Theoretical Economics 12 (2017), 957–978 1555-7561/20170957

Attaining efficiency with imperfect public monitoring and one-sided Markov adverse selection

Daniel Barron
Kellogg School of Management, Northwestern University

I prove an efficiency result for repeated games with imperfect public monitoring in which one player's utility is privately known and evolves according to a Markov process. Under certain assumptions, patient players can attain approximately efficient payoffs in equilibrium. The public signal must satisfy a "pairwise full rank" condition that is somewhat stronger than the monitoring condition required in the folk theorem proved by Fudenberg et al. (1994). Under stronger assumptions, the efficiency result partially extends to settings in which one player has private information that determines every player's payoff. The proof is partially constructive and uses an intuitive technique to mitigate the impact of private information on continuation payoffs.

Keywords. Repeated Bayesian games, efficiency.

JEL classification. C72, C73.

1. Introduction

Many economic environments entail both imperfect observability of actions and persistent, privately known types. Collusive firms have private information about their costs and imperfectly observe their rivals' pricing decisions; a boss motivates her workers without observing either effort or the opportunity cost of that effort; a regulator observes neither a monopolist's actions nor market demand. If formal contracts are unavailable, then repeated interaction must give players an incentive to both share their private information and choose efficiency-enhancing actions given that information. An established literature considers equilibria in repeated games for patient players if actions are imperfectly observed but types are either public or independent and identically distributed (i.i.d.) (Abreu et al. 1990, Fudenberg et al. 1994 (henceforth FLM), Fudenberg and Levine 1994, Fudenberg and Yamamoto 2011, Hörner et al. 2011). A smaller literature has considered games with persistent private information but observable actions (Jackson and Sonnenschein 2007, Escobar and Toikka 2013, Renault et al. 2013 (henceforth RSV)).

Daniel Barron: d-barron@kellogg.northwestern.edu

I am extremely grateful to Juuso Toikka, Johannes Hörner, Robert Gibbons, Glenn Ellison, and two anonymous referees for tremendous advice. Additional special thanks to Niko Matouschek, Jin Li, Mike Powell, and participants in the MIT Theory Lunch. I am indebted to the Cowles Foundation at Yale University for support as I wrote this paper.

DOI: 10.3982/TE1934

This paper takes one step toward analyzing settings with both imperfect monitoring and persistent private information by considering games in which one player is privately informed, this information follows a Markov process, and monitoring is public but imperfect. In this setting, I prove that patient players can attain approximately Pareto efficient payoffs in equilibrium. The proof of this result relies on a novel construction that limits the impact of a player's private information on her expected punishment from deviating. I first prove an efficiency result for a class of games in which the informed player's private information affects only her own payoff. Under stronger assumptions, I extend this argument to games in which the informed player is an "expert" whose type affects *every* player's payoff.

Consider an infinite-horizon game with imperfect public monitoring, and suppose that one player has private information about her own payoffs that evolves according to an irreducible Markov process. Two basic difficulties must be solved to attain efficient payoffs in equilibrium. First, the player with private information must reveal it in such a way that actions can be tailored to the state of the world. Second, players must be induced to choose the correct actions given the informed player's report.

To solve the revelation problem, Section 3 adapts the "quota mechanisms" developed by Jackson and Sonnenschein (2007) and substantially extended by Escobar and Toikka (2013). These mechanisms induce approximately efficient outcomes by requiring the informed player to report each type a fixed number of times in a block of T periods. This "quota" is determined by the invariant distribution of the private information. On the Pareto frontier, I adopt assumptions from RSV to ensure that the informed player cannot increase her utility by permuting her reports, so her gain from lying in such a mechanism is severely limited if T is large. If players were sufficiently patient and could commit to actions as a function of the informed player's message, then payoffs within a large T-period block would approximate those from truth-telling. Hence, a sequence of these T-period mechanisms would lead to approximately efficient payoffs in the infinite-horizon game with commitment.

Section 4 considers how to enforce actions in the game without commitment. Because monitoring is imperfect, players are sometimes punished on the equilibrium path, which implies that the "carrot-and-stick" incentives discussed in Abreu (1986) and used in the folk theorems by Fudenberg and Maskin (1986) and Escobar and Toikka (2013) would lead to substantial inefficiencies even for arbitrarily patient players. Therefore, my argument extends the tools developed in FLM, which proves a folk theorem for repeated games with imperfect public monitoring (but not Markov private information) by constructing rewards and punishments that are approximately efficient if players are sufficiently patient.

A natural adaptation of FLM's proof to the setting with Markov private information would be to construct a sequence of T-period quota mechanisms in which the actions played in each T-period block are used to reward or punish players for outcomes in past blocks. If not done carefully, however, such a construction would not induce efficient play because players have private information about their continuation payoffs. In particular, a player might have an incentive to deviate in the current block because that deviation is likely to lead to continuation play that he privately believes would yield a high payoff in the future. This problem is particularly severe for patient players because such players care mostly about their continuation payoffs, which depend on their current private information.

Because types evolve according to a Markov process, private information in each period is approximately uninformative about payoffs in the distant future. Rewards and punishments in these distant periods can be precisely targeted, so can be used to induce players to follow the equilibrium action without being inefficiently harsh. The proposed equilibrium breaks the infinite-horizon game into sets of K blocks of T periods each. Within each T period block, actions correspond to a fixed allocation rule and the player with private information is given a quota for the number of times she can report each type. One period is chosen at random from each block to determine continuation play. The chosen period affects play only in future blocks that are separated from that period by at least $K-1$ other blocks. Thus, the messages and signals in block 1 influence the actions played in blocks $1+K$, $1+2K$, and so on. Rewards and punishments in these distant blocks are approximately independent of today's private information if K is large. By choosing K appropriately, I can ensure that for very patient players—who would be very tempted to deviate based on their private information about future types—rewards and punishments are nearly uncorrelated with their current private information.

In this construction, rewards and punishments for actions in a block are "delayed" in the sense that they are enacted only after $(K-1)$ other blocks. This feature is similar to a construction in Ellison (1994), which uses delayed punishments to decrease the effective discount factor in a repeated game. While delay also decreases the effective discount factor in my construction, its main purpose is to limit the impact of today's private information on rewards or punishments for today's outcome. Because this outcome only affects the allocation rule in distant periods, and types in these distant periods are approximately independent of today's types, continuation payoffs vary with today's outcome in a way that is nearly independent of today's private information.

Finally, Section 5 extends this equilibrium construction to prove an efficiency result for games in which the informed player's information affects all players' payoffs. For example, the informed player might be an "expert" as in the canonical cheap-talk model of Crawford and Sobel (1982).[1] In such settings, stronger assumptions are required for quota mechanisms to be approximately efficient. RSV prove an efficiency result for cheap-talk games with Markov private types and observed actions, but their construction uses carrot-and-stick punishments and so does not directly extend to games with imperfect public monitoring. I adapt elements of RSV and extend my proof to show that patient players can attain some approximately efficient payoffs in equilibrium, though these payoffs are limited by the demands imposed by the informed player's truth-telling constraints. In particular, an uninformed player's maximum payoff is potentially constrained because the informed player must have an incentive to report truthfully.

In recent years, a growing body of research has focused on dynamic Bayesian games with Markov private types. In an independently developed paper, Hörner et al. (2015)

[1] In such settings, players might infer private information if they observe their past payoffs. My results hold regardless of whether players observe their own payoffs.

builds on FLM to develop techniques for general games with imperfect monitoring and multi-sided private information. Its tools apply to all the games I consider in this paper and many others. It also proves an efficiency result (their Theorem 3) that includes as a special case the class of games I consider in Sections 3 and 4. However, neither this result nor its folk theorem for correlated types (Theorem 5) includes environments in which the informed player's type affects *every* player's payoff, which I study in Section 5. My techniques apply to a more specialized class of games but allow the explicit (partial) construction of a class of equilibria that attain approximately efficient payoffs.

This paper proves an efficiency result for games with (i) one-sided Markov private information, in which (ii) utility is *not* transferable, (iii) players *cannot* commit to actions as a function of reports, and (iv) monitoring is *public* but *imperfect*. Both dynamic mechanism design and repeated games with adverse selection are established literatures; this paper's central contribution is to combine (i)–(iii) with (iv) (though see the discussion of Hörner et al. 2015 above). Pavan et al. (2014) provide very general tools to analyze settings if players can commit and utility is transferable, but its results are difficult to apply in a setting with neither commitment nor transfers. In work closely related to this paper, Escobar and Toikka (2013) consider a setting with neither commitment nor transferable utility and with multi-sided Markov private information; however, it also assumes perfect monitoring. Athey and Bagwell (2001, 2008) study a collusion model with neither commitment nor transferable utility, and with potentially persistent private information. However, these papers restrict attention to their respective applied settings and focus on results for players with fixed discount factors.

