Belief-Free" and shows that in some games they can achieve higher payoffs than any belief-free equilibrium. The definition of these equilibria can be translated to our language as follows: incentive constraints have to hold for initial beliefs and for all extreme beliefs obtained after *one iteration* of our operator on the set of *all* possible beliefs (in contrast, the belief-free equilibria check incentives for zero iterations and our CSE check them after infinite many iterations).

2. The Model

Consider the game, Γ^∞, defined by the infinite repetition of a stage game, Γ, with N players, $i = 1, \ldots, N$, each able to take actions $a_i \in A_i$. Assume that with probability $P(y|a)$, a vector of private outcomes $y = (y_1, \ldots, y_N)$ (each $y_i \in Y_i$) is observed conditional on the vector of private actions $a = (a_1, \ldots, a_N)$, where for all (a, y), $P(y|a) > 0$ (*full support*). Further assume that $A = A_1 \times \ldots A_N$ and $Y = Y_1 \times \ldots Y_N$ are both finite sets, and let $H_i = A_i \times Y_i$.

The current period payoff to player i is denoted $u_i : H_i \to R$. That is, player i's payoff is a function of his own current-period action and private outcome. If player i receives payoff stream $\{u_{i,t}\}_{t=0}^{\infty}$, his lifetime discounted payoff is $(1 - \beta) \sum_{t=0}^{\infty} \beta^t u_{i,t}$ where $\beta \in (0, 1)$. As usual, players care about the expected value of lifetime discounted payoffs.

Let $h_{i,t} = (a_{i,t}, y_{i,t})$ denote player i's private action and outcome at date $t \in \{0, 1, \ldots\}$, and $h_i^t = (h_{i,0}, \ldots, h_{i,t-1})$ denote player i's private history up to, but not including, date t. A (behavior) *strategy* for player i, $\sigma_i = \{\sigma_{i,t}\}_{t=0}^{\infty}$, is then, for each date t, a mapping from player i's private history h_i^t, to his probability of taking any given action $a_i \in A_i$ in period t. Let σ denote the joint strategy $\sigma = (\sigma_1, \ldots, \sigma_N)$ and σ_{-i} denote the joint strategy of all players other than player i, or $\sigma_{-i} = (\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \ldots, \sigma_N)$. (Throughout the paper we use notation $-i$ to refer to all players but player i.)

A. Finite State Strategies

In this paper, we restrict attention to equilibria in *finite state strategies*, or strategies which can be described as finite automata. (However, we allow deviation strategies to be unrestricted). A finite state strategy for player i is defined by four objects: 1) a finite private state space Ω_i (with D_i elements ω_i), 2) a function $p_i(a_i|\omega_i)$ giving the probability of each action a_i for each private state $\omega_i \in \Omega_i$, 3) a deterministic transition function $\omega_i^+ : \Omega_i \times H_i \to \Omega_i$ determining next period's private state as a function of this period's private state, player i's private action a_i, and his private outcome y_i, and 4) an initial state, $\omega_{i,0}$.[2] Given this setup, $\sigma_{i,0}(a_i) = p_i(a_i|\omega_{i,0})$, $\sigma_{i,1}(a_{i,0}, y_{i,0})(a_i) = p_i(a_i|\omega_i^+(\omega_{i,0}, a_{i,0}, y_{i,0}))$ and so on.[3]

Throughout the paper, we repeatedly make a distinction between a finite state strategy's *automaton* (objects 1 through 3) and object 4, player i's initial state, $\omega_{i,0}$. Let $\psi_i = (\Omega_i, p_i, \omega_i^+)$ denote agent i's automaton. The collection of automata over all players $\psi \equiv \{\psi_1, \ldots, \psi_N\}$ is referred to as the *joint automaton*. Finally, let the number of joint states $D = \Pi_{i \leq N} D_i$, and the number of joints states for players other than player i, $D_{-i} = \Pi_{j \neq i} D_j$.

B. Beliefs

Since our solution concept will be *Correlated Sequential Equilibrium*, allow player i's initial beliefs over the initial state of his opponents, $\omega_{-i,0}$, to be possibly nondegenerate. In particular, let player i's beliefs about the initial state of his opponents, $\mu_{i,0}$, be a point in the $(D_{-i} - 1)$-dimensional unit-simplex, denoted $\Delta^{D_{-i}}$. Taking as given $\mu_{i,0}$, the assumption of full support ($P(y|a) > 0$ for all (a, y)) implies that the beliefs of player i regarding his opponents' private histories, h_{-i}^t, are also always pinned down by Bayes' rule. But since the continuation strategies of players $-i$ depend only on their current joint state, $\omega_{-i,t}$, to verify player i's incentive constraints after any given private history h_i^t, we need not directly consider player i's beliefs regarding $\omega_{-i,0}$ and h_{-i}^t. Instead, we need focus only on player

[2]The restriction to deterministic transitions is for notational convenience only. All of our methods and results apply to automata with non-deterministic transitions.

[3] For a useful discussion of the validity of representing strategies as finite state automata in the context of games with private monitoring, see Mailath and Morris (2002) and Mailath and Samuelson (2006).

i's beliefs regarding his opponents' current state, $\omega_{-i,t}$. This is a much smaller object, and, importantly, its dimension does not grow over time.

For a particular initial belief, $\mu_{i,0}$, and private history, h_i^t, player i's belief over $\omega_{-i,t}$ is, like $\mu_{i,0}$, simply a point in the $(D_{-i}-1)$-dimensional unit-simplex. Let $\mu_{i,t}(\mu_{i,0}, h_i^t)$ denote player i's belief at the beginning of period t about $\omega_{-i,t}$ after private history h_i^t given initial beliefs $\mu_{i,0}$. Let $\mu_{i,t}(\mu_{i,0}, h_i^t)(\omega_{-i})$ denote the probability assigned to the particular state ω_{-i}.

Beliefs $\mu_{i,t}(\mu_{i,0}, h_i^t)$ can be defined recursively using Bayes' rule. Let $B_i(m_i, h_i|\psi_{-i}) \in \Delta^{D_{-i}}$ denote the belief of player i over the state of his opponents at the beginning of period t, if his beliefs over his opponents' state at period $t-1$ were $m_i \in \Delta^{D_{-i}}$ and he subsequently observed $h_i = (a_i, y_i)$. This posterior belief can be written out explicitly (from Bayes' rule) as:

$$B_i(m_i, h_i|\psi_{-i})(\omega'_{-i}) = \frac{\sum_{\omega_{-i}} m_i(\omega_{-i}) H_i(\omega_{-i}, \omega'_{-i}, h_i|\psi_{-i})}{\sum_{\omega_{-i}} m_i(\omega_{-i}) F_i(\omega_{-i}, h_i|\psi_{-i})}$$

where

$$\begin{aligned} F_i(\omega_{-i}, h_i|\psi_{-i}) &= \sum_{(a_{-i}, y_{-i})} p_{-i}(a_{-i}|\omega_{-i}) P(y_i, y_{-i}|a_i, a_{-i}), \\ H_i(\omega_{-i}, \omega'_{-i}, h_i|\psi_{-i}) &= \sum_{h_{-i} \in G_{-i}(\omega_{-i}, \omega'_{-i}|\psi_{-i})} p_{-i}(a_{-i}|\omega_{-i}) P(y_i, y_{-i}|a_i, a_{-i}) \end{aligned}$$

and

$$G_{-i}(\omega_{-i}, \omega'_{-i}|\psi_{-i}) = \{h_{-i} = (a_{-i}, y_{-i}) | \omega^{+}_{-i}(\omega_{-i}, a_{-i}, y_{-i}) = \omega'_{-i}\}$$

or G_{-i} is the set of (a_{-i}, y_{-i}) pairs which cause players $-i$ to transit from state ω_{-i} to state

ω'_{-i}.

To define beliefs recursively, let $B_i^s(m_i, h_i^s|\psi_{-i}) = B_i(B_i^{s-1}(m_i, h_i^{s-1}|\psi_{-i}), h_{i,s-1}|\psi_{-i})$ where $B_i^1(m_i, h_i|\psi_{-i}) = B_i(m_i, h_i|\psi_{-i})$. Then, $\mu_{i,t}(\mu_{i,0}, h_i^t) = B_i^t(\mu_{i,0}, h_i^t|\psi_{-i})$. Note that $B_i(m_i, h_i|\psi_{-i})$ does not depend on σ_i at all, and thus player i's beliefs are the same regardless of whether or not player i is playing a finite state strategy.

C. Equilibrium

Consider player i following an arbitrary strategy σ_i while players $-i$ follow a finite state strategy σ_{-i} defined by $(\omega_{-i,0}, \psi_{-i})$. That is, players $-i$ are restricted to finite state strategies, but player i is not. Let $V_{i,t}(h_i^t, \omega_{-i}|\sigma_i, \psi_{-i})$ denote the lifetime expected discounted payoff to player i conditional on his private history h_i^t, and players $-i$ being in state ω_{-i}. Thus,

$$\begin{aligned} V_{i,t}(h_i^t, \omega_{-i}|\sigma_i, \psi_{-i}) \;\; = \;\; & \sum_{a=(a_i,a_{-i})} (\sigma_{i,t}(h_i^t)(a_i) p_{-i}(a_{-i}|\omega_{-i})) \Big(\sum_y P(y|a)[(1-\beta)u_i(a_i, y_i) + \\ & \beta V_{i,t+1}((h_i^t, (a_i, y_i)), \omega_{-i}^{+}(\omega_{-i}, a_{-i}, y_{-i})|\sigma_i, \psi_{-i})] \Big). \end{aligned}$$

For arbitrary beliefs $m_i \in \Delta^{D_{-i}}$, let

$$EV_{i,t}(h_i^t, m_i|\sigma_i, \psi_{-i}) = \sum_{\omega_{-i}} m_i(\omega_{-i}) V_{i,t}(h_i^t, \omega_{-i}|\sigma_i, \psi_{-i}).$$

Player i's expected payoff given correct beliefs $\mu_{i,t}(\mu_{i,0}, h_i^t)$ is then $EV_{i,t}(h_i^t, \mu_{i,t}(\mu_{i,0}, h_i^t)|\sigma_i, \psi_{-i})$.

If σ_i is a finite state strategy (defined by $(\omega_{i,0}, \psi_i)$), let $\omega_{i,t}(\omega_{i,0}, h_i^t)$ denote the private state for player i at date t implied by initial state $\omega_{i,0}$, transition rule $\omega_i^{+}(\omega_i, a_i, y_i)$, and history $h_i^t = ((a_{i,0}, y_{i,0}), \ldots, (a_{i,t-1}, y_{i,t-1}))$. Then, for all $(h_i^t, \hat{h}_i^t)$ such that $\omega_{i,t}(\omega_{i,0}, h_i^t) = \omega_{i,t}(\omega_{i,0}, \hat{h}_i^t)$,

$V_{i,t}(h_i^t, \omega_{-i}|\sigma_i, \psi_{-i}) = V_{i,t}(\hat{h}_i^t, \omega_{-i}|\sigma_i, \psi_{-i})$. Given this, we can write player i's lifetime payoff, conditional on ω_{-i}, as a function of his current private state ω_i as opposed to depending directly on his private history, h_i^t. Thus we define $v_i(\omega_i, \omega_{-i}|\psi_i, \psi_{-i}) \equiv V_{i,t}(h_i^t, \omega_{-i}|\sigma_i, \psi_{-i})$ for any h_i^t such that $\omega_i = \omega_{i,t}(\omega_{i,0}, h_i^t)$. Then we denote player i's expected payoff, now a function of his current state, ω_i, and his beliefs over his opponents' state, ω_{-i}, as

$$Ev_i(\omega_i, m_i|\psi_i, \psi_{-i}) = \sum_{\omega_{-i}} m_i(\omega_{-i}) v_i(\omega_i, \omega_{-i}|\psi_i, \psi_{-i}).$$

Definition 1. *A probability distribution over initial states , $x \in \Delta^D$, and joint automaton, ψ, form a Correlated Sequential Equilibrium (CSE) of Γ^∞ if for all i, t, h_i^t, $\omega_{i,0}$ such that $\sum_{\omega_{-i,0}} x(\omega_{i,0}, \omega_{-i,0}) > 0$, and arbitrary $\hat{\sigma}_i$,*

$$Ev_i(\omega_{i,t}(\omega_{i,0}, h_i^t), \mu_{i,t}(\mu_{i,0}(x, \omega_{i,0}), h_i^t)|\psi_i, \psi_{-i}) \geq EV_{i,t}(h_i^t, \mu_{i,t}(\mu_{i,0}(x, \omega_{i,0}), h_i^t)|\hat{\sigma}_i, \psi_{-i})$$

where $\mu_{i,0}(x, \omega_{i,0})(\omega_{-i,0}) = x(\omega_{i,0}, \omega_{-i,0}) / \sum_{\overline{\omega}_{-i,0}} x(\omega_{i,0}, \overline{\omega}_{-i,0})$.

There are two difficulties in verifying whether a given (x, ψ) for a CSE. First, there are infinitely many deviation strategies. Second, to verify the IC constraints we need to know the beliefs players have on and off path after each element of the infinite set of possible private histories. The first difficulty is shared by all repeated game models and, as usual, it is solved by using the one-shot deviation principle. The resolution of the second difficulty is the main focus of this paper.

Lemma 1. *(One-shot Deviation Principle) Suppose a correlation device x and joint automa-*

ton ψ satisfy for all i, h_i^t, $\hat{a}_i$, and $\omega_{i,0}$ such that $\sum_{\omega_{-i,0}} x(\omega_{i,0}, \omega_{-i,0}) > 0$,

$$
\begin{aligned}
&Ev_i(\omega_{i,t}(\omega_{i,0}, h_i^t), \mu_{i,t}(\mu_{i,0}(x, \omega_{i,0}), h_i^t)|\psi_i, \psi_{-i}) \geq \\
&\sum_{\omega_{-i}} \mu_{i,t}(\mu_{i,0}(x, \omega_{i,0}), h_i^t)(\omega_{-i})[\sum_{a_{-i}} p_{-i}(a_{-i}|\omega_{-i}) \sum_{y} P(y|\hat{a}_i, a_{-i}) \\
&[(1-\beta)u_i(\hat{a}_i, y_i) + \beta v_i(\omega_i^+(\omega_{i,t}(\omega_{i,0}, h_i^t), \hat{a}_i, y_i), \omega_{-i}^+(\omega_{-i}, a_{-i}, y_{-i})|\psi_i, \psi_{-i})]].
\end{aligned}
$$

Then, (x, ψ) form a CSE. That is, it is sufficient to check that player i does not wish to deviate once and then revert to playing according to his automaton ψ_i.

Proof. Mailath and Samuelson (2006), page 397. ∎

3. Verifying Equilibria

We now turn to the main methodological contribution of the paper: set based methods delivering first, necessary and sufficient conditions for when a joint automaton ψ, when coupled with *particular* correlation device x, forms a Correlated Sequential Equilibrium (Theorem 1), and second, necessary and sufficient conditions for whether there exists *any* correlation device x such that a joint automaton ψ, when coupled with x, forms a CSE. That is, our second main result (Theorem 2) regards whether the joint automaton ψ itself is consistent with equilibrium.

