

# *Unattainable Payoffs for Repeated Games of Private Monitoring\**

Joshua Cherry<sup>†</sup>

Economics Department  
University of Michigan

Lones Smith<sup>‡</sup>

Economics Department  
University of Wisconsin

June 13, 2011

## **Abstract**

We tightly bound from the outside the set of sequential equilibrium payoffs in repeated games of private monitoring. To do this, we develop a tractable new solution concept for standard repeated games with perfect monitoring: *Markov Perfect Correlated Equilibrium* generalizes the operator approach of Abreu, Pearce, and Stacchetti (1990), but instead takes correlated equilibrium of the auxiliary game. We show that *for any private monitoring structure*, the set of sequential equilibrium payoffs of a repeated game is contained within the set of Markov Perfect Correlated Equilibrium payoffs of the repeated expected game. This bound can be made tight with a simple two-stage procedure.

The techniques we develop are tractable and shed light on all economic settings with imperfectly observed actions, like dynamic oligopoly, long-term partnerships, and relational contracting. In all cases, they yield the sharpest equilibrium payoff prediction that is agnostic about the monitoring structure.

---

\*We are grateful for helpful comments from Pavlo Prokopovych (whom Lones advised in his 2006 Michigan PhD) and seminar participants at the Western Economic Association Meetings, Midwest Theory Conference, NSF/NBER/CEME Conference, and the University of Michigan.

<sup>†</sup>email: jscherry@umich.edu. Joshua is thankful for the financial support of NSF IGERT grants 0114368 and 0654014.

<sup>‡</sup>email: lones@ssc.wisc.edu. Lones is grateful for the ongoing financial support of the NSF that has supported this work.

# 1 Introduction

A repeated game is a stylized model of a long-term relationship. The most common solution concepts for repeated games are Subgame Perfect Equilibrium (SPE) and its extension to imperfect public monitoring, Perfect Public Equilibrium (PPE). In both cases, equilibrium strategies depend only on commonly observed histories. This yields a recursive property that *every continuation game is equivalent to the entire game*. Abreu, Pearce, and Stacchetti (APS) pursued this logic in 1986 and 1990, and so characterized equilibrium payoffs using methods inspired by dynamic programming.

APS built on Green and Porter’s 1984 seminal exposition of dynamic Cournot oligopoly — who in turn took inspiration from Stigler’s influential (1964) theory of dynamic Bertrand oligopoly. To sustain collusion in a world with hidden pricing, Stigler had proposed that firms initiate a price war if standard statistical tests suggested cartel cheating. Stigler struggled with a problem that afflicts much of economic theory — i.e, *any* dynamic setting with unobserved payoff-relevant actions that do not just feed into an observable stochastic aggregate, like an observed price. It is arguably important in all long-term partnerships ranging from relational contracting to international political relations. Restricting attention to public signals intuitively ignores strategically relevant information, and misses the potential richness of the dynamic structure. Upon reflection, public monitoring should only be justified as a tractable *approximation* of this richer private monitoring “reality”. So then, how restrictive is it? Exactly how much does private monitoring expand the scope for collusion in oligopoly, eg? Our finding is: substantially, in some cases.

Unfortunately, not only is private monitoring an interesting problem, it is also a difficult one. And thus Stigler’s question remains unsolved after so much time. For as is well-known, private monitoring in repeated games induces correlated private histories, and this frustrates attempts to use recursive methods, as in APS. On a sequential equilibrium path, continuation play in any period constitutes a correlated equilibrium, where the private histories act as endogenous correlation devices. And computing a best response in a non-trivial sequential equilibrium may well require an impossibly complicated probabilistic inference. It is no surprise that Kandori (2002) calls this “one of the best known long-standing open questions in economic theory.”

Though incentives are harder to provide with unobserved actions, the induced correlation may facilitate coordination (as in Aumann 1974, 1987), and augment the sequential equilibrium payoff set. So motivated, we explicitly incorporate correlated

private histories, as first studied by Lehrer (1992). But our approach admits arbitrary correlation each period. First, we develop a new solution concept for infinitely repeated games with perfect monitoring that reflects these correlation possibilities. Whereas APS defined an operator that took the Nash equilibria of the ‘auxiliary game’ at the start of every subgame, we take correlated equilibria. This is a well-defined exercise since we publicize past correlated recommendations. The largest fixed point of the resulting operator yields the *Markov Perfect Correlated Equilibrium* (MPCE) payoff set, and is therefore recursive like PPE. Notably, not only is our solution concept tractable, it is arguably easier to compute than is the PPE set. For unlike Nash equilibrium, the set of correlated equilibria can be found by linear methods.

We then explore the implications of MPCE for repeated games of private monitoring. We show that for any monitoring structure, the set of sequential equilibrium payoffs is contained within the MPCE payoff set for the corresponding *expected stage game*. This helps us deduce the tightest bound on repeated game equilibrium payoffs that is independent of the monitoring structure.

Our paper has two parts. We begin with an infinitely repeated game of observed actions, and embellish it with an extensive-form correlation device that can generate any (possibly) history-dependent private messages every period. Since messages are made public after players act, a recursive structure emerges. Unlike Prokopovych (2006) who first took this road, we then show that a Markovian device suffices to describe all equilibrium payoffs. This yields our MPCE solution concept. Theorem 1 characterizes the resulting payoff set — it is compact, convex, and nondecreasing in the discount factor. Also, it contains all subgame perfect payoffs. Theorem 2 describes a tractable, recursive algorithm for computing it.

In the second thrust, we turn to a repeated game of private monitoring, and relate its sequential equilibria to the MPCE of the corresponding repetition of the expected stage game. Theorem 3 asserts that this set serves as an upper bound for the sequential equilibrium payoffs. We thereby identify the certainly unattainable sequential equilibrium payoffs for a repeated game of private monitoring for any fixed discount factor. Notably this bound holds for *all* monitoring structures, as well as private strategies in public monitoring games. In other words, we precisely compute the set of payoffs potentially added by the richer information structure introduced by private monitoring — one possible completion of Stigler’s original thought.

Theorem 4 explores how our payoff upper bound can be made tight. For unlike MPCE, a standard repeated game of private monitoring with an initial period does

not allow any pre-play signals. So motivated, we augment the MPCE concept. We compute the Nash equilibrium payoffs of all auxiliary games using continuation payoffs drawn from the MPCE set. Put differently, this applies the APS operator to the MPCE payoff set. Any payoff in the resulting set can be supported as a sequential equilibrium in a repeated game with some monitoring structure. We therefore obtain the *tightest possible bound* that makes no reference to the monitoring structure.

Research on repeated games with private monitoring has been driven by the folk theorem, and so proceeded by finding computable classes of sequential equilibria. In contrast, we provide a *superset* of the equilibrium payoff set. The earliest work found nearly efficient equilibria that dispense with all but a simple summary of past play. Loosely, these “belief-based” approaches focus on the chance of misleading private messages. This is possible when the monitoring is sufficiently accurate (e.g. Sekiguchi 1997, and Bhaskar and Obara 2002). A clever and recursive set of non-trivial equilibria in which players’ beliefs are irrelevant was later identified by Piccione (2002) and Ely and Valimaki (2002), and greatly extended by Ely, Horner, and Olszewski (2005). While this belief-free approach constitutes a strict subset of all sequential equilibrium payoffs and requires sufficiently patient players, it often secures a folk theorem.

Our paper is not intended *in any way* as a contribution to the folk theorem literature. For we shift from characterizing what is a sequential equilibrium, to what is *not*. Abreu, Milgrom, and Pearce (1991) call into question the relevance of a folk theorem in this setting. Since a discounted repeated game unjustifiably entwines time preference and the frequency of monitoring, the discrete time folk theorem logic yields more informative monitoring with more rapid play. A large discount factor is an appropriate modeling choice only if opportunities to observe others’ actions are frequent. Though Coca Cola and Pepsi can change prices arbitrarily often, without similarly (and implausibly) frequent reports of their rivals’ actions, they will change behavior only as often as information arrives. The analysis of dynamic oligopoly in Green and Porter (1984) was meaningful precisely because of the fixed discount factor. Our analysis sheds light on equilibrium payoffs when the folk theorem does not apply — such as when interaction is not very frequent, or when information revelation about unobserved actions inherently cannot be accelerated. Instead our paper offers definitive insights on payoffs for those applications with a fixed discount factor.

In arguing that the Cournot-Nash outcome was the wrong benchmark for deducing collusion Porter (1983) wrote: “Industrial organization economists have recognized for some time that the problem of distinguishing empirically between collusive and

noncooperative behavior, in the absence of a ‘smoking gun’, is a difficult one.” Firms can achieve higher payoffs in a fully compliant, noncooperative fashion. Combining this insight with our approach, we allow that firms might avail themselves of correlated information, and potentially achieve more outcomes. Our MPCE solution concept is agnostic about the details of who knows what and when. In this way, MPCE is a better litmus test of cheating for regulators to rule out the possibility of collusion; otherwise, one might mistakenly assert an antitrust violation.

