A FOLK THEOREM WITH MARKOVIAN PRIVATE INFORMATION

JUAN F. ESCOBAR AND JUUSO TOIKKA

Abstract. We study repeated Bayesian two-player games in which the players' privately known types evolve according to an irreducible Markov chain, type transitions are independent across players, and players have private values. The main result shows that, with communication, any Pareto-efficient payoff vector above a minmax value can be approximated arbitrarily closely in a perfect Bayesian equilibrium as the discount factor goes to one. As an intermediate step we construct a dynamic mechanism (without transfers) that is approximately efficient for patient players given sufficiently long time horizon. Our results apply to several games of interest, including models of relational contracting, collusion among asymmetrically informed firms, and insurance with private income shocks.

1. Introduction

Repeated Bayesian games, also known as repeated games of adverse selection, have applications ranging from oligopolistic competition and repeated auctions to relational contracting and voting in international organizations.[1] It is well known that if each player's payoff-relevant private information, or *type*, is independently and identically distributed (iid) over time, repeated play can facilitate cooperation beyond what is achievable in one-shot games.[2] In particular, the folk theorem of Fudenberg, Levine, and Maskin (1994) implies that first-best efficiency can be approximately achieved as the discount factor tends to one.

The assumption of iid types appears restrictive in many applications. For example, firms in an oligopoly or bidders in a series of procurement auctions may have private information about production costs. The costs do vary over time in response to new information and changing firm-specific conditions, but tend to be autocorrelated. In this paper we allow for such serial dependence by assuming that

---

Date: September 24, 2010.
The authors thank Manuel Amador, Kyle Bagwell, Aaron Bodoh-Creed, Matt Elliot, Johannes Hörner, Matt Jackson, Carlos Lever, Jon Levin, Romans Pancs, Ilya Segal, Andy Skrzypacz, and the seminar participants at Stanford, Caltech, UCLA, Yale, Duke, MIT, Harvard, Northwestern, and Princeton for useful discussions and comments. Toikka thanks Ilya Segal, Andy Skrzypacz, Kyle Bagwell, and Matt Jackson for time, advice, and encouragement.
Contact: `jescobar@dii.uchile.cl`, `toikka@mit.edu`.

[1]See Athey and Bagwell (2001), Athey, Bagwell, and Sanchirico (2004), Skrzypacz and Hopenhayn (2004), Levin (2003), and Maggi and Morelli (2006).

[2]See, e.g., Mailath and Samuelson (2006, chap. 11).

the players' types follow an irreducible Markov chain as in the repeated Bertrand oligopoly of Athey and Bagwell (2008).

A player with a Markovian type has private information about the distribution of his future payoffs. This introduces the possibility of signaling, which makes the game inherently dynamic and raises tractability issues that prevent using the standard tools from Abreu, Pearce, and Stacchetti (1990) and Fudenberg, Levine, and Maskin (1994). For instance, in contrast to the iid case, finding any equilibrium can be a nontrivial task as the repetition of a stage-game equilibrium is in general not an equilibrium of the repeated game. As a result, little is known about the set of equilibrium payoffs or behaviors in games with Markovian types. Athey and Bagwell (2008) have shown that under certain conditions it is possible to attain first-best collusion in a symmetric Bertrand duopoly with irreducible two-state Markov costs. However, the results of Myerson and Satterthwaite (1983) imply that in the limiting case where types are perfectly persistent there are games where equilibrium payoffs are bounded away from the first best even when players are arbitrarily patient.[3] Hence it is a priori unclear to what extent the positive results of Athey and Bagwell (2008) on the feasibility of cooperation generalize.

We study the problem of sustaining efficient cooperation among patient players in a class of repeated Bayesian two-player games with Markovian private information. In particular, we consider games in which the players' privately known types affect only their own payoffs (i.e., values are private). The players' types evolve according to an irreducible Markov chain, whose transitions are assumed to be independent across players. Before each round of play, the players privately observe their current types. Then they exchange (cheap-talk) messages. Finally, the players take public actions (i.e., monitoring is perfect).

Our main result shows that any (first-best) Pareto-efficient payoff profile v above a "minmax value" can be approximately attained as a perfect Bayesian equilibrium (PBE) payoff profile, provided that the players are sufficiently patient and a mild restriction on the Pareto frontier is satisfied. Moreover, this can be done so that not only is the expected payoff profile close to v at the start of the game, but the expected continuation payoff profiles are close to v at all histories on the equilibrium path.[4]

[3]See Model 2 in Athey and Bagwell (2008) for an example. Starting with the seminal work of Aumann and Maschler (1995) on zero-sum games, there is a literature on repeated games with perfectly persistent types. See, e.g., Fudenberg and Yamamoto (2009a), Hörner and Lovo (2009), Peski (2008), and Watson (2002). Such models are found also in the reputation literature, notably Kreps and Wilson (1982) and Milgrom and Roberts (1982).

[4]Because of serial dependence of types, the set of feasible payoffs $V(\delta)$ is a function of the discount factor δ in the games we study. Thus, it is impossible to fix an efficient payoff profile independent of δ. Instead, we fix v on the Pareto frontier of the limit set $V = \lim_{\delta\to 1} V(\delta)$ (where the convergence is in the Hausdorff metric) and show that any such v can be approximately attained.

While our result generalizes those of Athey and Bagwell (2008), our techniques are quite different. Whereas they use a constructive argument tailored to the two-type Bertrand duopoly, our proof is only partially constructive and combines mechanism design ideas with repeated game arguments.

We start the proof by considering an auxiliary finite-horizon mechanism design problem in which the players send messages as in the game, but a mechanism enforces actions. We construct an *indirect* dynamic mechanism in which, in each period, players publicly report types, and a fixed efficient choice rule maps the players' reports to actions. Instead of using transfers, incentives are provided by means of history-dependent message spaces. The message spaces allow a player to report a particular type in the current period only if the type is "credible" with respect to the true joint type process given both players' past reports. More precisely, each player is forced to report so that the empirical conditional distributions of his messages—where conditioning is on both players' previous period messages—converge to the true conditional distributions. We show that given any efficient payoff profile v, the mechanism can be constructed so that, by reporting honestly—that is, by reporting as truthfully as possible given the restrictions—each player can secure himself an expected payoff approximately equal to his payoff in v regardless of the other player's strategy, provided that the horizon is long enough and the players are sufficiently patient.

We then consider a "block mechanism," in which the finite-horizon mechanism is repeated indefinitely. We show that in all of its sequential equilibria, continuation payoffs are approximately equal to v at all histories. This is established by bounding the continuation payoffs from below by applying the finite-horizon security payoff result to each block, and bounding them from above using efficiency of v. An existence result by Fudenberg and Levine (1983) for infinite-horizon games of incomplete information implies that the block mechanism has a sequential equilibrium. Together the results imply that, for any Pareto-efficient payoff profile v, there exists a block mechanism that has a sequential equilibrium in which the continuation payoff profile is approximately equal to v at all histories.[5]

Finally, we construct a PBE of the game for patient players that has payoffs close to an efficient target payoff v by "decentralizing" a sequential equilibrium of the block mechanism. On the equilibrium path the players send messages as in the equilibrium of the block mechanism, and mimic the mechanism's actions. This behavior is supported by player-specific stick-and-carrot punishment equilibria similar to the ones used in the proof of the perfect monitoring folk theorem

[5]By the Revelation Principle of Myerson (1986) for multi-stage games, for any equilibrium of our block mechanism, there exists a direct mechanism that has an outcome-equivalent equilibrium with truthful reporting. However, the Revelation Principle requires in general that reports to the mechanism are *confidential*. Thus it is of limited value for the purposes of constructing equilibria of the game in which communication is not mediated.

by Fudenberg and Maskin (1986). The stick phase consists of minmaxing the player who deviated; the carrot phase has the players mimic an approximately efficient equilibrium of a block mechanism that rewards the punisher for following through with the punishment. However, to deal with the possibility of signaling, we obtain the punishment equilibria by decentralizing equilibria of "punishment mechanisms" rather than by direct construction. Our argument proceeds by bounding payoffs during the stick and carrot phases of the mechanism uniformly across equilibria, and then appealing to an existence result to obtain the desired punishment.

Our mechanism is of independent interest in that it gives an approximately efficient dynamic mechanism for patient players without assuming transferable utility. It is inspired by the linking mechanism of Jackson and Sonnenschein (2007), which requires that, over time, the distribution of each player's messages matches his true type distribution.[6] With iid types the linking mechanism approximately implements efficient choice rules given long enough horizon and sufficient patience. However, when types are Markovian, a player can use his opponent's past reports to predict her future types. This gives rise to contingent deviations, which undermine the linking mechanism. Our mechanism rules out these deviations by requiring convergence of *conditional* distributions.

If the set of stage-game actions includes the possibility to make budget-balanced transfers and payoffs are quasilinear, then Pareto efficiency can be achieved in our auxiliary mechanism design problem by using the Balanced Team Mechanism of Athey and Segal (2007). Our contribution to the literature on dynamic mechanism design is thus showing how approximate efficiency can be achieved by patient players even when there are wealth effects or utility is not transferable.[7]

We conclude the Introduction by discussing the relationship of our work to the literature on repeated games. A natural starting point is provided by the recursive tools developed by Abreu, Pearce, and Stacchetti (1990) and Fudenberg, Levine, and Maskin (1994) for the characterization of the equilibrium payoff set in repeated games with imperfect public monitoring. In particular, Fudenberg, Levine, and Maskin (1994) prove a folk theorem for such games under certain identifiability restrictions on the monitoring technology. They further observe that a repeated Bayesian game in which types are independent across periods and players can be

[6]Given identical and independent copies of a social choice problem, the linking mechanism of Jackson and Sonnenschein (2007) assigns each player a budget of messages to be used over the problems. The budget forces the distribution of the player's reports over the problems to match the true distribution from which the player's types are drawn. For earlier work using the idea, see for instance Radner (1981) and Townsend (1982).