Rather than considering separately the beliefs $m_i \in \Delta^{D_{-i}}$ that a player will have after some private history, it is useful to consider *sets* of beliefs. In particular, let $M_i(\omega_i) \subset \Delta^{D_{-i}}$ denote a closed, convex, set of beliefs, and M_i be a collection of D_i sets $M_i(\omega_i)$, one for each ω_i. Let $\mathcal{M}$ denote the space of such collections of sets M_i. To define the distance between two elements M_i and $M_i' \in \mathcal{M}$, first let the distance between two beliefs m_i

and $m_i' \in \Delta^{D_{-i}}$ be defined by the sup norm (or Chebyshev distance) denoted $|m_i, m_i'| = \max_{\omega_{-i}} |m_i(\omega_{-i}) - m_i'(\omega_{-i})|$. Next, for a belief m_i and a non-empty closed set $A \subset \Delta^{D_{-i}}$ the distance between them (the Hausdorff distance) be defined as $|m_i, A| = \min_{m_i' \in A} |m_i, m_i'|$. For two non-empty, closed sets $(A, A') \subset \Delta^{D_{-i}}$, the Hausdorff distance between them is defined as $|A, A'| = \max\left\{\max_{m_i \in A} |m_i, A'|, \max_{m_i' \in A'} |m_i', A|\right\}$. If A is non-empty let $|A, \emptyset| = 1$ and $|\emptyset, A| = 1$. Finally, let $|\emptyset, \emptyset| = 0$. (Note that for non-empty A and A', $|A, A'| \leq 1$.) Then the distance between two collections of belief sets M_i, $M_i' \in \mathcal{M}$ is defined as $|M_i, M_i'| = \max_{\omega_i} |M_i(\omega_i), M_i'(\omega_i)|$.

We begin by constructing two related operators from $\mathcal{M}$ to $\mathcal{M}$ where fixed points of these operators will be a focus of our main results. Let the one-step operator $T(M_i)$ be defined as[4]

$$T(M_i) = \{T(M_i)(\omega_i') | \omega_i' \in \Omega_i\}$$

where

$$\begin{aligned} T(M_i)(\omega_i') &= \text{co}(\{m_i' | \text{ there exists } \omega_i \in \Omega_i, m_i \in M_i(\omega_i) \text{ and } (a_i, y_i) \in G_i(\omega_i, \omega_i' | \psi_i) \\ &\qquad \text{such that } m_i' = B_i(m_i, a_i, y_i | \psi_{-i})\}), \end{aligned}$$

where co() denotes the convex hull and $G_i(\omega_i, \omega_i' | \psi_i)$ is the set of (a_i, y_i) such that $\omega_i^+(\omega_i, a_i, y_i) = \omega_i'$. The T operator works as follows: Suppose one takes as given the sets of "allowable" beliefs player i can have over the private state of the other players, ω_{-i}, last period. For any given

[4]The T operator depends on Ψ_{-i} and varies across players (as does $\mathcal{M}$), but to conserve notation we write $T(M_i)$ rather than $T_i(M_i | \Psi_{-i})$.

such allowable belief, Bayesian updating then implies what player i should believe about ω'_{-i} this period for each realization of (a_i, y_i), generating a collection of allowable belief sets. That is, if there exists a way to choose player i's state last period, ω_i, the beliefs of player i over the private states of his opponents last period consistent with $m_i \in M_i(\omega_i)$, and a new realization of (a_i, y_i) such that Bayesian updating delivers beliefs m'_i, then $m'_i \in T(M_i)(\omega_i^+(\omega_i, a_i, y_i))$. In effect, the T operator gives, for a particular collection of belief sets M_i, the belief sets associated with all possible successor beliefs generated by new data and interpreted through σ_{-i} (as well as all convex combinations of such beliefs). Note that since B_i and G_i depend only on the joint automaton ψ, as opposed to starting conditions, x, the T operator retains the property as well.

Next, let the operator $T^U(M_i)$ (^{U} for union) be:

$$T^U(M_i) \quad = \quad \{T^U(M_i)(\omega_i) | \omega_i \in \Omega_i\} \text{ where } T^U(M_i)(\omega_i) = \text{co}(T(M_i)(\omega_i) \cup M_i(\omega_i)).$$

In words, the T^U operator calculates for every state ω_i, the convex hull of the union of the prior beliefs player i could hold last period, $M_i(\omega_i)$, and all the posterior beliefs he can hold in that same state, $T(M_i)(\omega_i)$.

We note here that the T and T^U operators are relatively easy to operationalize. In particular, the following lemma implies that the extreme points of the collection of sets $T(M_i)$ and $T^U(M_i)$ can be calculated using only the extreme points of the collection of sets M_i.

Lemma 2. *If $M_i(\omega_i)$ is closed and convex for all ω_i, then $T(M_i)(\omega_i)$ and $T^U(M_i)(\omega_i)$ are both closed and convex for all ω_i. Next, if m_i is an extreme point of $T^U(M_i)(\omega_i)$ but not $T(M_i)(\omega_i)$, then m_i is an extreme point of $M_i(\omega_i)$. Finally, if m_i is an extreme point of* *both $T(M_i)(\omega_i)$ and $T^U(M_i)(\omega_i)$ then there exists $\hat{m}_i$, $\hat{\omega}_i$, h_i such that $m_i = B_i(\hat{m}_i, h_i|\psi_{-i})$, $h_i \in G_i(\hat{\omega}_i, \omega_i|\psi_i)$ and $\hat{m}_i$ is an extreme point of $M_i(\hat{\omega}_i)$.*

Proof. See Appendix. ∎

A. Fixed Points of T and T^U

Our results rely on properties of the fixed points of T and T^U. We write $M_i^0 \subset M_i^1$ if $M_i^0(\omega_i) \subset M_i^1(\omega_i)$ for all ω_i. Furthermore, we write M_i is non-empty if there exists a private state ω_i such that $M_i(\omega_i)$ is non-empty.

Both T and T^U are monotonic operators (that is, if $M_i^0 \subset M_i^1$, then $T(M_i^0) \subset T(M_i^1)$ and $T^U(M_i^0) \subset T^U(M_i^1)$). By construction, $M_i \subset T^U(M_i)$ for all $M_i \in \mathcal{M}$. Since $M_i \subset T^U(M_i)$, and $T^U(M_i) \subset T^U(T^U(M_i))$ (from monotonicity), the sequence $\{M_i, T^U(M_i), T^U(T^U(M_i)), \ldots\}$ converges. That B_i is continuous implies T^U is continuous and thus this limit is a fixed point of T^U. Call this fixed point $M_i^{*U}(M_i)$. Next note that if $M_i \subset T(M_i)$, then $T(M_i) = T^U(M_i)$. This implies if $M_i \subset T(M_i)$, the sequence $\{M_i, T(M_i), T(T(M_i)), \ldots\}$ also converges to $M_i^{*U}(M_i)$.

B. When is a pair (x, ψ) a Correlated Sequential Equilibrium?

For an arbitrary correlation device, x, let the belief sets $M_{i,0}(x, \omega_i) \in \Delta^{D_i}$ be defined such that

$$M_{i,0}(x, \omega_i) = \{\mu_{i,0}(x, \omega_i)\}$$

for all ω_i such that $\sum_{\omega_{-i}} x(\omega_i, \omega_{-i}) > 0$. Otherwise, let $M_{i,0}(x, \omega_i) = \emptyset$. That is, for all ω_i, if ω_i occurs with positive probability under distribution x, $M_{i,0}(x, \omega_i)$ is the single point belief set consisting of what player i believes about ω_{-i} when his initial state is ω_i. Let $M_{i,0}(x)$ be a collection of D_i sets $M_{i,0}(x, \omega_i)$, one for each ω_i, and (with some abuse of notation) $M_i^{*U}(x) \equiv M_i^{*U}(M_{i,0}(x))$.

Theorem 1. *A correlation device x and a joint automaton ψ form a Correlated Sequential Equilibrium if and only if the incentive compatibility conditions*

$$
\begin{aligned}
Ev_i(\omega_i, m_i|\psi_i, \psi_{-i}) \quad \geq \quad & \sum_{\omega_{-i}} m_i(\omega_{-i})[\sum_{a_{-i}} p_{-i}(a_{-i}|\omega_{-i}) \sum_y P(y|\hat{a}_i, a_{-i}) \\
& [(1-\beta)u_i(\hat{a}_i, y_i) + \beta v_i(\omega_i^+(\omega_i, \hat{a}_i, y_i), \omega_{-i}^+(\omega_{-i}, a_{-i}, y_{-i})|\psi_i, \psi_{-i})]]
\end{aligned} \tag{1}
$$

hold for all i, $\hat{a}_i$, ω_i and m_i such that m_i is an extreme point of $M_i^{*U}(x)$.

Proof. If: Since incentive compatibility conditions (1) are linear in beliefs, then if they hold for the extreme beliefs of $M_i^{*U}(x)$, they hold for all beliefs in these sets. By monotonicity, $(T^U)^t(M_{i,0}(x)) \subset M_i^{*U}(x)$ for all $t \geq 0$, so incentives hold in the first period for all initial signals, and in all subsequent periods for all possible continuation histories.

Only if: Suppose that incentive compatibility conditions (1) are violated for some state ω_i and extreme belief $m_i \in M_i^{*U}(x)(\omega_i)$. Since the incentive conditions (1) are continuous in beliefs and are weak inequalities, there exists an $\varepsilon > 0$ such that for all beliefs m_i' such that $|m_i', m_i| < \varepsilon$, incentives are violated in state ω_i with beliefs m_i'.

Now, by definition of T^U, for every t and ω_i, every extreme point of $(T^U)^t(M_{i,0}(x))(\omega_i)$ is either an extreme point of $(T^U)^{t-1}(M_{i,0}(x))(\omega_i)$ or an extreme point of $T\left((T^U)^{t-1}(M_{i,0}(x))\right)(\omega_i)$. Therefore, we can find an initial state $\omega_{i,0}$ and a private history h_i^t such that player i after h_i^t is in state ω_i and his beliefs $\mu_{i,t}(\mu_{i,0}, h_i^t)$ satisfy $|\mu_{i,t}(\mu_{i,0}, h_i^t), m_i| < \varepsilon$ (using that $(T^U)^n(M_{i,0}(x)) \to M^{*U}(x)$). Thus (x, ψ) are not a CSE. ∎

C. When does there exist an x such that (x, ψ) is a Correlated Sequential Equilibrium?

For a joint automaton $\psi = (\Omega, p, \omega^+)$, denote the Markov transition matrix on the joint state $\omega \in \Omega$ by

$$(2) \quad \tau(\omega, \omega')(\psi) = \sum_{(a,y) \text{ s.t. } (a_i,y_i)\in G_i(\omega_i,\omega_i'|\psi_i) \text{ for all } i} P(y|a) \prod_i p_i(a_i|\omega_i).$$

Since $\tau(\psi)$ defines a finite state Markov chain, it has at least one invariant distribution, $\pi \in \Delta^D$.

Lemma 3. *Let π be an invariant distribution of the Markov process $\tau(\psi)$. Then for all i, $M_{i,0}(\pi) \subset T(M_{i,0}(\pi))$.*

Proof. See Appendix. ∎

The basic idea behind the proof of Lemma 3 is that beliefs drawn from an invariant distribution are an average, and thus a convex combination, of beliefs which condition on additional information. Since the T operator is the convex hull of all possible posteriors from given priors, and the average posterior belief is the prior belief, the convex hull of the set of possible posterior beliefs must contain the prior belief. Lemma 3 then implies that $T(M_{i,0}(\pi)) = T^U(M_{i,0}(\pi))$ and that $\{M_{i,0}(\pi), T(M_{i,0}(\pi)), T(T(M_{i,0}(\pi))), \ldots\}$ converges to $M_i^{*U}(\pi)$.

Lemma 4. *For a given (x, ψ) let $\pi = \lim_{t\to\infty} \frac{1}{t+1} \sum_{n=0}^{t} x\tau(\psi)^n$. Then*

a) The limit exists and is an invariant distribution of $\tau(\psi)$.

b) $M_{i,0}(\pi) \subset M_i^{*U}(x)$.

Proof. See Appendix. ∎

Part b) of Lemma 4 states that the *initial* beliefs player i can have if initial states are drawn from the invariant distribution of $\tau(\psi)$ defined in part a) (the sets $M_{i,0}(\pi)$) are always contained in the set of *all* beliefs player i can have over all dates when starting with the arbitrary correlation device x, $M_i^{*U}(x)$. The intuition of Lemma 4 is similar to Lemma 3: the beliefs $M_{i,0}(\pi)$ correspond to drawing initial states from a random time from the Markov chain $\tau(\psi)$ and hence are a convex combination of beliefs that condition both on calendar time and the realized history, which in turn are contained in $M_i^{*U}(x)$.

Theorem 2. *For a given joint automaton, ψ, there exists a correlation device x such that (x, ψ) form a CSE if and only if for some invariant distribution π of $\tau(\psi)$, incentives hold (i.e. condition (1) from Theorem 1) for all i, ω_i and m_i which is an extreme point of $M_i^{*U}(\pi)(\omega_i)$.*[5]

Proof. If: Let $x = \pi$. From Lemma 3 (and the monotonicity of T), the time zero beliefs of each player i, $M_{i,0}(\pi, \omega_{i,0}) \in M_i^{*U}(\pi)(\omega_{i,0})$ for each $\omega_{i,0}$ drawn with positive probability. Moreover, the subsequent beliefs for each player i are elements of $M_i^{*U}(\pi)(\omega_{i,t})$ for each date t and private history h_i^t, where $\omega_{i,t}$ is player i's state at date t after private history h_i^t.

Suppose condition (1) holds, for all i, $\hat{a}_i$, ω_i, and extreme points of $M_i^{*U}(\pi)(\omega_i)$, where m_i and $\hat{m}_i$ are two such points. Then since (1) is linear in these beliefs, for all $\alpha \in [0, 1]$, condition (1) holds for beliefs $\alpha m_i + (1 - \alpha)\hat{m}_i$, again for all i, $\hat{a}_i$, and ω_i. Thus incentives hold for all dates t and private histories h_i^t, if initial states are drawn according to π.

Only if. Suppose there exists a correlation device x such that (x, ψ) form a CSE, but

[5]The authors thank an anonymous referee for correctly suggesting that one of our sufficient conditions from a previous version of this paper — that incentives hold for all extreme points of $M_i^{*U}(\pi)$ for an invariant distribution π — was most likely also necessary.

for all invariant distributions π of $\tau(\psi)$, (π, ψ) does not form a CSE. That (x, ψ) forms a CSE implies, by Theorem 1, that incentives hold for all i, ω_i and m_i which is an extreme point of $M_i^{*U}(x)$. Let

$$\pi = \lim_{t \to \infty} \frac{1}{t} \sum_{n=0}^{t-1} x\tau^n$$

By Lemma 4, π is an invariant distribution of $\tau(\psi)$ and $M_{i,0}(\pi) \subset M_i^{*U}(x)$.

Since T^U is a monotone operator:

$$\left(T^U\right)^n \left(M_{i,0}(\pi)\right) \subset \left(T^U\right)^n \left(M_i^{*U}(x)\right) = M_i^{*U}(x)$$

and so in the limit:

$$M_i^{*U}(\pi) \subset M_i^{*U}(x)$$

Applying Theorem 1, this implies that (π, ψ) is also a CSE, a contradiction. ∎

D. Strategies with Unique Invariant Distributions

In the previous section, we showed that a joint automaton ψ is consistent with equilibrium if and only if it is a CSE to have initial private states drawn from an invariant distribution of $\tau(\psi)$. Verifying for a particular invariant distribution π of $\tau(\psi)$ whether (π, ψ) form a CSE then involves calculating $M_i^{*U}(\pi) \equiv \lim_{s\to\infty} T^s(M_{i,0}(\pi))$ and checking incentives at its extreme points. A second method involves calculating $\overline{M}_i \equiv \lim_{s\to\infty} T^s(\overline{\Delta}_i)$ (where $\overline{\Delta}_i$ denotes the collection of D_i, $D_{-i} - 1$-dimensional unit simplexes) and checking incentives at its extreme points. Since the set inclusion relationship, $\subset$, defines a complete lattice on the space of D_i closed subsets of $\Delta^{D_{-i}}$, $\overline{M}_i$ is the largest fixed point of T and all other fixed points of T are subsets of it (by Tarski's fixed point theorem). Thus if incentives hold at the extreme points of $\overline{M}_i$ (for all i), or incentives hold at the extreme points of *any* point in the sequence $\{T^s(\overline{\Delta}_i)\}_{s=0}^{\infty}$, (π, ψ) is a CSE for *any* invariant distribution π of $\tau(\psi)$. But this only establishes a sufficient condition for equilibrium. Here we show that if $\tau(\psi)$ is a *regular matrix* (i.e., there exists an s such that $\tau(\psi)^s$ has all non-zero entries) then incentives holding at the extreme points of $\overline{M}_i$ is necessary as well.