The paper is organized as follows. We gently begin with a motivational example. Next, we discuss infinitely repeated games of perfect monitoring with an extensive form correlation device, and develop our new MPCE solution concept. We illustrate it by returning to our example. We then formally describe infinitely repeated games with private monitoring, and compare their sequential equilibrium payoffs with the MPCE payoffs of standard repeated games. Here, we establish our payoff upper bound and show that it can be tight. All proofs are in the Appendix.

## 2 Motivational Example

### A. Analysis of a Repeated Prisoners’ Dilemma.

Consider an infinitely repeated two player game of perfect monitoring with payoffs given by Figure 1. The players share the discount factor  $3/4$ , and so are not patient enough to support the cooperative outcome in a subgame perfect equilibrium. Stahl (1991) shows that even with public correlation, the set of SPE payoffs is the convex hull of  $\{(0, 0), (7, 0), (0, 7)\}$ , and thus the highest symmetric subgame perfect equilibrium payoff is  $(7/2, 7/2)$ . If instead we have imperfect public monitoring, then from Kandori (1992) the PPE payoff set is even smaller.<sup>1</sup>

	<i>C</i>	<i>D</i>
<i>C</i>	(4,4)	(-13,20)
<i>D</i>	(20,-13)	(0,0)

Figure 1: Example Stage Game

Next, suppose that players privately observe a payoff irrelevant signal from  $\{g, b\}$

---

<sup>1</sup>Kandori (1992) shows that the PPE set is monotone in the informativeness (in the sense of Blackwell (1953)) of the public signal. Specifically, the PPE payoff set weakly shrinks when the public signal is garbled.

before play each period. The signal profiles  $\{(g, g), (g, b), (b, g)\}$  occur with probabilities  $(1/2, 1/4, 1/4)$ , respectively, independently of the past. After actions are chosen, the private signal profile is commonly revealed to both players. To simplify matters, assume players can access a public correlation device that draws a number  $z$  from a uniform distribution on  $[0, 1]$ .

Consider the strategy profile: “In phase 1, play  $C$  after observing  $g$ , and  $D$  after  $b$ . If agents play the same action, then repeat phase 1. Otherwise, if player  $i = 1, 2$  alone plays  $D$ , then proceed to phase 2- $i$ , where player  $i$  plays  $C$  and player  $-i$  mixes so that  $i$  gets an expected payoff of 0. If both players play  $C$ , then stay in phase 2- $i$ . Otherwise, return to phase 1.”

When the repeated game is enriched by the signal process, these strategies constitute a sequential equilibrium. The equilibrium payoff for each player is

$$v = (1/4)(4(1/2) - 13(1/4) + 20(1/4)) + (3/4)(v(1/2) + 2v(1/4) + 0(1/4))$$

i.e.  $v = 15/4$ . When called upon to play  $C$ , a player will acquiesce because

$$(1/4)(4(1/2) - 13(1/4)) + (3/4)((15/4)(1/2) + 2(15/4)(1/4)) \geq (1/4)20(1/2)$$

At the start of phase 1, both players expect the payoff  $15/4$ . In phase 2- $i$ , player  $i$  expects a payoff of 0 and player  $-i$  expects  $15/2$ . The payoff  $(15/4, 15/4)$  Pareto dominates the highest symmetric subgame perfect equilibrium payoff  $(7/2, 7/2)$  attainable without any signals. In fact,  $(15/4, 15/4)$  dominates any symmetric PPE payoff attainable under any imperfect public monitoring structure. Nevertheless,  $(15/4, 15/4)$  can be attained in an MPCE because both the information and strategies depend only on the most recent period.

This example reflects two truths: (a) relative to public monitoring, private monitoring may greatly expand the set of sequential equilibrium payoffs, and (b) MPCE captures these richer information structures and the larger payoff set.

For a bigger picture insight, consider the intuition in Kandori (2011). Although correlation cannot enhance play in the one-shot prisoners’ dilemma, the repeated game instead confronts players with a game of chicken. This auxiliary game admits nontrivial correlated equilibria. Thus, imperfectly correlated signals can have a meaningful dynamic strategic effect.

More specifically, in this game the gain to defecting is higher when the other

player cooperates than when he defects since  $20 - 4 = 16 > 0 - (-13) = 13$ . But our correlating signal confuses the players about what action profile is played in any period. Consequently, the temptation to cheat is a weighted average of 16 and 13, and so smaller than if no correlation were available. This correlation is not without a cost, since the equilibrium prescribes the most efficient payoff (4, 4) less often.

## B. Economic Settings

We now argue that this example captures a wide range of economic settings.

**REPEATED PARTNERSHIP.** A theorist and an empiricist seek to write a paper together. At the start of each day, they independently choose whether to exert high effort or low effort (actions  $C$  and  $D$  in the example, respectively). They meet at the end of every day to demonstrate their accomplishments. Suppose, however, that they entertain subjective interpretations of their colleague's effort (as in MacLeod (2003) and Fuchs (2007)). Each colleague entertains either a good ( $g$ ) or bad ( $b$ ) subjective interpretation, corresponding to high or low effort by his colleague, respectively. For example, the empiricist's regression output is commonly observed, but the theorist cannot accurately gauge the effort required to produce the results. A key additional source of discounting here is that the partnership might end.

**PRINCIPAL-AGENCY.** An employee chooses each period to exert either high or low effort (actions  $C$  and  $D$  in the example, respectively). His manager simultaneously chooses one of two compensation schemes: pay a bonus for high output, or never pay a bonus (actions  $C$  and  $D$  in the example). The private signals can take one of the following two interpretations. In the first, private signals are non-binding recommendations to managers and employees made by a board of directors. The board's fiduciary duty to maximize shareholder value would justify influencing the relational contracts implemented by the firm. In the second, the private signals are subjective evaluations of output made by the agents. MacLeod (2003) characterizes the optimal contract when the joint density of the subjective evaluations is given. With MPCE one can study this context while being agnostic about the exact structure of the agents' subjectivity.

**DYNAMIC QUALITY CHOICE.** A single product firm has one long-run customer and can use higher or lower quality inputs (actions  $C$  and  $D$ ). A product with better inputs yields higher performance. Without observing the firm's input choice, the

customer decides whether or not to purchase the item (actions  $C$  and  $D$ ). After each period, the firm and the customer each observe a private signal indicating a good ( $g$ ) or bad ( $b$ ) performing product.

SECRET PRICE CUTS. Thus the actions  $C$  and  $D$  in the inspirational example from Stigler (1964) represent high and low prices, while the private signals  $g$  and  $b$  may correspond to high and low demand.

### 3 A Mediated Repeated Game

We begin with a repeated game of perfect monitoring  $G(\delta)$ , played in periods  $1, 2, \dots$ , and payoffs discounted by the factor  $0 < \delta < 1$ . Each period, every player  $i \in N = \{1, 2, \dots, n\}$  chooses an action  $a_i$  from a finite action set  $A_i$ . An action profile  $a$  is thus an element of  $A = \prod_i A_i$ , the set of pure action profiles.<sup>2</sup> Payoffs given the action profile  $a$  are  $u(a) = (u_1(a), \dots, u_n(a))$ . Let  $\alpha_i$  denote the mixed action for player  $i$  that chooses action  $a_i \in A_i$  with chance  $\alpha_i(a_i)$ . Abusing notation,  $u(\alpha) = (u_1(\alpha), \dots, u_n(\alpha))$  denotes the expected payoffs from the mixture  $\alpha$ . As usual, this stage game has a Nash equilibrium. Let  $V$  be its set of feasible and individually rational payoffs.

We embellish the infinitely repeated game  $G(\delta)$  with a correlation device that sends private messages to players each period conditional on the action history. The device makes public the private message profile *after* play concludes each period. Before each period (including the first), each player privately receives a message  $\tilde{a}_i \in A_i$ , which we interpret as a recommendation to play action  $a_i$ . By Aumann (1987), restricting messages to recommendations is without loss of generality.<sup>3</sup> Players commonly observe the null history  $\mathfrak{h}^1 = \emptyset$  before play begins. A history  $\mathfrak{h}^t = (a^1, \tilde{a}^1, \dots, a^{t-1}, \tilde{a}^{t-1})$  is a complete record of all past *outcomes* in periods  $1, 2, \dots, t-1$ , i.e. pairs of action and recommendation profiles. The history  $\mathfrak{h}^t$  is commonly observed by all players at the start of period  $t$ . Let  $\mathfrak{H}^t$  be the set of all histories  $\mathfrak{h}^t$ , and  $\mathfrak{H} = \bigcup_{t=1}^{\infty} \mathfrak{H}^t$  the set of all histories of any length.

A (direct) correlation device  $\mu$  is a probability measure on the set of action profiles  $A$ . An *extensive form correlation device* is a sequence of functions  $\lambda = (\lambda^t)_{t=1}^{\infty}$

---

<sup>2</sup>Throughout, subscripts will denote players and superscripts will denote periods. Let  $|X|$  denote the cardinality of  $X$ . Also, we parse any vector  $x \equiv (x_i, x_{-i})$ . Since we consider finite action and signal sets, all functions thereon are measurable.