[7]Athey and Segal (2007) show further that in some settings their mechanism can be made self-enforcing if the players are sufficiently patient. However, the result maintains the assumptions about transfers, and assumes in addition that there exists a "static" punishment equilibrium. We dispense with both transfers and the assumption about the punishment equilibrium.

converted to a repeated game with imperfect public monitoring in which the public signal has product structure.[8] As a result, they obtain a Nash-threat folk theorem for such games. In the special case of iid types neither result implies the other: Our approach uses the pure-action minmax value as the threat point. Hence there are games such as the incomplete information versions of the Bertrand and Cournot duopolies in which we (asymptotically) achieve a superset of the efficient points achieved by Fudenberg, Levine, and Maskin (1994), but if the stage game has a non-degenerate mixed-strategy equilibrium the inclusion can be reversed.[9]

Constructing equilibria in repeated Bayesian games requires dealing with the problem of tracking the evolution of beliefs over private histories. Our approach sidesteps this difficulty by using the auxiliary mechanism design problem to prove the existence of strategies that result in bounds on equilibrium payoffs that hold uniformly in public and private histories and, thus, uniformly in beliefs. This is a key element in our proof and uses the irreducibility of the type process. However, even if payoffs are bounded uniformly in beliefs, the players' best responses still depend on the beliefs. This is unlike the "belief-free" approach of Hörner and Lovo (2009), Hörner, Lovo, and Tomala (2009), and Fudenberg and Yamamoto (2009a)—who study games with perfectly persistent types—in which attention is restricted to equilibria in strategies that are best responses regardless of beliefs.

Our approach is reminiscent of the "review strategies" of Radner (1985). However, our incentive problem is one of adverse selection rather than moral hazard.[10] In particular, contrary to the signals in the case of moral hazard, a player has full control over the messages he sends. As a result, under moral hazard the inefficiency in the approximately efficient equilibria comes from the fact that, with small probability, the agent fails the review which triggers the punishment. In contrast, the inefficiency in our equilibria stems from the fact that the players are sometimes forced to lie to avoid triggering the punishment.

Finally, a repeated game with Markovian private information can be viewed as a stochastic games where players have asymmetric information about the state. Dutta (1995), Fudenberg and Yamamoto (2009b), and Hörner, Sugaya, Takahashi, and Vielle (2010) prove increasingly general versions of the folk theorem for stochastic games with a public irreducible state.

---

[8]For examples of papers using this approach, see footnote 1, and also Abdulkadiroglu and Bagwell (2007) and Hauser and Hopenhayn (2008) who study trading of favors as in Mobius (2001).

[9]Cole and Kocherlakota (2001) extend the recursive characterization by Abreu, Pearce, and Stacchetti (1990) to a class of games that includes ours (see also Fernandes and Phelan, 2000). Their method operates on pairs of type-dependent payoff profiles and beliefs. The inclusion of beliefs makes the operators hard to manipulate, and, as a result, the characterization is difficult to put to work. In particular, extending the techniques of Fudenberg, Levine, and Maskin (1994) to this case appears difficult.

[10]Review strategies have also been used in adverse selection context with iid types. See, e.g., Radner (1981), and Hörner and Jamison (2007).

The rest of the paper is organized as follows. In the next section we sketch our argument in the context of a simple Bertrand game. We then set up the model in Section 3 and present the main result in Section 4. We consider the auxiliary mechanism design problem in Section 5. In Section 6 we prove the main theorem building on the results from Section 5. We conclude in Section 7. Two appendices collect the proofs and auxiliary results we omit from the main text. A reader who is mainly interested in the mechanism design part can read only Sections 2, 3.1–3.2, 3.4, and 5 without loss of continuity.

2. An Example

Consider repeated price competition between two firms, 1 and 2, whose privately known marginal costs are $\theta_1 \in \{L, H\}$ and $\theta_2 \in \{M, V\}$, respectively, with $L < M < H < V$ (i.e., "low, medium, high, and very high"). Firm 1's cost evolves according to a symmetric Markov chain in which with probability $p \geq \frac{1}{2}$ the cost in period $t+1$ is the same as in period t; firm 2's costs are iid and equiprobable. The cost draws are independent across firms. In each period there is one buyer with reservation value $r > V$. Having first privately learned their current cost realizations, the firms exchange cheap-talk messages, and then quote prices.

This duopoly game is a special case of the general model introduced in Section 3, and hence our main result (Theorem 4.1) applies. To illustrate our general proof strategy, we sketch here the argument showing that, given sufficiently little discounting, there exists equilibria with profits arbitrarily close to the profits from the collusive scheme where in each period the firm with the lowest cost makes the sale at the monopoly price r.

2.1. A Mechanism Design Problem.

We start by considering the case where the horizon T is large but finite, and the firms do not discount profits. We assume further that the firms only exchange messages and some mechanism automatically sets prices as a function of these messages.

It is instructive to consider first using the linking mechanism of Jackson and Sonnenschein (2007), which would (approximately) yield the desired profits if both firms' costs were drawn iid over time. In the linking mechanism each firm is only allowed to report each cost in $\frac{1}{2}$ of the periods as this is the long-run distribution of costs for each firm. In every period the mechanism then sets the price to r for the firm who reported the lowest cost, and to $r+1$ for the other firm.

Suppose first that both firms report honestly, i.e., as truthfully as they can. Then, for T large, firm 2 makes a sale in approximately $\frac{T}{4}$ periods. The resulting (average) profits are approximately $\frac{r-L}{2} + \frac{r-H}{4}$ for firm 1, and $\frac{r-M}{4}$ for firm 2.

Suppose then that instead of reporting honestly, firm 2 sends message M if and only if firm 1 reported H in the previous period. (This strategy is feasible as it

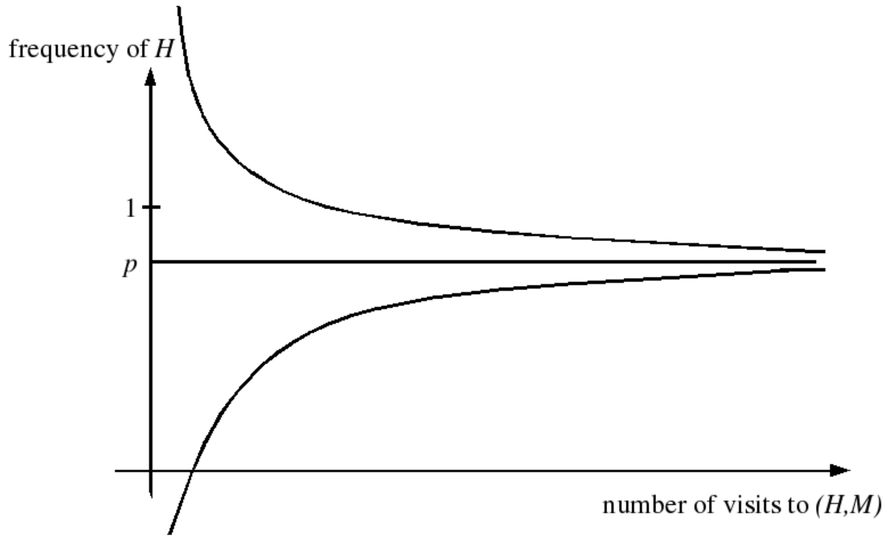

FIGURE 1. Reporting restrictions.

results in firm 2 reporting M in $\frac{1}{2}$ of the periods.) For T large, firm 2 makes a sale in approximately $p\frac{T}{2}$ periods earning approximately $p\frac{r-M}{4} + p\frac{r-V}{4}$. So for p large enough, firm 2 is strictly better off matching its reports to firm 1's reports rather than to its own costs, and hence approximate truth-telling is not an equilibrium of the linking mechanism. Moreover, the misrepresentation by firm 2 hurts the honest firm 1, whose payoff drops from $\frac{r-L}{2} + \frac{r-H}{4}$ to $\frac{r-L}{2} + (1-p)\frac{r-H}{2}$.

Note that the deviation by firm 2 introduces strong correlation between the firms' reports whereas the true cost processes are independent. In particular, conditional on firm 1's cost being H at t, the true frequency of firm 2's cost M at $t+1$ is $\frac{1}{2}$, not 1. This suggests ruling out contingent deviations by forcing the firms' reports to match the true distribution conditional on past reports.

Motivated by the above observation, we construct the following mechanism. Fix a message profile (θ_1, θ_2) and consider the (random) set of all periods in which the (not necessarily truthful) previous period message profile was (θ_1, θ_2). Our mechanism requires that the frequencies of firm i's reports over these periods converge to the corresponding true conditional cost frequencies, where the conditioning is on θ_i. For example, over the periods that follow reports (H, M), the frequency with which firm 1 reports H must converge to p as the number of visits to (H, M) tends to infinity. Similarly, over the said periods, the frequency with which firm 2 reports M must converge to $\frac{1}{2}$.

This restriction on reporting is schematically illustrated in Figure 1. Imagine plotting on the picture the frequency at which firm 1 has reported H over the periods that follow (H, M) in the previous period. The mechanism allows firm 1 to report only in such a way that this frequency as a function of the total number of visits to (H, M) stays within the bounds given by the two curves converging to the true frequency p.

Our mechanism tracks the frequency with which firm 1 reports H following each possible message profile, and the frequency with which firm 2 reports M following each message profile. So in total it tracks eight different frequencies and requires the firms to report such that all of them stay within acceptable bounds.

Consider firm 1 that is reporting honestly. For simplicity, assume that it can report truthfully in every period. (We show in the proof of the main result that the bounds on reporting can be chosen such that an honest player is very likely to be truthful.) Then firm 1's reports over the periods in which the previous period messages were (H, M) are independent draws from $(1-p)[L] + p[H]$. Since the firms report simultaneously, this implies that the joint distribution of their messages over these periods converges to the product distribution

$$\begin{pmatrix} (1-p)\frac{1}{2} & (1-p)\frac{1}{2} \\ p\frac{1}{2} & p\frac{1}{2} \end{pmatrix},$$

regardless of firm 2's strategy. Note that this is in fact the true conditional distribution of period-$t+1$ costs given cost profile (H, M) in period t.