Lemma 5. *Suppose $\tau(\psi)$ is a regular matrix. Then $\overline{M}_i$ is the unique non-empty fixed point of T and for all non-empty $M_i \in \mathcal{M}$, $\lim_{n\to\infty} T^n(M_i) = \overline{M}_i$.*

Proof. See Appendix. ∎

Corollary 1. (of Theorem 2) *If $\tau(\psi)$ is a regular matrix, then there exists a correlation device x such that (x, ψ) form a CSE if and only if incentives hold (i.e. condition (1) from Theorem 1) for all i and m_i such that m_i is an extreme point of $\overline{M}_i$.*

Proof. Lemma 5, Lemma 3 and that for all M_i such that $M_i \subset T(M_i)$, $T(M_i) = T^U(M_i)$ imply $M_i^{*U}(\pi) = \overline{M}_i$, where π is the unique invariant distribution of $\tau(\psi)$. Theorem 2 then implies the result. ∎

E. Which starting conditions work?

For a given joint automaton ψ, Theorem 2 gives us necessary and sufficient conditions for the existence of a correlation device x such that (x, ψ) form a CSE. Suppose we find a ψ that satisfies these conditions. A natural question is then, what x can be used to start the strategies without violating incentive constraints? From the proof of Theorem 2 we know that at least one of the invariant distributions of $\tau(\psi)$ can be used.

One can use Theorem 1 to verify for any x, whether (x, ψ) is a CSE. That requires computing a fixed point of T^U for every such x. We now show that one can compute once a fixed point of a related operator and use it to evaluate any x.

In particular, define $M_i^I(\omega_i)$ to be the set of beliefs such that incentives hold in the current period for all beliefs $m_i \in M_i^I(\omega_i)$ if player i is in state ω_i and plans to follow the strategy in the future. Clearly, a necessary condition for (x, ψ) to be a CSE is that $M_{i,0}(x) \subset M_i^I$ since otherwise incentives would be violated in the first period. We need to insure, however, that incentives are satisfied not only for a particular belief generated by the correlation device, but also for all possible successors of that belief, and successors of those beliefs, and so on.

Define the operator $T^I(M_i)$ (I for incentives) as

$$
\begin{aligned}
T^I(M_i) &= \left\{T^I(M_i)(\omega_i) | \omega_i \in \Omega_i\right\} \text{ where} \\
(3) \quad T^I(M_i)(\omega_i) &= \text{co}(\{m_i | m_i \in M_i(\omega_i) \text{ and for all } (a_i, y_i), \\
&\qquad B_i(m_i, a_i, y_i | \psi_{-i}) \in M_i\left(\omega^+(\omega_i, a_i, y_i)\right)\}).
\end{aligned}
$$

In words, T^I eliminates an element of $M_i(\omega_i)$ if there exists a private history (a_i, y_i) and a successor belief which is not in $M_i(\omega_i^+(\omega_i, a_i, y_i))$.

Clearly, T^I is monotone and $T^I(M_i) \subset M_i$ for any M_i. Thus the sequence $\{(T^I)^n(M_i^I)\}_{n=0}^\infty$ (starting with the set of beliefs such that incentives hold in the first period), represents a sequence of (weakly) ever smaller collection of sets, guaranteeing that the limit, denoted M_i^{*I}, exists. Importantly, M_i^{*I} can be computed independently of x, allowing us to then evaluate all correlation devices to this benchmark:

Corollary 2. (of Theorem 1) *A correlation device x and a joint automaton ψ form a Correlated Sequential Equilibrium if and only if for all i, $M_{i,0}(x) \subset M_i^{*I}$*

Proof. For any M_i, by the definition of T^I, we have

$$M_i \subset M_i^{*I} \Longleftrightarrow M_i^{*U}(M_i) \subset M_i^{*I}$$

hence by Theorem 1, (x, ψ) form a CSE if and only if $M_{i,0}(x) \subset M_i^{*I}$. ∎

Since the set of correlated equilibria is convex, if (x, ψ) and (x', ψ) are CSE, so is (x'', ψ) for any x'' which is a convex combination of x and x'. Finally, for belief-free equilibria (such as those in Ely and Välimäki (2002)), the conditions of the corollary hold automatically since $M_i^{*I} = \overline{\Delta}_i$, or that incentives hold, by construction, for all beliefs.

4. Applications

In this section we attempt to show that these methods are useful in analyzing interesting economic applications. In it, we first analyze two strategies for the partnership game of Mailath and Morris (2002). Here we show that this game can be (relatively easily) analyzed computationally: particular strategies for particular game parameters can be shown to be (or not to be) equilibria. In our second application, we use our methods to *analytically* analyze the optimal bang-bang equilibrium of Abreu, Pearce, and Stacchetti (1986) and show it is generically never an equilibrium with any amount of private monitoring. That is, these methods can be used to prove new analytical results. In our third application, we consider a duopoly model of secret (or private) price cuts. In this application, we show that one period price wars are more robust to private monitoring than two-period price wars. Finally, we present a coordination game which allows us to compute an equilibrium with strategies which depend not only on a player's past private signals, but also his past private actions.

A. A Repeated Partnership Game (Mailath and Morris (2002))

In this example, we use the repeated partnership game of Mailath and Morris (2002) to show that; a) one can use our methods to easily compute the relevant belief sets to verify incentive conditions for both two and three state strategies; b) analyze which starting conditions work; c) do comparative statics regarding model parameters; and d) investigate which histories are problematic when parameters are such that a strategy is not an equilibrium.

We also highlight two somewhat surprising results. First, we show that sometimes tit-for-tat coordination works if both players start in the bad state but not when both players start in the good state. Second, we compute an example that suggests a new intuition (confirmed by later examples). In particular, the established intuition about moving from public to private monitoring is that the noise in monitoring makes it difficult for players to coordinate to be in matching states. These examples show a new effect: once we are in private monitoring, knowing too well the state of the opponent can be also bad for incentives. If a player has less knowledge about the state of his opponent (because of stochastic starting conditions or less predictable consumers or less correlated private signals) it can make it easier to satisfy incentives.

The Partnership Game

Consider the two player partnership game in which each player $i \in \{1, 2\}$ can take action $a_i \in \{C, D\}$ (cooperate or defect) and each can realize a private outcome $y_i \in \{G, B\}$ (good or bad). The $P(y|a)$ function is such that if m players cooperate, then with probability $p_m(1 - \epsilon)^2 + (1 - p_m)\epsilon^2$, both players realize the good private outcome. With probability $(1 - \epsilon)\epsilon$, player 1 realizes the good outcome while player 2 realizes the bad. (Likewise, with this same probability, player 2 realizes the good outcome and player 1 the bad.) Finally, with probability $p_m\epsilon^2 + (1 - p_m)(1 - \epsilon)^2$, both players realize the bad outcome. Essentially, this game is akin to one in which p_m determines the probability of an unobservable common outcome and ϵ is the probability that player i's outcome differs from the common outcome. Thus when $\epsilon = 0$, outcomes are public, and when ϵ approaches zero, outcomes are almost public. Payoffs are determined by specifying β and for each player i the vector $\{u_i(C, G), u_i(C, B), u_i(D, G), u_i(D, B)\}$.

Tit-for-Tat

Next consider perhaps the simplest non-trivial pure strategy: tit-for-tat. That is, let each player i play C if his private outcome was good in the previous period and D otherwise. This is a two-state strategy with $\Omega_i = \{R, P\}$, for "reward" and "punish." For $i \in \{1, 2\}$, $p_i(C|R) = 1$, $p_i(D|P) = 1$, $\omega_i^+(\omega_i, a_i, G) = R$, $\omega_i^+(\omega_i, a_i, B) = P$ for $\omega_i \in \{R, P\}$ and $a_i \in \{C, D\}$. Since every joint state can be reached from every other joint state with positive probability, $\tau(\psi)$ is a regular matrix and Corollary 1 of Theorem 2 applies and thus tit-for-tat is compatible with equilibrium if and only if incentives hold for the extreme points of the unique non-empty fixed point of T, $\overline{M}_i$. Since the number of states of i's opponent $D_{-i} = 2$, the set $\overline{M}_i(\omega_i)$ is simply a closed interval specifying the range of probabilities that player $-i$ is in state R, given that player i is in state $\omega_i \in \{R, P\}$. Operator T maps a collection of two intervals (one for each ω_i) to a collection of two intervals.

For $\beta = 0.9$, $p_0 = 0.3$, $p_1 = 0.55$, and $p_2 = 0.9$ and a payoff of 1 for receiving a good outcome and a payoff of -0.4 for cooperating, we can easily verify that the static game is a prisoner's dilemma and that tit-for-tat is an equilibrium of the public outcome ($\epsilon = 0$) game, starting from either both players in state R or both players in state P. For $\epsilon > 0$, beliefs matter and to check equilibrium conditions one must construct the intervals $\overline{M}_i(\omega_i)$. The procedure of iterating the T mapping is relatively easily implemented on a computer.[6] For $\epsilon = 0.025$ the procedure converges (in less than a second) to these intervals: $\overline{M}_i(R) = [0.923, 0.972]$, and $\overline{M}_i(P) = [0.036, 0.189]$ (see Figure 1).

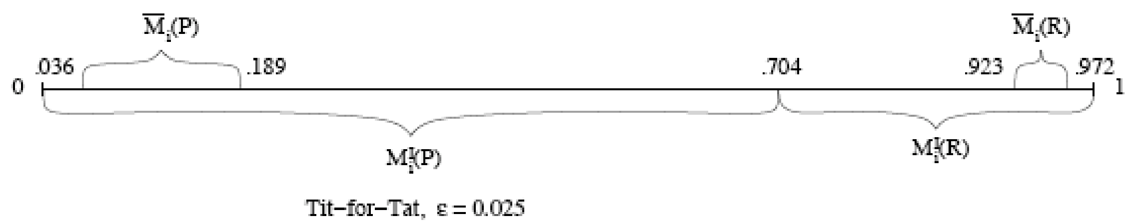

Figure 1

Again, tit-for-tat is compatible with equilibrium if and only if each player indeed wishes to play C when he believes his opponent is in state R with either probability 0.923 or 0.972, and indeed wishes to play D when he believes his opponent is in state R with either probability 0.036 or 0.189 (assuming a reversion to path play after a deviation). This is a matter of simply checking equation 1 for each of these four beliefs, and it holds in this case, thus there exist starting conditions such that tit-for-tat is an equilibrium.

[6]The Matlab code for checking arbitrary finite state strategies for arbitrary games can be found on the authors' websites.

In particular, Theorem 2 delivers one such starting condition. If both players follow the equilibrium, the transition matrix $\tau(\psi)$ between joint state $\omega \in \Omega = \{RR, RP, PR, PP\}$ and $\omega' \in \Omega$ implies a unique invariant distribution $\pi = (0.659, 0.038, 0.038, 0.264)$. If one chooses the correlation device $x = \pi$, then if player $i \in \{1, 2\}$ has R as his initial recommended state, he believes his opponent's initial recommended state is R with probability $0.945 = 0.659/(0.659 + 0.038)$. Likewise, if his initial recommended state is P, he believes his opponent's initial recommended state is R with probability $0.127 = 0.038/(0.038 + 0.264)$. Note that Lemma 4 implies the belief of player i after recommendation R, $\mu_{i,0}(R) = 0.945 \in \overline{M}_i(R)$ and likewise, $\mu_{i,0}(P) = 0.127 \in \overline{M}_i(P)$. Thus the correlation device $x = \pi$ and tit-for-tat form a CSE.

Are there any other starting conditions for which tit-for-tat is an equilibrium? Using the T^I operator, one can also readily calculate the sets M_i^{*I} for players $i \in \{1, 2\}$. In this example, $M_i^{*I}(R) = [0.704, 1]$ and $M_i^{*I}(P) = [0, 0.704]$. Corollary 2 then implies any correlation device x which delivers conditional beliefs $\mu_{i,0}(R) \in [0.704, 1]$ and $\mu_{i,0}(P) \in [0, 0.704]$, together with tit-for-tat, forms a CSE. Thus starting each player off in state $\omega_i = R$ with certainty (or x puts all mass on $\omega = RR$) and following tit-for-tat is a *sequential* equilibrium since $M_{i,0}(x, R) = \{1\} \subset M_i^{*I}(R)$ and $M_{i,0}(x, P) = \emptyset \subset M_i^{*I}(P)$. Likewise, starting each player off in state P (x puts all weight on $\omega = PP$) is also a sequential equilibrium since $M_{i,0}(x, R) = \emptyset \subset M_i^{*I}(R)$ and $M_{i,0}(x, P) = \{0\} \subset M_i^{*I}(P)$. Finally, letting x be such that one player starts off in state R and his opponent starts off in state P (with certainty) is *not* a sequential equilibrium since $M_{i,0}(x, R) = \{0\} \not\subset M_i^{*I}(R)$. Note by calculating $\overline{M}_i$ and M_i^{*I}, we have evaluated *all* deterministic starting conditions and thus all potential sequential equilibria associated with tit-for-tat.

If ϵ is increased to $\epsilon = 0.04$, then the intervals $\overline{M}_i(\omega_i)$ shift toward the middle and widen: $\overline{M}_i(R) = [0.883, 0.955]$ and $\overline{M}_i(P) = [0.057, 0.262]$. Further, we can calculate the sets $M_i^{*I}(R) = [0.918, 1]$ and $M_i^{*I}(P) = [0, 0.918]$. Now, if $\omega_i = R$ and player i believes that his opponent is in state R with probability 0.883, he wishes to deviate and play D rather than C. Thus, with $\epsilon = 0.04$, tit-for-tat is not an equilibrium for any starting conditions. Simply put, being only 88% sure your opponent saw the same good outcome as you (and thus will cooperate along with you) is an insufficient inducement for cooperation in this repeated prisoner's dilemma. Further, from all starting conditions, there exist histories where a player is supposed to cooperate, but is arbitrarily close to being only 88% sure that the other player is also cooperating. (See Figure 2).

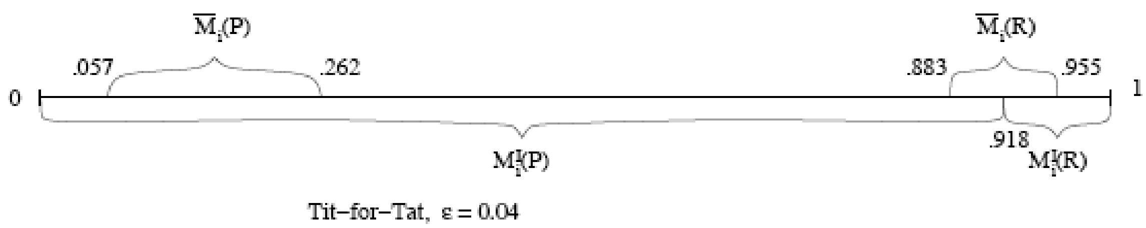

Figure 2

From Mailath and Morris (2002) we know that in this example, for sufficiently small ϵ, tit-for-tat is an equilibrium, and obviously for sufficiently high ϵ it is not. Our analysis of this example allows us to go further: to establish exactly for which ϵ's the profile is an equilibrium. That is, our methods allow us to consider whether any proposed strategy is an equilibrium strategy, regardless of whether the outcomes are nearly public.