<sup>3</sup>This can equivalently be justified by the Revelation Principle. In our finite model, the Revelation Principle holds since there cannot be issues with the measurable composition of functions.

such that  $(\lambda^t : \mathbb{H}^t \rightarrow \Delta(A))_{t=1}^\infty$ , and  $\Lambda$  is the space of all such functions.<sup>4</sup> The interpretation is that after history  $\mathbb{h}^t$ , the correlation device selects an action profile  $\tilde{a} = (\tilde{a}_1, \dots, \tilde{a}_n) \in A$  according to the distribution  $\lambda(\mathbb{h}^t)$  and privately informs each player  $i$  of his recommended action  $\tilde{a}_i$ . Players then simultaneously choose actions. Finally, the recommendations are revealed to all players, and they become part of the next history  $\mathbb{h}^{t+1}$ . Finally, let  $G^\lambda(\delta)$  be the infinitely repeated mediated game with stage game  $G$ , extensive form correlation device  $\lambda \in \Lambda$ , and discount factor  $0 < \delta < 1$ .

A (behavior) strategy  $\mathfrak{s}_i$  for player  $i$  is a sequence  $(\mathfrak{s}_i^t)_{t=1}^\infty$ , where  $\mathfrak{s}_i^t : \mathbb{H}_i^t \times A_i \rightarrow \Delta(A_i)$  for every period  $t = 1, 2, \dots$ . So a strategy assigns a mixed action to every pair of history and recommendation. For any strategy profile  $(s_1, \dots, s_n) = \mathfrak{s} \in \mathbb{S} = \prod_{i \in N} \mathbb{S}_i$ , correlation device  $\lambda$ , and history  $\mathbb{h}^t$ , the payoff for player  $i$  is the present value of future payoffs:

$$v_i^t(\mathfrak{s} | \mathbb{h}^t, \lambda) = (1 - \delta) E \left[ \sum_{r=t}^{\infty} \delta^{r-t} u_i(a^r) \middle| \lambda, \mathfrak{s}, \mathbb{h}^t \right]$$

A strategy profile  $\mathfrak{s}$  is a sequential equilibrium of  $G^\lambda(\delta)$  if in every period  $t$ , history  $\mathbb{h}^t$ , and alternative strategy  $\tilde{\mathfrak{s}}_i$ ,

$$v_i^t(\mathfrak{s} | \mathbb{h}^t, \lambda) \geq v_i^t(\tilde{\mathfrak{s}}_i, \mathfrak{s}_{-i} | \mathbb{h}^t, \lambda)$$

## 4 Markov Perfect Correlated Equilibrium

If  $\mathfrak{s} \in \mathbb{S}$  is a sequential equilibrium strategy profile of  $G^\lambda(\delta)$ , then Prokopovych (2006) calls the pair  $(\mathfrak{s}, \lambda)$  a *Perfect Correlated Equilibrium (PCE)* of  $G(\delta)$ . The correlation device assumed in a PCE may depend arbitrarily on history. We now introduce a simpler solution concept that yields the same payoff prediction. A correlation device  $\lambda$  is *Markovian* if its recommendations depend solely on the outcome  $(a, \tilde{a})$  of the most recent period. Denote by  $\Lambda_M$  the space of all such devices  $\lambda : A^2 \rightarrow \Delta(A)$ . Similarly, a strategy  $\mathfrak{s}$  is Markovian if it depends only on the most recent outcome and currently recommended action  $\tilde{a}_i$ , i.e.  $\mathfrak{s}_i : A^2 \times A_i$ . If the device  $\lambda$  is Markovian, then there is a Markovian best response to a Markovian strategy (cf. Hernandez-Lerma, 1989 Theorem 2.2). Thus, a pair  $(\mathfrak{s}, \lambda)$  is a *Markov Perfect Correlated Equilibrium*

---

<sup>4</sup>The notion of an extensive form correlation device is attributable to Forges (1986), who provided the canonical representation and geometric properties of extensive form correlation devices.

(MPCE) of  $G(\delta)$  if it is a PCE of  $G(\delta)$  and both the correlation device  $\lambda$  and the strategy profile  $s$  are Markovian.

Let  $V^\lambda$  be the set of all sequential equilibrium payoff vectors of  $G^\lambda(\delta)$ . The *MPCE payoff set*  $V^*$  is the set of all payoff vectors attainable in an MPCE. Namely,

$$V^* \equiv \bigcup_{\lambda \in \Lambda_M} V^\lambda$$

The Appendix exploits self-generation methods to prove:

**Lemma 1** *Any PCE payoff is attainable in an MPCE.*

Because every MPCE is a PCE by definition, Lemma 1 implies that both concepts yield the same equilibrium payoff sets.

Let  $\mu \in \Delta(A)$  be a probability distribution on the set of action profiles  $A$  — as realized in a PCE as  $\mu = \lambda(h^t)$ , or in an MPCE as  $\mu = \lambda(a, \tilde{a})$ . Fix a compact convex set of payoff vectors  $W \subset \mathbb{R}^n$ . A *continuation value function*  $\mathbb{k} : A^2 \rightarrow W$  describes discounted future (equilibrium) payoffs for each current period outcome. Given the stage game payoffs, the mapping  $\mathbb{k}$  completely describes the *auxiliary game*  $G_{\mathbb{k}}$ . This game is (the agent normal form of) a one-shot Bayesian game whose type profile  $(\tilde{a}_1, \dots, \tilde{a}_n) \in A$  is drawn from the distribution  $\mu$ . Each player's type  $\tilde{a}_i$  has the action set  $A_i$ , but the revised payoff function  $E_\mu[(1 - \delta)u_i(a) + \delta \mathbb{k}_i(a, \tilde{a}) | \tilde{a}_i]$  for the recommended action  $\tilde{a}_i$ .

If the distribution  $\mu$  is a correlated equilibrium of  $G_{\mathbb{k}}$ , then the pair  $(\mu, \mathbb{k})$  is *admissible w.r.t. W*, where  $W$  is the co-domain of  $\mathbb{k}$ . In this case,

$$E_\mu[(1 - \delta)u_i(a) + \delta \mathbb{k}_i(a, \tilde{a}) | \tilde{a}_i] \geq E_\mu[(1 - \delta)u_i(a'_i, a_{-i}) + \delta \mathbb{k}_i(a'_i, a_{-i}, \tilde{a}) | \tilde{a}_i] \quad (1)$$

for all players  $i$ , actions  $a'_i \in A_i$ , and recommendations  $\tilde{a}_i \in A_i$  and  $\tilde{a} \in A$ . The *value w* of a pair  $(\mu, \mathbb{k})$  is the (ex-ante) expected payoff  $E_\mu[(1 - \delta)u(a) + \delta \mathbb{k}(a, a)]$ . Inversely, we write that the admissible pair  $(\mu, \mathbb{k})$  *enforces* the payoff  $w$  on the set  $W$  if  $w$  is the value of the pair, and  $W$  is the co-domain of  $\mathbb{k}$ .

Let the set  $B(W)$  be the union of all payoffs enforced on  $W$ , so that

$$B(W) = \{v = E_\mu[(1 - \delta)\pi(a) + \delta \mathbb{k}(a, \tilde{a})] \mid (\mu, \mathbb{k}) \text{ is admissible w.r.t. } W\}$$

Equivalently,  $B(W)$  is the union of all correlated equilibrium payoffs in the auxiliary game  $G_{\mathbb{k}}$ , as  $\mathbb{k}$  ranges over all continuation value functions with co-domain  $W$ .

The operator  $B(\cdot)$  has some convenient properties. First, it is monotone: If  $W \subseteq W'$ , then  $B(W) \subseteq B(W')$ . Intuitively, the right side consists of the correlated equilibria of a larger set of auxiliary games. Secondly,  $B(\cdot)$  is convex-valued: If  $(\mu^1, \mathbb{k}^1)$  supports  $w^1$  and  $(\mu^2, \mathbb{k}^2)$  supports  $w^2$ , then for all weights  $\theta \in [0, 1]$ , the payoff  $\theta w^1 + (1 - \theta)w^2$  is supported by  $(\theta\mu^1 + (1 - \theta)\mu^2, \theta\mathbb{k}^1 + (1 - \theta)\mathbb{k}^2)$ .

As usual, we call a set  $W \subset \mathbb{R}^n$  is *self-generating* if  $W \subseteq B(W)$ .

**Theorem 1 (MPCE Payoffs)** *The MPCE payoff set  $V^*$  has the properties:*

- (a) *It is the largest fixed point of  $B(\cdot)$ .*
- (b) *It is a compact convex subset of  $V$ .*
- (c) *It contains the convex hull of the set of SPE payoffs of  $G(\delta)$*
- (d) *It is nondecreasing in  $\delta$ .*

The proof is in the Appendix, but here we offer some intuition. First, part (a) captures the recursive structure of MPCE, which is analogous to factorization of PPE. If a set  $W$  is self-generating, then there exists an admissible pair with co-domain  $W$ . For any  $w \in W$ , a sequential equilibrium with payoff  $w$  can be constructed period-by-period by replacing every continuation value with a pair admissible w.r.t.  $W$ . This is always possible since  $W$  is self-generating.

Next, compactness in (b) follows since weak inequalities define incentive compatibility. Public randomization can always be created using a correlation device, and so the MPCE payoff set is convex. To publicly randomize between outcomes, let us step outside the space of direct devices and consider a new device that generates two messages for each player: the original message and a second that indicates the outcome of the public randomization. By the Revelation Principle, there exists an equivalent direct device.