Similar calculations for the other three cost profiles in place of (H, M) show that if firm 1 is truthful, then the empirical transition distributions for the sequence of message profiles converge to the true transition distributions for the joint cost process regardless of the strategy of firm 2.[11] We may then use the fact that convergence of transitions implies convergence of the empirical distribution to the invariant distribution[12] to conclude that the distribution of messages converges to the invariant distribution for the joint cost process. In particular, this implies that the truthful firm 1 faces the same distribution of firm 2's costs as it would if firm 2 was reporting honestly regardless of 2's actual reporting strategy. But given private values, firm 1's profit must be approximately equal to its profit under mutual truth-telling, i.e., $\frac{r-L}{2} + \frac{r-H}{4}$. Note that this is firm 1's profit in the collusive scheme we are trying to sustain.

The above heuristic argument shows that given the history-dependent restrictions on messages, firm 1 can secure a profit approximately equal to the target collusive profits regardless of the strategy of firm 2 by simply reporting honestly. The symmetric argument for firm 2 then implies that in any equilibrium the firms' profits are bounded from below by approximately the target profits $\frac{r-L}{2} + \frac{r-H}{4}$ and $\frac{r-M}{4}$, respectively. But then the profits have to actually be close to these numbers by feasibility, even if in general the firms do not report honestly in equilibrium.

[11]For the general model this result is established in the proof of Proposition 5.1. The actual formal argument has to consider all past message profiles simultaneously, since part of the problem is to show that each of them is visited often enough for law-of-large-numbers arguments to apply.

[12]See Lemma A.1 in Appendix A.

2.2. **From the Mechanism to Game Equilibria.** In order to construct equilibria of the original repeated pricing game based on the mechanism sketched above, we need to (1) introduce discounting, (2) extend the result to an infinite horizon, and (3) make the mechanism self-enforcing.

Discounting can be introduced simply by continuity since the mechanism design problem has a finite horizon.

We cover the infinite horizon by having the firms repeatedly play the finite horizon mechanism over T-period blocks. This serves to guarantee that the firms' continuation profits are close to the target at all histories—a result used in establishing (3). It is worth noting that because of the autocorrelation in firm 1's costs, it is not possible to simply assume that the firms treat adjacent blocks independently of each other. However, the lower bound on profits from honest reporting does apply to each block, and this can be used to bound the continuation profits in all equilibria of the block mechanism.

Finally, the block mechanism can be made self-enforcing by using the following stick-and-carrot scheme. If firm 1 deviates by making a cost report that would have been infeasible in the mechanism or by naming a price different from what the mechanism would have chosen for it, then firm 2 prices at L for a finite number of periods. Firm 1 best responds during this stick phase. Since continuation play in general depends on firm 2's beliefs about firm 1's costs, firm 1's best response in general involves signaling in order to impact continuation play upon entering the carrot phase. The construction of the carrot phase is analogous to that of the equilibrium path, but it builds on a mechanism that is based on a collusive scheme that has firm 2 selling—still at the monopoly price—more frequently than in the target collusive scheme. This is done to reward firm 2 for the losses it incurred during the stick phase. Firm 2's deviations are punished analogously.

3. The Model

We consider dynamic two-player games where a fixed Bayesian stage game is played in each period over an infinite horizon.

3.1. **The Stage Game.** The stage game is a finite Bayesian two-player game in normal form. Let $I = \{1, 2\}$ denote the set of players. It is convenient to identify the stage game with the payoff function

$$u : A \times \Theta \to \mathbb{R}^2,$$

where $A = A_1 \times A_2$ is a finite set of action profiles, and $\Theta = \Theta_1 \times \Theta_2$ is a finite set of possible type profiles. The interpretation is that each player $i \in I$ has a privately known type $\theta_i \in \Theta_i$ and chooses an action $a_i \in A_i$. We allow for (correlated) mixed actions by extending u to $\Delta(A) \times \Theta$ by taking expectations.

We assume throughout that the stage game u has private values, stated formally as follows.

Assumption 3.1 (Private Values)**.** For all $i \in I$, $a \in A$, $\theta \in \Theta$ and $\theta' \in \Theta$,

$$\theta_i = \theta_i' \quad \Rightarrow \quad u_i(a,\theta) = u_i(a,\theta').$$

Given the assumption, we write $u_i(a,\theta_i) = u_i(a,\theta)$. Under private values a player is concerned about the other player's type only in so far as it influences the action chosen by the other player.

3.2. **The Dynamic Game.** The dynamic game has the stage game u played over an infinite horizon with communication allowed in each period $t = 1, 2, \ldots$ Player i's current type θ_i^t evolves according to a Markov chain (λ_i, P_i) on Θ_i, where λ_i is the initial distribution, and P_i is the transition matrix. The timing within period t is as follows:

- $t.1$ Each player $i \in I$ privately learns $\theta_i^t \in \Theta_i$.
- $t.2$ The players send simultaneous public messages $m_i^t \in \Theta_i$.
- $t.3$ The players observe the outcome of a public randomization device.
- $t.4$ The stage game u is played with the realized actions $a_i^t \in A_i$ perfectly monitored by all players.

We do not introduce notation for the public randomization device in order to economize on notation.[13]

Let (λ, P) denote the joint type process, i.e., a Markov chain on Θ induced by the Markov chains (λ_i, P_i), $i \in I$, for the players. We make two assumptions about the joint type process.

Assumption 3.2 (Irreducible Types)**.** P is irreducible.[14]

Irreducibility of P implies that the dynamic game is stationary, or repetitive, in a particular sense. It also implies that each P_i is irreducible, and hence for each chain (λ_i, P_i) there exists a unique invariant distribution π_i.

Assumption 3.3 (Independent Transitions)**.** For all $\theta \in \Theta$ and $\theta' \in \Theta$,

$$P(\theta,\theta') = P_1(\theta_1,\theta_1')P_2(\theta_2,\theta_2').$$

The assumption of independent transitions imposes conditional independence across players. That is, the players' types in period $t+1$ are independent conditional on the types in period t. However, no restrictions are put on the joint initial

[13]Since we allow for communication, there is a sense in which allowing for a public randomization device is redundant. Namely, provided that the set of possible messages is large enough, the players can conduct jointly-controlled lotteries to generate public randomizations (see Aumann and Maschler, 1995).

[14]Under Assumption 3.3, a sufficient (but not necessary) condition for P to be irreducible is that each P_i is irreducible and aperiodic (i.e., that each P_i is ergodic).

distribution λ. Thus, unconditionally, types are not necessarily independent across players. Independence of the transitions implies that the invariant distribution for the joint process, denoted π, is the product of the π_i.

Player i's dynamic game payoff is the discounted average of his stage game payoffs. That is, given a sequence $(x_i^t)_{t=1}^{\infty}$ of stage game payoffs, player i's dynamic game payoff is given by

$$(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}x_i^t,$$

where the discount factor $\delta \in [0,1[$ is assumed common for all players.

3.3. **Histories, Assessments, and Equilibria.** A public history in the game can be of two sorts. For each $t \geq 1$, some public histories contain all the messages and actions taken up to and including period $t-1$, whereas others contain all the messages and actions taken up to period $t-1$ together with the message sent at the beginning of period t. The first type of history takes the form $(m^1, a^1, \dots, m^{t-1}, a^{t-1})$, whereas the second type takes the form $(m^1, a^1, \dots \dots m^{t-1}, a^{t-1}, m^t)$. The set of all public histories at t is thus $H^t = (\Theta^{t-1} \times A^{t-1}) \cup (\Theta^t \times A^{t-1})$ and the set of all public histories is $H = \cup_{t \geq 1} H^t$.

A private history of length t for player i consists of the sequence of private types drawn up to and including t. Formally, the set of private histories of length t for player i is $H_i^t = \Theta_i^t$ and the set of all private histories is simply $H_i = \cup_{t \geq 1} H_i^t$.

A (behavior) strategy for player i is a sequence of functions $\sigma_i = (\sigma_i^t)_{t \geq 1}$ such that $\sigma_i^t \colon H^t \times H_i^t \to \Delta(A_i) \cup \Delta(\Theta_i)$ with $\sigma_i^t(\cdot \mid h^t, h_i^t) \in \Delta(A_i)$ if $h^t \in \Theta^t \times A^{t-1}$, while $\sigma_i^t(\cdot \mid h^t, h_i^t) \in \Delta(\Theta_i)$ if $h^t \in \Theta^{t-1} \times A^{t-1}$.

A belief system for player i is a sequence $\mu_i = (\mu_i^t)_{t \geq 1}$ such that $\mu_i^t \colon H^t \times \Theta_i \to \Delta(\Theta_{-i}^t)$. Note that the belief player i forms about his rival, $\mu_i^t(\cdot \mid h^t, \theta_i^1)$, depends on his private history of types only through his first type θ_i^1. This is so since transitions are independent while initial types can be correlated. Thus it is natural to rule out beliefs that condition on irrelevant information, namely the own private history of types beyond the initial type.[15]

An assessment is a pair (σ, μ) where $\sigma = (\sigma_i)_{i \in I}$ is a strategy profile and $\mu = (\mu_i)_{i \in I}$ is a belief system profile. Given any assessment (σ, μ), let $u_i^{\mu_i}(\sigma \mid h^t, h_i^t)$ denote player i's continuation value at history (h^t, h_i^t), i.e., the expected sum of discounted average payoffs for player i after history (h^t, h_i^t), given the strategy profile σ and taking expectations over i's rival's private histories according to $\mu_i^t(\cdot \mid h^t, h_i^t)$

An assessment (σ, μ) is sequentially rational if for any player i, any history (h^t, h_i^t) and any strategy σ_i' for i, $u_i^{\mu_i}(\sigma \mid h^t, h_i^t) \geq u_i^{\mu_i}(\sigma_i', \sigma_{-i} \mid h^t, h_i^t)$.