Next, rather than increasing ϵ from $\epsilon = 0.025$ to $\epsilon = 0.04$, instead consider keeping $\epsilon = 0.025$ and decreasing the cost to cooperating from 0.4 to 0.357. Since the belief sets $\overline{M}_i$ do not depend on payoffs, they are still represented by Figure 1. Further, for these new payoffs, incentives continue to hold at the extreme points of $\overline{M}_i(R)$ and $\overline{M}_i(P)$, ensuring that letting the correlation device x on initial recommended states be the invariant distribution $\pi = (0.659, 0.038, 0.038, 0.264)$ remains a correlated equilibrium. However, given this change in payoffs, letting x be such that both players start off in state R with certainty is now no longer a sequential equilibrium. In fact for these payoffs, the only sequential equilibrium associated with tit-for-tat is for both players to start off in state P with certainty, which delivers the worst payoff over all ways of starting up a tit-for-tat equilibrium. This seems quite surprising given that each player's payoff to cooperating in state R is increasing in his belief that his opponent is also in state R. How can starting off with too much certainty be a problem? The difficulty with starting each player off in the reward state with certainty is that while each player is willing to cooperate in the first period, each is unwilling to defect in the second period, as tit-for-tat calls for, if he sees a bad outcome in the first period. The problem is that the certainty that one's opponent was in state R in the first period makes the player in the second period (after a bad outcome in the first period) insufficiently confident that his opponent is also in state P. In particular, his belief in period 2 that his opponent is in state R, $B_i(m_{i,0} = 1, h_i = (C, B)|\psi_{-i}) = .203$, which is outside of $\overline{M}_i(P) = [0.036, 0.189]$. On the other hand, if the correlation device $x = (0.8, 0.03, 0.03, 0.14)$ on the initial states $\Omega = \{RR, RP, PR, PP\}$, then if player i receives recommended state $\omega_{i,0} = R$, he believes his opponent is in state R with probability $m_{i,0} = 0.8/0.83 = 0.964$. Then, $B_i(m_{i,0} = 0.964, h_i = (C, B)|\psi_{-i}) = 0.185$, which is sufficiently low such that tit-for-tat is again a correlated equilibrium. (In fact, one can use our methods to find the correlation device x which delivers the *best* symmetric equilibrium payoff associated with any given strategy. In this case, this is approximately $x = (0.8, 0.03, 0.03, 0.14)$).

Tit for Tat-Tat

For this same game, consider a more complicated strategy: Let each player i play C if his private outcome was good in the previous *two* periods and D otherwise. This is a three-state strategy with $\Omega_i = \{R, P1, P2\}$, for "reward", "punish 1" and "punish 2." For $i \in \{1,2\}$, $p_i(C|R) = 1$, $p_i(D|P1) = 1$, $p_i(D|P2) = 1$, $\omega_i^+(\omega_i, a_i, B) = P1$ (for all $a_i \in \{C, D\}$ and $\omega_i \in \Omega_i$), $\omega_i^+(R, a_i, G) = R$, $\omega_i^+(P1, a_i, G) = P2$, and $\omega_i^+(P2, a_i, G) = R$ for $a_i \in \{C, D\}$. Since $D_{-i} = 3$, the set $\overline{M}_i(\omega_i)$ (for a given ω_i) is a two-dimensional convex subset of the unit simplex. The mapping T from Section 3 then maps a collection of three such subsets (one for each ω_i) to a collection of three such subsets. Figure 3 displays these sets for the same parameters as the last example (with $\epsilon = .04$). (In Figure 3, the red polygon is the belief set $\overline{M}_i(R)$, the blue polygon is the belief set $\overline{M}_i(P1)$, and the green polygon is the belief set $\overline{M}_i(P2)$. The dotted red area represents set $M_i^I(R)$, the blue dotted area is the set $M_i^I(P1)$, and the green dotted area is the set $M_i^I(P2)$.) Again, this example took only seconds to compute. Of perhaps more interest, with $\epsilon = .04$, incentives hold at the extreme points of each set, thus there exist correlation devices such that this joint automaton (tit-for-tat-tat) is a CSE, while this is not the case for tit-for-tat. (In particular, both players starting in the same state with certainty is a sequential equilibrium.)

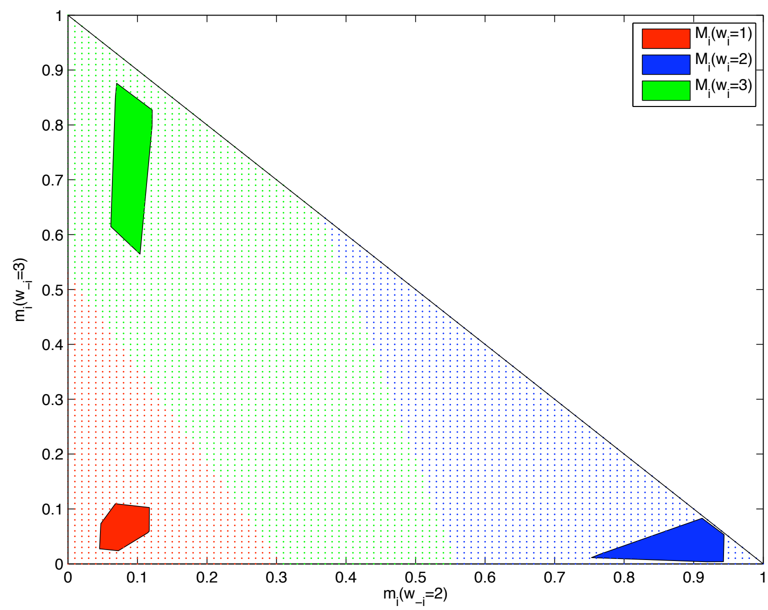

Figure 3: Belief sets $\overline{M}_i(\omega_i)$.

B. Abreu, Pearce, and Stacchetti (1986)

In this section we consider the classic repeated Cournot game and shows how to use our methods to prove that the stick-and-carrot class of strategies, which are optimal in public monitoring case, cannot be equilibria with any private monitoring.

Consider the repeated oligopoly quantity setting game of Abreu, Pearce, and Stacchetti (1986) (APS), where two firms set private quantities which affect the probability distribution over a publicly observed price. When prices and quantities have continuous support, they show that the best and worst strongly symmetric public equilibria are each implemented by a simple two-state (or bang-bang) strategy: In state R (Reward), both firms choose a low quantity (lower than the Cournot-Nash level) and in state P (Punish), both firms choose a high quantity (higher than the Cournot-Nash level). In the best equilibrium, both firms start out in state R. If the observed price is in some strict subset, Y_R, of the set of all prices, they stay in state R, otherwise they transit to state P (starting from which is the worst public equilibrium). In state P, they move to state R if the observed price is in some other strict subset, Y_P, otherwise they stay in state P. Further, $Y_R \neq Y_P$. That is, the set of prices this period which causes the players to be in state R next period depends on whether the players are in state R or P this period. (Generally, reaching state R from state R requires a sufficiently *high* price, while reaching state R from state P requires a sufficiently *low* price.) When prices and quantities have finite support, as we assume in this paper, such a bang-bang strategy can arbitrarily well approximate the best and worst public equilibria as one allows a finer and finer finite grid of prices and quantities.

Now instead of assuming that there is a single public price, as in APS, suppose each firm draws a privately observed price and the joint distribution of prices is affected by both firms' quantity decisions. Further, allow this joint distribution to be arbitrary except for an assumption of full support. This allows for, but does not require, almost perfect correlation between the prices seen by the two firms and makes the above example consistent with the assumptions of this paper. With this change from public to private monitoring, is the APS bang-bang strategy still a sequential equilibrium? Here, we show, generically, it is not. To show this, we first prove the following simple, and surprisingly useful lemma.

Lemma 6. *Suppose ψ is such that there exists a player i, private states $\omega_{i,1}$ and $\omega_{i,2}$ and a one-period private history $h_i = (a_i, y_i)$ such that $\omega_i^+(\omega_{i,1}, h_i) = \omega_{i,1}$ and $\omega_i^+(\omega_{i,2}, h_i) = \omega_{i,2}$, or $\omega_i^+(\omega_{i,1}, h_i) = \omega_{i,2}$ and $\omega_i^+(\omega_{i,2}, h_i) = \omega_{i,1}$. Let $h_i^{(s)}$ be an s-period repetition of h_i. Then, for all $m_i \in \Delta_i^{D_{-i}}$, all accumulation points of the sequence $\{B_i^s(m_i, h_i^{(s)}|\psi_{-i})\}_{s\in\{0,2,4,\dots\}}$ (of which there is at least one) are elements of both $\overline{M}_i(\omega_{i,1})$ and $\overline{M}_i(\omega_{i,2})$.*

Proof. That $\{B_i^s(m_i, h_i^{(s)}|\psi_{-i})\}_{s\in\{0,2,4,\dots\}}$ has at least one accumulation point comes from the compactness of Δ_i^{D-i}. Next, since $\omega_i^+(\omega_{i,1}, h_i) = \omega_{i,1}$ and $\omega_i^+(\omega_{i,2}, h_i) = \omega_{i,2}$, or $\omega_i^+(\omega_{i,1}, h_i) = \omega_{i,2}$ and $\omega_i^+(\omega_{i,2}, h_i) = \omega_{i,1}$, for all even s and $m_i \in \Delta_i^{D-i}$, $B_i^s(m_i, h_i^{(s)}|\psi_{-i}) \in T^s(\overline{\Delta}_i)(\omega_{i,1})$ and $B_i^s(m_i, h_i^{(s)}|\psi_{-i}) \in T^s(\overline{\Delta}_i)(\omega_{i,2})$. Further, since $\{T^s(\overline{\Delta}_i)(\omega_{i,1})\}_{s\in\{0,2,4,\dots\}}$ and $\{T^s(\overline{\Delta}_i)(\omega_{i,2})\}_{s\in\{0,2,4,\dots\}}$ are each decreasing sequences (in terms of set inclusion) of nonempty compact sets (with non-empty limits $\overline{M}_i(\omega_{i,1})$ and $\overline{M}_i(\omega_{i,2})$), if $B_i^s(m_i, h_i^{(s)}|\psi_{-i}) \in T^s(\overline{\Delta}_i)(\omega_{i,1})$ and $B_i^s(m_i, h_i^{(s)}|\psi_{-i}) \in T^s(\overline{\Delta}_i)(\omega_{i,2})$, then any accumulation point of the sequence $\{B_i^s(m_i, h_i^{(s)}|\psi_{-i})\}_{s\in\{0,2,4,\dots\}}$ is in both limiting sets $\overline{M}_i(\omega_{i,1})$ and $\overline{M}_i(\omega_{i,2})$. ∎

First, note that the conditions of Lemma 6 above hold for the bang-bang strategy of APS since $Y_R \neq Y_P$. (That is, if $Y_R \neq Y_P$, there must exist a price y_i such that $y_i \in Y_R$ and $y_i \notin Y_P$, or vice-versa. If $y_i \in Y_R$ and $y_i \notin Y_P$, then, by definition, $\omega_i^+(R, (a_i, y_i)) = R$ and $\omega_i^+(P, (a_i, y_i)) = P$ for all a_i. If $y_i \notin Y_R$ and $y_i \in Y_P$, then $\omega_i^+(R, (a_i, y_i)) = P$ and $\omega_i^+(P, (a_i, y_i)) = R$ for all a_i.)

Next, note that full support implies that every joint state to state transition has positive probability on path, thus the conditions of Lemma 5 hold. Corollary 1 (of Theorem 2) then implies that incentives holding on the extreme points of the sets $\overline{M}_i(R)$ and $\overline{M}_i(P)$ for each player i is not only sufficient but also necessary for this two-state strategy to be an equilibrium. In particular, this implies for all $m_i \in \overline{M}_i(P) \cap \overline{M}_i(R)$, the continuation strategy associated with state R and the continuation strategy associated with state P must *both* be optimal for player i when he has this belief regarding the state of his opponent. That is, player i must be indifferent between the two continuation strategies for **all** $m_i \in \overline{M}_i(P) \cap \overline{M}_i(R)$.

But note that since the expected payoff for a given strategy is linear in beliefs (and a belief is a scalar between zero and one), if $\overline{M}_i(P) \cap \overline{M}_i(R)$ has more than one point, then the payoff from following the continuation strategy associated with state R must be equal to the payoff associated with state P not only for all beliefs in $\overline{M}_i(P) \cap \overline{M}_i(R)$ but for all beliefs in $[0, 1]$. At least for generic parameter values, this is not the case for the bang-bang strategy of APS.

Does $\overline{M}_i(P) \cap \overline{M}_i(R)$ have more than one point? Consider a particular price y_i such that $y_i \in Y_R$ and $y_i \notin Y_P$, or vice-versa, and two different quantities, $a_{i,0}$ and $a_{i,1}$. Since $\tau(\psi)$ is a regular matrix, $\lim_{s\to\infty} B_i^s(m_{i,0}, (a_i, y_i)^{(s)}|\psi_{-i})$ exists for all (a_i, y_i) and is independent of initial belief $m_{i,0}$.[7] Further, for each $a_i \in \{a_{i,0}, a_{i,1}\}$, Lemma 6 implies $\lim_{s\to\infty} B_i^s(m_{i,0}, (a_i, y_i)^{(s)}|\psi_{-i}) \in \overline{M}_i(P) \cap \overline{M}_i(R)$. But this limit is generically **not** independent of a_i. A player's own action affects his updating. Thus generically, $\overline{M}_i(P) \cap \overline{M}_i(R)$ will contain more than one point and the bang-bang strategy of APS is not an equilibrium.

Similar reasoning can be applied to any game that fits our framework and any strategy where $\tau(\psi)$ is a regular matrix and the condition of Lemma 6 holds: a necessary condition for equilibrium is indifference over continuation strategies at all points in the nonempty intersection of belief sets $\overline{M}_i(\omega_{i,1})$ and $\overline{M}_i(\omega_{i,2})$. Further, for two state strategies where neither state is absorbing on path, if $\omega_i^+(\omega_i, h_i)$ is a nontrivial function of ω_i, or the state a player goes to tomorrow depends on the state he is in today, then the condition of Lemma 6 indeed holds. More generally, depending on the dimensionality of $\overline{M}_i(\omega_{i,1}) \cap \overline{M}_i(\omega_{i,2})$, state dependent transitions may imply indifference for possibly all beliefs, essentially ruling out all but belief-free equilibria.

[7] The limit exists by the reasoning in the proof of Lemma 5.

C. Secret Price Cuts

In this section we study a secret price cutting game with a rich action and signal space. First, we show that a natural strategy from the public monitoring game, namely Taking Turns is not going to work with private monitoring. Second, we show that one-period price wars can support collusion, but they may require random correlated starting conditions. Finally, we show an example with two-period price wars that support collusion, while one-period ones are not enough. In that example, if customer behavior is more predictable, it is more difficult to sustain collusion in the private monitoring case. It also suggests that strategies with two-period punishments are much more fragile to private monitoring than one-period punishments.

A Bertand Pricing Game

Consider a repeated Bertrand duopoly game. At each date, each of two players (firms) privately chooses a price $a_i \in \{0, 0.01, 0.02, \ldots, 4\}$. A player's private outcome is his number of customers $y_i \in Y_i = \{0, 1, 2, 3, 4, 5\}$. With probability $(1-\epsilon)$, the total number of customers, $y_1 + y_2 = 5$, and with probability $\epsilon/10$, the total number of customers is any particular element of $\{0, 1, 2, 3, 4, 6, 7, 8, 9, 10\}$. If both players choose the same price, each customer flips a fair coin to determine from which firm he buys. If the firms choose different prices, each customer chooses the lower price firm with probability $1-\delta$. (If the total number of customers is more than five, and these coin flips imply one player selling to more than five customers, that player is assumed to have exactly five customers, with the other player selling to the other customers.) Production is assumed to have a constant marginal cost $c \geq 0$ so $u_i(a_i, y_i) = (a_i - c) * y_i$. If $\delta = 0$, and as the grid on prices gets infinitely fine, the unique stage game Nash equilibrium is for both firms to choose price $a_i = c$. If ϵ and δ are each strictly positive, all joint outcomes (y_1, y_2) occur with positive probability for all (a_1, a_2) and this game fits in our framework.