For insight into part (c), consider the extensive form correlation device that recommends the subgame perfect equilibrium behavior after every history. By construction, this device constitutes a PCE, and Lemma 1 guarantees that this payoff is attainable in an MPCE. Part (c) in particular implies that the folk theorem holds for MPCE.

Part (d) follows from the well-known principle that dynamic incentives can induce any behavior in patient players that it can in their less patient counterparts.

The MPCE payoff set can be obtained by iterating the  $B$  operator on a seed set  $W^0 \subseteq \mathbb{R}^n$  containing the feasible and individually rational payoffs  $V$ . The algorithm starts by observing that  $V^* \subseteq V \subseteq W^0$ . Then either  $W^0$  is self-generating or  $B(W^0) \subseteq W^0$ . Repeatedly applying  $B(\cdot)$  to the inequality  $V^* \subseteq W^k$ , where

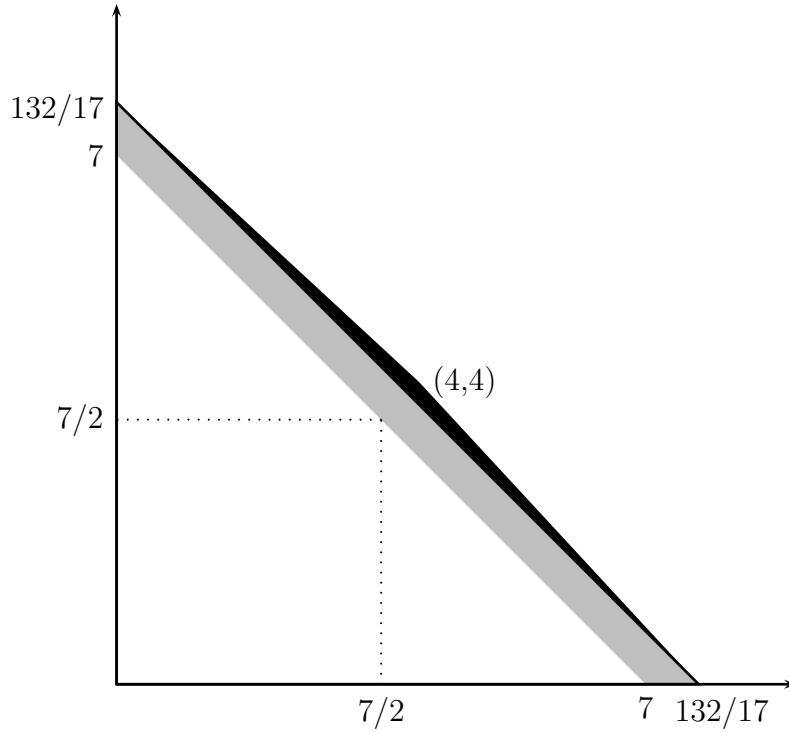


Figure 2: **Payoffs in the Repeated Game in Figure 1.** The white area is the SPE payoff set; MPCE payoffs also include the grey area, so that these are MPCE payoffs unattainable in an SPE; the black area represents feasible and individually rational payoffs that are not MPCE, and thus unattainable in any sequential equilibrium.

$W^k = B(W^{k-1})$ , produces a strictly decreasing sequence of nested sets that converges to the MPCE set  $V^*$ .

**Theorem 2 (Algorithm)** *The MPCE payoff set is  $V^* = \lim_{j \rightarrow \infty} W^j$ , where the payoff set  $W^0$  obeys  $V^* \subseteq W^0$ , and  $W^{j+1} = B(W^j)$  for  $j = 1, 2, 3, \dots$*

To implement the algorithm, we employ methods similar to those introduced by Judd, Yeltekin, and Conklin (2003). Compactness and convexity allow us to represent a set by its extreme points, and they imply that  $B(W) = B(\text{ext } W)$ . This makes the algorithm computationally tractable.

Let's return to the repeated game of Section 2. In Figure 2, one can see that the MPCE payoff set is significantly larger than that of subgame perfect equilibrium. The extreme feasible and individually rational payoffs  $(132/17, 0)$  and  $(0, 132/17)$  are also the highest single player payoff vectors. So by convexity, the symmetric payoff  $(66/17, 66/17)$  is also an MPCE, and in fact the highest symmetric MPCE payoff. This payoff is a convex combination of two extremal MPCE payoffs.

We now justify these claims. First, let us construct the device that delivers the highest payoff to one player. Let  $(p, q, r, 1 - p - q - r) \in \Delta(A)$  be the chances of  $\{(C, D), (C, D), (C, D), (C, D)\}$ , respectively, and  $w_1, w_2 \in \mathbb{R}^2$  the continuation payoffs for players 1, 2. Given the stage game of Figure 1, the highest MPCE payoff for player 1 solves

$$\max_{p, q, r, (w_1, w_2) \in V} (1 - \delta)(4p - 13q + 20r) + \delta w_1$$

given: (i)  $p, q, r \geq 0$  and  $p + q + r \leq 1$ , and (ii) payoffs are feasible and individually rational, and in particular  $0 \leq w_1, w_2 \leq 132/17$ , and (iii) two self-generation feasibility constraints that players not be promised payoffs higher than can be delivered:

$$w_1 \leq (1 - \delta)(4p - 13q + 20r) + \delta w_1 \quad \text{and} \quad w_2 \leq (1 - \delta)(4p + 20q - 13r) + \delta w_2$$

and (iv) two incentive constraints, for when players are told to play  $C$ :

$$(1 - \delta)(4p - 13q) + \delta w_1 \geq (1 - \delta)20p \quad \text{and} \quad (1 - \delta)(4p - 13r) + \delta w_2 \geq (1 - \delta)20r$$

Solving this program yields

$$132/17 = w_1 = 4p - 13q + 20r \quad \text{and} \quad 0 = 4p + 20q - 13r \quad \text{and} \quad p + q + r = 1$$

So  $(p, q, r) = (13/17, 0, 4/17)$ . Then the payoff  $(132/17, 0)$  is attainable in an MPCE. By symmetry, so too is the payoff  $(0, 132/17)$ . By convexity, the payoff  $(66/17, 66/17)$  is an MPCE.

One can verify that imposing symmetry of the form  $q = r$  yields a lower constrained maximum — i.e. a symmetric device does not yield the highest symmetric payoff. This implies that  $(66/17, 66/17)$  is the highest symmetric MPCE payoff.

This effect is not limited to this example. In fact, a sufficient condition for correlation to be helpful in an infinitely repeated prisoner's dilemma is that: (i) mutual cooperation is efficient but not a subgame perfect equilibrium outcome, and (ii) the gain to defecting is higher when the other player cooperates than when he defects.

## 5 Repeated Games of Private Monitoring

A. **THE STAGE GAME.** The structure here is standard, following closely the set-up of Ely, Horner, and Olszewski (2005). As in Section 3, a repeated game is played in periods  $1, 2, \dots$ . Each period, every player  $i \in N = \{1, 2, \dots, n\}$  chooses an action  $a_i$  from a finite action set  $A_i$ . But now, after play any period, each player receives a private message  $m_i$  from a finite set  $M_i$ . A *monitoring structure*  $\psi$  is a collection of  $|A|$  probability distributions  $\{\psi(\cdot|a) \in \Delta(M) \mid a \in A\}$  on the message profile set  $M = \prod_i M_i$ . Let the set of all monitoring structures be  $\Psi$ . After an action profile  $a$  is realized, a message profile  $m = (m_1, \dots, m_n)$  is drawn with chance  $\psi(m|a)$ , and each player  $i$  is then privately informed of his component message  $m_i$ .

A player's realized payoff  $\pi_i(a_i, m_i)$  following action  $a_i$  and message  $m_i$  depends on the other actions only through their effect on the private messages. In other words, observing one's payoff does not confer additional information. Player  $i$ 's expected payoff from the action profile  $a$  is then

$$u_i(a) = \sum_{m_i \in M_i} \psi_i(m_i|a) \pi_i(a_i, m_i) \quad (2)$$

We shall consider different monitoring structures  $\psi$  consistent with the same “expected stage game”. This requires that the payoffs  $u(a) = (u_1(a), \dots, u_n(a))$  not depend on the monitoring structure. Since payoffs depend on  $\psi$  in (2), this exercise implies a corresponding change in the stochastic payoff structure  $\pi$ . Such a choice is possible provided (2) is solvable in  $\pi_i$  for any  $\psi_i$ , and for all players  $i$ . This is feasible if and only if the matrix  $(\psi_i(m_i|a_i, a_{-i}), m_i \in M_i, a_{-i} \in A_{-i})$  has full rank for every player  $i$ , and every action  $a_i$ . This requires that each player can statistically identify the actions of his opponents.<sup>5</sup> This generically holds when, for instance, everyone has at least as many messages as there are players. We assume that this condition is met by any monitoring structure in  $\Psi$  under consideration. Our results do not explicitly depend on this; it simply allows us to meaningfully consider a fixed stage game.