2. Model and statement of main result

Consider an infinite-horizon dynamic game with N players. Player 1 ("she"; other players are "he") has a private type $\theta_t \in \Theta = \{\theta^1, \dots, \theta^{|\Theta|}\}$ with $|\Theta| < \infty$ that evolves according to a Markov process with initial distribution $\nu \in \Delta(\Theta)$ and transition probability $P(\theta_{t+1}|\theta_t)$. Player 1 sends a costless public message $m_t \in M_t$ after learning θ_t. After m_t is observed, players simultaneously choose unobserved actions $a_{i,t} \in A_i$ with profile $a_t = (a_{1,t}, \dots, a_{N,t}) \in A = A_1 \times \cdots \times A_N$, which together determine the distribution $F(y_t|a_t)$ of a public signal $y_t \in Y$, $|Y| < \infty$. The utility of player i depends on his private action $a_{i,t}$, the public signal y_t, and—in the case of player 1—the type θ_t: $u_1(a_{1,t}, y_t, \theta_t)$ and $u_i(a_{i,t}, y_t)$ for $i \neq 1$. Denote the vector of period-t payoffs by $u_t \in \mathbb{R}^N$. Players discount at rate δ, with normalized discounted payoffs $\sum_{t=0}^{\infty}(1-\delta)\delta^t u_t$.

The stage game in each period is outlined as follows:

i. A public randomization device $\xi_t \sim U[0,1]$ is realized.

ii. Type $\theta_t \in \Theta$ is drawn according to $P(\theta_t|\theta_{t-1})$.

iii. Player 1 observes θ_t and sends public message $m_t \in M$.

iv. After observing m_t, each player i simultaneously chooses private action $a_{i,t} \in A_{i,t}$.

v. Public signal y_t is realized according to $F(y_t|a_t)$.

To simplify notation, I typically include θ_t as an argument in u_i even if $i \neq 1$. The *expected stage-game payoff for player i* is $g_i(a, \theta) = E_y[u_i(a_i, y, \theta)|a]$, where $g_i(a, \theta)$ is constant in θ for $i \neq 1$. Section 5 relaxes this restriction and instead assumes that θ affects every player's payoff. Let $g(a, \theta) = [g_i(a, \theta)]_{i=1}^N$ denote the vector of expected payoffs. An (pure-strategy) *allocation rule* is a mapping from player 1's private type to an action vector, $\alpha : \Theta \to A$.

All players observe the realization of the randomization device, the message, and the public signal. Player 1 privately observes her type θ_t and every player i privately observes his own action $a_{i,t}$. An *outcome* in period t is $(\xi_t, \theta_t, m_t, a_t, y_t)$, and a *public outcome* is (ξ_t, m_t, y_t). Denote the corresponding histories by $h^T = (\xi_t, \theta_t, m_t, a_t, y_t)_{t=0}^T$ and $h^T_{\text{pub}} = (\xi_t, m_t, y_t)_{t=0}^T$. A strategy for player i, $\sigma_i \in \Sigma_i$, is a mapping from player i's information to feasible actions, with strategy profile $\sigma = (\sigma_1, \dots, \sigma_N) \in \Sigma$. The goal of this paper is to construct a class of perfect Bayesian equilibria that approximate Pareto efficient payoffs.[2]

I assume that $P(\cdot|\cdot)$ converges to an unique invariant distribution to ensure that long-run average payoffs are uniquely defined.

Assumption 1. *Probability $P(\cdot|\cdot)$ is irreducible and aperiodic, with unique invariant distribution π.*

In equilibrium, player i's payoff is bounded below by the payoff he can guarantee himself after a deviation. Following Escobar and Toikka (2013), the *stationary min-max payoff for player i* is his minimum possible payoff if all other players play a constant, pure-strategy action, player i best responds to these actions, and θ is drawn from π. The stationary min-max is not the weakest notion of min-max, since play can neither depend on θ nor be mixed. It is also convenient to define the *max-max payoff for player i*, which is i's maximum expected payoff if θ is drawn from π.

Definition 1. For each player i, define the *stationary min-max* payoff

$$g_i^m = \min_{a_{-i} \in A_{-i}} \max_{\alpha_i : \Theta \to A_i} E_{\theta \sim \pi}\big[g_i\big(\alpha_i(\theta), a_{-i}, \theta\big)\big]$$

with corresponding allocation rule $\alpha^{m,i} : \Theta \to A$. The *max-max* payoff for player i is

$$g_i^M = \max_{\alpha : \Theta \to A} E_{\theta \sim \pi}\big[g_i\big(\alpha(\theta), \theta\big)\big]$$

with allocation rule $\alpha^{M,i} : \Theta \to A$. Define the set of min-max and max-max allocation rules $\mathcal{A}^m$ and $\mathcal{A}^M$, respectively.[3]

[2]Following Fudenberg and Tirole (1991, pp. 331–333), in a perfect Bayesian equilibrium, players update using Bayes' rule whenever possible and best respond given their beliefs. For any a, the distribution over y has full support by Assumption 2. Therefore, Bayes' rule applies unless player 1 has deviated from the equilibrium m. In that case, beliefs may update arbitrarily.

[3]In principle, multiple min-max or max-max allocation rules could exist for a given player i. It suffices that at least one of these allocation rules satisfies the assumptions below; with abuse of notation, let $\alpha^{m,i}$ (or $\alpha^{M,i}$) denote one such allocation rule (if it exists).

For player $i \neq 1$, $\alpha^{m,i}$ and $\alpha^{M,i}$ are constant in θ_t because player i's payoff is constant in θ_t. Player 1's min-max action $\alpha_1^{m,1}$ can vary in θ_t, but the other players' actions $\alpha_{-1}^{m,1}$ cannot. Every element of $\alpha^{M,1}$ can vary in θ_t.

The quota mechanism in my construction induces player 1 to trade off a favorable report *today* for the possibility of a favorable report in the future. If the mechanism implements an allocation rule such that player 1 would be unwilling to permute her reports, then her gains from lying are also sharply limited. This restriction, which is satisfied by any Pareto efficient allocation rule, is closely akin to Rochet's (1987) necessary and sufficient condition for implementability in a static mechanism design problem with transferable utility. It is also identical to Condition C′1 in RSV.

Definition 2. An allocation rule $\alpha : \Theta \to A$ satisfies *cyclical monotonicity* if for any permutation $\psi : \Theta \to \Theta$,

$$\sum_{\theta \in \Theta} g_1\big(\alpha(\theta), \theta\big) \geq \sum_{\theta \in \Theta} g_1\big(\alpha\big(\psi(\theta)\big), \theta\big). \tag{1}$$

Let $\mathcal{A}^{\text{CM}}$ be the set of allocation rules satisfying (1).

I consider the set of *feasible and individually rational stationary payoffs*. This payoff set is constructed in the following way: take all allocation rules $\alpha : \Theta \to A$ that either (i) min-max or max-max some player, or (ii) are cyclically monotone. The target payoff set V^* is the convex hull of expected payoffs from these allocation rules under the invariant distribution, such that every player earns at least his min-max payoff.

Definition 3. Define the set of *implementable* allocation rules as the set of all rules that either satisfy cyclical monotonicity, or min- or max-max a player: $\mathcal{A}^I = \mathcal{A}^{\text{CM}} \cup \mathcal{A}^m \cup \mathcal{A}^M$. Let

$$V = \text{co}\big\{E_{\theta \sim \pi}\big[g\big(\alpha(\theta), \theta\big)\big] | a \in \mathcal{A}^I\big\},$$
$$V^* = \big\{v \in V | \text{for all } i, v_i \geq g_i^m\big\}.$$

As in FLM, the distribution over outcomes F must statistically distinguish between deviations by different players. This property holds if F satisfies a *pairwise full rank* condition. While not required for the result, I also assume for convenience that $F(y|a)$ has full support.