Taking Turns

Consider the following three state strategy: In state Me, player i chooses $a_i = 3.99$, while in state You, player i chooses $a_i = 4$. In state P (Punishment), player i chooses $a_i = 0$. If in state Me, player i receives 3 or more customers, he transits to state You, otherwise he transits to state P. If in state You, player i receives 2 or fewer customers, he transits to state Me, otherwise he transits to state P. Finally, if in state P, player i receives 0, 1, 4, or 5 customers, he stays in state P, if he receives 2 customers, he transits to state Me and if he receives 3 customers he transits to state You.

If $\beta = .95$, $\delta = .05$ and $c = 1$, for the game with public monitoring ($\epsilon = 0$), this strategy is a perfect public equilibrium when one player starts in state Me and the other in state You. As long as the lower price firm gets a majority of the customers, (a high probability event), both players choose a high price (with one slightly undercutting the other) and take turns regarding which one receives most of the customers. In the unlikely event a firm receives a majority of the customers out of turn, a price war ensues. In a price war, each firm has the incentive to charge $a_i = 0$ since this maximizes the probability that customers will be split as evenly as possible, causing the price war to end.

But note that the conditions for Lemmas 5 and 6 holds in this example. If player 1 is in state Me and receives $y_1 = 2$ customers, he transits to state P, while if he is in state P and receives $y_1 = 2$ customers, he transits to state Me. Thus for each player i, $\overline{M}_i(Me) \cap \overline{M}_i(You)$

is nonempty. Thus for the incentive conditions to be satisfied, each player must be indifferent between following the continuation strategy associated with state Me and the continuation strategy associated with state You for all points in this nonempty intersection, which will generically not be the case.

Such state dependence transitions appear to be essential to any turn-taking equilibrium with public monitoring. That is, which outcomes require a transition to a punishment state must rely on whose turn it was win the majority of customers or whether the players are currently in the punishment state. But these state dependent transitions make the strategy not an equilibrium with private monitoring.

High Equal Prices with Price Wars

Now instead consider a different two state strategy. In state R (Reward), each firm chooses $a_i = 4$ and in state P (Punish), each firm chooses $a_i = 0$. From any state, if $y_i \in \{0, 5\}$ (a firm sells to either zero or five customers), it transits to state P in the next period, regardless of its price a_i. If $y_i \in \{1, 2, 3, 4\}$, from any state, it transits to state R tomorrow. In words, each firm sets a price of four unless last period it had an extreme number of customers. If $\epsilon = 0$, or the total number of customers is certain to be five, this is a game of public monitoring, and this strategy is a public equilibrium as long as δ, the probability that a customer chooses the high price firm, is not too high (or for β near 1, $\delta \leq .06$).

If $\epsilon \leq .04$, (with $\beta = .95$, $\delta = .05$, and $c = 1$), unlike taking turns, there exists a correlation device such that this strategy is also an equilibrium of the private monitoring game (specifically, drawing initial states from the unique invariant distribution, where joint state $\omega \in \{RR, RP, PR, PP\}$ is drawn with probability $(0.90, .01, .01, .08)$.) Interestingly, however, for these parameters there exists no *deterministic* correlation device such that this is an equilibrium. Starting one player in state R and the other in state P is obviously not an equilibrium. However, for less obvious reasons, starting both in state R or both in state P is also not an equilibrium. For $\epsilon = .04$, $\overline{M}_i(R) = [0.263, .994]$ and $\overline{M}_i(P) = [0.016, 0.124]$, relatively wide but non-overlapping belief sets, and incentives hold on their extreme points. However, if both players start off in state R with certainty, while $M_i^{*U}(P) = \overline{M}_i(P)$, $M_i^{*U}(R) = [0.104, 1.000]$, which has not only a higher upper bound but also a smaller lower bound. At this reduced lower bound, incentives do not hold.

Specifically, the lower bound of $M_i^{*U}(R)$ is generated by assuming player i believes his opponent is in state R with probability 1, sets $a_i = 0$ and receives one customer. (That is, $B_i(m_i = 1, h_i = (0,1)|\psi_{-i}) = 0.104$). Bayesian updating essentially depends on reconciling the player's observations with its possible explanations and the most likely explanation for player i receiving only one customer when he undercut his opponent is that the total number of customers was actually only one and this customer chose the lower price, putting player $-i$ in state P (which happens with probability $1 - 0.104$). On the other hand, if player i is only 99.4% certain that his opponent is in state R (the upper bound of $\overline{M}_i(R)$) then if he sets $a_i = 0$ and receives one customer, he now believes his opponent is in state R with probability $0.265 \in \overline{M}_i(R)$ and incentives hold. This change in updating occurs since the small amount of doubt leaves another explanation for player i receiving only one customer — his opponent was actually in state P and thus both set a price of zero, and thus it is more likely his opponent received a positive number of customers.

A similar explanation rules out both players starting out in state P with certainty. If both players start out in state P with certainty, $M_i^{*U}(P) = [0, 0.124]$, (where incentives continue to hold), but $M_i^{*U}(R) = [0.008, 0.994]$, where incentives don't hold at the lower bound. Here, the lower bound of $M_i^{*U}(R)$ is generated by assuming a player i believes his opponent is in state R with probability 0, sets $a_i = 4$ and receives four customers. The most likely explanation for player i receiving four customers when he charged a higher price than his opponent is that there were nine customers and his opponent received five, implying he is in state P with high probability. However, if player i puts probability 1.6% on his opponent being in state R (the lower bound of $\overline{M}_i(P)$) then if he sets $a_i = 4$ and receives four customers, he now believes his opponent is in state R with probability $0.384 \in \overline{M}_i(R)$ and incentives hold. As before, this change in updating occurs since the small amount of doubt leaves another explanation for player i receiving four customers — his opponent was actually in state R and thus charged the same price as him, and thus it is more likely he received between one and four customers.

Two-period price wars

For this game, if the marginal cost of production $c = 0$, one can show analytically that the two-state strategy considered in the previous section is not an equilibrium of the $\epsilon = 0$ public game. A price war of possibly only one period of zero profits (as opposed to negative profits if $c > 0$) is an insufficient punishment to hinder slightly undercutting one's opponent. In this section, we show that a minimum two period punishment can be an equilibrium, but that the coordination necessary for two-period punishments implies that the number of customers must be very close to public information.

Consider the following three state strategy: In state R, each firm chooses $a_i = 4$ and in states $P1$ and $P2$, each firm chooses $a_i = 0$. From any state, if $y_i \in \{0, 5\}$ (a firm sells to either zero or five customers), it transits to state $P1$ in the next period, regardless of its price a_i. On the other hand, if $y_i \in \{1, 2, 3, 4\}$, it transits to state R tomorrow if today's state was R or $P2$, and transits to state $P2$ tomorrow if today's state was $P1$. In words, each firm sets a price of zero unless in each of the last two periods it had an interior number of customers. If $\epsilon = 0$, or the total number of customers is certain to be five, this is a game of public monitoring, and this strategy is a public equilibrium as long as δ, the probability that a customer chooses the high price firm, is not too high (or for β near 1, $\delta \leq .16$).

From Mailath and Morris (2002) we then know for any given β and δ, there exists an $\bar{\epsilon} > 0$ such that for all $0 < \epsilon \leq \bar{\epsilon}$, this strategy is also an equilibrium of the private monitoring game with an uncertain number of customers. However, assuming $\beta = 0.95$, if $\delta = 0.1$ (or the customer chooses the lower price with probability 0.9), our computation method shows that for the above strategy to be an equilibrium, one needs $\epsilon < 0.0000004$, or there must be less than four chances in ten million that the number of customers differs from five. For smaller δ, (or for higher probabilities that consumers choose the lower price), ϵ must be even *lower*. If $\delta = 0.05$ (or the customer chooses the lower price with probability 0.95), equilibrium requires $\epsilon < 0.000000004$, or there must be less than four chances in a billion that the number of customers differs from five.

The reason ϵ must be so small (and small relative to δ) again comes from a player's off path Bayesian updating. For instance, suppose ϵ and ϵ/δ are both positive but infinitesimal. Then, regardless of a player's action, and regardless of his beliefs regarding his opponent's state (and thus his action) if he receives zero or five customers, he concludes his opponent also received zero or five customers, and if he receives one through four customers, he concludes his opponent did as well. This guarantees regarding of starting states and actions taken, within two periods each player is convinced the other player is in the same state he is. (More formally, in the limit as $\epsilon \to 0$ for a given $\delta > 0$, $\overline{M}_i(R) = \{(1,0,0)\}$, $\overline{M}_i(P1) = \{(0,1,0)\}$, and $\overline{M}_i(P2) = \{(0,0,1)\}$.) On the other hand, if ϵ and δ/ϵ are both positive and not infinitesimal, very different Bayesian updating occurs.

Suppose $\delta = .1$ and $\epsilon = 0.00000001$ (which is too high for this strategy to be an equilibrium). What goes wrong? Again, one feature of our computation method is that it points out at exactly which state, ω_i, and which extreme belief in $\overline{M}_i(\omega_i)$ incentives fail to hold. For these parameters, incentives fail to hold for state $P2$, when player i believes his opponent is in state R with (approximately) 50% probability and state $P2$ with approximately 50% probability. Here, with this level of doubt, player i is unwilling to play $a_i = 0$, preferring a higher price.

Further, as in the previous example, our methods allow one to trace how such an extreme belief can be supported. This particular extreme belief (player i is in state $P2$ but believes his opponent is 50/50 in R or $P2$) is generated as follows: Suppose player i is in state R, believes his opponent is also in state R (with certainty), deviates and plays $a_i = 0$, and receives zero customers, putting him in state $P1$ tomorrow. One possibility is that the number of customers was five, but each of them chose the higher price firm. This happens with probability $\delta^5 * (1 - \epsilon)$ which is about .0000003125, or one in 3.2 million. In this scenario, player i's opponent had five customers and is in state $P1$ tomorrow. A second possibility is that the number of customers was one and this single customer chose the higher price firm. This happens with probability $\delta * (\epsilon/10)$ which is .00000125, or one in eight hundred thousand. In this second scenario, player 1's opponent had one customer and is in state R tomorrow. The ratio of these events is 0.00016 (or one in 625) which closely matches the actual posterior of player i given this scenario. And given he is in state $P1$ and believes his opponent is in state $P1$ with probability 0.99984 and state R with probability 0.00016, he wishes to follow the strategy and play $a_i = 0$.

But from this state and belief, suppose player i then chooses an intermediate price $a_i \in \{.01, \ldots, 3.99\}$ and receives three customers, putting player i in state $P2$ the following period. How does he account for this event? One possibility is that his opponent was in state $P1$ (and thus played $a_{-i} = 0$) and four out of five customers chose the higher price firm, putting player $-i$ in state $P2$ tomorrow. This happens with probability $0.999984 * 5 * \delta^4 * (1 - \delta) * (1 - \epsilon)$ which is about .00003. Another possibility is that his opponent was in state R (and thus played $a_{-i} = 4$) and only one out of five customers chose the higher price firm, putting player $-i$ in state R tomorrow. This happens with probability $.0.00016 * 5 * (1 - \delta)^4 * \delta * (1 - \epsilon)$ which is also about .00003. Since the ratio of these two events is near one, from state P2, player i now believes player $-i$ is in state R with (about) 50% probability and state $P2$ with 50% probability.

D. A Coordination Game

The following is an example of a game and strategy where equilibrium depends on information *not* being almost public, and thus the ability to analyze general private monitoring environments is crucial. Further, to our knowledge, it is the first example of a finite state automaton equilibrium in a private monitoring game where transitions depend on ones own actions. (These are not covered by Mailath and Morris (2002) since their method is to alway start with a public equilibrum, in which transitions depend only on public histories. It differs from Bhaskar and Obara (2002) since their strategy has an infinite number of states.)

Consider a two player battle of the sexes game where each player $i \in \{1, 2\}$ can take action $a_i \in \{Ballet, Hockey\}$ and each can realize a private outcome $y_i \in \{G, B\}$ (good or bad). If both players take the same action, they both realize a good outcome with probability 0.9, both receive a bad outcome with probability 0.08 and player i realizes a good outcome while player $-i$ receives a bad outcome with probability 0.01. If the players take differing actions, they both realize a good outcome with probability 0.05, both receive a bad outcome with probability 0.05 and player i realizes a good outcome while player $-i$ receives a bad outcome with probability 0.45. If player 1 realizes a bad outcome, her payoff is zero, and if she realizes a good outcome, her payoff is 1.1 if she played *Ballet* and 1 if she played *Hockey*. Likewise, if player 2 realizes a bad outcome, his payoff is zero, and if he realizes a good outcome, his payoff is 1.1 if he played *Hockey* and 1 if he played *Ballet*. As in the previous example, $\beta = 0.9$.

Our methods can be used to check if the following simple strategy is an equilibrium: if a player's private outcome was good, repeat last period's play regardless of whether it was on or off path. If his (or her) private outcome was bad, switch away from last period's play regardless of whether it was on or off path. This strategy is a two state automaton $\omega_i = ([PlayBallet], [PlayHockey])$ and belief sets are intervals specifying the probability that the other player is in state *PlayBallet*. For these parameters, the intervals are $\overline{M}_i(PlayBallet) = [0.890, 0.988]$ and $\overline{M}_i(PlayHockey) = [0.012, 0.110]$, and incentives hold on the boundaries of these two intervals. But note they hold precisely because this is *not* a game with almost public outcomes. That is, suppose player 1 is in state *PlayHockey* and deviates by playing *Ballet*, while believing (with high probability) that player 2 is in state *PlayHockey*. If she realizes a bad outcome, the function P above implies she believes player 2 most likely received a good outcome (and thus will not switch states) and thus it is in her interest to follow the equilibrium by playing *Hockey* next period. If P were such that she believed player 2 also had a bad outcome, as would be the case if outcomes were almost public, after this deviation, player 1 would no longer be willing to follow the strategy.

5. Equilibrium Paths

The strategies we analyzed specify behavior both on and off the path of play. However, if we are interested in the observable predictions of a theory, only the on-path behavior is relevant. In this section we discuss how one can extend our methods to examine when a path of play is consistent the underlying model.

Let the path of play of a correlation device x and a joint automaton ψ be defined as the stochastic process on (a, y) implied by (x, ψ). If one uses our methods to verify that a particular (x, ψ) constitute a CSE, then the path of play predicted by (x, ψ) is consistent with the model. However, if a particular (x, ψ) is not a CSE, it does not immediately follow that the path of play of (x, ψ) is inconsistent with the model. There may exist a collection of possibly non-finite-state strategies $(\hat{\sigma}^1, \ldots, \hat{\sigma}^J)$ and a probability distribution over these strategies $\hat{x}$ such that $(\hat{x}, (\hat{\sigma}^1, \ldots, \hat{\sigma}^J))$ has the same path of play as (x, ψ) but where $(\hat{x}, (\hat{\sigma}^1, \ldots, \hat{\sigma}^J))$ is a Correlated Sequential Equilibrium.[8]

The following Lemma shows, under the full support condition, that if (x, ψ) is not a Correlated *Nash* Equilibrium (where optimality is checked only at the initial node) then all pairs $(\hat{x}, (\hat{\sigma}^1, \ldots, \hat{\sigma}^{\hat{J}}))$ with the same path of play as (x, ψ) are also not Correlated Nash

[8]For non-finite state strategies, we say $(x, (\sigma^1, \ldots, \sigma^J))$ is a Correlated Sequential Equilibrium if for each recommended strategy σ^j, player i and each private history h_i^t, player i finds it optimal to play $\sigma_i^j(h_i^t)$ given his uniquely defined beliefs over σ_{-i} and h_{-i}^t conditional on σ_i and h_i^t.