B. **THE REPEATED GAME.** Let  $G_\psi(\delta)$  denote the infinitely repeated game of private monitoring with monitoring structure  $\psi$ , played in periods  $t = 1, 2, 3, \dots$ . Payoffs are discounted as usual by the factor  $0 < \delta < 1$ . The game reduces to a standard repeated game with perfect monitoring when private messages are action profiles, i.e.

---

<sup>5</sup>This is somewhat analogous to the pairwise full rank condition of Fudenberg, Levine, and Maskin (1994), which requires that each player be able to statistically identify the actions of another player.

if  $M_i = A$  and  $\psi_i(m_i|a) = 1$  when  $m_i = a$  and 0 otherwise, for all players  $i$ . Similarly, the game reduces to a standard repeated game with public monitoring if  $M_i = M$  for all players  $i$ , and  $\psi_i(m|a) = 1$  if and only if  $\psi_j(m|a) = 1$  for every pair of players  $i, j$ .

In each period, a player observes his realized action  $a_i \in A_i$  and private message  $m_i$ . Let the null history  $h_i^1$  be player  $i$ 's history before play begins. A *private history*  $h_i^t$  is the complete record of player  $i$ 's past actions  $(a_i^1, \dots, a_i^{t-1})$  and past private messages  $(m_i^1, \dots, m_i^{t-1})$ , including the null history. Let  $H_i^t$  be the set of all possible private histories  $h_i^t$  for player  $i$ , and  $H_i = \bigcup_{t=1}^{\infty} H_i^t$  the set of all such histories of any length. A (behavior) strategy  $s_i$  is a sequence of functions  $\{s_i^t\}_{t=1}^{\infty}$ , where  $s_i^t : H_i^t \rightarrow \Delta(A_i)$  for every period  $t = 1, 2, 3, \dots$ . In other words, it maps every private history into a mixed action. Let  $\mathcal{S}$  be the space of all such strategy profiles  $s = (s_1, \dots, s_n)$ .

Given the strategy profile  $s \in \mathcal{S}$ , Bayes' rule and the Law of Total Probability naturally imply beliefs and behavior at all future information sets. Let  $v_i : \mathcal{S} \rightarrow \mathbb{R}$  be the discounted average payoff for player  $i$  in the repeated game  $G_\psi(\delta)$ . While more precisely presented in the Appendix, here we write that player  $i$ 's discounted average payoff starting in period  $t$  from the strategy profile  $s$  is  $v_i^t(s|h_i^t)$ . Then a strategy profile  $s$  is a *sequential equilibrium* of  $G_\psi(\delta)$  if and only if no player can ever profitably deviate, i.e.  $v_i(s|h_i^t) \geq v_i(\tilde{s}_i, s_{-i}|h_i^t)$  for every private history  $h_i^t$  and strategy  $\tilde{s}_i : H_i \rightarrow \Delta(A_i)$  of every player  $i$ . Since playing a Nash equilibrium of  $G$  after every history is a sequential equilibrium, existence is guaranteed. Let  $V_\psi$  be the set of sequential equilibrium payoff vectors of the mediated game  $G_\psi(\delta)$ .

## 6 Unattainable Private Monitoring Payoffs

A. AN UPPER BOUND. We bound the sequential equilibrium payoffs by the MPCE payoff set  $V^*$ . This inclusion might at first blush appear surprising: For the repeated game  $G_\psi(\delta)$  has no proper subgames, whereas  $G^\lambda(\delta)$  introduces a new subgame every period. So while continuation play in  $G^\lambda(\delta)$  is common knowledge, it is not so in  $G_\psi(\delta)$ . We proceed by associating outcomes in  $G_\psi(\delta)$  with those of  $G^\lambda(\delta)$ . To do so, we replace the endogenous correlated beliefs in  $G_\psi(\delta)$  with those from a fixed correlation device  $\lambda$ . Also, we do so in an incentive compatible fashion.

**Theorem 3 (Upper Bound)** *For any monitoring structure  $\psi$ , every sequential equilibrium payoff of the repeated game  $G_\psi(\delta)$  is attained in an MPCE of  $G(\delta)$ .*

This implies that MPCE captures the payoffs in many studied subclasses of equilibria. It contains all PPE payoffs for any public monitoring structure, as well as all sequential equilibrium payoffs in private strategies (Kandori and Obara, 2006), as well as all belief-free and weakly-belief-free equilibrium payoffs (Kandori, 2011).

The proof in the Appendix first deduces this result for PCE, and then appeals to Lemma 1. The proof for PCE involves two steps. We show that for any strategy profile  $s \in \mathcal{S}$ , there exists a correlation device  $\lambda \in \Lambda$  and strategy profile  $\mathfrak{s} \in \mathfrak{S}$  that induce in  $G^\lambda(\delta)$  the same outcome as does  $s$  in  $G_\psi(\delta)$ . After the history  $\mathfrak{h}^t$  in the mediated game  $G^\lambda(\delta)$ , the correlation device draws a “fictitious private history”  $h_i^t$  for each player  $i \in N$  according to the true posterior probability of that history conditional on the actions of history  $\mathfrak{h}^t$ . The device then recommends the actions prescribed at that private history profile  $h^t$  by the continuation strategy profile  $\mathfrak{s}(h^t)$ . By induction on the period  $t$ , we show that the distribution over recommendations in the mediated game coincides with the distribution of actions in  $G_\psi(\delta)$ . In our next step, we argue that if  $s$  is a sequential equilibrium strategy profile of  $G_\psi(\delta)$ , then  $\lambda$  constitutes a PCE. For if some player has a profitable deviation in  $G^\lambda(\delta)$ , then we argue that he must also have one in  $G_\psi(\delta)$ . The argument turns on the equivalence of beliefs about continuation play in  $G^\lambda(\delta)$  and  $G_\psi(\delta)$ .

**B. A TIGHT UPPER BOUND.** Since this upper bound is independent of the monitoring structure  $\psi$ , one might think that the inclusion in Theorem 3 could not be tight. In fact, this is true, but only because correlated play in a private monitoring game starts no earlier than the second period. So inspired, we now exploit the MPCE payoffs to deduce a tight upper bound for equilibrium payoffs of private monitoring games.

For a standard repeated game played in periods  $1, 2, 3, \dots$ , we can remove the first period correlation from MPCE. An admissible pair  $(\mu, \mathbb{k})$  is called *Nash admissible* if  $\mu$  is the result of independent mixtures, i.e.  $\mu \in \prod_i \Delta(A_i)$ . We then obtain the operator from APS, here denoted by  $B_{NE}$ :

$$B_{NE}(W) = \{v = E_\mu[(1 - \delta)\pi(a) + \delta \mathbb{k}(a, \tilde{a})] \mid (\mu, \mathbb{k}) \text{ is Nash admissible w.r.t. } W\}$$

This collects the Nash equilibrium payoffs of all auxiliary games formed with continuation value functions mapping into  $W$ . Since first period strategies are uncorrelated in  $G_\psi(\delta)$ , we use a two-stage procedure. First, we compute the MPCE payoff set, and then use this set  $W = V^*$  as continuation payoffs in  $B_{NE}(W)$ .

**Theorem 4 (Tightness)** *A payoff is Nash admissible w.r.t. the MPCE set of  $G(\delta)$  if and only if it is a sequential equilibrium payoff of  $G_\psi(\delta)$  for some monitoring structure  $\psi$ , so that*

$$\bigcup_{\psi \in \Psi} V_\psi = B_{NE}(V^*)$$

Without reference to the monitoring structure, there exists no tighter bound on the sequential equilibrium payoffs in a repeated game of private monitoring.

In the example of Section 2, Theorem 3 demonstrates that  $(66/17, 66/17)$  is the highest symmetric sequential equilibrium payoff in the infinitely repeated game *with any monitoring structure*, and so all symmetric payoffs in  $(66/17, 4]$  are unattainable. In fact, except for the payoffs  $(132/17, 0)$  and  $(0, 132/17)$ , all efficient payoff vectors are unattainable in a sequential equilibrium.

**C. HOW RESTRICTIVE IS PERFECT PUBLIC EQUILIBRIUM?** We are finally able to address one of our key motivations, and ask how restrictive is the public monitoring assumption that the literature has settled on. Does it ever greatly understate the sequential equilibrium payoff set? Since the set of PPE payoffs is a subset of subgame perfect payoffs, the demonstrated gap between the MPCE and SPE payoff sets implies that a public solution concept may fail to capture the potential outcomes of environments with richer information in which the folk theorem is silent. Thus a regulator attempting to detect antitrust violations may, upon observing payoffs inconsistent with some PPE, draw the wrong conclusion about collusion. These efforts ought to keep in mind the strategic opportunities afforded by complex information structures; This is done precisely by using MPCE in the place of a public monitoring solution concept. Furthermore, in many applications the relevant monitoring structure is difficult to determine, and thus PPE is difficult to use. Thus unlike PPE, MPCE enables one to study equilibrium payoffs while being agnostic about the monitoring structure.

We will now demonstrate that *for a generic class of prisoner's dilemma games, if the discount factor is high, but not too high, correlation improves upon subgame perfect equilibrium, and hence perfect public equilibrium*. Consider the infinite repetition of the following prisoner's dilemma.