[15]When the initial types are independently drawn, it is natural to restrict attention to belief systems such that for any j, j's rival forms beliefs about j's private history using a map $\mu_j^t \colon H^t \to \Delta(\Theta_j^t)$.

We say that the belief system profile $\mu = (\mu_i)_{i\in I}$ is computed using Bayes rule given strategy profile $\sigma = (\sigma_i)_{i\in I}$ if $\mu_i^1(\theta_{-i} \mid \theta_i^1) = \lambda_{-i}(\theta_{-i} \mid \theta_i^1)$ and if for any h^t and $\theta_i^1 \in \Theta_i$ with $\sigma_{-i}(x_{-i} \mid h^t, h^t_{-i}) > 0$ and $\mu_i^t(h^t_{-i} \mid h^t, \theta_i^1) > 0$ for some $x_{-i} \in A_{-i} \cup \Theta_{-i}$ and $h^t_{-i} \in \Theta^t_{-i}$, the belief player i forms at history $((h^t, x), \theta_i^1)$ is computed using Bayes rule (i.e., Bayes rule is used wherever possible both on and off the path of play).

An assessment (σ, μ) is a perfect Bayesian equilibrium if it is sequentially rational and μ is computed using Bayes rule given σ.

3.4. **Feasible Payoffs.** We now consider sequences $f = (f^t)_{t\geq 1}$ of arbitrary decision rules $f^t \colon \Theta^t \times A^{t-1} \to \Delta(A)$ mapping histories consisting of types and actions into distributions over actions. We define the set of feasible discounted payoffs attained using all such sequences, given the discount factor δ, as

$$V(\delta) = \Big\{ v \in \mathbb{R}^2 \mid \text{ for some } f = (f^t)_{t\geq 1}, \\ v_i = (1-\delta)\mathbb{E}_f \big[\sum_{t=1}^{\infty} \delta^{t-1} u_i(a^t, \theta_i^t) \big] \text{ for all } i \in I \Big\},$$

where the expectation $\mathbb{E}_f$ is with respect to the probability measure induced over the set of histories by the decision rules $f = (f^t)_{t\geq 1}$ and the joint type process (λ, P).

It is useful to consider the set of all payoffs attainable using randomized rules in a one-shot interaction in which types $\theta \in \Theta$ are drawn according to the invariant distribution π, formally defined as

$$V = \big\{ v \in \mathbb{R}^2 \mid \text{ for some } f \colon \Theta \to \Delta(A),\ v_i = \mathbb{E}_\pi \big[u_i(f(\theta), \theta_i) \big] \text{ for all } i \in I \big\}.$$

Note that V depends neither on λ nor on δ.

Using the irreducibility of P, the following result shows that for discount factors close to 1, $V(\delta)$ is approximately equal to V.

Lemma 3.1 (Dutta, 1995)**.** *As $\delta \to 1$, $V(\delta) \to V$ in the Hausdorff metric. Moreover, the convergence is uniform in the initial distribution λ.*

Heuristically, the result follows from noting that for δ close to 1 only the invariant distribution of types matters, and hence the limit is independent of the initial distribution. Moreover, given the stationarity of the environment, stationary (but in general randomized) decision rules are enough to generate all feasible payoffs. Consult Dutta (1995) for details.[16]

In what follows we investigate what payoffs $v \in V$ can be attained in equilibrium when the discount factor is arbitrarily large, keeping in mind that in this case $V(\delta)$ is arbitrarily close to V.

[16]Dutta (1995) studies dynamic games where the state is publicly observable. However, as feasibility is defined without reference to incentives, the result applies verbatim.

3.5. **Minmax Values.** We define player i's *(pure action) minmax value* as

$$\underline{v}_i = \min_{a_{-i} \in A_{-i}} \mathbb{E}_{\pi_i}\Big[\max_{a_i \in A_i} u_i((a_i, a_{-i}), \theta_i)\Big].$$

Our motivation for this definition comes from observing that $\underline{v}_i$ is approximately the lowest payoff that can be imposed on a very patient player i if player $-i$ is restricted to playing a fixed pure strategy for a long time, and player i best responds to that action knowing his current type.[17] This in turn is motivated by the practical concern to be able to construct punishment equilibria that generate payoffs close to the minmax value. In particular, no claim is made that $\underline{v}_i$ would in general be the harshest punishment, or that all equilibria would need to give i at least this payoff. There are games such as the Cournot and Bertrand oligopoly examples below where this minmax value indeed corresponds to the worst possible punishment, but there are also games where randomization by player $-i$ would allow for a harsher punishment (e.g., standard matching pennies).[18] Furthermore, since there is serial correlation, it is conceivable that player $-i$ could try to tailor the punishment to the information he learns about i's type during the punishment rather than simply play a fixed action.[19] However, constructing punishment equilibria that deal with these two extensions appears complicated and is left for future research.

Despite the possible limitations discussed above, our notion of minmax (and the punishment equilibria that we construct based on it) provides an effective punishment that facilitates sustaining good outcomes in a large class of games. We note that in the special case of a repeated Bayesian game with iid types our definition of the minmax value reduces to the standard pure action minmax value.

[17]To see this, note that when player i is patient, his long-run payoff from any stationary decision rule—such as the one where $-i$ plays a fixed action and i myopically best responds knowing his type—is approximately equal to the expectation of his payoff from the stationary decision rule under the invariant distribution.

[18]Fudenberg and Maskin (1986) prove a mixed minmax folk theorem by adjusting the continuation payoffs in the carrot phase so that during the stick phase players are indifferent over all actions in the support of a mixed minmax profile, and hence willing to mix. Extending this logic to our setting faces two problems: (1) variation in payoffs during the stick phase is private information, and (2) our methods characterize continuation values only "up to an ε." Alternatively, following Gossner (1995), one could try to use a statistical test to make sure that the punisher is mixing in the right proportions (triggering a punishment against him if he fails the test). Again, some difficulties arise. The construction of our punishment mechanism (Section 6.1) would need to allow the minmaxing player to pick actions, for in the repeated game his strategy will be the minmaxing distribution up to some ε. This opens the door for signaling, and in addition, one of the players may want to correlate his play with that of the rival. The test could accommodate this through restrictions on conditional distributions, but the constructions are clearly more involved. These generalizations are left for future work.

[19]When types are perfectly persistent, Blackwell's approachability theorem yields the proper minmax values (Blackwell, 1956; Hörner and Lovo, 2009). The extension to games with imperfectly persistent types seems promising, but far from obvious.

	S	O
S	$\theta_1, 2$	$0, 0$
O	$0, 0$	$2, \theta_2$

FIGURE 2. Battle of sexes with taste shocks.

3.6. **Examples.** This subsection illustrates our dynamic game model and some of the definitions already introduced.

Example 3.1 (Cournot competition)**.** *Each player i is a firm that chooses a quantity $a_i \in A_i$. We assume that $A_i \subseteq [0, \infty[$, $0 \in A_i$ and there is $\bar{a}_{-i} \in A_{-i}$ such that $\bar{a}_{-i} \geq 1$. The market price is given by $p_i = \max\{1 - \sum_{i \in I} a_i, 0\}$. Firm i's cost function takes the form $c_i(a_i, \theta_i) \geq 0$, where $\theta_i \in \Theta_i$, $c_i(a_i, \theta_i)$ is nondecreasing in a_i, and $c_i(0, \theta_i) = 0$. Period payoffs are given by $u_i(a, \theta_i) = \max\{1 - \sum_{i \in I} a_i, 0\} a_i - c_i(a_i, \theta_i)$. Then $\underline{v}_i = 0$ since i's rival can flood the market by setting $a_{-i} = \bar{a}_{-i}$ and drive the price to zero.*

Example 3.2 (Bertrand competition/First-price auction)**.** *As in the previous example, but now firms fix prices $a_i \in \{0, \frac{1}{N}, \frac{2}{N}, \ldots, 1\}$, where $N \geq 2$. There is only one buyer demanding one unit of the good each period, having a reservation value of one, and buying from the firm setting the lowest price (randomizing uniformly if multiple firms choose the lowest price). Firm i's cost is $\theta_i \in \Theta_i \subseteq [0, \infty[$ and therefore its period payoffs are $u_i(a, \theta_i) = (a_i - \theta_i)\mathbf{1}_{\{a_i = \min_{j \in I} a_j\}} \frac{1}{|\{k | a_k = \min_{j \in I} a_j\}|}$. Firm i's minmax equals $\underline{v}_i = 0$ and is attained when $a_{-i} = 0$.*

The above examples are special in two ways: First, all types can be punished as hard as possible with the same action. Second, player i's payoff is equal to the minmax value type by type, not just in expectation. The next example is one in which the second property is false; player i can only be punished on average.

Example 3.3 (Insurance without commitment)**.** *Each of two players faces an income shock $\theta_i \in \Theta_i \subset \mathbb{R}_+$. After receiving the shock (and communicating), player i chooses an amount $a_i \in \{0, \ldots, \theta_i\}$ to give to the other player. Player i's payoff is $\bar{u}_i(\theta_i - a_i + a_{-i})$, where $\bar{u}_i$ is nondecreasing and concave. Firm i's minmax value equals $\underline{v}_i = \mathbb{E}_{\pi_i}[\bar{u}_i(\theta_i)]$; it is attained by living in autarky (i.e., $a_1 = a_2 = 0$).*

As a final example we consider a game in which neither of the special properties shared by the Cournot and Bertrand games is true.