Equilibria. Thus if a particular (x, ψ) is not a Correlated Nash Equilibrium, its path of play is inconsistent with the model.

Lemma 7. *Suppose $(\sigma^1, \ldots, \sigma^J)$ is a finite collection of (possibly non-finite state) strategies and $\hat{x}$ is a probability distribution on those strategies. Suppose ψ is a joint automaton and x is a probability distribution over initial states of ψ. Further assume $(\hat{x}, (\sigma^1, \ldots, \sigma^J))$ and (x, ψ) have the same path of play. Then $(\hat{x}, (\sigma^1, \ldots, \sigma^J))$ is a Correlated Nash Equilibrium only if (x, ψ) is a Correlated Nash Equilibrium.*

Proof. See Appendix. ∎

From results by Sekiguchi (1997) and Kandori and Matsushima (1998), (again assuming full support) if a correlation device x and a joint automaton ψ together form a Correlated Nash Equilibrium, then there exists a Correlated Sequential Equilibrium with the same path of play as (x, ψ). Thus, from this and Lemma 7, the path of play of (x, ψ) is consistent with the underlying model if and only if (x, ψ) form a Correlated Nash Equilibrium. Can we determine whether (x, ψ) is a Correlated Nash Equilibrium? To this end, we can provide partial results.

The operators $T(M_i)$ and $T^U(M_i)$ compute sets of posterior beliefs given the sets of priors and *all possible* new data (y_i, a_i). We can define two related operators, $\mathbb{T}(M_i)$ and $\mathbb{T}^U(M_i)$ in which updating is done only using actions consistent with the equilibrium play. First, let $\mathbb{G}_i(\omega_i, \omega_i' | \psi_i)$ be the set of current-period data (y_i, a_i) such that a) after observing them player i transits from state ω_i to ω_i' and b) the action a_i is played by player i with positive probability in state ω_i given automaton ψ_i (which is the new restriction). Second, let:

$$\mathbb{T}(M_i)(\omega_i') = \text{co}(\{m_i' | \text{ there exists } \omega_i, m_i \in M_i(\omega_i) \text{ and } (a_i, y_i) \in \mathbb{G}_i(\omega_i, \omega_i'|\psi_i)$$
$$\text{such that } m_i' = B_i(m_i, a_i, y_i|\psi_{-i})\})$$

and as usual, $\mathbb{T}(M_i) = \{\mathbb{T}(M_i)(\omega_i') \,|\omega_i' \in \Omega_i\}$. Third, let

$$\mathbb{T}^U(M_i)(\omega_i) = \text{co}(\mathbb{T}(M_i)(\omega_i) \cup M_i(\omega_i))$$

The operators $\mathbb{T}$ and $\mathbb{T}^U$ are analogous to T and T^U with the only difference being that updating is restricted to actions on the equilibrium path. As a result, these new operators have the same properties as the old operators. In particular, if $\tau(\psi)$ is a regular matrix, $\mathbb{T}$ has a unique fixed point and no matter what set of beliefs we start with, $\mathbb{T}^n$ converges to this fixed point. Denote this fixed point by $\mathbb{M}_i$. Moreover, starting at any collection of sets of beliefs, M_i, the sequence $\left(\mathbb{T}^U\right)^n(M_i)$ is increasing and converging to some limit, denoted by $\mathbb{M}^{*U}(M_i)$.

Theorem 3. *For a given joint automaton* ψ

1. *Suppose* $\tau(\psi)$ *is regular. Then there exists a correlation device* x *such that* (x, ψ) *is a CNE, if and only if for all players* i*, private states* ω_i*, belief points* m_i *such that* m_i *is an extreme point of* $\mathbb{M}_i$ *and all deviation strategies* $\hat{\sigma}_i$*,* $Ev_i(\omega_i, m_i|\psi_i, \psi_{-i}) \geq EV_{i,0}(\emptyset, m_i|\hat{\sigma}_i, \psi_{-i})$.
2. *For a given correlation device* x*,* (x, ψ) *is a CNE, if and only if for all players* i*, private states* ω_i*, belief points* m_i *such that* m_i *is an extreme point of* $\mathbb{M}^{*U}(M_{i,0}(x))$ *and all*

deviation strategies $\hat{\sigma}_i$, $Ev_i(\omega_i, m_i|\psi_i, \psi_{-i}) \geq EV_{i,0}(\emptyset, m_i|\hat{\sigma}_i, \psi_{-i})$.

The proof, similar to our earlier proofs, is omitted.

But how useful is Theorem 3? Our earlier results, Theorems 1 and 2 concerned Correlated *Sequential* Equilibria and used Lemma 1, the one-step deviation principle. Thus the result that it was sufficient to check incentives only at extreme beliefs could be readily operationalized since only a finite number of deviation strategies need to be checked. With Correlated *Nash* Equilibria, the one-step deviation principle does not hold, and thus one must check incentives for an infinite number of possible deviation strategies.

But Theorem 3 is still progress. For instance, one can still check incentives for extreme beliefs against a finite subclass of deviation strategies, say one-step deviations. If such a deviation is profitable, then (x, ψ) is not a CNE, and its path of play is inconsistent with the underlying model. Further, we know that $\mathbb{M}_i \subseteq \overline{M}_i$ and $\mathbb{M}^{*U}(M_{i,0}(x)) \subseteq M^{*U}(M_{i,0}(x))$ because in the construction of the new sets we update on a restricted set of events and one cannot rule out that the inclusion will be strict. However, if the sets turn out to coincide, then our results regarding Correlated Sequential Equilibrium also hold for Correlated Nash Equilibria and thus have implications for the path of play.

The difficulty is when $\mathbb{M}_i$ is strict subset of $\overline{M}_i$, or for a particular x, $\mathbb{M}^{*U}(M_{i,0}(x))$ is a strict subset of $M^{*U}(M_{i,0}(x))$, and one-step deviations are profitable for the extreme points of the larger set, but not the smaller set. Then we cannot say if the path of play is consistent with the model without looking at more complicated deviation strategies, an approach suggested by Kandori and Obara (2007).

6. Concluding Remarks

Beyond using our methods directly to compute equilibria, one can extend and apply these methods in several ways.

First, as shown in a recent paper by Kandori and Obara (2007) one can use set based methods similar to ours to study strategies that can be represented by finite automata on the equilibrium path but can be much more complicated off the equilibrium path. For example, they allow the strategy off the equilibrium path to be a function of beliefs over other players' states, which implies an infinite number of the automaton states (since players believe that others are always on the equilibrium path, the beliefs are still manageable).

Second, one can prove that if incentives hold strictly (uniformly bounded) for all extreme beliefs of the fixed point operator T^U, then this CSE is robust to small perturbations of the stage game payoffs or the discount factor. The reasoning is as follows: first, the T^U operator and the initial belief sets $M_{i,0}(x)$ are independent of the payoffs. Hence the fixed point is independent. Second, the incentive constraints are continuous in the stage-game payoffs and the discount factor. Hence, if for the given game the incentives hold strictly for all extreme beliefs of the fixed point of the T^U operator, they also hold weakly for small perturbations of the payoffs or the discount factor. Then, Theorem 1 implies that for the perturbed game the same (x, ψ) are a CSE. Similar arguments can be used for perturbations of the monitoring technology (the $P(y|a)$ function) to study robustness to changes in monitoring.

References

[1] Bhaskar, V. and I. Obara (2002): "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 102(1), 40-69.

[2] Compte, O. (2002): "On Failing to Cooperate when Monitoring Is Private," *Journal of Economic Theory*, 102(1, Jan.), pp. 151–188.

[3] Cripps, M., G. Mailath and L. Samuelson (forthcoming): "Disappearing Private Reputations in Long-Run Relationships," *Journal of Economic Theory.*

[4] Ely, J. C. (2002): "Correlated Equilibrium and Trigger Strategies with Private Monitoring,"manuscript, Northwestern University.

[5] Ely, J. C., J. Hörner, and W. Olszewski (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica*, 73(2, Mar.), pp. 377-415.

[6] Ely, J. C., and J. Välimäki (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102(1,Jan.), pp.84-105.

[7] Kandori, M. (2002): "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory*, 102(1, Jan.), pp. 1-15.

[8] Kandori, M., and I. Obara (2006): "Efficiency in Repeated Games Revisited: The Role of Private Strategies," *Econometrica,* 74(2, Mar.), pp. 499-519.

[9] Kandori, M. and I. Obara (2007): "Finite State Equilibria in Dynamic Games," manuscript.

[10] Mailath, G. J., and S. Morris (2002): "Repeated Games with Almost-Public Monitoring," *Journal of Economic Theory*, 102(1, Jan.), pp. 189-228.

[11] Mailath, G. J., and S. Morris (2006): "Coordination Failure in Repeated Games with Almost-Public Monitoring," *Theoretical Economics*, 1(3, Sept.), pp. 311-340.

[12] Mailath, G. J., and L. Samuelson (2006): *Repeated Games and Reputations: Long-Run Relationships*, New York: Oxford University Press.

[13] Piccione, M. (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102(1, Jan.), pp. 70–83.

[14] Sekiguchi, T. (1997): "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory*, 76(2,Oct.), pp. 345-361.

Appendix

Proof of Lemma 2

Proof. First, recall that $T(M_i)(\omega_i)$ is convex from the definition of T. Next, from its definition, we can express $T(M_i)(\omega'_i)$ as

$$T(M_i)(\omega'_i) = \text{co}(\cup_{\omega_i, h_i \in G_i(\omega_i, \omega'_i|\psi_i)} T(M_i)(\omega_i, h_i)(\omega'_i)),$$

where $T(M_i)(\omega_i, h_i)(\omega'_i) = \{m'_i| \text{ there exists } m_i \in M_i(\omega_i) \text{ such that } m'_i = B_i(m_i, h_i|\psi_{-i})\}$. Next, note that $B_i(m_i, h_i|\psi_{-i})(\omega'_i)$ is continuous in m_i on the whole domain $m_i \in \Delta^{D_{-i}}$ and $M_i(\omega_i)$ is closed (and bounded). Since $T(M_i)(\omega_i, h_i)(\omega'_i)$ is an image of a closed and bounded set under a continous mapping, it is closed (and bounded) as well. As a finite union of closed sets, $T(M_i)(\omega'_i)$ is closed as well. The same reasoning applies to the T^U operator. The observation that if m_i is an extreme point of $T^U(M_i)(\omega_i)$ but not $T(M_i)(\omega_i)$, then m_i is an extreme point of $M_i(\omega_i)$ follows directly from the definition of T^U.

For the last part of the lemma, we use an important property of the non-linear function $B_i(m_i, h_i|\psi_{-i})(\omega_{-i})$. For all ω'_{-i}, m_i^1, m_i^2, h_i and $\alpha \in (0, 1)$,

$$B_i(\alpha m_i^1 + (1-\alpha)m_i^2, h_i|\psi_{-i})(\omega'_{-i}) = \alpha' B_i(m_i^1, h_i|\psi_{-i})(\omega'_{-i}) + (1-\alpha')B_i(m_i^2, h_i|\psi_{-i})(\omega'_{-i})$$

for some $\alpha' \in (0, 1)$. That is, the posterior of a convex combination of beliefs m_i^1 and m_i^2 is a convex combination of their posteriors, albeit with different weights. To see this, algebraic manipulation delivers

$$
\begin{aligned}
&B_i(\alpha m_i^1 + (1-\alpha)m_i^2, h_i|\psi_{-i})(\omega'_{-i}) = \\
&\frac{\alpha \sum_{\omega_{-i}} m_i^1(\omega_{-i})F_i(\omega_{-i}, h_i|\psi_{-i})}{\sum_{\omega_{-i}}(\alpha m_i^1(\omega_{-i}) + (1-\alpha)m_i^2(\omega_{-i}))F_i(\omega_{-i}, h_i|\psi_{-i})} B_i(m_i^1, h_i|\psi_{-i})(\omega'_{-i}) + \\
&\frac{(1-\alpha) \sum_{\omega_{-i}} m_i^2(\omega_{-i})F_i(\omega_{-i}, h_i|\psi_{-i})}{\sum_{\omega_{-i}}(\alpha m_i^1(\omega_{-i}) + (1-\alpha)m_i^2(\omega_{-i}))F_i(\omega_{-i}, h_i|\psi_{-i})} B_i(m_i^2, h_i|\psi_{-i})(\omega'_{-i}).
\end{aligned}
$$

Note

$$
\begin{aligned}
&\frac{\alpha \sum_{\omega_{-i}} m_i^1(\omega_{-i})F_i(\omega_{-i}, h_i|\psi_{-i})}{\sum_{\omega_{-i}}(\alpha m_i^1(\omega_{-i}) + (1-\alpha)m_i^2(\omega_{-i}))F_i(\omega_{-i}, h_i|\psi_{-i})} + \\
&\frac{(1-\alpha) \sum_{\omega_{-i}} m_i^2(\omega_{-i})F_i(\omega_{-i}, h_i|\psi_{-i})}{\sum_{\omega_{-i}}(\alpha m_i^1(\omega_{-i}) + (1-\alpha)m_i^2(\omega_{-i}))F_i(\omega_{-i}, h_i|\psi_{-i})} = 1.
\end{aligned}
$$

Further, examination of the first quotient has the numerator strictly positive and strictly less than the denominator. So indeed

$$
\alpha'\left(\alpha, m_i^1, m_i^2\right) = \frac{\alpha \sum_{\omega_{-i}} m_i^1(\omega_{-i})F_i(\omega_{-i}, h_i|\psi_{-i})}{\sum_{\omega_{-i}}(\alpha m_i^1(\omega_{-i}) + (1-\alpha)m_i^2(\omega_{-i}))F_i(\omega_{-i}, h_i|\psi_{-i})} \in (0,1).
$$

Now take any m_i which is an extreme point of $T(M_i)(\omega_i)$ and suppose that for all collections (m'_i, ω'_i, h'_i) such that $m_i = B_i\left(m'_i, h'_i|\psi_{-i}\right)$, $m'_i \in M_i(\omega'_i)$ and $h'_i \in G_i(\omega'_i, \omega_i)$, the belief m'_i is not an extreme point of $M_i(\omega'_i)$. That implies that there exist two priors (m_i^0, m_i^1) that are extreme points of $M_i(\omega_i)$ such that m'_i is a strict convex combination of them. There are three possibilities: 1) $B_i\left(m'_i, h'_i|\psi_{-i}\right) = B_i\left(m_i^0, h'_i|\psi_{-i}\right)$ or 2) $B_i\left(m'_i, h'_i|\psi_{-i}\right) = B_i\left(m_i^1, h'_i|\psi_{-i}\right)$ or 3) $B_i\left(m'_i, h'_i|\psi_{-i}\right)$ is a strict convex combination of $B_i\left(m_i^0, h'_i|\psi_{-i}\right)$ and $B_i\left(m_i^1, h'_i|\psi_{-i}\right)$. In the first two cases, we have then found the priors that lead to the posterior m_i, a contradiction. In the third case, m_i is not an extreme point of $T(M_i)(\omega_i)$, again a contradiction. ∎