Assumption 2. *For any a, $F(\cdot|a)$ has full support. For any $\alpha \in \mathcal{A}^I$ and any $\theta \in \Theta$, $F(\cdot|\alpha(\theta))$ satisfies pairwise full rank: for any $i, j \in \{1, \dots, N\}$,* $\text{rank}(\Pi_{ij}^y(\alpha(\theta))) = |A_i| + |A_j| - 1$, *where*

$$\Pi_{ij}^y\big(\alpha(\theta)\big) = \begin{bmatrix} \big[F\big(y_r | a_i = a_i^k, \alpha_{-i}(\theta)\big)\big]_{k \leq |A_i|, r \leq |Y|} \\ \big[F\big(y_r | a_j = a_j^k, \alpha_{-j}(\theta)\big)\big]_{k \leq |A_j|, r \leq |Y|} \end{bmatrix}.$$

Pairwise full rank ensures that unilateral deviations by different players are statistically distinguishable from one another. Without this assumption, the equilibrium might require multiple players to be simultaneously punished, which could lead to substantial inefficiencies even if players are patient. Assumption 2 is somewhat stronger than the condition in FLM, which shows that if there exists one action profile with pairwise full rank, then a dense subset of (mixed-strategy) action profiles also satisfies pairwise full rank. My construction uses pure strategies, so pairwise full rank must hold at each allocation rule.[4]

Finally, I assume that if player 1 is being min- or max-maxed, then she has a strict incentive to report a type that leads to the same action as her true type. Therefore, player 1 would only be willing to lie to change the action if she expected a nontrivial gain in the continuation game from doing so.

Assumption 3. *Define*

$$L^m = \min_{\theta,\theta'|\alpha^{m,1}(\theta)\neq\alpha^{m,1}(\theta')} g_1\big(\alpha^{m,1}(\theta),\theta\big) - g_1\big(\alpha^{m,1}(\theta'),\theta\big).$$

If $\alpha^{m,1}(\theta) = \alpha^{m,1}(\theta')$ for all $\theta, \theta' \in \Theta$, then set L^m to some strictly positive number. Let L^M be the corresponding value for $\alpha^{M,1}$. Then $L \equiv \min\{L^m, L^M\} > 0$.

If α min- or max-maxes player 1, then she has a weak myopic incentive to tell the truth. Assumption 3 requires this incentive to be strict, and is implied by two relatively weak conditions: (i) π has full support on Θ and (ii) for every $\theta \in \Theta$, if $a \neq \tilde{a}$, then $g_1(a,\theta) \neq g_1(\tilde{a},\theta)$.

The next two sections prove the following efficiency result.

Theorem 1. *Suppose Assumptions 1–3 hold, and let $W \subseteq \text{int}(V^*)$ be a smooth*[5] *set. For all $\epsilon > 0$, there exists a $\delta^* < 1$ such that if $\delta \geq \delta^*$, for any $w \in W$ there exists an equilibrium payoff $v \in \mathbb{R}^N$ satisfying $\|w - v\| < \epsilon$.*

The proof of Theorem 1 consists of two steps. Section 3 considers a T-period mechanism design problem in which players can *commit* to an allocation rule as a function of player 1's messages. It introduces the central construction of the proof: the (T,K)*-recurrent mechanism*, which implements a sequence of T-period mechanisms in the infinite-horizon game. Section 4 then extends FLM's technique to (T,K)-recurrent mechanisms so as to enforce actions in the game without commitment.

[4]Some variants of the folk theorem from FLM require players to strictly prefer their equilibrium actions, which implies that equilibria must be in pure strategies. These versions typically assume that all extremal points of the payoff frontier satisfy pairwise full rank. See, e.g., Mailath and Samuelson (2006, Proposition 9.2.2).

[5]A set $W \subseteq \mathbb{R}^N$ is *smooth* if (i) it is closed and convex, (ii) it has a nonempty interior, and (iii) the boundary is a C^2 submanifold of $\mathbb{R}^N$: at each boundary point v, there exists a unique tangent hyperplane P_v that varies continuously with v.

3. The mechanism design problem

In this section, I consider a T-period game with commitment. At the start of the game, players commit to a mechanism that specifies actions as a function of public histories. The game is played T times, with normalized discounted payoffs $(\sum_{t=0}^{T-1} \delta^t (1-\delta) u_t)/(1-\delta^T)$.

I consider two classes of mechanisms in this setting. The first is the quota mechanism, used for allocation rules $\alpha \in \mathcal{A}^I \setminus \{\mathcal{A}^m \cup \mathcal{A}^M\}$. Player 1 is given a quota for the maximum number of times she can report each $\theta \in \Theta$. The mechanism implements $\alpha(m)$ following message m.

Definition 4. Fix $\alpha \in \mathcal{A}^I$ in the T-period game with commitment. For any $\theta^j \in \Theta$, define

$$Q(\theta^j) = \begin{cases} \lfloor T\pi(\theta^j) \rfloor + 1 & \text{if } j \leq T - \sum_{\theta \in \Theta} \lfloor T\pi(\theta) \rfloor, \\ \lfloor T\pi(\theta^j) \rfloor & \text{otherwise.} \end{cases}$$

The *T-period quota mechanism implementing* α is, for any $t \in \{0, \dots, T-1\}$, $a_t = \alpha(m_t)$, where

$$m_t \in M_t \equiv \left\{ \theta \in \Theta \,\middle|\, \sum_{t' < t} 1\{m_{t'} = \theta\} < Q(\theta) \right\}.$$

In this mechanism, player 1 faces an intertemporal trade-off: she foregoes reporting a type in the future if she reports that type today.[6] A second important class of mechanisms is *unrestricted* mechanisms. In each period, player 1 reports $m \in \Theta$ and the mechanism plays $\alpha(m)$.

Definition 5. Fix an allocation rule $\alpha \in \mathcal{A}^I$ in the T-period game with commitment. The *T-period unrestricted mechanism implementing* α is, for any $t \in \{0, \dots, T-1\}$, $m_t \in \Theta$ and $a_t = \alpha(m_t)$.

Player 1 must have a myopic incentive to report truthfully in an unrestricted mechanism. If α min- or max-maxes player 1, then by definition she finds it optimal to report truthfully. If α min- or max-maxes some other player, then the allocation rule α is constant in θ and so player 1 again has a (weak) incentive to tell the truth.

The first step of the argument is to define a set of *invariant payoffs*. These are the equilibrium payoffs as $\delta \to 1$ of either (i) a T-period unrestricted mechanism that implements $\alpha \in \mathcal{A}^m \cup \mathcal{A}^M$ or (ii) a T-period quota mechanism that implements some other $\alpha \in \mathcal{A}^I$.

[6] Note that I can use a simpler mechanism than Jackson and Sonnenschein (2007) and Escobar and Toikka (2013) because private information is one-sided in my setting.

Definition 6. If $\alpha \in \mathcal{A}^m \cup \mathcal{A}^M$, consider an unrestricted mechanism. If $\alpha \in \mathcal{A}^I \setminus \{\mathcal{A}^m, \mathcal{A}^M\}$, consider a T-period quota mechanism. Let $\sigma^*_\delta(\alpha)$ be a strategy that maximizes player 1's payoff in this mechanism, and define

$$v^{(\delta,T)}(\alpha) = E\left[\frac{1-\delta}{1-\delta^T}\sum_{t=0}^{T-1}\delta^t u_t \middle| \sigma^*_\delta(\alpha), \nu = \pi\right],$$

where ν is the distribution over types in $t = 0$. Define the *invariant payoff* for α as $v^T(\alpha) = \lim_{\delta\to 1} v^{(\delta,T)}(\alpha)$ whenever this limit exists. The set of *individually rational invariant payoffs* is

$$V^{T*} = \text{co}\{v^T(\alpha) | \alpha \in \mathcal{A}^I, \text{and for any } i, v_i^T(\alpha) \geq v_i^T(\alpha^{m,i})\}.$$

The remainder of this section proceeds as follows. First, I show that as $T \to \infty$, the set of invariant payoffs V^{T*} approximates the set of stationary payoffs V^*. Second, I consider a class of games in which players' payoffs are perturbed and show that as long as these perturbations are small and δ is close to 1, then player 1 will report as if her payoffs are unperturbed. Finally, I construct a class of mechanisms in the infinite-horizon game by concatenating T-period mechanisms. It turns out that under some conditions, player 1's optimal strategy in this infinite-horizon mechanism is simply the concatenation of optimal strategies from the corresponding T-period games.

First, I prove that as $T \to \infty$, invariant payoffs approximate payoffs under truth-telling.

Proposition 1. *For any $\alpha \in \mathcal{A}^I$ and any $T \geq 1$, $v^T(\alpha)$ exists, with*

$$\lim_{T\to\infty} v^T(\alpha) = E_{\theta\sim\pi}[g(\alpha(\theta), \theta)].$$

Define σ^{truth} as the strategy in which $m_t = \theta_t$ for all t. If $\alpha \in \mathcal{A}^m \cup \mathcal{A}^M$, then $\sigma^*_\delta(\alpha) = \sigma^{\text{truth}}$.