Example 3.4 (Battle of Sexes with Taste Shocks)**.** *Each player decides whether to go to a soccer game (S) or the opera (O). Payoffs are as shown in Figure 2. Assume $\theta_i \in \{1, 3\}$ so that each player may prefer (S, S) over (O, O) or vice versa. If $\theta_1 = \theta_2$, then one prefers (S, S) while the other prefers (O, O). If $\pi_i(1) = \pi_1(3) = 1/2$, then the minmax value is given by $\underline{v}_i = 2$ and is attained by all $a_{-i} \in \{S, O\}$. Any a_{-i}, however, is i's favorite action for one of the types.*

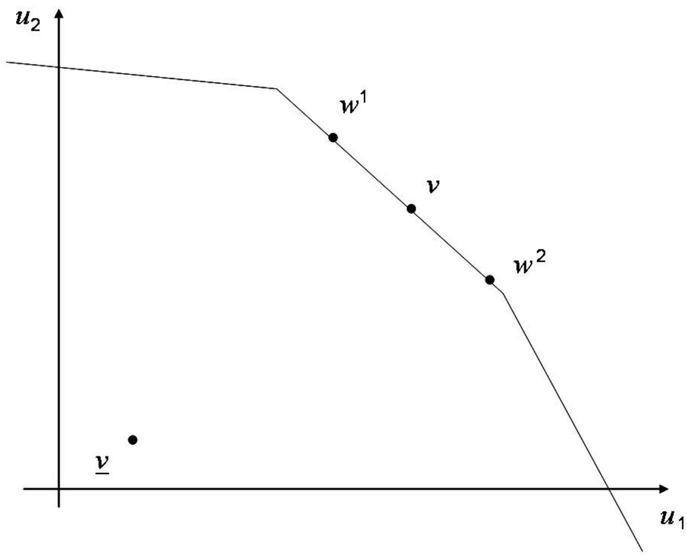

FIGURE 3. Pareto frontier.

4. The Main Result

Let $\mathbf{Fr}(V)$ denote the Pareto frontier of V. The following is the main result of the paper.

Theorem 4.1. *Assume $\boldsymbol{Fr}(V)$ is not a singleton. Let $v \in \boldsymbol{Fr}(V)$ be such that for all i*

$$\underline{v}_i < v_i.$$

Then, for all $\varepsilon > 0$, there exists $\bar{\delta} < 1$ such that for all $\delta > \bar{\delta}$, there exists a perfect Bayesian equilibrium of the dynamic game such that for all on-path histories, the expected continuation payoffs are within distance ε of v.

Theorem 4.1 shows that any payoff v that is Pareto-efficient in V and dominates the minmax value $\underline{v}$ can be virtually attained in an equilibrium of the dynamic game, provided that the players are patient enough. Moreover, this can be done so that the continuation payoffs are close to v at all on-path histories.

The result requires the Pareto frontier of V to have more than one point. As discussed in Section 6 and illustrated by Figure 3, this condition implies the existence of points w^1 and w^2, dominating the minmax profile $\underline{v}$, that are used to build player specific punishments. This assumption plays a role similar to, but is slightly stronger than, the full dimensionality condition usually imposed in repeated games with perfect monitoring (see, e.g., Fudenberg and Maskin, 1986).[20]

[20]If the Pareto-frontier of V is not a singleton, then V has full dimension except in the non-generic case where all $u(a,\theta)$, $(a,\theta) \in A \times \Theta$, lie on the same downward-sloping line. (Note that in that case everything is efficient in the stage game.)

The assumption appears unrestrictive: All that is required is that in a one-shot interaction, when types are drawn according to the invariant distribution, there is more than one Pareto-efficient decision rule. This is trivially satisfied in any game with linear transfers where each player prefers more of the numeraire. More generally, multiple Pareto-optima are the norm in situations with competing interests including all of the examples of Section 3.6. In fact, in the special case of a repeated game without types (i.e., when Θ is a singleton), if the Pareto-frontier of V is a singleton, then the one-shot game has a Nash-equilibrium with these payoffs. However, for completeness we sketch at the end of this section a weaker version of the result that dispenses with the assumption at the expense of weakening the definition of the minmax value.

The assumption about the possibility of communication in every period can not in general be dispensed with. To see this, it suffices to consider the Cournot example of Section 3.6 in the special case of iid types. Without communication, firms cannot coordinate to achieve payoffs close to a collusive scheme where the firm with the lowest cost always produces the monopoly output given its costs. In contrast, Theorem 4.1 shows that such payoffs can be approximated arbitrarily closely when communication is allowed.

The rest of the paper is essentially devoted to the proof of Theorem 4.1. As parts of it are relatively heavy on the details, we outline here the general proof strategy and discuss its implications for the equilibrium behavior. The construction of the equilibrium attaining payoffs near v in Theorem 4.1 has two main parts. We start in Section 5 by considering the problem of designing a (dynamic) mechanism that virtually attains the target payoff $v \in V$ given a sufficiently long finite horizon and enough patience. In each period the mechanism implements actions as functions of the messages sent by each player about his current type. Rather than using transfers, the mechanism uses history-dependent sets of feasible messages that allow the players to only send messages that are "credible" given the true underlying type process and history of messages. Given a message profile m, the mechanism then implements the action $f(m)$, where $f\colon \Theta \to \Delta(A)$ is the decision rule for which $v = \mathbb{E}_\pi[u(f(\theta), \theta)]$.

We show in Theorem 5.1 that—as a result of the restrictions imposed by the history-dependent message spaces—by reporting honestly player i can secure a discounted expected payoff bounded from below by v_i (up to an arbitrarily small approximation term) regardless of the other player's strategy. We then cover the infinite horizon with a "block mechanism" that has the players play the finite-horizon mechanism over and over again. The security-payoff result can then be applied to each repetition to get a lower bound arbitrarily close to v_i on player i's continuation values in the block mechanism. This combined with the efficiency of the target payoff v implies that, at any on-path history in any Nash equilibrium of the block mechanism, the continuation payoffs are arbitrarily close to v even if in the Nash equilibrium the players do not report honestly (Corollary 5.2). The existence of a PBE in the block mechanism then gives us a candidate for the equilibrium path of the PBE for the game.

Theorem 4.1 is then finally proved in Section 6 where we decentralize a PBE of the block mechanism by constructing stick-and-carrot punishment equilibria. The main concern there is ruling out observable deviations from the PBE of the mechanism (i.e., sending a message that would not have been feasible in the mechanism, or deviating from the actions specified by f). Such deviations are punished by reverting for finitely many periods to a stick phase where the deviator is minmaxed, followed by a carrot phase rewarding the non-deviator for following through with the punishment. In order to deal with issues such as manipulation of beliefs during the punishment by the player being punished, these punishment equilibria are also constructed by decentralizing a PBE of a mechanism.[21] The "punishment mechanism" is a modification of the block mechanism, where initially the deviating player i is minmaxed, and then a block mechanism approximating w^i, the reward profile for player $-i$, ensues.

While on the face of it the construction of the punishment equilibria appears to follow familiar lines (say, of Fudenberg and Maskin, 1986), incomplete information introduces its own complications. Beyond the technical intricacies of constructing equilibria for a dynamic Bayesian game, there is a qualitative difference involving the minmax value. As already discussed in the examples, our minmax value can in general be imposed only as an average payoff over a sufficiently long block of periods. In particular, player i can have a type θ_i such that, at the time of choosing actions (i.e., at $t.4$), conditional on his own current type being θ_i, player i's expected current period payoff when he is being minmaxed may well be even higher than his payoff in the efficient target profile v we are trying to sustain. This is in stark contrast with standard repeated games, in which the minmax value can be imposed as the expected payoff (at the time of choosing the actions) period-by-period.[22] This observation explains why our construction with two players features player-specific punishments even though in repeated games with perfect monitoring the two-player case can be handled using a stick phase where the players mutually minmax each other (see Fudenberg and Maskin, 1986).[23]

[21] In equilibrium, the player being punished best responds to the punishment. However, given serial correlation, this is in general not achieved by myopic maximization during the stick phase as it is in the interest of the player to manipulate the other player's beliefs about his type in order to have a higher payoff in the carrot phase.

[22] As discussed, the Bertrand and Cournot examples are special in that they have this feature also in the incomplete-information version.

[23] To see in more detail what goes wrong, suppose the stick phase consisted of mutual minmaxing, and consider the first period of the stick phase. As usual, the punishment for not playing along with the mutual minmaxing is that the stick phase restarts in the next period. Now, if player

As is typical in the literature on repeated games, Theorem 4.1 focuses on payoffs. It is also interesting to ask what kind of behavior sustains (approximately) efficient outcomes. However, our proof is non-constructive, and thus it does not yield a characterization of the equilibrium behavior. (In fact, this is the very reason why the proof strategy succeeds.) Nevertheless, certain things can still be inferred about behavior. First, on the equilibrium path, for any message profile sent at stage $t.2$ the actions at stage $t.4$ are those prescribed by the decision rule f. In this sense the equilibrium actions are stationary. Second, since expected payoffs are close to efficient payoffs from mutual truth-telling, it must be the case that "players report truthfully in a large fraction of periods with high probability." Hence in equilibrium misrepresentation of private information is limited. Third, given that the equilibrium path mimics the equilibrium of the mechanism, players' messages must respect the restrictions of the history-dependent message spaces. This puts relatively strong bounds on the equilibrium strategies. One simple qualitative implication is that the players will sometimes have to lie in order for cooperation to continue.

We conclude this section by considering briefly the case where the Pareto frontier of V is a singleton. In this case Theorem 4.1 is vacuous. However, we can recover a weaker result by weakening the minmax value to

$$\min_{a_{-i} \in A_{-i}} \max_{\theta_i \in \Theta_i} \max_{a_i \in A_i} u_i((a_i, a_{-i}), \theta_i).^{24}$$

In this case the stick-and-carrot punishment can be taken to consist of mutual minmaxing followed by the return to the cooperative phase, and hence there is no need to reward the punisher. Essentially the same proof then shows that the unique Pareto efficient point can be approximated with perfect Bayesian equilibrium payoffs as δ goes to 1 provided that it dominates the weaker minmax value.

5. An Approximately Efficient Dynamic Mechanism

We assume in this section that the players can write a contract, also known as a *mechanism*, which specifies for each period a (possibly randomized) action profile to be played as a function of the public messages sent by the players, and which can be enforced by a third party such as a court of law. Such a mechanism induces a dynamic game that differs from the original game defined above in that, in each period, the players only send public messages from some set of feasible messages (at $t.2$); the actions are automatically implemented by the mechanism

i happens to draw the favorable type θ_i discussed above, he is happy to delay the start of the carrot phase by one period. With enough persistence in the type process he may actually prefer to do so for a while.