As before, we assume that  $(C, C)$  is the efficient action profile (i.e.  $b - c < 2$ ), and that the gain to defecting when the opponent plays  $C$  is larger than the gain when he plays  $D$  (i.e.  $b - 1 < c$ ). Stahl (1991) characterizes the subgame perfect equilibrium payoff correspondence with respect to the discount factor. If  $\delta < c/b$  then the only such payoff vector is the stage game Nash equilibrium payoff  $(0, 0)$ . When  $\delta \geq 1 - 1/b$ ,

	$C$	$D$
$C$	$(1,1)$	$(-c,b)$
$D$	$(b,-c)$	$(0,0)$

Figure 3: Prisoner's Dilemma ( $b > 1, c > 0$ )

every feasible and individually rational payoff vector is a subgame perfect equilibrium payoff vector. But for intermediate discount factors  $\delta \in [c/b, 1 - 1/b)$  then the set of subgame perfect payoffs is the triangle  $T = \{(0, 0), (b - c, 0), (0, b - c)\}$ .

Observe next that for intermediate levels of discounting in this range, correlation can be used to support a payoff vector  $(v^*, v^*) \notin T$ . To see this, let us exploit the logic at the end of section 2.A and the intuition of Kandori (2011), which suggests there are gains to confusing the players about the history. Consider the continuation value function:

$$\mathbb{k}(a, \tilde{a}) = \begin{cases} (v^*, v^*) & \text{if } a = \tilde{a} = (C, C) \\ (b - c, 0) & \text{if } a = \tilde{a} = (D, C) \\ (0, b - c) & \text{if } a = \tilde{a} = (C, D) \\ (0, 0) & \text{if } a = \tilde{a} = (D, D) \\ (0, 0) & \text{if } a \neq \tilde{a} \end{cases}$$

We now check whether the auxiliary game implied by this continuation value function induces the game of “chicken”, and therefore has non-trivial correlated equilibria. We will then check that there exists a correlated equilibrium with an expected payoff outside of the triangle  $T$ . “Chicken” requires that

$$(1 - \delta) + \delta v^* \leq (1 - \delta)b$$

$$(1 - \delta)(-c) + \delta(b - c) \geq 0$$

Both expressions are satisfied for a payoff vector  $v^* \notin T$  over the entire range of parameters considered.

Since the continuation payoffs to all outcomes except a recommended  $(C, C)$  are subgame perfect, we can focus on supporting the current MPCE payoff  $v^*$ . Consider

the distribution over recommendations

$$\mu(a) = \begin{cases} p & \text{if } a = (C, C) \\ \frac{1-p}{2} & \text{if } a = (C, D) \\ \frac{1-p}{2} & \text{if } a = (D, C) \end{cases}$$

Then the expected payoff satisfies the equation

$$v^* = p[(1 - \delta) + v^*] + \frac{1-p}{2}[(1 - \delta)b] + \frac{1-p}{2}[(1 - \delta)(-c) + \delta(b - c)]$$

which implies that the target payoff

$$v^* = \frac{b(1-p) + (2+c-2\delta)p - c}{2(1-\delta p)}$$

Incentive compatibility requires

$$2p[(1 - \delta) + v^*] + (1 - p)[(1 - \delta)b] \geq 2p[(1 - \delta)(-c) + \delta(b - c)]$$

Since the auxiliary game is “chicken”, there exists a probability  $p > 0$  that satisfies incentive compatibility. When  $p > 0$  the target payoff  $v^*$  is larger than  $(b - c)/2$  and thus outside of  $T$ . This establishes our earlier claim.

## 7 Conclusion

Understanding the equilibria of repeated games with private monitoring has long been the next frontier in game theory. Yet finding sequential equilibria here has been hard, because existing recursive methods only capture subsets of them. In this paper, we have developed a new solution concept for repeated games, Markov Perfect Correlated Equilibrium, with a recursively computable payoff set. This is the smallest set that contains all equilibrium payoffs of the analogous repeated game endowed with any monitoring structure. It therefore provides insights into important economic environments while being agnostic about specific, possibly unobservable, informational aspects of the game. We also hope our bound will offer a rebirth to the recursive methods of Abreu, Pearce, and Stacchetti (1990) in settings with richer information structures than they had envisioned. Finally, we have shown that the restriction to perfect public equilibrium can be misleading at times — for instance in

games with a prisoner's dilemma structure.

## A Omitted Proofs

### A.1 Any PCE Payoff is an MPCE Payoff: Proof of Lemma 1

Let  $W \subset \mathbb{R}^n$  be a compact, convex set with extreme points denoted  $\text{ext } W$ . The continuation value function  $\mathbb{k} : A^2 \rightarrow W \subset \mathbb{R}^n$  has the *bang-bang* property if  $\mathbb{k}(a, \tilde{a}) \in \text{ext } W$  for all action profiles  $a \in A$  and recommendation profiles  $\tilde{a} \in A$ .

We first argue that any continuation value function can be replaced with one that takes values in  $\text{ext } W$ .

**Claim 1 (Bang-Bang)** *Any continuation value function is equivalent to one with the bang-bang property.*

*Proof of Claim 1:* We adapt the proof of Theorem 3 in APS, accounting for correlation and a finite domain of the continuation value function.<sup>6</sup> For a bounded set  $W \subset \mathbb{R}^n$ , let  $\mathcal{K}(W)$  be the set of all functions from  $A \times A$  to  $W$ , and  $\mathcal{K}(W|w) \subseteq \mathcal{K}(W)$  the set of continuation value functions that support  $w$  on  $W$ . Since  $\mathcal{K}(W) = W^{|A|^2}$ , and  $W$  is compact, it is compact in the product topology, by Tychonov's Theorem. Next, since a convex combination of admissible pairs is also an admissible pair,  $\mathcal{K}(W|w)$  is a convex set. As a closed subset of a compact set, it is compact. By the Krein-Milman Theorem, any  $\mathbb{k} \in \mathcal{K}(W|w)$  can be written as a convex combination of extreme points of  $\mathcal{K}(W|w)$ . Finally, linearity of incentives and payoffs implies that  $\hat{\mathbb{k}}$  is a convex combination of extreme points of  $\mathcal{K}(\text{ext } W|w)$ , and consequently has the bang-bang property.  $\square$

*Proof of Lemma 1:* Let  $[x_1, \dots, x_m]$  be the convex hull of the points  $(x_1, \dots, x_m)$ . Let  $V_{PCE}$  be the set of PCE payoffs. Fix a PCE  $\lambda \in \Lambda$  with payoff  $w \in V_{PCE}$ . Define the product space  $\mathcal{V} \equiv V_{PCE}^{|A|^2}$ . To prove that the payoff  $w$  is attainable in an MPCE, we show that there exists a correlation device  $\lambda_M \in \Lambda_M$  that delivers the payoff  $w$  and is incentive compatible. Thus, we want to show that the payoff vector  $w$  is supported by the convex hull of a self-generating set of  $|A|^2$  payoff vectors.

---

<sup>6</sup>In APS, an equilibrium prescribes continuation behavior for each of a continuum of possible public signals. This required an appeal to Aumann (1965) for technical reasons. In our context, a continuation value function is defined on a finite set. The set of continuation value functions, therefore, is a simpler object that can be treated with simpler mathematical tools.

Any continuation value function can be written as an ordered  $|A|^2$ -tuple of payoff vectors, one for each action profile and recommendation. Define the correspondence on  $|A|^2$ -tuples of payoff vectors  $\phi : \mathcal{V} \rightarrow 2^{\mathcal{V}}$  by

$$\phi(v_1, \dots, v_{|A|^2}) = [\mathcal{K}(V_{PCE}|v_1), \dots, \mathcal{K}(V_{PCE}|v_{|A|^2})] \cap [v_1, \dots, v_{|A|^2}]$$

The correspondence  $\phi$  maps  $|A|^2$ -tuples of payoff vectors to the convex hull of supporting sets of  $|A|^2$ -tuples of payoff vectors.

We now claim that the correspondence  $\phi$  satisfies the hypotheses of the Kakutani Fixed Point Theorem. Since  $V_{PCE}$  is non-empty, compact and convex,  $\mathcal{V}$  is non-empty, compact and convex. The correspondence has non-empty values: since  $v_j$  is a PCE payoff,  $\mathcal{K}(V_{PCE}, v_j)$  is not empty. Since by Claim 1 continuation payoffs can equivalently be taken from  $\text{ext } V_{PCE}$ , the intersection with the convex hull of an arbitrary set of PCE payoffs is non-empty. Furthermore,  $\phi$  takes compact convex values as the intersection of two compact, convex sets. By Claim 1 and the Theorem of the Maximum,  $\mathcal{K}(V_{PCE}, v_j)$  is upper hemi-continuous in  $v_j$ . Similarly,  $[v_1, \dots, v_{|A|^2}]$  is upper hemi-continuous. Then  $\phi$  is the intersection of upper hemi-continuous correspondences and therefore also upper hemi-continuous. Thus, by the Kakutani Fixed Point Theorem there exists a fixed point  $(v_1^*, \dots, v_{|A|^2}^*)$ .