All proofs are provided in the Appendix, available in a supplementary file on the journal website, http://econtheory.org/supp/1934/supplement.pdf.

Consider $\alpha \in \mathcal{A}^m \cup \mathcal{A}^M$ and an unrestricted mechanism. Player 1 has a myopic incentive to reveal her type in each period, so $\sigma^*_\delta(\alpha) = \sigma^{\text{truth}}$. For any prior ν, the distribution of θ_t converges to the invariant distribution π as $t \to \infty$. Therefore, if players are patient and T is sufficiently large, then payoffs in the unrestricted mechanism approximate $E_{\theta\sim\pi}[g(\alpha(\theta), \theta)]$.

The proof is a little more complicated if α is implemented using a quota mechanism. For each θ, the normalized quota $Q(\theta)/T$ converges to $\pi(\theta)$ as $T \to \infty$. Players $i \neq 1$ care only about the action taken and, moreover, care little about the timing of those actions if δ is close to 1. Each action $\alpha(\theta)$ is played exactly $Q(\theta)$ times, so as $T \to \infty$, player $i \neq 1$'s invariant payoff approximates $E_{\theta\sim\pi}[g_i(\alpha(\theta), \theta)]$.

Player 1 can always report truthfully in each period and, moreover, it can be shown that she would be unlikely to exceed her quota until the final periods by so doing. Her payoff from reporting truthfully in each period approximates $E_{\theta\sim\pi}[g_1(\alpha(\theta), \theta)]$. Because

$\alpha \in \mathcal{A}^{\text{CM}}$, player 1 cannot gain from permuting her reports. But misreporting her type in each period is analogous to permuting her reports if T is large and δ is close to 1. Therefore, player 1's gains from misreporting are limited and her payoff from her optimal strategy also approximates $E_{\theta\sim\pi}[g_1(\alpha(\theta), \theta)]$.

Truth-telling (or even approximate truth-telling) does *not* necessarily maximize player 1's payoff in the quota mechanism. Consequently, my construction does not exactly pin down equilibrium payoffs and instead uses $v^T(\alpha)$ as an approximation. To formalize this uncertainty about continuation payoffs, I introduce a version of the T-period game with *perturbed* payoffs. Payoffs are similar to the T-period game with commitment, except players also receives some utility $(\sum_{t=0}^{T-1} d((m_t, y_t), \theta_T))/T$ that depends on the history of messages and outcomes in each period and player 1's private type at the end of the game. Perturbed games are characterized by a maximum perturbation size $d \geq 0$, which bounds how much this extra utility can vary.

Definition 7. Fix $d \geq 0$, $T \in \mathbb{N}$, and $\alpha \in \mathcal{A}^I$. A game is (d, T)-*perturbed* if it is a T-period game with commitment and payoffs equal

$$\frac{1-\delta}{1-\delta^T}\sum_{t=0}^{T-1}\delta^t u_t + \frac{1}{T}\sum_{t=0}^{T-1} d\big((m_t, y_t), \theta_T\big). \tag{2}$$

The function $d : M \times Y \times \Theta \to \mathbb{R}^N$ satisfies, for all $m, \hat{m}, \theta \in \Theta$,

$$\big\|E_y\big[d\big((m, y), \theta\big)|\alpha(m)\big] - E_y\big[d\big((\hat{m}, y), \theta\big)|\alpha(\hat{m})\big]\big\| < d.$$

For all $y \in Y$, $\theta \in \Theta$, and $m \in \Theta$ such that $\alpha(m) = \alpha(\hat{m})$, $d((m, y), \theta) = d((\hat{m}, y), \theta)$.

In the next section, the perturbed payoff $\frac{1}{T}\sum_{t=0}^{T-1} d((m_t, y_t), \theta_T)$ reflects the fact that payoffs are not exactly pinned down in a quota mechanism. These perturbations are additively separable in the equilibrium I construct because only a single period from each T-period block (drawn uniformly at random) affects continuation play. The next result shows that if d is small, then player 1's reporting strategy in this perturbed game is identical to her strategy in the unperturbed game.

Proposition 2. *Fix $\alpha \in \mathcal{A}^I$. If $\alpha \in \mathcal{A}^m \cup \mathcal{A}^M$, consider an unrestricted mechanism. Otherwise, consider a T-period quota mechanism. Define*

$$d(\delta, T) = 2\frac{1-\delta}{1-\delta^T}\delta^T TL.$$

Then the following statements hold:

(i) *If $d < d(\delta, T)$, then $\sigma_\delta^*(\alpha)$ maximizes player 1's payoff in any (d, T)-perturbed game.*

(ii) *For any* $\epsilon > 0$, *there exists some* $\chi > 0$ *and* $\delta^* < 1$ *such that if* $\|\nu - \pi\| < \chi$, $\delta > \delta^*$, *and* $d < d(\delta, T)$, *then*

$$E\left[\frac{1-\delta}{1-\delta^T}\sum_{t=0}^{T-1}\delta^t u_t \middle| \sigma_\delta^*(\alpha)\right] \in B\big(v^T(\alpha), \epsilon\big). \tag{3}$$

In a quota mechanism, each action is played a fixed number of times. Hence, the perturbation term $E[\sum_{t=0}^{T-1} d_1((m_t, y_t), \theta_T)|\sigma]$ is constant in σ, so the unperturbed reporting strategy $\sigma_\delta^*(\alpha)$ continues to be an optimal strategy.[7] In an unrestricted mechanism, player 1 could send any message any number of times. However, if α min-maxes or max-maxes player $i \neq 1$, then α is constant in θ and so player 1's message is irrelevant. If α min-maxes or max-maxes player 1, then by Assumption 3 she has a strict incentive to report a type that induces the same action as her true type. If d is small, then player 1 cannot gain too much from the perturbation term by misreporting θ_t, so she reports truthfully.

The final goal of this section is to construct an infinite-horizon mechanism in a way that preserves player 1's incentives from the T-period game. The resulting (T, K)-*recurrent mechanism* is the major building block of the equilibrium construction in Section 4 and is designed to limit the effect of player 1's private information about future payoffs on her temptation to deviate today.

Consider the infinite-horizon game with commitment, defined as the infinitely repeated game in which players can commit *ex ante* to actions as a function of the public history. A (T, K)-recurrent mechanism breaks this infinite-horizon game into blocks of T periods each. In each block, an allocation rule $\alpha \in \mathcal{A}^I$ is implemented using a T-period unrestricted (if $\alpha \in \mathcal{A}^M \cup \mathcal{A}^m$) or quota (if $\alpha \in \mathcal{A}^I \setminus \{\mathcal{A}^M \cup \mathcal{A}^m\}$) mechanism. The allocation rule in each block is determined by messages and outcomes from previous blocks, but only in sharply limited ways as a function of K.

More precisely, a (T, K)-recurrent mechanism groups the first K blocks together and labels them $j = 0$. Similarly, the next K blocks are $j = 1$ and so on. Within each group of K blocks, the first block is labeled $k = 0$, the second $k = 1$, and so on to $k = K - 1$. Hence, each T-period block is identified with a unique label (k, j). The allocation rule implemented in block (k, j) depends *only* on messages and signals from blocks (k, j') for $j' < j$. As illustrated in Figure 1, player 1's message in block (k, j) affects continuation play only in periods at least $T(K - 1)$ periods in the future. Player 1's private information about her type today gives her little information about her type in these far distant periods.

Definition 8. Consider the infinite-horizon dynamic game and fix $T, K \in \mathbb{N}$ and $\delta \in (0, 1)$. For any $j \in \{0, 1, \ldots\}$, $k \in \{0, \ldots, K - 1\}$, define *block* (k, j) as the periods

$$T^{(k,j)} = \big\{(Kj + k)T, (Kj + k)T + 1, \ldots (Kj + k + 1)T - 1\big\}.$$

[7]This argument requires that $F(y|a)$ is independent of θ. Otherwise, the perturbation term would depend on the joint distribution of (y_t, θ_t) and so would not be constant in σ.

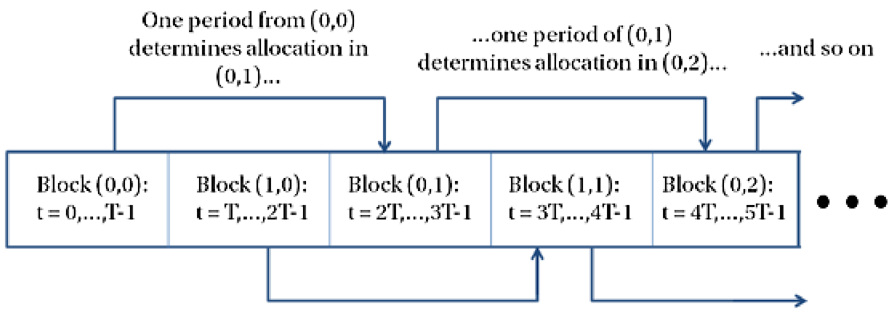

FIGURE 1. A $(T, 2)$-recurrent mechanism. Each block represents T periods. Arrows indicate which periods influence the targeted allocation rule in each block.