Proof of Lemma 3

Proof. For ω_i such that $\sum_{\overline{\omega}_{-i}} \pi(\omega_i, \overline{\omega}_{-i}) > 0$, let $m_i^0(\omega_i)(\omega_{-i}) = \frac{\pi(\omega_i,\omega_{-i})}{\sum_{\overline{\omega}_{-i}} \pi(\omega_i,\overline{\omega}_{-i})}$. That is, $m_i^0(\omega_i)$ is the single point in the set $M_{i,0}(\pi, \omega_i)$. Since π is an invariant distribution, for all $\omega = (\omega_i, \omega_{-i})$

$$\begin{aligned} m_i^0(\omega_i)(\omega_{-i}) &= \frac{\sum_{\omega^0} \pi(\omega^0) \sum_{h_i \in G_i(\omega_i^0, \omega_i|\psi_i)} \sum_{h_{-i} \in G_i(\omega_{-i}^0, \omega_{-i}|\psi_{-i})} p_i(a_i|\omega_i^0) p_{-i}(a_{-i}|\omega_{-i}^0) P(y|a)}{\sum_{\omega^0} \pi(\omega^0) \sum_{h_i \in G_i(\omega_i^0, \omega_i|\psi_i)} \sum_{h_{-i}} p_i(a_i|\omega_i^0) p_{-i}(a_{-i}|\omega_{-i}^0) P(y|a)} \\ &= \frac{\sum_{\omega_i^0} \sum_{h_i \in G_i(\omega_i^0, \omega_i|\psi_i)} p_i(a_i|\omega_i^0) \sum_{\omega_{-i}^0} \pi(\omega^0) H_i(\omega_{-i}^0, \omega_{-i}, h_i|\psi_{-i})}{\sum_{\omega_i^0} \sum_{h_i \in G_i(\omega_i^0, \omega_i|\psi_i)} p_i(a_i|\omega_i^0) \sum_{\omega_{-i}^0} \pi(\omega^0) F_i(\omega_{-i}^0, h_i|\psi_{-i})}. \end{aligned}$$

Next, note that

$$B_i(m_i^0(\omega_i^0), h_i|\psi_{-i})(\omega_{-i}) = \frac{\sum_{\omega_{-i}^0} \pi(\omega_i^0, \omega_{-i}^0) H_i(\omega_{-i}^0, \omega_{-i}, h_i|\psi_{-i})}{\sum_{\omega_{-i}^0} \pi(\omega_i^0, \omega_{-i}^0) F_i(\omega_{-i}^0, h_i|\psi_{-i})}.$$

We wish to show for all ω_i, $m_i^0(\omega_i)$ is a convex combination of $B_i(m_i^0, h_i|\psi_{-i})$ over all (ω_i^0, h_i) such that $h_i \in G_i(\omega_i^0, \omega_i|\psi_i)$. For all (ω_i^0, h_i) such that $h_i \in G_i(\omega_i^0, \omega_i|\psi_i)$, let

$$\alpha(\omega_i^0, h_i|\omega_i) = \frac{p_i(a_i|\omega_i^0) \sum_{\omega_{-i}^0} \pi(\omega_i^0, \omega_{-i}^0) F_i(\omega_{-i}^0, h_i|\psi_{-i})}{\sum_{\overline{\omega}_i^0} \sum_{\overline{h}_i \in G_i(\overline{\omega}_i^0, \omega_i|\psi_i)} p_i(\overline{a}_i|\overline{\omega}_i^0) \sum_{\omega_{-i}^0} \pi(\overline{\omega}_i^0, \omega_{-i}^0) F_i(\omega_{-i}^0, \overline{h}_i|\psi_{-i})}.$$

Since the denominator of $\alpha(\omega_i^0, h_i|\omega_i)$ is the sum of the numerators over all (ω_i^0, h_i) such that $h_i \in G_i(\omega_i^0, \omega_i|\psi_i)$, it is clear that $\sum_{\omega_i} \sum_{h_i \in G_i(\omega_i^0, \omega_i|\psi_i)} \alpha(\omega_i^0, h_i|\omega_i) = 1$.

Next, for a given ω_i and ω_{-i}, consider

$$\begin{aligned}
&\sum_{\omega_i^0} \sum_{h_i \in G_i(\omega_i^0, \omega_i|\psi_i)} \alpha(\omega_i^0, h_i|\omega_i) B_i(m_i^0(\omega_i), h_i|\psi_{-i})(\omega_{-i}) \\
= &\sum_{\omega_i^0} \sum_{h_i \in G_i(\omega_i^0, \omega_i|\psi_i)} \frac{p_i(a_i|\omega_i^0) \sum_{\omega_{-i}^0} \pi(\omega_i^0, \omega_{-i}^0) F_i(\omega_{-i}^0, h_i|\psi_{-i}) B_i(m_i^0(\omega_i), h_i|\psi_{-i})(\omega_{-i})}{\sum_{\overline{\omega}_i^0} \sum_{\overline{h}_i \in G_i(\overline{\omega}_i^0, \omega_i|\psi_i)} p_i(\overline{a}_i|\overline{\omega}_i^0) \sum_{\omega_{-i}^0} \pi(\overline{\omega}_i^0, \omega_{-i}^0) F_i(\omega_{-i}^0, \overline{h}_i|\psi_{-i})} \\
= &\frac{\sum_{\omega_i^0} \sum_{h_i \in G_i(\omega_i^0, \omega_i|\psi_i)} p_i(a_i|\omega_i^0) \sum_{\omega_{-i}^0} \pi(\omega^0) H_i(\omega_{-i}^0, \omega_{-i}, h_i|\psi_{-i})}{\sum_{\omega_i^0} \sum_{h_i \in G_i(\omega_i^0, \omega_i|\psi_i)} p_i(a_i|\omega_i^0) \sum_{\omega_{-i}^0} \pi(\omega^0) F_i(\omega_{-i}^0, h_i|\psi_{-i})} \\
= &m_i^0(\omega_i)(\omega_{-i}).
\end{aligned}$$

∎

Proof of Lemma 4

Proof. First, that the limit exists and π is a stationary distribution of τ is a standard result on Markov chains (see for example Theorem 11.1 in Stokey and Lucas).

Next, define

$$\pi_t = \frac{1}{t+1} \sum_{n=0}^{t} x\tau^n$$

Note that π_t is a probability distribution over joint states for any t (it is the distribution over joint states given starting correlation device x and the transition matrix τ, averaged over periods $\{0, ...t\}$).

We prove by induction that for all t, $M_{i,0}(\pi_t) \subset \left(T^U\right)^t (M_{i,0}(x))$ and $M_{i,0}(x\tau^t) \subset \left(T^U\right)^t (M_{i,0}(x))$ (where $\left(T^U\right)^0 (M) = M$).

For $t = 0$ the all these collections of sets are equal, so the claim is true. Now, suppose the claim is true for $t-1$.

Let $m_i^t(\omega_i)(\omega_{-i}) = \frac{\pi_t(\omega_i,\omega_{-i})}{\sum_{\overline{\omega}_{-i}} \pi_t(\omega_i,\overline{\omega}_{-i})}$ be the belief player i assigns to players $-i$ being in state ω_{-i} conditional on observing that the correlation device π_t puts him in state ω_i. Also let $\hat{m}_i^t(\omega_i)(\omega_{-i}) = \frac{(x\tau^t)(\omega_i,\omega_{-i})}{\sum_{\overline{\omega}_{-i}}(x\tau^t)(\omega_i,\overline{\omega}_{-i})}$ (analogous belief for correlation device $x\tau^t$). Note that

$$\pi_t = \frac{\sum_{n=0}^t x\tau^n}{t+1} = \frac{t\pi_{t-1} + x\tau^t}{t+1}$$

that is, π_t is a weighted average of a distributions π_{t-1} and $x\tau^t$.

By the same calculation as in Lemma 3, $\hat{m}_i^t(\omega_i)(\omega_{-i})$ is a convex combination of posterior beliefs $B_i(\hat{m}_i^{t-1}, h_i|\psi_{-i})$ over all $\left(\omega_i^{t-1}, h_i\right)$ such that $h_i \in G_i(\omega_i^{t-1}, \omega_i|\psi_{-i})$. The intuition is that $\hat{m}_i^t(\omega_i)(\omega_{-i})$ can be thought of as beliefs player i has after learning that at time t he is in state ω_i but not knowing his history of the game so far. If he knew that his belief last period was $\hat{m}_i^{t-1}$ he could then compute his posterior using that prior and averaging over all one-period histories that according to the equilibrium path could have brought him to the current state ω_i.

Since by the inductive hypothesis all priors $\hat{m}_i^{t-1}(\omega_i) \in \left(T^U\right)^{t-1}(M_{i,0}(x))(\omega_i)$, all such posteriors $\hat{m}_i^t(\omega_i) \in T\left(\left(T^U\right)^{t-1}(M_{i,0}(x))\right)(\omega_i) \subset \left(T^U\right)^t(M_{i,0}(x))(\omega_i)$.

Finally, since the correlation device π_t draws joint states either according to π_{t-1} (with probability $\frac{t}{t+1}$) or $x\tau^t$ (with probability $\frac{1}{t+1}$), the posterior satisfies

$$\begin{aligned} m_i^t(\omega_i)(\omega_{-i}) &= \frac{\pi_t(\omega_i,\omega_{-i})}{\sum_{\overline{\omega}_{-i}}\pi_t(\omega_i,\overline{\omega}_{-i})} \\ &= \frac{\frac{t}{t+1}\pi_{t-1}(\omega_i,\omega_{-i})+\frac{1}{t+1}(x\tau^t)(\omega_i,\omega_{-i})}{\sum_{\overline{\omega}_{-i}}\pi_t(\omega_i,\overline{\omega}_{-i})} \\ &= \frac{t}{t+1}\frac{\sum_{\overline{\omega}_{-i}}\pi_{t-1}(\omega_i,\overline{\omega}_{-i})}{\sum_{\overline{\omega}_{-i}}\pi_t(\omega_i,\overline{\omega}_{-i})}m_i^{t-1}(\omega_i)(\omega_{-i}) \\ &\quad +\frac{1}{t+1}\frac{\sum_{\overline{\omega}_{-i}}(x\tau^t)(\omega_i,\overline{\omega}_{-i})}{\sum_{\overline{\omega}_{-i}}\pi_t(\omega_i,\overline{\omega}_{-i})}\hat{m}_i^t(\omega_i)(\omega_{-i}) \end{aligned}$$

Since the coefficients on the two beliefs are positive and add up to one, $m_i^t(\omega_i)(\omega_{-i})$ is a convex combination of the beliefs $m_i^{t-1}(\omega_i)(\omega_{-i})$ and $\hat{m}_i^t(\omega_i)(\omega_{-i})$. Since we have shown that $\hat{m}_i^t(\omega_i) \in (T^U)^t(M_{i,0}(x))(\omega_i)$ and by the inductive hypothesis,

$$m_i^{t-1}(\omega_i) \in (T^U)^{t-1}(M_{i,0}(x))(\omega_i) \subset (T^U)^t(M_{i,0}(x))(\omega_i)$$

we conclude that $m_i^t(\omega_i) \subset (T^U)^t(M_{i,0}(x))(\omega_i)$, which finishes the proof of induction.

As $M_{i,0}(\pi_t) \subset (T^U)^t(M_{i,0}(x))$ for all t, it also holds in the limit, so indeed $M_{i,0}(\pi) \subset M_i^{*U}(M_{i,0}(x))$. ∎

Proof of Lemma 5

Proof. That $\tau(\psi)$ is a regular matrix implies that there exists an L such that for any joint states ω and ω' the players on equilibrium path move with a positive probability from state ω to ω' in exactly L periods. That implies that for any non-empty M_i (i.e. that there exists at least one ω_i such that $M_i(\omega_i)$ is non-empty), the set $T^n(M_i)(\omega_i)$ is non-empty for all $\omega_i \in \Omega_i$ for any $n \geq L$.

Next, let $\mathcal{H}(h_i)$ denote the $D_{-i} \times D_{-i}$ matrix $H_i(\omega_{-i},\omega'_{-i},h_i|\psi_{-i})$ where rows correspond to ω_{-i} and the columns to ω'_{-i}. We note that the matrix $\mathcal{H}(h_i)$ has all entries between 0 and 1 and that the rows add up to at most 1, so that if some element is positive, all other elements are strictly bounded away from 1.

Since $\tau(\psi)$ is a regular matrix and we have assumed that the set of signals players $-i$ observe with positive probability does not depend on player i actions (full support), for all $h_{i,1}...h_{i,L}$ all elements of the matrix $\mathcal{H}(h_{i,L}) * ... * \mathcal{H}(h_{i,1})$ contain no zeros (since player i assigns positive probability to the other players moving from any state to any state in L periods on the equilibrium path). Let $\varepsilon > 0$ be the lower bound on the elements of that matrix (it exists since L and the set of h_i are finite).

The rest of the proof has two steps. Let beliefs m_i^{E0} and m_i^{E1} be such that $m_i^{E0}(\omega_{-i}^0) = 1$ and $m_i^{E1}(\omega_{-i}^1) = 1$. That is, m_i^{E0} puts all probability on state ω_{-i}^0 and m_i^{E1} puts all weight on state ω_{-i}^1. First, we show that for all $\{h_{i,n}\}_{n=0}^{\infty}$, $\lim_{n\to\infty} |B_i^n(m_i^{E0}, h_i^n|\psi_{-i}), B_i^n(m_i^{E1}, h_i^n|\psi_{-i})| = 0$. Next we show this implies $\lim_{n\to\infty} T^n(M_i) = \overline{M}_i$ for all non-empty $M_i \in \mathcal{M}$.

Step 1:

Recall from Lemma 1 that

$$B_i(m_i, h_i|\psi_{-i})(\omega'_{-i}) = \frac{\sum_{\omega_{-i}} m_i(\omega_{-i}) H_i(\omega_{-i}, \omega'_{-i}, h_i|\psi_{-i})}{\sum_{\omega_{-i}} m_i(\omega_{-i}) F_i(\omega_{-i}, h_i|\psi_{-i})}$$

Let $B_i(m_i, h_i|\psi_{-i})$ denote the vector $B_i(m_i, h_i|\psi_i)(\omega'_{-i})$ and $F_i(h_i|\psi_{-i})$ denote the vector $F_i(\omega_{-i}, h_i|\psi_{-i})$. We can then re-write Bayes' rule in the matrix form as:

$$\text{(A1)} \quad B_i(m_i, h_i|\psi_{-i}) = \underbrace{\frac{1}{m_i \cdot F_i(h_i|\psi_{-i})}}_{\text{scalar}} m_i \mathcal{H}(h_i)$$

where m_i is a row vector with elements $m_i(\omega_{-i})$.

If player i starts with prior m_i^0 and observes $(h_{i,L}, ...h_{i,1})$ (with $h_{i,1}$ being the most recent observation), then his posterior beliefs after L periods are:

$$\begin{aligned} & B_i^L\left(m_i^0, h_{i,L}, \ldots, h_{i,1}|\psi_{-i}\right) \\ = & \frac{1}{B_i^{L-1}\left(m_i^0, h_{i,L}, \ldots, h_{i,2}|\psi_{-i}\right) \cdot F_i(h_{i,1}|\psi_{-i})} B_i^{L-1}\left(m_i^0, h_{i,L}, \ldots, h_{i,2}|\psi_{-i}\right) \mathcal{H}\left(h_{i,1}\right) \\ = & \frac{1}{\left(m_i^0 \mathcal{H}\left(h_{i,L}\right) \ldots \mathcal{H}\left(h_{i,2}\right)\right) \cdot F_i(h_{i,1}|\psi_{-i})} m_i^0 \mathcal{H}\left(h_{i,L}\right) \ldots \mathcal{H}\left(h_{i,1}\right) \end{aligned}$$

This implies for $j \in \{0, 1\}$, $B_i^L(m_i^{E_j}, h_{i,L}, \ldots, h_{i,1}|\psi_{-i})$ is equal to the ω_{-i}^j row of matrix

$$\frac{1}{\left(m_i^{E_j} \mathcal{H}\left(h_{i,L}\right) \ldots \mathcal{H}\left(h_{i,2}\right)\right) \cdot F_i\left(h_{i,1}|\psi_{-i}\right)} \mathcal{H}\left(h_{i,L}\right) \ldots \mathcal{H}\left(h_{i,1}\right)$$

For a matrix Q let $R_l^Q = \sum_k q_{lk}$ be the sum of the elements of row l of this matrix. Denote by $R(Q)$ a matrix obtained by dividing each element of matrix Q by the corresponding R_l^Q, that is if $B = R(Q)$ then $b_{lk} = \frac{q_{lk}}{R_l^Q}$. By definition the rows of $R(Q)$ add up to 1. Hence, $R(\mathcal{H}\left(h_{i,L}\right) ...\mathcal{H}\left(h_{i,1}\right))$ is a probability matrix and the posterior belief $B_i^L\left(m_i^{E0}, h_{i,L}, \ldots, h_{i,1}|\psi_{-i}\right)$ is equal to the ω_{-i}^0 row of $R(\mathcal{H}\left(h_{i,L}\right) ...\mathcal{H}\left(h_{i,1}\right))$.