For each element  $v_j^*$  of the fixed point,  $j = 1, \dots, |A|^2$ , there exists a probability distribution  $\mu_j^*$  on  $A^2$  used to enforce it, since each is a PCE payoff. Then the device  $\lambda_M$  making recommendations according to  $\mu_j^*$  for  $j = 1, \dots, |A|^2$  is Markovian and incentive compatible by construction.  $\square$

## A.2 Characterization of MPCE: Proof of Theorem 1

Part (a) FACTORIZATION: First we show that if  $W$  is self-generating, then  $B(W) \subseteq V^*$ . For any payoff vector  $w \in B(W)$  there exists a pair  $(\mu, \mathbb{k})$  that enforces  $w$  on  $W$ . Since  $W$  is self-generating,  $\mathbb{k}(a, \tilde{a}) \in W$  for all outcomes  $(a, \tilde{a})$ . Each payoff  $\mathbb{k}(a, \tilde{a})$  is enforced on  $W$ . In this way, we can (by the Axiom of Choice) recursively define a PCE by constructing admissible pairs ad infinitum. By Lemma 1, the PCE payoff  $w$  is an MPCE payoff. Thus,  $W \subseteq V^*$ . Next, we prove that  $V^*$  is a fixed point of  $B(\cdot)$ . Since  $V^*$  contains every self-generating set, we need only show that  $V^*$  is self-generating. Consider an MPCE payoff  $w \in V^*$ . There exists a pair  $(\mu, \mathbb{k})$  such that  $\mathbb{k}(a, \tilde{a}) \in V^*$  for each pair of action and recommendation profiles  $(a, \tilde{a})$ . Hence,  $w$  is admissible w.r.t.  $V^*$ , or equivalently that  $w \in B(V^*)$ .

Finally, suppose that there exists a fixed point  $W$  of  $B(\cdot)$  that strictly contains  $V^*$ . Then  $W$  is self-generating, and so is contained in the MPCE set  $V^*$ . This contradicts the premise that  $W$  strictly contains  $V^*$ . So  $V^*$  is the largest fixed point of  $B(\cdot)$ .  $\square$

Part (b) **COMPACT AND CONVEX**: First, we want to show that  $B(W)$  is compact if  $W$  is compact. Since  $B(W)$  is bounded, by the Heine-Borel Theorem it is compact if it is also closed. Consider a sequence  $\{b_j\}$  in  $B(W)$  that converges to some  $b \in \mathbb{R}^n$ . Each  $b_j \in B(W)$  is supported on  $W$  by an admissible pair  $(\mu_j, \mathbb{k}_j)$ . Endow the space of such functions that map  $A \times A^2$  into  $\Delta(A) \times W$  with the weak-\* topology (i.e. pointwise convergence). The sequence is bounded, and so by the Bolzano-Weierstrass Theorem it has a convergent subsequence  $\{\mu_l, \mathbb{k}_l\}$ . The weak inequalities that define incentives are satisfied pointwise in the sequence  $\{\mu_l, \mathbb{k}_l\}$ , and hence are also by the limit  $(\mu, \mathbb{k})$ , which thus enforces  $b \in \mathbb{R}^n$ . Then  $b \in B(W)$ , and so  $B(W)$  is closed.  $\square$

Part (c) **CONTAINS SPE PAYOFFS**: Since the mediated game has perfect monitoring of actions, players may ignore the correlation device, and instead play the subgame perfect equilibrium behavior after every history.  $\square$

Part (d) **NONDECREASING  $\delta$** : The proof is very similar to that of APS, Theorem 6.  $\square$

### A.3 Algorithm: Proof of Theorem 2

We extend the methods of Judd, Yeltekin, and Conklin (2003) to allow for correlation. Let  $\mathbb{W}$  be the set of all convex subsets of  $V$ , partially ordered by set inclusion. Then the operator  $B(\cdot)$  is monotone on the complete lattice  $\mathbb{W}$ . By Tarski's Fixed Point Theorem,  $B(\cdot)$  has a largest fixed point  $V^*$ . Let  $W^0 = V$  and recursively define  $W^k = B(W^{k-1})$  for  $k = 1, 2, \dots$ . First, by monotonicity  $V^* = B(V^*) \subseteq B(W^0) = W^1$ . Next, suppose that  $V^* \subseteq W^k$ . Monotonicity again yields  $V^* = B(V^*) \subseteq B(W^k) = W^{k+1}$ . By induction,  $V^* \subseteq W^k$  for all  $k = 1, 2, \dots$ ,

The sequence  $\{W^k\}_{k=0}^\infty$  is bounded and monotone, and therefore converges (in the Hausdorff topology) to a point in the complete lattice  $\mathbb{W}$ . Let  $W^\infty = \lim_{k \rightarrow \infty} W^k$ . This limit is a fixed point of  $B(\cdot)$ , and by construction contains  $V^*$ . But  $V^*$  cannot be a strict subset of  $W^\infty$ , since that would imply that  $V^*$  is not the largest fixed point of  $B(\cdot)$ , contrary to Theorem 1.  $\square$

## A.4 MPCE as an Upper Bound: Proof of Theorem 3

At the information set  $h_i^t$ , player  $i$  believes that the other players' private history profile is  $h_{-i}^t$  with posterior probability  $\mu_{i,s}^t(h_{-i}^t|h_i^t)$ , and that their period  $t$  action profile is  $a_{-i}$  with posterior probability

$$\beta_i^t(a_{-i}|h_i^t, s) = \sum_{h_{-i}^t \in H_{-i}^t} \mu_{i,s}^t(h_{-i}^t|h_i^t, s) s_{-i}(a_{-i}|h_{-i}^t)$$

Player  $i$ 's *continuation payoff* under the strategy profile  $s$  at the private history  $h_i^t$  is therefore

$$\kappa_i^t(h_i^t|s) = (1 - \delta)E \left[ \sum_{r=t+1}^{\infty} \delta^{r-t-1} u_i(\beta_i^r) \middle| h_i^t, s \right] \quad (3)$$

where  $u_i(\beta_i^t|h_i^t, s) = \sum_{a_{-i} \in A_{-i}} u_i(s_i(h_i^t), a_{-i}) \beta_i^t(a_{-i}|h_i^t, s)$ . Then player  $i$ 's expected payoff under the strategy profile  $s$  at the private history  $h_i^t$  is

$$v_i^t(s|h_i^t) = (1 - \delta)u_i(\beta_i^t|h_i^t, s) + \delta\kappa_i^t(h_i^t|s)$$

As is well-known, a strategy profile  $s$  is a sequential equilibrium if and only if there are no profitable one-shot deviations. This is equivalent to

$$(1 - \delta)u_i(\beta_i^t|h_i^t, s) + \delta\kappa_i^t(h_i^t|s) \geq (1 - \delta)u_i(\beta_i^t|h_i^t, \tilde{s}_i, s_{-i}) + \delta\kappa_i^t(h_i^t|\tilde{s}_i, s_{-i}) \quad (4)$$

for all players  $i$ , private histories  $h_i^t$ , and strategies  $\tilde{s}_i \neq s_i$ .

Recall that  $s$  and  $v$  denote, respectively, the strategy profiles and payoffs in  $G_\psi(\delta)$ , and  $\tilde{s}$  and  $\tilde{v}$  denote, respectively, the strategy profiles and payoffs in  $G^\lambda(\delta)$ .

**Claim 2 (The Correlation Device)** *For any strategy profile  $s \in \mathcal{S}$  of  $G_\psi(\delta)$ , there exists a correlation device  $\lambda_s \in \Lambda$  and strategy  $\tilde{s} \in \mathcal{S}$  in the mediated game that induces the same outcome in  $G^{\lambda_s}(\delta)$  as  $s$  does in  $G_\psi(\delta)$ .*

*Proof of Claim 2:* For any strategy profile  $s \in \mathcal{S}$ , let  $\beta^t(a^t|(a^1, \dots, a^{t-1}), s)$  be the induced posterior probability of the action profile  $a^t$  in period  $t$  given the action history  $(a^1, \dots, a^{t-1})$ . The action mixture in period 1 is simply  $\beta^1(a^1) = \alpha^1(a)$ . Given the realized action profile  $a^1$ , action profile  $a^2$  occurs with chance  $\beta^2(a^2|a^1) =$

$\sum_{m^1 \in M} \psi(m^1 | a^1) s(a^2 | a^1, m^1)$  using the joint density of signals  $\psi(\cdot | a^1)$ . In general,

$$\beta^t(a^t | s, (a^1, \dots, a^{t-1})) = \sum_{(m^1, \dots, m^{t-1}) \in M^{t-1}} s(a^t | (a^1, \dots, a^{t-1}), (m^1, \dots, m^{t-1})) \prod_{k=1}^{t-1} \psi(m^k | a^k)$$

For all action histories  $h^t \in \mathbb{H}^t$ , define  $\lambda_s(h^t) = \beta^t(\cdot | s, h^t)$ . Then, the recommendation distribution of  $\lambda_s$  coincides with the distribution of actions in  $G_\psi(\delta)$ . Call  $\bar{s}$  the *obedient strategy* in  $G^{\lambda_s}(\delta)$  — namely, where players obey the recommendation of the correlation device  $\lambda$  after all histories. Since  $\lambda_s$  recommends the same outcome as  $w$ , the obedient strategy  $\bar{s}$  in  $G^{\lambda_s}(\delta)$  delivers the same outcome as  $s$ .  $\square$

We must prove that obeying  $\lambda_s$  is a mutual best response, or  $v_i^t(\bar{s} | \lambda_s) \geq v_i^t(s'_i, \bar{s}_{-i} | \lambda_s) \forall s'_i \in \mathcal{S}_i$ . We'll argue that for every deviation  $s'_i \in \mathcal{S}_i$  in the mediated game, there is a corresponding strategy  $s'_i \in \mathcal{S}_i$  with  $v_i^t(s'_i, \bar{s}_{-i} | \lambda_s) = v_i^t(s'_i, s_{-i})$ . Namely, any deviation in the mediated game yields the same payoff as some strategy in the repeated game of private monitoring; this cannot be a profitable deviation against the sequential equilibrium profile  $s_{-i}$ . So  $v_i^t(\bar{s} | \lambda_s) = v_i^t(s) \geq v_i^t(s'_i, s_{-i}) = v_i^t(s'_i, \bar{s} | \lambda_s)$ , as required.