Define the (k, j) *block public history* $h^{(k,j)} = (\xi_t, m_t, y_t)_{t\in T^{(k,j)}}$. A mechanism is (T, K)-*recurrent* if it satisfies the following conditions:

(i) For any (k, j), the mechanism is either a T-period quota mechanism implementing $\alpha \in \mathcal{A}^I \setminus \{\mathcal{A}^m \cup \mathcal{A}^M\}$ or an unrestricted mechanism implementing $\alpha \in \mathcal{A}^m \cup \mathcal{A}^M$. Let $\alpha^{(k,j)} \in \mathcal{A}^I$ be the implemented allocation rule.

(ii) At the end of block (k, j), a single period $t^{(k,j)} \in T^{(k,j)}$ is chosen uniformly at random using the public randomization device.

(iii) The rule $\alpha^{(k,j)}$ depends only on $(\alpha^{(k,j')}(m_{t^{(k,j')}}), y_{t^{(k,j')}})_{j'=0}^{j-1}$.

(iv) Let σ^* denote the following strategy: in each block (k, j), player 1 plays the optimal T-period strategy $\sigma^{(k,j)} \equiv \sigma^*_\delta(\alpha^{(k,j)})$ for the mechanism in block (k, j). For any (k, j) and any public history h^t_{pub} such that $t = \min T^{(k,j)}$, let $H^{(k,j)}(h^t_{\text{pub}}, m)$ denote the set of histories in period $t' = \max T^{(k,j)}$ that follow h^t_{pub} for which $m_{t^{(k,j)}} = m$. Then there exists a $\bar{w}^{(k,j)}_1(h^t_{\text{pub}}) \in \mathbb{R}$ such that for any $m \in \Theta$,

$$\sum_{j'=j+1}^{\infty} \delta^{TK(j'-j)}(1-\delta^{TK})E\big[v^T_1\big(\alpha^{(k,j')}\big)|\sigma^*, H^{(k,j)}\big(h^t_{\text{pub}}, m\big)\big] = \bar{w}^{(k,j)}_1\big(h^t_{\text{pub}}\big).$$

A *block* (k, j) consists of T adjacent periods and is denoted $T^{(k,j)}$. For example, in a $(T, 2)$-recurrent mechanism, block $(0, 0)$ consists of the first T periods, block $(1, 0)$ is the next T periods, block $(0, 1)$ is the third set of T periods, and so on.

The first property of a (T, K)-recurrent mechanism says that the T-period mechanism played in each block (k, j) must be either a quota mechanism (if $\alpha^{(k,j)} \in \mathcal{A}^I \setminus \{\mathcal{A}^m \cup \mathcal{A}^M\}$) or an unrestricted mechanism (if $\alpha^{(k,j)} \in \mathcal{A}^m \cup \mathcal{A}^M$). The second property says that at the end of block (k, j), a single period $t^{(k,j)}$ from that block is chosen uniformly at random. Only these chosen periods have any impact on the continuation mechanism. The third property requires the allocation rule in block (k, j) to depend only on the public history in periods $t^{(k,j')}$ with the *same* index k. In particular, the allocation rule implemented in block (k, j) is independent of play in any block (k', j') with $k' \neq k$. The fourth property says that agent 1's expected invariant payoff in future blocks is independent of her messages in the current block. That is, player 1 cannot change her expected future invariant payoffs by changing her reporting strategy.

The final proposition in this section argues that if K and δ are large and δ^{TK} is bounded away from 1, then equilibrium payoffs in a (T, K)-recurrent mechanism are close to the corresponding discounted sum of invariant payoffs. In that case, player 1's payoff in each T-period block can rewritten as a (d, T)-perturbed game with d small, so by Proposition 2 she treats each block as a T-period mechanism.

Proposition 3. *For any $\epsilon > 0$, there exists $K^* < \infty$ and $\delta^* < 1$ such that if $K \geq K^*$, $\delta \geq \delta^*$, and $\delta^{TK} \leq 1 - \epsilon$, then σ^* is an optimal strategy in any (T, K)-recurrent mechanism. For any history h^t in block (k, j) and $j' > j$,*

$$\begin{aligned}&\sum_{t' \in T^{(k,j')}} (1-\delta)\delta^{t'} E\big[u_{t'} | \sigma^*, h^t\big] \\ &\quad \in B\Big(E\big[(1-\delta^T)\delta^{KTj'+kT} v^T\big(\alpha^{(k,j')}\big) | \sigma^*, h^t\big], (1-\delta^T)\delta^{KTj'+kT}\epsilon\Big).\end{aligned} \tag{4}$$

In a (T, K)-recurrent mechanism, player 1's messages in block (k, j) influence her payoffs in two ways: (i) they constrain her messages in the remainder of that block and (ii) they affect the allocation rule in blocks (k, j') for $j' > j$. The strategy $\sigma_\delta^*(\alpha^{(k,j)})$ is optimal given (i). Property (iv) of Definition 8 implies that player 1's message cannot affect her future expected invariant payoff, which implies that (ii) can be captured as a perturbation in a (d, T)-perturbed game.

More precisely, consider the impact of a message in block (k, j) on block $(k, j + 1)$. We can bound the difference between player 1's expected payoff in $(k, j + 1)$ and the invariant payoff $v_1^T(\alpha^{(k,j+1)})$ by $\eta > 0$. Because player 1's private information affects her payoffs in every future block, we can put a similar bound (for simplicity, also $\eta > 0$) on the difference between her expected and invariant payoffs in each (k, j') with $j' > j$. The sum of player 1's gains in every future period is therefore no larger than $\eta/(1 - \delta^{TK})$. The bound $\eta > 0$ is decreasing in the number of periods between (k, j) and $(k, j+1)$ because players have less private information about payoffs in the more distant future. Holding δ^{TK} bounded away from 1 as K and δ increase ensures that this sum approaches 0 (because $\eta \to 0$ as $K \to \infty$). The resulting payoffs can be represented as a (d, T)-perturbed game with d small. By Proposition 2, $\sigma_\delta^*(\alpha^{(k,j)})$ is an optimal strategy in such a game and, hence, σ^* is an optimal strategy in the (T, K)-recurrent mechanism. Expression (4) follows from the definition of invariant payoffs. Note that the error term of (4), $(1 - \delta^T)\delta^{KTj'+kT}\epsilon$, decays exponentially in j', so its sum across j' is small if $\epsilon > 0$ is small.

To understand why (T, K)-recurrent mechanisms facilitate the proof of Proposition 3, suppose instead that actions in each T-period block affect allocation rules in *all* future blocks, but blocks are separated by a fixed number of "burnt" periods T_A in which play is arbitrary. Given a fixed T_A, expected and invariant payoffs in the next block differ by no more than a *fixed* $\eta > 0$. As $\delta \to 1$, player 1's potential gain from misreporting, $\eta/(1 - \delta^{T_A+T})$, becomes large relative to her stage-game payoff, so player 1 might deviate if she is very patient. Increasing T_A as $\delta \to 1$ mitigates this problem but might lead to inefficient payoffs in the limit. The (T, K)-recurrent mechanisms solve this problem by replacing the T_A burnt periods with $K-1$ blocks, each of which implements an approximately efficient T-period mechanism.[8]

4. Enforcing actions in equilibrium

In this section, I modify the techniques used by FLM to construct (T, K)-recurrent mechanisms that are equilibria of the game without commitment. This construction is flexible enough to approximate any payoff in V^* if players are sufficiently patient.

Consider a $(T, 2)$-recurrent mechanism. Intuitively, odd $(1, j)$ and even $(2, j)$ blocks form two distinct games: outcomes observed in block $(1, 1)$ only affect the targeted allocation rule for blocks $(1, 2)$, $(1, 3)$, and so on, and similarly for blocks $(2, j)$. Players' beliefs in blocks $(1, j)$ and $(2, j)$ are still related. However, if continuation play can induce player i to play $\alpha_i(m_t)$ for *any* possible on-path beliefs about θ, then $\alpha(m_t)$ can be enforced in equilibrium.