[24] While this is in general a higher payoff than the minmax defined above, the two coincide for the Cournot and Bertrand games discussed in the examples. This not the case for the insurance game nor the Battle of Sexes.

as a function of these messages (at $t.4$). In what follows we introduce a particular dynamic mechanism that is approximately efficient—in a sense to be made precise later—even with a sufficiently long finite horizon provided that the players are sufficiently patient.

While the mechanism is constructed as an intermediate step towards Theorem 4.1, it is also of independent interest. In particular, as we do not assume transferable utility, the mechanism is applicable to settings such as decision making within organizations, or allocation of tasks within a firm, in which it is possible to write enforceable contracts, but transfers are typically not available or not used. Moreover, it gives a new efficiency result for settings such as dynamic insurance problems, in which utility is transferable, but wealth effects prevent the use of the dynamic VCG mechanisms proposed by Athey and Segal (2007) and Bergemann and Välimäki (2007).

5.1. **A Preliminary Result.** The following lemma, which relies on Massart's (1990) result about the rate of converge in the Glivenko-Cantelli theorem, motivates the message spaces we construct for the mechanism in the next section. Throughout $\|\cdot\|$ denotes the sup-norm.

Lemma 5.1. *Let Θ be a finite set, and let g be a probability measure on $(\Theta, 2^{\Theta})$. Given an infinite sequence of independent draws from g, let g^n denote the empirical measure obtained by observing the first n draws. (I.e., for all $n \in \mathbb{N}$ and all $\theta \in \Theta$, $g^n(\theta) = \frac{1}{n}\sum_{l=1}^{n} 1_{\{\theta^l=\theta\}}$.) Fix $\alpha > 0$ and construct a decreasing sequence $(b_n)_{n\in\mathbb{N}}$, $b_n \to 0$, by setting*

$$b_n = \sqrt{\frac{2}{n}\log\frac{\pi^2 n^2}{3\alpha}}.$$

Then

$$\mathbb{P}(\forall n \in \mathbb{N}\ \|g^n - g\| \leq b_n) \geq 1 - \alpha.$$

The proof of the lemma can be found in Appendix A. For a suggestive interpretation of the result, consider an honest player who observes a sequence of independent draws from Θ that are distributed according to a probability measure g. Suppose that the player is asked to report the realized value of each draw subject to the constraint that the empirical distribution of his reports after n observations, g^n, be within b_n of g (in the sup-norm) for all n. Then with probability at least $1-\alpha$ the player can truthfully report the entire sequence. It is the existence rather than the exact form of the sequence (b_n) that is important for our argument.

5.2. **The Mechanism.** The environment is a T-period truncation of the game for some $T \in \mathbb{N}$:

- two players: $i = 1, 2$,

- discrete time: $t = 1, 2, \ldots, T$,
- set of feasible action profiles in each period: A,
- player i's periodic payoff function: $u_i : A \times \Theta_i \to \mathbb{R}$,
- player i's type θ_i^t follows a Markov process (λ_i, P_i),
- discounted average payoffs.

By the maintained Assumptions 3.1– 3.3 we have private values and irreducible types, and transitions are independent across players.

We construct a collection of history-dependent message spaces for the T-period environment as follows. Let $\alpha > 0$. (The interesting case is where α is small.) Define the decreasing sequence $(b_n^\alpha)_{n\in\mathbb{N}}$, $b_n^\alpha \to 0$, as in Lemma 5.1 by setting

$$b_n^\alpha = \sqrt{\frac{2}{n} \log \frac{\pi^2 n^2}{3\alpha}}.$$

For all i, and all $\theta_i \in \Theta_i$, define

$$\Xi_i^\alpha(\theta_i) = \left\{ \xi_i \in \Theta_i^\infty : \forall n \in \mathbb{N} \max_{\hat{\theta}_i \in \Theta_i} \left| \frac{|\{l \leq n : \xi_i^l = \hat{\theta}_i)\}|}{n} - P_i(\theta_i, \hat{\theta}_i) \right| \leq b_n^\alpha \right\}.$$

In words, $\Xi_i^\alpha(\theta_i)$ is the set of sequences on Θ_i such that the distribution of types in the sequence converges to $P_i(\theta_i, \cdot)$ in the sup-norm at a rate specified by the sequence (b_n^α). By Lemma 5.1 the set $\Xi_i^\alpha(\theta_i)$ is non-empty for all i and all θ_i. Indeed, if we have an iid sequence of random variables distributed according to $P_i(\theta_i, \cdot)$, then the realized sequence lies in the set $\Xi_i^\alpha(\theta_i)$ with probability at least $1 - \alpha$.

We can now define the sets of feasible messages. For $t = 1$ player i's set of feasible messages is simply Θ_i. Consider then $t > 1$. Let $h_m^t = (m^1, \ldots, m^{t-1})$ be a history of message profiles before period t. Let $\phi_i(h_m^t)$ denote the (possibly null) subsequence of player i's messages in history h_m^t in periods $\tau \leq t - 1$ such that $m^{\tau-1} = m^{t-1}$. (I.e., $\phi_i(h_m^t)$ is a record of i's messages in periods where the previous period message profile was the same as in period $t - 1$.) Player i's set of feasible messages at message history h_m^t is

$$M_i^\alpha(h_m^t) = \left\{ \theta_i \in \Theta_i : \exists \xi_i \in \Theta_i^\infty \ (\phi_i(h_m^t), \theta_i, \xi_i) \in \Xi_i^\alpha(m_i^{t-1}) \right\}.$$ [25]

By construction the set $M_i^\alpha(h_m^t)$ is non-empty at any history h_m^t at which all past messages in $\phi_i(h_m^t)$ have been chosen from the appropriate feasible sets. Since other kinds of histories are by definition infeasible, it follows that player i always has some feasible message that he can send.

Let H_m^t denote the set of all period t message histories h_m^t (both feasible and infeasible) with $H_m^1 = \{h_m^1\}$ an arbitrary singleton. Letting $M_i^\alpha(h_m^1) = \Theta_i$ the message spaces are then determined by the function $M^{\alpha,T} : \cup_{t=1}^T H_m^t \to 2^\Theta$ defined by $M^{\alpha,T}(h_m^t) = M_1^\alpha(h_m^t) \times M_2^\alpha(h_m^t)$.

[25] Here, $(\phi_i(h_m^t), \theta_i, \xi_i)$ denotes the concatenation of $\phi_i(h_m^t)$, θ_i, and ξ_i.

Given the message spaces, the mechanism is defined as follows.

Definition 5.1. A *mechanism* is a pair $(f, M^{\alpha,T})$, where $f : \Theta \to \Delta(A)$ is a decision rule, and $M^{\alpha,T} : \cup_{\tau=1}^{T} H_m^{\tau} \to 2^{\Theta}$ is a collection of history-dependent message spaces. At each message history $h_m^t \in \cup_{\tau=1}^{T} H_m^{\tau}$ each player $i \in I$ sends a simultaneous public message $m_i^t \in M_i^{\alpha}(h_m^t)$ and the mechanism implements the (possibly randomized) action $f(m^t) \in \Delta(A)$.

For a given horizon T and a decision rule f, there is a family of mechanisms parameterized by the constant α. Any such mechanism induces a T-period dynamic game between the players. A pure (reporting) strategy for player i in the game induced by the mechanism $(f, M^{\alpha,T})$ is a sequence of mappings $\rho_i = (\rho_i^t)_{t=1}^{T}$ where each ρ_i^t is a mapping from the set of feasible histories of the mechanism's actions, messages, and player i's true types into Θ_i such that the type chosen by ρ_i^t is feasible given the message history. Let $R_i^{\alpha,T}$ denote the set of player i's pure strategies. The set of player i's mixed strategies is denoted $\Delta(R_i^{\alpha,T})$.

As the construction of the mechanism is somewhat involved, we offer here an informal discussion (see also the example in Section 2). We start by noting that by construction, at any history, whether a message is feasible or not depends only on the history of messages. Furthermore, the construction of the message spaces uses only the transition matrix P of the joint type process. That is, it is independent of the joint initial distribution λ, the payoff function u, and the decision rule f. Finally, the construction is independent of the time horizon in the sense that for any S and T, $S < T$, the S-period mechanism is simply an S-period truncation of the T-period mechanism.[26]

The general idea behind the message spaces is that they amount to keeping track of $|\Theta|$ empirical message distributions for each player i. These empirical distributions are indexed by the previous period message profile θ, which determines to which empirical distribution player i's current message is to be added. The message spaces then allow i to report a type θ_i' given previous period message profile $\theta = (\theta_i, \theta_{-i})$ only if this has the relevant empirical distribution converging fast enough to $P_i(\theta_i, \cdot)$—the conditional distribution of θ_i^t given $\theta^{t-1} = \theta$.

To see the restriction on player i's reporting in more detail, fix a type profile $\theta = (\theta_i, \theta_{-i}) \in \Theta$ and consider the (random) set of periods $\tau(\theta) = \{t = 2, \ldots, T : m^{t-1} = \theta\}$, i.e., the periods where the realized message profile in the previous period was θ. The message spaces force player i to report in such a way that the empirical distribution of i's reports along the periods in $\tau(\theta)$ converge to $P_i(\theta_i, \cdot)$ at a rate specified by the sequence (b_n^{α}). The set $M_i^{\alpha}(h_m^t)$ captures this by allowing

[26]This last property is for the sake of convenience. For any given $T < \infty$ it is possible to improve on the bounding sequence (b_n^{α}), which is chosen here to work for all T. While the bounds matter for the rate of convergence, qualitatively the results are unaffected.

player i to report type θ'_i given reporting history h^t_m ending in θ only if, after the addition of θ'_i, it is still possible to continue the sequence of i's reports over $\tau(\theta)$ in a way that preserves the rate of convergence. More precisely, it must be possible to continue the sequence to an element of $\Xi^\alpha_i(\theta_i)$.