**Claim 3 (Verifying Incentives)** *If  $s \in \mathcal{S}$  is a sequential equilibrium strategy of  $G_\psi(\delta)$ , then the correlation device  $\lambda_s \in \Lambda$  is a PCE of  $G(\delta)$ .*

*Proof of Claim 3:* By the one-shot deviation principle, the obedient strategy is a best reply to itself iff there is no history after which a player would choose to disobey his recommendation once, and return to the obedient strategy thereafter. So, it suffices to restrict attention to alternative strategies that differ from the obedient strategy in one history. Consider a history  $h^t \in \mathbb{H}^t$  at which strategy  $s'_i$  instead plays the action  $a'_i$  in period  $t$ . Let  $H(h^t) \subseteq H^t$  be the set of private histories consistent with the action history portion of  $h^t$  in the mediated game. At any private history  $h_i^t \in H(h^t)$ :

$$\begin{aligned} v_i^t(s_i, \bar{s}_{-i}) &= (1 - \delta) E_\lambda [u_i(a'_i, a_{-i}^t) | a'_i, h^t] + \delta E_\lambda \left[ \sum_{r=t+1}^{\infty} \delta^{r-t-1} u_i(a^r) \middle| (a'_i, h^t) \right] \\ &= (1 - \delta) u_i(a'_i, s_{-i}(\alpha | h_i^t)) + \delta \kappa_i^{t+1}((h_i^t, a'_i) | (s'_i, s_{-i})) \\ &= v_i(s'_i, s_{-i}) \end{aligned}$$

Thus, if  $s'_i$  is a profitable deviation from that recommended by the device  $\lambda_s$  in the mediated game, then there exists a profitable deviation in  $G_\psi(\delta)$ . This would contradict the premise that  $s$  is a sequential equilibrium profile in  $G_\psi(\delta)$ . Since any

strategy in  $G^{\lambda_s}$  is equivalent to some non-profitable deviation in  $G_\psi(\delta)$ , the correlation device  $\lambda_s$  and the obedient strategy  $\bar{s}$  constitute a PCE of  $G(\delta)$ .  $\square$

## A.5 MPCE Inclusion is Tight: Proof of Theorem 4

( $\subseteq$ ): Fix a game  $G_\psi(\delta)$ , and consider a sequential equilibrium strategy profile  $s$  with payoff  $v$ . Construct a PCE that induces the same outcome as  $s$ . Absent a pre-play signal, first period actions owe to independent mixtures, and so the PCE recommends an independent mixture in the first period. Next, by Lemma 1, the continuation values prescribed by the PCE are in  $V^*$ . Thus, the payoff  $v$  is Nash enforced on  $V^*$ .  $\square$

( $\supseteq$ ): We want to show that for every payoff  $w$  in  $B_{NE}(V^*)$ , there exists a monitoring structure  $\psi$  and a sequential equilibrium  $s$  of  $G_\psi(\delta)$  with the same payoff  $w$ . Consider one such payoff and the pair  $(\mu, \mathbb{k})$  that Nash enforces it on  $V^*$ . Thus, for every action profile  $a^j \in A$ ,  $j = 1, \dots, |A|$ , there is a payoff  $w^j \in V^*$  that is enforced on  $V^*$  by the admissible pair  $(\mu^j, \mathbb{k}^j)$ . Let the monitoring structure  $\psi$  provide perfect monitoring of actions, as well as a vector of private signals for each player. In particular, after the action profile  $a^j$ , each player privately observes his component of a draw from  $\mu^j$ . So defined, consider the following strategy profile  $s$  in  $G_\psi(\delta)$ . “In the first period, mix according to  $\mu$ . Following every subsequent history, choose the action corresponding to the most recently received message.” Since the private messages are MPCE recommendations,  $s$  constitutes a Nash equilibrium. Then there exists a sequential equilibrium with the same path as  $s$ . So, there exists a private monitoring sequential equilibrium with the payoff  $w$ .  $\square$

## References

- ABREU, D., P. MILGROM, AND D. PEARCE (1991): “Information and Timing in Repeated Partnerships,” *Econometrica*, 59(6), 1713–1733.
- ABREU, D., D. PEARCE, AND E. STACCHETTI (1986): “Optimal Cartel Equilibria With Imperfect Monitoring,” *Journal of Economic Theory*, 39, 251–269.
- (1990): “Toward a Theory of Discounted Repeated Games with Imperfect Monitoring,” *Econometrica*, 58(5), 1041–1063.
- AUMANN, R. (1974): “Subjectivity and Correlation in Randomized Strategies,” *Journal of Mathematical Economics*, 1, 67–96.
- (1987): “Correlated Equilibrium as an Expression of Bayesian Rationality,” *Econometrica*, 55(1), 1–18.
- BHASKAR, V., AND I. OBARA (2002): “Belief-Based Equilibria in the Repeated Prisoners’ Dilemma with Private Monitoring,” *Journal of Economic Theory*, 102, 40–69.
- BLACKWELL, D. (1953): “Equivalent Comparisons of Experiments,” *Annals of Mathematical Statistics*, 24(2), 265–272.
- ELY, J., J. HORNER, AND W. OLSZEWSKI (2005): “Belief-free Equilibria in Repeated Games,” *Econometrica*, 73(2), 377–415.
- ELY, J., AND J. VALIMAKI (2002): “A Robust Folk Theorem for the Prisoner’s Dilemma,” *Journal of Economic Theory*, 102(1), 84–105.
- FORGES, F. (1986): “An Approach to Communication Equilibrium,” *Econometrica*, 54(6), 1375–1385.
- FUCHS, W. (2007): “Contracting with Repeated Moral Hazard and Private Evaluations,” *American Economic Review*, 97(4), 1432–1448.
- FUDENBERG, D., D. K. LEVINE, AND E. MASKIN (1994): “The Folk Theorem with Imperfect Public Information,” *Econometrica*, 62(5), 997–1039.
- GREEN, E. J., AND R. PORTER (1984): “Noncooperative Collusion Under Imperfect Price Information,” *Econometrica*, 52(1), 87–100.

- HERNANDEZ-LERMA, O. (1989): *Adaptive Markov Control Processes*. Springer-Verlag.
- JUDD, K., S. YELTEKIN, AND J. CONKLIN (2003): “Computing Supergame Equilibria,” *Econometrica*, 71(4), 1239–1254.
- KANDORI, M. (1992): “The Use of Information in Repeated Games with Imperfect Monitoring,” *The Review of Economic Studies*, 59(3), 581–593.
- KANDORI, M. (2002): “Introduction to Repeated Games with Private Monitoring,” *Journal of Economic Theory*, 102, 1–15.
- KANDORI, M. (2011): “Weakly Belief-Free Equilibria in Repeated Games With Private Monitoring,” *Econometrica*, 79(3), 877–892.
- KANDORI, M., AND I. OBARA (2006): “Efficiency in Repeated Games Revisited: The Role of Private Strategies,” *Econometrica*, 74(3), 499–519.
- LEHRER, E. (1992): “Correlated Equilibria in Two-Player Repeated Games with Nonobservable Actions,” *Mathematics of Operations Research*, 17(1), 175–199.
- MACLEOD, W. B. (2003): “Optimal Contracting with Subjective Evaluation,” *The American Economic Review*, 93(1), 216–240.
- PICCIONE, M. (2002): “The Repeated Prisoners Dilemma with Imperfect Private Monitoring,” *Journal of Economic Theory*, 102, 70–83.
- PORTER, R. H. (1983): “A Study of Cartel Stability: The Joint Executive Committee, 1880-1886,” *The Bell Journal of Economics*, 14(2), 301–314.
- PROKOPOVYCH, P. (2006): “Perfect Correlated Equilibria in Repeated Games,” University of Michigan Manuscript.
- SEKIGUCHI, T. (1997): “Efficiency in the Prisoners’ Dilemma with Private Monitoring,” *Journal of Economic Theory*, 76, 345–361.
- STAHL, D. (1991): “The Graph of Prisoners’ Dilemma Supergame Payoffs as a Function of the Discount Factor,” *Games and Economic Behavior*, 3, 368–384.
- STIGLER, G. (1964): “A Theory of Oligopoly,” *Journal of Political Economy*, 72(1), 44–61.