As a first step, I define the notions of *decomposability* and *enforceability* that I will use in the rest of the proof.

Definition 9. Let $T \in \mathbb{N}$, $\zeta \geq 0$, $W \subseteq \mathbb{R}^N$, and $\delta \in [0, 1)$. A payoff $v \in \mathbb{R}^N$ is (T, ζ, W, δ)-decomposable if there exists $\bar{w} \in \mathbb{R}^N$ such that for every realization of the public randomization device $\xi \in [0, 1]$, there exists some $\alpha^\xi \in \mathcal{A}^I$ and $w^\xi : Y \times \Theta \to W$ satisfying the following conditions:

(i) For any $\xi \in [0, 1]$ and $m \in \Theta$,

$$E_y\left[w^\xi(y, m) | \alpha^\xi(m)\right] = \bar{w}. \tag{5}$$

(ii) The *adding up* constraint holds:

$$v = (1-\delta) E_\xi\left[v^T\left(\alpha^\xi\right)\right] + \delta \bar{w}. \tag{6}$$

(iii) For all $\xi \in [0, 1]$, α^ξ is (T, ζ, W, δ)-*enforceable*: either of the following statements holds.

(a) For every player $i \in \{1, \dots, N\}$, message $m \in \Theta$, and type $\theta \in \Theta$,

$$\begin{aligned} \max_{a_i \in A_i \setminus \alpha_i^\xi(m)} \Big\{ & (1-\delta) \min\left\{0, g_i\left(\alpha^\xi(m), \theta\right) - g_i\left(a_i, \alpha_{-i}^\xi(m), \theta\right)\right\} \\ & + \frac{\delta}{T}\left(\bar{w} - E_y\left[w_i^\xi(y, m) | a_i, \alpha_{-i}^\xi(m)\right]\right)\Big\} \geq \zeta. \end{aligned} \tag{7}$$

(b) The variable α^ξ min- or max-maxes player 1, (7) holds for $i \neq 1$ and all m, θ, and (7) holds for $i = 1$ if $m = \theta$.

[8]Appendix D discusses this alternative mechanism in more detail. I thank two anonymous referees for suggesting this exposition of the intuition.

(iv) For all $\xi \in [0,1]$ and for any $m, m' \in \Theta$, if $\alpha^\xi(m) = \alpha^\xi(m')$, then for any $y \in Y$, $w^\xi(y,m) = w^\xi(y,m')$.

A set W is (T, ζ, δ)-decomposable if every $w \in W$ is (T, ζ, W, δ)-decomposable.

Decomposability is defined with respect to the number of periods in each block T, a set of payoffs $W \subseteq \mathbb{R}^N$, a discount factor δ, and a positive number $\zeta > 0$ that describes the slack in each player's incentive constraints. A payoff $v \in \mathbb{R}^N$ is decomposable if there exists a convex combination of allocation rules and continuation payoffs satisfying four conditions. First, assuming that players follow the allocation rule, the expected invariant continuation payoff is constant in the reported type m. Second, the (appropriately discounted) invariant and continuation payoffs add up to v. Third, each player loses at least $\zeta > 0$ by deviating from the specified action, even if the current period's outcome affects continuation play with probability $\frac{1}{T}$. Importantly, the first term in (7) is the *maximum* of player i's myopic gain from deviating and 0. Finally, messages that induce the same action lead to the same distribution over continuation payoffs.

For $\alpha \in \mathcal{A}^I$ to be (T, ζ, W, δ)-enforceable, players must have strict incentives to play according to α, regardless of their beliefs about θ. The sole exception to this is if the allocation rule min- or max-maxes player 1, in which case player 1's incentive constraint must hold only if $m_t = \theta_t$. Note that (7) requires player 1 to have a strictly smaller continuation payoff following a deviation, even if she is being min- or max-maxed. This is a substantial departure from the argument in FLM, which requires player 1's continuation payoffs to be constant in y_t if she is min-maxed. I discuss this complication after Proposition 5.

Definition 9 superficially resembles the notion of enforceability in FLM (Definition 4.1), but I apply it in a very different way. In my argument, each "period" is in fact a T-period block in a (T, K)-recurrent mechanism. Let $G(k)$ denotes the blocks $((k,0),(k,1),\ldots)$. Since equilibrium play in $G(k)$ depends only on the outcomes in those periods, each $G(k)$ can be treated as a separate game. Of course, these games are not completely independent because types follow a Markov process. Consequently, my equilibrium construction must deter players from deviating for any on-path type and for a range of beliefs about that type.

For a fixed k, (6) can be interpreted as the payoff decomposition for the game $G(k)$. Fix k and define $U_j = \sum_{t=0}^{T-1} \delta^t u_{TKj+Tk+t}$ as the discounted payoff from block (k, j). From the perspective of the first period in $G(k)$, payoffs from the periods in $G(k)$ may be written

$$U_0 + \delta^{TK} U_1 + \delta^{2TK} U_2 + \cdots = \sum_{j=0}^{\infty} \delta^{TKj} U_j. \tag{8}$$

Setting $\tilde{U}_j = ((1-\delta)/(1-\delta^T))U_j$, we can rescale (8) by $(1-\delta)(1-\delta^{TK})/(1-\delta^T)$ to yield $(1-\delta^{TK})\sum_{j=0}^{\infty} \delta^{TKj}\tilde{U}_j$, which is equivalent to the average payoff in a repeated game with discount factor δ^{TK} and stage-game payoffs $\tilde{U}_j$. We can decompose this payoff in the standard way by setting $v = (1-\delta^{TK})\tilde{U}_0 + \delta^{TK}\bar{w}$, where v is the rescaled payoff in $G(k)$, $\tilde{U}_0$ is the average payoff from the first block, and $\bar{w} = (1-\delta^{TK})\sum_{j=1}^{\infty} \delta^{TK(j-1)}\tilde{U}_j$.

Replacing $\tilde{U}_j$ with the expected invariant payoff $E_\xi[v^T(\alpha^\xi)]$ and replacing δ^{TK} with δ yields the payoff decomposition (6).[9]

The next result shows that any (T, ζ, δ)-decomposable set can be approximated by a set of equilibrium payoffs. After stating the result, I discuss how I use (T, K)-recurrent mechanisms to construct an equilibrium from Definition 9.

Proposition 4. *Let $W \subseteq \mathbb{R}^N$ be a closed, convex, bounded set. Suppose there exists a $\hat{\delta} < 1$ and a continuous function $\zeta : [\hat{\delta}, 1) \to (0, \infty)$ such that for all $\delta \geq \hat{\delta}$, W is $(T, \zeta(\delta), \delta)$-decomposable. Then for any $\epsilon > 0$, there exists a $\delta^* < 1$ such that for all $\delta \geq \delta^*$ and $w \in W$, there exists an equilibrium with payoff $v \in B(w, \epsilon)$.*

Let W be (T, ζ, δ^{TK})-decomposable and suppose $w \in W$. By Definition 9, w can be decomposed into the sum of an invariant payoff from an allocation rule and a continuation payoff in W. Iterating this process results in a history-dependent sequence of allocation rules $(\alpha^0, \alpha^1, \ldots)$. A (T, K)-recurrent mechanism can be constructed from this sequence: α^0 is played in all blocks $(k, 0)$, while the allocation rule in block (k, j) is determined by a single period from each of $(k, 0), \ldots, (k, j-1)$. Conditions (i) and (iv) of Definition 9 ensure that the resulting mechanism is (T, K)-recurrent.

Fix the effective discount factor between blocks, δ^{TK}, and note that δ^{TK} can be kept (approximately) constant as δ increases by also increasing K. For K and δ large and δ^{TK} bounded away from 1, Proposition 3 proves that payoffs in block (k, j) would approximate $v^T(\alpha^{(k,j)})$ if players could commit to their actions. Then (6) implies that expected payoffs approximate w. Hence, it suffices to show that players do not wish to deviate from equilibrium actions.

Suppose the allocation rule in the current block does not min- or max-max a player. Then (7) implies that no player has a profitable deviation from the action profile, regardless of that player's belief about θ_t. If the strategy min- or max-maxes player i, then player i cannot myopically gain by either deviating in her action (for any i) or reporting the wrong type (for $i = 1$). One subtlety complicates the argument in this case: (7) holds with respect to the effective discount factor δ^{TK}, while the true gain from deviating in a period is weighted by $(1 - \delta)$. So (7) overweights player i's myopic loss if he deviates from $\alpha(m_t)$. This is why the first term in (7) is constrained to be nonnegative; otherwise, this overweighted myopic loss might be large enough to satisfy the enforceability constraint but not large enough to actually deter deviations.