Let $d_k(Q)$ be the difference between the largest and smallest elements of $Q's$ column k : $d_k(Q) = \max_{l,j}\left(q_{lk} - q_{jk}\right))$. Let $d\left(Q\right)$ be the vector of these differences. Then $\max_{\omega'_{-i}} d(R(\mathcal{H}\left(h_{i,L}\right) ...\mathcal{H}\left(h_{i,1}\right)))(\omega'_{-i})$ is the maximum distance of the posterior beliefs $B_i^L\left(m_i^{E0}, h_{i,L}, \ldots, h_{i,1}|\psi_{-i}\right)$ and $B_i^L\left(m_i^{E1}, h_{i,L}, \ldots, h_{i,1}|\psi_{-i}\right)$ over all extreme priors, m_i^{E0} and m_i^{E1}. To continue, we invoke the following technical lemma (proven below):

Technical Lemma:

Suppose that $\{Q_n\}_{n=1}^{\infty}$ *is a sequence of square matrices with all elements* $q_{nij} \in (\varepsilon, 1-\varepsilon)$

for some $\varepsilon > 0$. *Then there exists a* $\delta \in (0,1)$ *such that for every* n :

$$d\left(R(Q_n...Q_1)\right) \leq \delta d\left(R(Q_{n-1}...Q_1)\right) \leq \delta^{n-1} d\left(R(Q_1)\right)$$

i.e. the distance between the normalized rows of $Q_n...Q_1$*contracts by a factor at least* δ *as we left-multiply it by another matrix from the sequence.*

Now, since there exists $L \geq 1$ and $\varepsilon > 0$ such that for all $(h_{i,L}, ...h_{i,1})$ all elements of $\mathcal{H}\left(h_{i,L}\right)...\mathcal{H}\left(h_{i,1}\right)$ are bounded between $(\varepsilon, 1-\varepsilon)$, this technical lemma implies that there exists a $\delta \in (0,1)$ such that for any integer n :

$$d(R(\mathcal{H}\left(h_{i,nL}\right)...\mathcal{H}\left(h_{i,1}\right))) \leq \delta\ d(R(\mathcal{H}\left(h_{i,(n-1)L}\right)...\mathcal{H}\left(h_{i,1}\right))) \leq \delta^{n-1}\mathbf{1}$$

where $\mathbf{1}$ is a vector of ones (of length D_{-i}). Therefore, for any ε' we can find n large enough so that for any history of length nL and any two extreme priors, m_i^{E0} and m_i^{E1}, the distance between the posteriors will be less than ε'. So, for every history h_i^n, as $n \to \infty$, the posteriors converge to the same belief for all extreme priors.

Step 2:

As we have shown in the proof of Lemma 3, beliefs $B_i\left(m_i^0, h_i|\psi_{-i}\right)$ are a convex combination of beliefs $B_i\left(m_i^E, h_i|\psi_{-i}\right)$ of all extreme priors m_i^E. Applying this reasoning iteratively (that if prior belief m_i is a convex combination of priors m_i' and m_i'', then after applying B_i the posterior of m_i is a convex combination of the posteriors of m_i' and m_i''), we get that for any history sequence, the posteriors after all possible beliefs are convex combinations of posteriors $B_i^L\left(m_i^E, h_{i,L}, \ldots, h_{i,1}|\psi_{-i}\right)$. Since for any sequence $\left\{h_i^L\right\}_{L=1}^{\infty}$, for all m_i^E the posteriors $B_i^L\left(m_i^E, h_{i,L}, \ldots, h_{i,1}|\psi_{-i}\right)$ converge, the same is true for posteriors after arbitrary priors. In other words, after long enough histories, the posteriors depend (almost) only on the history and not on the prior.

As we described in the text, by the Tarski's fixed point theorem T has at least one fixed point, $\overline{M}_i$. Now, suppose there exists a collection of sets M_i^0 such that $\lim_{n\to\infty} T^n\left(M_i^0\right) \neq \overline{M}_i$ (either because the sequence$\{T^n\left(M_i^0\right)\}_{n=0}^{\infty}$ converges to something else or does not converge at all).

By monotonicity of T, for all n, $T^n\left(M_i^0\right) \subset T^n\left(\overline{\Delta}_i\right)$. Since $T^n\left(\overline{\Delta}_i\right)$ converges to $\overline{M}_i$, for any $\varepsilon > 0$ we can find n large enough so that for all $\omega_i \in \Omega_i$ and all $m_i \in T^n\left(M_i^0\right)(\omega_i)$, $\left|m_i, \overline{M}_i(\omega_i)\right| < \varepsilon$. That is, the sets $T^n\left(M_i^0\right)$ cannot "stick out" of $\overline{M}_i$ in the limit.

So the only remaining possibility for $\lim_{n\to\infty} T^n\left(M_i^0\right) \neq \overline{M}_i$ is that there exists $\varepsilon > 0$ such that for all n' we can $n \geq n'$ and a state ω_i^n such that $\max_{m_i \in \overline{M}_i(\omega_i^n)} \left|T^n\left(M_i^0\right)(\omega_i^n), m_i\right| > \varepsilon$ (in words, that the set $\overline{M}_i(\omega_i^n)$ strictly "sticks out" of the set $T^n\left(M_i^0\right)(\omega_i^n)$ even for arbitrary large n). If so, then we can find an extreme belief $m_i^n \in \overline{M}_i(\omega_i^n)$ that satisfies $\left|m_i^n, T^n\left(M_i^0\right)(\omega_i^n)\right| > 0$. Fix n' such that the distance between $B_i^n\left(m_i^{E0}, h_i^n|\psi_{-i}\right)$ and $B_i^n\left(m_i^{E1}, h_i^n|\psi_{-i}\right)$ is uniformly bounded by $\varepsilon/2$ for all histories h_i^n (for all $n > n'$) and all extreme points m_i^{E0}, m_i^{E1}. Since $\lim_{n\to\infty} T^n\left(\overline{\Delta}_i\right) = \overline{M}_i$, we can find a history h_i^n and a prior m_i^{E0} such that $\left|B_i^n\left(m_i^{E0}, h_i^n|\psi_{-i}\right), m_i^n\right| \leq \varepsilon/2$ and a starting state ω_i^0 such that after that history, player i is in the state ω_i^n. Now, take any prior $m_i^0 \in M_i^0\left(\omega_i^0\right)$. It is a convex combination of the priors m_i^E. Moreover, after the history h_i^n, the posterior $B_i^n\left(m_i^0, h_i^n|\psi_{-i}\right) \in T^n\left(M_i^0\right)(\omega_i^n)$ and it is a convex combination of the posteriors $B_i^n\left(m_i^E, h_i^n|\psi_{-i}\right)$. (The last claim follows from inspection of (A1) - see also Lemma 2). Therefore:

$$\left|B_i^n\left(m_i^0, h_i^n|\psi_{-i}\right), B_i^n\left(m_i^{E0}, h_i^n|\psi_{-i}\right)\right| \leq \max_{m_i^{E1}, m_i^{E2}} \left|B_i^n\left(m_i^{E1}, h_i^n|\psi_{-i}\right), B_i^n\left(m_i^{E2}, h_i^n|\psi_{-i}\right)\right| \leq \varepsilon/2$$

Using the triangle inequality, $\left|B_i^n\left(m_i^0, h_i^n|\psi_{-i}\right), m_i^n\right| \leq \varepsilon$, but that contradicts that $|m_i^n, T^n(M_i^0)(\omega_i^n)| > \varepsilon$. ∎

Proof of Technical Lemma.

Proof. Consider a general multiplication: $Q = Q_n...Q_1$. Let $C = Q_n$, $F = Q_{n-1}$, $B = Q_{n-2}...Q_1$. Also, let $G = FB$, so that $Q = CG = CFB$. By assumption all the elements of C and F are bounded from below by $\varepsilon > 0$, but we do not know that about B or G.

For arbitrary matrix A, let R_k^A be the sum of elements in row k of that matrix. Then:

$$R_i^Q = \sum_j q_{ij} = \sum_j \left(\sum_k c_{ik} g_{kj}\right) = \sum_k c_{ik} \sum_j g_{kj} = \sum_k c_{ik} R_k^G$$

Moreover,

$$\frac{q_{ij}}{R_i^Q} = \sum_k \Gamma_k^i \frac{g_{kj}}{R_k^G}$$

where

$$\Gamma_k^i = \frac{c_{ik} R_k^G}{\sum_l c_{il} R_l^G}$$

In words, the elements of $R(Q_n G)$ are a weighted average of elements of $R(G)$ (note that $\sum_k \Gamma_k^i = 1$).

We now bound the weights Γ_k^i uniformly away from zero for all G. To this end, bound

$$\Gamma_k^i = \frac{c_{ik} R_k^G}{\sum_l c_{il} R_l^G} > c_{ik} \frac{R_k^G}{\sum_l R_l^G}$$

Next,

$$\begin{aligned}
\frac{R_i^G}{\sum_l R_l^G} &= \frac{\sum_k f_{ik} R_k^B}{\sum_l \sum_k f_{lk} R_k^B} = \frac{\sum_k f_{ik} R_k^B}{\sum_k \sum_l f_{lk} R_k^B} = \frac{\sum_k f_{ik} R_k^B}{\sum_k R_k^B L_k^F} \\
&= \sum_k \frac{f_{ik}}{L_k^F} \frac{L_k^F R_k^B}{\sum_k R_k^B L_k^F} = \sum_k \frac{f_{ik}}{L_k^F} \gamma_k
\end{aligned}$$

where L_k^F is the sum of elements of column k of matrix F and

$$\gamma_k = \frac{L_k^F R_k^B}{\sum_k R_k^B L_k^F} \in [0, 1] \, .$$

Note that for any matrices F and B, $\sum_k \gamma_k = 1$.

Therefore we can find a bound $\varepsilon_L \in \left(0, \frac{1}{2}\right)$ that depends only on F and C :

$$\Gamma_k^i \geq c_{ik} \frac{R_k^G}{\sum_l R_l^G} \geq \varepsilon \min_k \frac{f_{ik}}{L_k^F} > \varepsilon_L$$

where ε_L can be chosen independently of i and k.

To finish the proof we show how to choose δ. Consider any column k. Any element of column k of matrix $R\left(Q_n ... Q_1\right)$ is a weighted average of elements in the same column of $R\left(Q_{n-1} ... Q_1\right)$, with the weights bounded uniformly away from zero by ε_L. Suppose that the largest and smallest elements of column k of $R\left(Q_{n-1} ... Q_1\right)$ are equal to q_h and q_l respectively.

Then

$$d_k\left(R\left(Q_n...Q_1\right)\right) \leq \left(1-\varepsilon_L\right)q_h + \varepsilon_L q_l - \left(\varepsilon_L q_h + \left(1-\varepsilon_L\right)q_l\right) = \left(1-2\varepsilon_L\right)d_k\left(R\left(Q_{n-1}...Q_1\right)\right).$$

So we can pick $\delta = (1-2\varepsilon_L)$. ∎

Proof of Lemma 7.

Proof. Let $\overline{V}_{i,0}(\sigma_i, \sigma_{-i})$ be the discounted expected payoff to player i when he plays σ_i and his opponents σ_{-i}. Let $\sigma_i(\omega_i, \psi_i)$ denote the strategy associated with following automaton ψ_i from initial state ω_i. That (x, ψ) is not a CNE implies there exists a player i, strategy $\hat{\sigma}_i$, and initial state $\overline{\omega}_i$ such that $\sum_{\omega_{-i}} x(\overline{\omega}_i, \omega_{-i}) > 0$, such that

(A2) $E[\overline{V}_{i,0}(\hat{\sigma}_i, \sigma_{-i}(\omega_{-i}, \psi_{-i})) - \overline{V}_{i,0}(\sigma_i(\overline{\omega}_i, \psi_i), \sigma_{-i}(\omega_{-i}, \psi_{-i}))|\overline{w}_i] > 0.$

Partition $(\sigma^1, \ldots, \sigma^J)$ as follows. Let $\Lambda(\omega)$ be the set of $\sigma^j \in (\sigma^1, \ldots, \sigma^J)$ such that the path of play of σ^j is the same as the path of play implied by the joint automaton ψ starting from the joint state ω. (Likewise let $\Lambda_i(\omega_i)$ be the set of $\sigma_i^j \in (\sigma_i^1, \ldots, \sigma_i^J)$ such that the path of play of σ_i^j is the same as the path of play implied by automaton ψ_i starting from private state ω_i.) Since $\sigma^j \in \Lambda(\omega)$ has the same path play as $\sigma(\omega, \psi)$, for all $\sigma^j \in (\sigma^1, \ldots, \sigma^J)$ such that $\sigma^j \in \Lambda(\omega)$

$$\begin{aligned}
& [\overline{V}_{i,0}(\hat{\sigma}_i, \sigma_{-i}^j) - \overline{V}_{i,0}(\sigma_i^j, \sigma_{-i}^j)] \\
= \; & [\overline{V}_{i,0}(\hat{\sigma}_i, \sigma_{-i}(\omega_{-i}, \psi_{-i})) - \overline{V}_{i,0}(\sigma_i^j, \sigma_{-i}(\omega_{-i}, \psi_{-i})] \\
= \; & [\overline{V}_{i,0}(\hat{\sigma}_i, \sigma_{-i}(\omega_{-i}, \psi_{-i})) - \overline{V}_{i,0}(\sigma_i(\overline{\omega}_i, \psi_i), \sigma_{-i}(\omega_{-i}, \psi_{-i})].
\end{aligned}$$

Thus for all $\sigma_i^j \in \Lambda_i(\overline{\omega}_i)$,

(A3) $E[\overline{V}_{i,0}(\hat{\sigma}_i, \sigma_{-i}^j) - \overline{V}_{i,0}(\sigma_i^j, \sigma_{-i}^j)|\Lambda_i(\overline{w}_i)] > 0.$

That is, if player i plays $\hat{\sigma}_i$ as opposed to σ_i^j, but conditions his expectation regarding σ_{-i}^j only on the fact that $\sigma_i^j \in \Lambda(\overline{\omega}_i)$, rather than σ_i^j itself, his deviation is profitable. From the law of iterated expectations,

$$E[\overline{V}_{i,0}(\hat{\sigma}_i, \sigma_{-i}^j) - \overline{V}_{i,0}(\sigma_i^j, \sigma_{-i}^j)|\Lambda_i(\overline{w}_i)] = E[E[\overline{V}_{i,0}(\hat{\sigma}_i, \sigma_{-i}^j) - \overline{V}_{i,0}(\sigma_i^j, \sigma_{-i}^j)|\sigma_i^j]|\Lambda_i(\overline{w}_i)] > 0.$$

From the mean value theorem, there must exist $\sigma_i^j \in \Lambda_i(\overline{\omega}_i)$ such that the inner expectation is positive, implying $(\hat{x}, (\sigma^1, \ldots, \sigma^J))$ is not a CNE. ∎