The motivation for using the set $\Xi^\alpha_i(\theta_i)$ as the basis for the message spaces is that, since the true joint type process is Markovian, player i's types in periods in $\tau(\theta)$ are independent draws from $P_i(\theta_i,\cdot)$ provided that he reported truthfully in the previous periods. Thus in this case i's true types would indeed converge at the rate imposed by (b^α_n) along the periods in $\tau(\theta)$ with probability at least $1-\alpha$ (by Lemma 5.1). Hence if α is small, the constraint is unlikely to bind for a player who tries to report truthfully.

Finally, the mechanism can be compared to the linking mechanism of Jackson and Sonnenschein (2007). In the linking mechanism each player i is assigned a budget of messages—to be used over T independent and identical copies of a collective choice problem—that forces the distribution of player i's reports to match his true type distribution. Conceptually, the key difference is that our mechanism has "conditional budgets," i.e., the set of feasible messages depends on the history.[27] This is what allows us to deal with the serial correlation of types in a dynamic setting. In particular, the dependence of player i's set of feasible messages on player $-i$'s past message is what effectively prevents i from systematically matching $-i$'s messages with particular messages of his own. There are also important differences in how the "budgeting" of messages is implemented (by bounding the convergence of the message distributions rather using fixed budgets), but—while a crucial part of our proof—these are somewhat more technical in nature.

5.3. **Approximate Efficiency for Patient Players.** We now show that the mechanism defined in the previous section can be used to approximate Pareto-efficient payoff profiles arbitrarily closely if the horizon T is long enough and if the players are sufficiently patient. In fact, we show a stronger result: Under the said conditions, for any payoff profile v in V, there is a mechanism in which each player has a strategy that secures a lower bound on his expected payoff that is approximately equal to his payoff in the target payoff profile v regardless of the strategy of the other player.

We say that player i is *honest* in the mechanism $(f, M^{\alpha,T})$ if he reports his type truthfully whenever he can. None of the results depend on the specification of an honest player's behavior at histories where the set of feasible messages forces him

[27]Our mechanism can be thought of as an attempt to link together periods where the previous period types were identical and where—by the Strong Markov property—the current types are independently and identically distributed. However, this interpretation is only suggestive as the mechanism can base the set of feasible messages only on the players' past messages which need not be truthful.

to lie. However, to fix ideas, we assume that at such histories an honest player i always reports the smallest feasible message with respect to some fixed ordering of Θ_i. It is worth noting that when player i is honest, his strategy conditions only on his own current type.

Fix a mechanism $(f, M^{\alpha,T})$ and let ρ_i^* denote the honest strategy for player i. We say that *player i can secure the expected payoff $\bar{v}_i$ in the mechanism $(f, M^{\alpha,T})$ by being honest* if

$$\min_{\rho_{-i}\in\Delta(R_{-i}^{\alpha,T})} \mathbb{E}_{(\rho_i^*,\rho_{-i})}\Big[\frac{1-\delta}{1-\delta^T}\sum_{t=1}^{T}\delta^{t-1}u_i(f(m^t),\theta_i^t)\Big] \geq \bar{v}_i,$$

where the expectation is with respect to the distribution induced by the strategy profile (ρ_i^*, ρ_{-i}). That is, honesty secures expected payoff $\bar{v}_i$ to player i if, regardless of the reporting strategy of the other player, player i's expected payoff from honest reporting is at least $\bar{v}_i$.

The following "security payoff theorem" is the main result about our mechanisms.

Theorem 5.1. *Let $v \in V$ and let $\varepsilon > 0$. Then there exists a decision rule f, a constant $\alpha > 0$, and a time T^* such that for all $T \geq T^*$ there exists a discount factor $\delta^* < 1$ such that for all $\delta \geq \delta^*$ and all initial distributions λ, each player i can secure the expected payoff $v_i - \varepsilon$ in the mechanism $(f, M^{\alpha,T})$ by being honest.*

Note that the result is independent of "initial conditions" in that the same mechanism and critical discount factor work for all initial distributions.

The proof, which is presented in the next subsection, can be sketched as follows. By definition any $v \in V$ can be generated under truth-telling by a decision rule f when the expectation over types is with respect to the invariant distribution of the joint type process. This f is the decision rule used in the mechanism. By construction of the message spaces the honest player i can be taken to be truthful in all periods with an arbitrarily high probability regardless of $-i$'s strategy by choosing α small enough. So suppose this is the case and consider the problem of choosing player $-i$'s strategy to minimize the payoff to a truthful player i in the mechanism $(f, M^{\alpha,T})$. When player i is truthful, his payoff depends only on the joint distribution of his own true type θ_i and the other player's message m_{-i} since we have private values by Assumption 3.1. If there is no discounting, then even the timing is irrelevant and only the long-run distribution of (θ_i, m_{-i}) matters. Since the minimization problem is continuous in the discount factor, the Maximum theorem implies that this remains approximately true given sufficiently little discounting. Hence the proof boils down to showing that the joint long-run distribution of i's types and $-i$'s messages converges to the invariant distribution of the joint type process.

For T large the distribution of player i's types θ_i is close to the invariant distribution π_i by the law of large numbers since the type process is irreducible by Assumption 3.2. Similarly, for T large the distribution of player $-i$'s messages m_{-i} can be shown to be close to the invariant distribution π_{-i} by construction of the message spaces.[28] Furthermore, since the message spaces condition on both players' messages from the previous period and the players send their current messages simultaneously, we can use the independence of transitions (Assumption 3.3) to show that the joint distribution of (θ_i, m_{-i}) is close to the product distribution $\pi_i \times \pi_{-i}$. But this is precisely the invariant distribution for the true joint type process, which is what we wanted to show.

We now turn to the implications of Theorem 5.1. Let $V(\delta, T)$ denote the set of feasible expected discounted average payoffs in the T-period truncation of the game. Formally,

$$V(\delta,T) = \Big\{ v \in \mathbb{R}^I \mid \text{for some } f = (f^t)_{t=1}^T \\ v_i = \frac{1-\delta}{1-\delta^T}\mathbb{E}_f\big[\sum_{t=1}^{T}\delta^{t-1}u_i(a^t,\theta^t)\big] \text{ for all } i \in I \Big\},$$

where $f^t \colon \Theta^t \times A^{t-1} \to \Delta(A)$ and $\mathbb{E}_f$ is the expectation induced by the decision rules $f = (f^t)_{t=1}^T$ and the process (λ, P).

Lemma 5.2. *For all $\varepsilon > 0$, there exists a time T^* such that for all $T > T^*$, there exists a discount factor $\delta^* < 1$ such that for all $\delta > \delta^*$*

$$dist(V, V(\delta,T)) < \varepsilon,$$

where dist is the Hausdorff distance.

This lemma shows that for δ and T large enough $V(\delta, T)$ is well approximated by V (see Appendix A for the proof). This motivates the following approximate efficiency result.

Corollary 5.1. *Let v be a point on the Pareto-frontier of V. Let $\varepsilon > 0$. There exists a decision rule f, a constant $\alpha > 0$ and a time T^* such that for all $T \geq T^*$ there exists a discount factor $\delta^* < 1$ such that, for all $\delta \geq \delta^*$, the expected payoff profile is within distance ε of v in all Nash equilibria of the mechanism $(f, M^{\alpha,T})$.*

Proof sketch. Clearly all Nash equilibria (and hence all of its refinements such as PBE) of the mechanism must yield each player i an expected payoff at least as

[28]This is the only step that uses the fact that there are only two players. With more than two players the mechanism still forces the marginal distribution of each player's messages to converge to the invariant distribution, but the joint distribution of messages by players $-i$ need not converge to the product of these distributions. (An analogous problem arises already in a static model with iid types; see Jackson and Sonnenschein, 2007.) Handling the n-player case requires an augmented mechanism, the details of which are work in progress.

great as the lower bound secured by honesty. By Theorem 5.1 this lower bound can be taken to be arbitrarily close to v_i by choosing the parameters appropriately. Pareto efficiency of v in V then implies that the players' payoffs must in fact be approximately equal to v: $V(\delta, T)$ is close to V for δ close to 1 and T large by Lemma 5.2. Thus player i receiving substantially more than v_i when player $-i$ receives at least v_{-i} is infeasible for large δ and T by efficiency of v in V. Details can be found in Appendix A. □

Since the game induced by the mechanism is finite, all standard dynamic refinements of Nash equilibria such as a PBE or a sequential equilibrium exist. Hence Corollary 5.1 implies that the mechanism can be used to virtually implement Pareto-efficient payoffs in, say, a sequential equilibrium provided that the horizon is long enough and the players are sufficiently patient. As the proof is non-constructive, we do not have a characterization of the behavior in such an equilibrium. It appears prohibitively difficult to solve for it analytically as the players face complicated non-stationary dynamic optimization problems. The security payoff theorem does imply that honesty is an ε-Nash equilibrium, but in general honesty is not a best-response. However, since the payoffs are close to the efficient payoffs from mutual truth-telling, any equilibrium must have the players reporting truthfully in a "large fraction of periods with high probability." Formal results along these lines are left for future work.

In order to cover the original dynamic game defined in Section 3 we now extend the efficiency result to an infinite horizon. Rather than using the infinite horizon version of the above mechanism, we construct a "block mechanism" in which the players repeatedly play a fixed finite horizon mechanism $(f, M^{\alpha,T})$.[29] We then apply the security payoff of Theorem 5.1 to each repetition. This serves to guarantee that the players not only have approximately efficient expected payoffs at the beginning of the mechanism, but also their expected continuation payoffs are approximately efficient. This is of interest in settings with "participation constraints." In particular, it is needed in the proof of Theorem 4.1 where we essentially decentralize an equilibrium of the block mechanism. There participation constraints arise from the players' ability to "opt out" by choosing not to play the actions that would have been implemented by the mechanism.