The penultimate step in the proof holds $T < \infty$ fixed and considers the sets of payoffs that are (T, ζ, δ)-decomposable.

Proposition 5. *Fix a smooth $W \subseteq \operatorname{int}(V^{T*})$. For any $\epsilon > 0$, there exists a $\bar{\delta} < 1$ such that for any $\delta \geq \bar{\delta}$ and $w \in W$, there exists an equilibrium with payoff $v \in B(w, \epsilon)$.*

This discussion focuses on how my proof departs from the analogous argument in FLM. Fix T and consider v on the boundary of W. Let X be the unique plane that is tangent to W at v, and assume for now that X is not parallel to any of the axes. For

[9] I thank an anonymous referee for suggesting this discussion.

any δ, there exists $\alpha \in \mathcal{A}^I$ and a plane X_δ that is parallel to X such that (6) holds for appropriately chosen continuation payoffs in X_δ. Assumption 2 implies that any $\alpha \in \mathcal{A}^I$ is $(T, \zeta, X_\delta, \delta)$-enforceable for any $\zeta \geq 0$. So v is $(T, \zeta, X_\delta, \delta)$-decomposable.

Let $\{w^{\xi,\delta}(y, m)\}_{\xi,y,m} \subseteq X_\delta$ be the set of payoffs that enforce α. If we fix $\zeta = 0$, then an increase in δ has two effects. First, $X_\delta \to X$ as $\delta \to 1$ to satisfy (6). Second, we can decrease the distance between elements of $\{w^{\xi,\delta}(y, m)\}$ without violating (7). FLM show that for sufficiently high δ, the resulting $\{w^{\xi,\delta}(y, m)\}$ lie within W, which implies that v is $(T, 0, W, \delta)$-decomposable.

However, $\zeta > 0$ in the setting with Markov types: players must be given strict incentives to follow equilibrium actions because they have private information about their continuation payoffs. As $\delta \to 1$ and the distance between elements of $\{w^{\xi,\delta}(y, m)\}$ decreases, any fixed $\zeta > 0$ will overwhelm the incentives provided by continuation payoffs, violating (7). I solve this problem by increasing K in the (T, K)-recurrent mechanism as δ increases, which decreases the effect of player 1's private information on her gain from deviating and, hence, leads to a smaller $\zeta > 0$. If K and δ increase so that δ^{TK} is approximately constant, then ζ falls quickly enough to satisfy (7) as $\delta \to 1$. Consequently, any $v \in W$ such that X is not parallel to an axis can be sustained using continuation payoffs in W.

Suppose X is parallel to the ith axis. If continuation payoffs lie in X_δ, then player i's continuation payoff is constant in his action and so (7) cannot be satisfied for i. So continuation payoffs cannot lie in X_δ, which implies that these payoffs are inefficient even as $\delta \to 1$. However, if α min- or max-maxes player i, then the size of this inefficiency is proportional to $\zeta > 0$; if ζ is sufficiently small relative to a fixed δ^{TK}, then continuation payoffs are approximately efficient. I show that such $v \in W$ can be decomposed by an α that min- or max-maxes player i along with nearly efficient continuation payoffs if ζ is small, so v can again be approximated by an equilibrium. Any interior point in W can be attained by public randomization among boundary points. I conclude that W is (T, ζ, δ)-decomposable and so Proposition 4 implies that it can be approximated by equilibrium payoffs.

To prove Theorem 1, it suffices to note that V^{T*} converges to V^* as $T \to \infty$ by Proposition 1. The argument above applies for any fixed T, so any payoff in W can be approximated arbitrarily closely by an equilibrium payoff if δ is close to 1.

5. Extension: Games with an expert

In this section, I turn to games in which player 1's private information affects *every* player's payoff and I extend the argument developed above to these *games with an expert*.

Consider the model from Section 2, with the sole difference that u_i can depend on θ for $i \neq 1$. Intuitively, player 1 is an expert who privately observes a state of the world θ that affects everybody's payoffs. Importantly, player $i \neq 1$ might learn about θ from his realized payoff. For the exposition (and as in RSV), I assume that players do not observe their realized payoffs. However, this assumption is not required for the result to hold: the result in this section also holds (with minor modifications to the proof) if players privately observe their own realized payoffs at the end of each period.[10] I return to this point and discuss it further at the end of this section.

Because θ affects everyone's payoffs, player 1 must be motivated to report truthfully with high probability in most periods. The quota mechanism does not necessarily provide incentives for approximate truth-telling, but RSV have shown that it does if the allocation rule satisfies a strict form of (1). I define such *strictly cyclically monotone* allocation rules.

Definition 10. An allocation rule $\alpha \in \mathcal{A}^{\text{CM}}$ is *strictly cyclically monotone* if for any permutation $\psi : \Theta \to \Theta$, either $\alpha(\theta) = \alpha(\psi(\theta))$ for all $\theta \in \Theta$ or

$$\sum_{\theta \in \Theta} g_1\big(\alpha(\theta), \theta\big) > \sum_{\theta \in \Theta} g_1\big(\alpha\big(\psi(\theta)\big), \theta\big).$$

Define $\hat{\mathcal{A}}^{\text{CM}} \subseteq \mathcal{A}^{\text{CM}}$ as the allocation rules satisfying strict cyclical monotonicity.

Unlike the games studied in Sections 2–4, min- and max-max actions for players other than 1 now depend on θ. Player 1 must be induced to report θ truthfully if another player is being min- or max-maxed, which constrains the payoffs that can be sustained in equilibrium.

Definition 11. For player $i \neq 1$, the *strictly cyclical max-max* for player i solves

$$\hat{g}_i^M = \max_{\alpha \in \hat{\mathcal{A}}^{\text{CM}}} E_{\theta \sim \pi}\big[g_i\big(\alpha(\theta), \theta\big)\big]$$

subject to the condition that for any θ, $\alpha_i(\theta) \in \arg\max_{a_i} g_i(a_i, \alpha_{-i}(\theta), \theta)$. Let $\hat{\alpha}^{M,i}$ be the corresponding allocation rule. Similarly, define the *strictly cyclical min-max* for player i as

$$\hat{g}_i^m = \min_{\alpha \in \hat{\mathcal{A}}^{\text{CM}}} E_{\theta \sim \pi}\big[g_i\big(\alpha(\theta), \theta\big)\big]$$

subject to the condition that for any θ, $\alpha_i(\theta) \in \arg\max_{a_i} g_i(a_i, \alpha_{-i}(\theta), \theta)$. Let $\hat{\alpha}^{m,i}$ be the corresponding allocation rule.

For player 1, min-max $\hat{\alpha}^{m,1}$ and max-max $\hat{\alpha}^{M,1}$ are the same as in Definition 1. Define the set of min-max and max-max allocation rules $\hat{\mathcal{A}}^m$ and $\hat{\mathcal{A}}^M$, respectively.

The min- and max-max payoffs in Definition 11 are clearly not the weakest possible notions. For instance, the payoff $\hat{g}_i^m$ supposes that player i *observes* θ_t, but he could be punished more harshly without this information. However, in that case player i's actions would depend on his beliefs about past θ, which depend on player 1's reporting strategy in previous blocks. The proof of Proposition 4 requires that strategies in one block of a (T, K)-recurrent strategy be independent of past blocks.

[10] Indeed, the argument would work if payoffs were publicly observed. However, in that case players could infer a from these payoffs, which eliminates much of the difficulty of imperfect monitoring.

The set of payoffs that can be sustained in equilibrium is bounded *above* by max-max payoffs. This upper bound arises because a max-max allocation for player i must satisfy two conflicting requirements. First, it must be cyclically monotone to induce player 1 to report truthfully. Second, player i must have a myopic incentive to follow the allocation rule, since otherwise deviations could only be deterred using inefficient continuation play. While the upper bound presented here could almost certainly be relaxed—for instance, it assumes that player i knows the true state of the world—these conflicting requirements intuitively limit the set of allocation rules that could be sustained in equilibrium.

Finally, I assume that for $\alpha \in \{\hat{\alpha}^{m,i}, \hat{\alpha}^{M,i}\}_{i=2}^{N}$, there exist payments such that player 1 would strictly prefer to reveal θ in a static mechanism design problem. I defer discussion of this assumption until after I state the main theorem of this section.

Assumption 4. *There exists $\tilde{L} > 0$ such that for any $\alpha \in \{\hat{\alpha}^{m,i}, \hat{\alpha}^{M,i}\}_{i=2}^{N}$, there exists a function $\tau^{\text{truth}} : \Theta \to \mathbb{R}$ such that for any θ and θ', (i) if $\alpha(\theta) = \alpha(\theta')$, then $\tau^{\text{truth}}(\theta) = \tau^{\text{truth}}(\theta')$; (ii) if $\alpha(\theta) \neq \alpha(\theta')$, then*

$$u_1\big(\alpha(\theta), \theta\big) - \tau^{\text{truth}}(\theta) - \tilde{L} \geq u_1\big(\alpha\big(\theta'\big), \theta\big) - \tau^{\text{truth}}\big(\theta'\big).$$

The structure of the theorem and proof in this section is similar to Theorem 1. I state the theorem formally, then highlight some of the key differences in the proof.

Definition 12. Define $\hat{\mathcal{A}}^I = \hat{\mathcal{A}}^m \cup \hat{\mathcal{A}}^M \cup \hat{\mathcal{A}}^{\text{CM}}$, and let

$$\hat{V} = \text{co}\big\{E_{\theta\sim\pi}\big[g\big(\alpha(\theta), \theta\big)\big] | \alpha \in \hat{\mathcal{A}}^I\big\},$$
$$\hat{V}^* = \big\{v \in \hat{V} | \text{ for any } i, \hat{g}_i^m \leq v \leq \hat{g}_i^M\big\}.$$

Theorem 2. *Suppose Assumptions 1–4 hold and let $W \subseteq \text{int}\,\hat{V}^*$. Then for any $\epsilon > 0$, there exists $\delta^* < 1$ such that if $\delta \geq \delta^*$, for any $w \in W$ there exists an equilibrium payoff $v \in \mathbb{R}^N$ satisfying $\|w - v\| < \epsilon$.*

The proof of Theorem 2 modifies the argument from Theorem 1 in a few key ways. First, player 1 must be induced to report truthfully with high probability in each period, since the other players' payoffs now depend on θ. For $\alpha \in \hat{\mathcal{A}}^{\text{CM}}$, techniques from RSV can be adapted to show that quota mechanisms can induce "approximate truth-telling" among patient players. For allocation rules that min- or max-max player 1, Proposition 2 shows that player 1 has an incentive to report truthfully.

The real new difficulty arises for allocation rules that min- or max-max some player $i \neq 1$. Because $\alpha \in \hat{\mathcal{A}}^{\text{CM}}$, a quota mechanism would induce player 1 to report approximately truthfully in this case. Even if reporting is approximately truthful, however, player i might not have an incentive to follow his min- or max-max action. In particular, if player i believes that player 1 is lying with high probability *conditional on the public history*, then i might prefer to deviate. Because i is being min- or max-maxed, using continuation payoffs to deter such deviations would result in a substantial loss of efficiency even as $\delta \to 1$. Therefore, for such allocation rules, player 1 must be induced to report *exactly* truthfully, rather than just *approximately* truthfully.

My construction uses continuation payoffs to ensure that player 1 has the incentive to tell the truth. If player $i \neq 1$ is min- or max-maxed, then the construction already uses continuation payoffs to ensure she does not deviate in actions. I modify the construction so that, *before* players take actions, player 1's expected continuation payoff mimics the transfers τ^{truth} that induce truth-telling in a static mechanism design problem. Continuation payoffs *after* outcome y is observed are then used to deter deviations from α. Assumption 4 ensures that τ^{truth} strictly motivates truth-telling, so continuation payoffs that approximate these transfers sufficiently closely also induce truth-telling. Therefore, player 1 has the incentive to report truthfully in every period if $i \neq 1$ is being min- or max-maxed.

Finally, suppose that players privately observe their own payoffs. Player 1 already observes $(y_t, a_{i,t}, \theta_t)$ in each period t, so observing her payoff does not change her information or incentives. Any other player could potentially learn θ_{t-1} from his period-$(t-1)$ payoff, which would affect his prior in period t. However, in my equilibrium construction, either (i) player $i \neq 1$ has no incentive to deviate *for any* belief or (ii) player 1 reports truthfully, $m_t = \theta_t$. In either case, player i is willing to follow the equilibrium regardless of his prior at the start of period t. So the proof of Theorem 2 holds even if each player privately observes his own payoffs.[11]

6. Conclusion

My results rely on three important assumptions. First, only player 1 has private information. If multiple players have private information, then the mechanism design problem in Section 3 is not a decision problem and so invariant payoffs might depend on the prior ν. Second, the distribution of y is independent of θ. As noted in footnote 7, this assumption is required to ensure that player 1's optimal reporting strategy does not vary in a slightly perturbed game. It also implies that player 1 can be induced to follow the equilibrium action profile even after she lies about her type. Finally, the distribution over types does not depend on past actions or outcomes, which implies that actions in one block of a (T, K)-recurrent mechanism do not directly affect payoffs in an adjacent block.

In principle, payoffs from other blocks could be used both to induce truth-telling and to deter deviations in actions. Indeed, in Section 5, I use continuation payoffs to induce truth-telling when player $i \neq 1$ is min-maxed or max-maxed. Generalizing this technique might allow the argument to be extended in both Sections 4 and 5.[12]

References

Abreu, Dilip (1986), "Extremal equilibria of oligopolistic supergames." *Journal of Economic Theory*, 39, 191–225. [958]

[11] One minor complication: following a deviation, player $i \neq 1$'s payoff might be inconsistent with player 1's report. However, if player i's beliefs following this event have full support over the set of types, then he will have no incentive to deviate in the continuation game.

[12] Hörner et al. (2015) have an independent argument along these lines.

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990), "Toward a theory of discounted repeated games with imperfect monitoring." *Econometrica*, 58, 1041–1063. [957]

Athey, Susan and Kyle Bagwell (2001), "Optimal collusion with private information." *Rand Journal of Economics*, 32, 428–465. [960]

Athey, Susan and Kyle Bagwell (2008), "Collusion with persistent cost shocks." *Econometrica*, 76, 493–540. [960]

Crawford, Vincent P. and Joel Sobel (1982), "Strategic information transmission." *Econometrica*, 50, 1431–1451. [959]

Ellison, Glenn (1994), "Cooperation in the Prisoner's Dilemma with anonymous random matching." *Review of Economic Studies*, 61, 567–588. [959]

Escobar, Juan F. and Juuso Toikka (2013), "Efficiency in game swith Markovian private information." *Econometrica*, 81, 1887–1934. [957, 958, 960, 961, 964]

Fudenberg, Drew, David Levine, and Eric Maskin (1994), "The folk theorem with imperfect public information." *Econometrica*, 62, 997–1039. [957]

Fudenberg, Drew and David K. Levine (1994), "Efficiency and observability in games with long-run and short-run players." *Journal of Economic Theory*, 62, 103–135. [957]

Fudenberg, Drew and Eric Maskin (1986), "The folk theorem in repeated games with discounting or with incomplete information." *Econometrica*, 54, 533–554. [958]

Fudenberg, Drew and Jean Tirole (1991), *Game Theory*. MIT Press, Cambridge, Massachusetts. [961]

Fudenberg, Drew and Yuichi Yamamoto (2011), "The folk theorem for irreducible stochastic games with imperfect public monitoring." *Journal of Economic Theory*, 146, 1664–1683. [957]

Hörner, Johannes, Takuo Sugaya, Satoru Takahashi, and Nicolas Vieille (2011), "Recursive methods in discounted stochastic games: An algorithm for $\delta \to 1$ and a folk theorem." *Econometrica*, 79, 1277–1318. [957]

Hörner, Johannes, Satoru Takahashi, and Nicholas Vieille (2015), "Truthful equilibria in dynamic Bayesian games." *Econometrica*, 83, 1795–1848. [959, 960, 976]

Jackson, Matthew O. and Hugo Sonnenschein (2007), "Overcoming incentive constraints by linking decisions." *Econometrica*, 75, 241–257. [957, 958, 964]

Mailath, George J. and Larry Samuelson (2006), *Repeated Games and Reputations*. Oxford University Press, New York, New York. [963]

Pavan, Alessandro, Ilya Segal, and Juuso Toikka (2014), "Dynamic mechanism design: A Myersonian approach." *Econometrica*, 82, 601–653. [960]

Renault, Jérôme, Eilon Solan, and Nicholas Vieille (2013), "Dynamic sender–receiver games." *Journal of Economic Theory*, 148, 502–534. [957]

Rochet, Jean-Charles (1987), "A necessary and sufficient condition for rationalizability in a quasi-linear context." *Journal of Mathematical Economics*, 16, 191–200. [962]

Co-editor George J. Mailath handled this manuscript.

Manuscript received 11 August, 2014; final version accepted 1 October, 2016; available online 13 December, 2016.