Consider the infinite horizon environment defined by the dynamic game. Note that for any $T < \infty$ the T-period blocks $(k-1)T+1, \ldots, kT$, $k \in \mathbb{N}$, define a sequence of T-period environments, which differ from each other only because of the initial distribution of types. Since the construction of the mechanisms $(f, M^{\alpha,T})$ is independent of the initial distribution, any such mechanism can be

[29] As is evident from the construction, a mechanism $(f, M^{\alpha,T})$ can be extended to an infinite horizon by simply putting $T = \infty$. The results developed above for the finite horizon case have natural analogs in the infinite horizon case. We do not pursue the details.

applied to any of the T-period blocks. With this in mind, for any message history $h_m^t = (m^1, \ldots, m^{t-1}) \in \cup_{\tau=1}^{\infty} H_m^{\tau}$, let

$$\bar{h}_m^t = (m^{t-[(t-1) \bmod T]}, \ldots, m^{t-1}),$$

where $(t-1) \bmod T$ denotes the residue from the division of $t-1$ by T, and where we adopt the convention that $(m^s, ..., m^t) = h_m^1$ if $s > t$. (Recall that h_m^1 is an arbitrary constant.) Then $\bar{h}_m^t$ simply collects from h_m^t the messages that have been sent in the current block.

Definition 5.2. A *block mechanism* is an infinitely repeated mechanism $(f, M^{\alpha,T})^{\infty}$ in which the mechanism $(f, M^{\alpha,T})$ is applied to each T-period block $(k-1)T + 1, \ldots, kT$, $k \in \mathbb{N}$. At each message history $h_m^t \in \cup_{\tau=1}^{\infty} H_m^{\tau}$ each player $i \in I$ sends a simultaneous public message $m_i^t \in M_i^{\alpha}(\bar{h}_m^t)$ and the mechanism implements the (possibly randomized) action $f(m^t) \in \Delta(A)$.

The next corollary shows that block mechanisms can be used to approximate Pareto-efficient payoffs in equilibria that have approximately stationary continuation payoffs.

Corollary 5.2. *Let v be a point on the Pareto-frontier of V. Let $\varepsilon > 0$. There exists a block mechanism $(f, M^{\alpha,T})^{\infty}$ and a discount factor $\delta^* < 1$ such that, for all $\delta \geq \delta^*$ and all initial distributions λ, the expected continuation payoff profile is within distance ε of v at every on-path history in all Nash equilibria of the block mechanism $(f, M^{\alpha,T})^{\infty}$. Moreover, in all sequential equilibria the expected continuation payoff profile is within distance ε of v at all feasible histories.*

Proof Sketch. Fix v on the Pareto-frontier of V and let $\varepsilon > 0$. By Theorem 5.1 there exists a mechanism $(f, M^{\alpha,T})$ and a critical discount factor $\delta^* < 1$ such that for all $\delta \geq \delta^*$, at the beginning of each block, each player can secure the expected payoff $v_i - \frac{\varepsilon}{3}$ from the block by being honest given any distribution of types at the start of the block.[30] Now, to make the security result hold at all periods rather than just at the beginning of blocks, we choose a discount factor $\delta^{**} \geq \delta^*$ high enough so that, at any period t, the payoff from the remaining periods in the current block has a negligible impact on the total expected continuation payoff from period t onwards. (This is possible since payoffs are bounded.) Then for all $\delta \geq \delta^{**}$ each player i can secure the expected payoff $v_i - \frac{\varepsilon}{2}$ at every period t.

Since a player can always revert to playing honestly from now on, any on-path history in any Nash equilibrium must yield each player i an expected continuation

[30]While in the block mechanism the players in general have public and private histories that could act as a correlation device, this does not affect the security payoff from a given block. Indeed, suppose that at the beginning of a block player i reverts to playing honestly in the block, and consider choosing $-i$'s strategy to minimize i's payoff over the block. Since the honest strategy does not condition on the public nor the private history, having the payoff-irrelevant correlation device is of no value for this minimization problem.

payoff of at least $v_i - \frac{\varepsilon}{2}$. But then, analogously to the proof of Corollary 5.1, the efficiency of v in V implies that no player can receive more than $v_i + \varepsilon$ for δ large enough. The result for sequential equilibria follows by noting that there the strategy profile must be sequentially rational at all feasible histories, and hence at each history the players' expected continuation payoffs must be at least as high as the lower bound achieved by reverting to honest reporting from now on. □

A result by Fudenberg and Levine (1983) implies that a sequential equilibrium exists in the infinite horizon game induced by the block mechanism. Hence for any v on the Pareto frontier of V there exists a block mechanism that has a sequential equilibrium in which the expected continuation payoffs at all (feasible) histories are approximately efficient.[31]

5.4. **Proof of Theorem 5.1.** Let $v \in V$ and let $\varepsilon > 0$. By definition of V there exists a decision rule f such that $v = \mathbb{E}_\pi[u(f(\theta), \theta)]$. Consider the problem of minimizing the payoff of an honest player i in the mechanism $(f, M^{\alpha,T})$ for some $\alpha > 0$ and $T < \infty$. Without loss, assume that player 1 plays the honest strategy ρ_1^*, while player 2 chooses a strategy $\rho_2 \in \Delta(R_2^{\alpha,T})$ so as to minimize 1's payoff. We want to show that we can choose $\alpha > 0$ such that if T is large enough, then there exist $\delta^* < 1$ such that, for all $\delta \geq \delta^*$ and for all initial distributions λ,

$$\min_{\rho_2 \in \Delta(R_2^{\alpha,T})} \mathbb{E}_{(\rho_1^*, \rho_2)} \Big[\frac{1-\delta}{1-\delta^T} \sum_{t=1}^{T} \delta^{t-1} u_1(f(m^t), \theta_1^t) \Big] \geq v_1 - \varepsilon,$$

where the expectation is with respect to the distribution induced by the strategies (ρ_1^*, ρ_2).

We start by simplifying the minimization problem on the left-hand side of the above inequality. The minimum is attained by a pure strategy, so it suffices to consider pure strategies of player 2. Furthermore, since the mechanism's randomizations are independent across periods and the honest strategy ρ_1^* does not condition on the mechanism's actions, it is without loss to assume that ρ_2 does not condition on the mechanism's actions either. Finally, by Blackwell's theorem it suffices to consider the case where λ_1 is degenerate and puts probability one on some θ_1.[32] But then it is without loss to assume that ρ_2 does not condition on player 2's private history (i.e., on player 2's realized types), since transitions are independent between players by Assumption 3.3 and θ_1^1 is known. Thus we are

[31] As in the case of a T-period mechanism, honesty is an ε-equilibrium of the block mechanism.

[32] Recall that $M_1^1 = \Theta_1$ so that $m_1^1 = \theta_1^1$. Thus in general player 2 learns θ_1^1 after the first period, whereas with a degenerate λ_1 player 2 knows θ_1^1 before the first period. Since player 1 is honest and hence non-strategic, this extra information can only help player 2.

left with a problem of the form

$$w(\theta_1, \delta, T, \alpha) = \min_{\rho_2 \in \bar{R}_2^{\alpha,T}} \mathbb{E}_{(\rho_1^*, \rho_2)} \Big[\frac{1-\delta}{1-\delta^T} \sum_{t=1}^{T} \delta^{t-1} u_1(f(m^t), \theta_1^t) \Big],$$

where θ_1 refers to player 1's first period type, and $\bar{R}_2^{\alpha,T} \subset R_2^{\alpha,T}$ denotes the set of player 2's pure strategies that do not condition on player 2's private history nor the mechanism's actions. In other words, $\bar{R}_2^{\alpha,T}$ is the set of pure strategies that condition only on player 1's past messages. Note that the expectation is only over player 1's types, the messages being deterministic functions thereof.

We argue then that it suffices to consider the case of no discounting. Extend the problem to $\delta = 1$ by defining

$$w(\theta_1, 1, T, \alpha) = \min_{\rho_2 \in \bar{R}_2^{\alpha,T}} \mathbb{E}_{(\rho_1^*, \rho_2)} \Big[\frac{1}{T} \sum_{t=1}^{T} u_1(f(m^t), \theta_1^t) \Big].$$

It is then straightforward to check that for fixed θ_1, T, and α the objective function is continuous in (δ, ρ_2) on $[0,1] \times \bar{R}_2^{\alpha,T}$. Thus by Berge's Maximum theorem the value of the problem, $w(\theta_1, \delta, T, \alpha)$, is continuous in δ on $[0,1]$. Hence we can approximate $w(\theta_1, \delta, T, \alpha)$ for δ large by considering $w(\theta_1, 1, T, \alpha)$. Thus it suffices to show

$$(*) \qquad \exists \alpha > 0 \ \exists T^* < \infty : \forall T \geq T^* \ \forall \theta_1 \in \Theta_1 \quad w(\theta_1, 1, T, \alpha) \geq v_1 - \frac{\varepsilon}{2}.$$

As a first step towards condition $(*)$ we make precise the idea that an honest player is unlikely to be constrained by the message spaces.[33] Given realized sequences of player i's types $(\theta^t)_{t=1}^T$ and messages $(m^t)_{t=1}^T$, we say that player i is *truthful* if $m_i^t = \theta_1^t$ for all t.

Lemma 5.3. *Let λ_1 assign probability one to some $\theta_1 \in \Theta_1$. For all T and all $\alpha > 0$, if player 1 plays the honest strategy $\rho_1^* \in R_1^{\alpha,T}$ and player 2 plays a pure strategy $\rho_2 \in \bar{R}_2^{\alpha,T}$, then player 1 is truthful with probability at least $1 - |\Theta|\alpha$.*

This lemma follows essentially by construction of the message spaces. The proof strategy is to first assume that the honest player 1 is not subject to any restrictions in his reporting (i.e., set $M_1^\alpha(h_m^t) = \Theta_1$ for all h_m^t) and hence is always truthful. Then we argue that the truthful messages would remain feasible with probability at least $1 - |\Theta|\alpha$ even if player 1 was subject to the history-dependent message spaces. As the proof is not particularly illuminating, we leave the details to Appendix A.

The following proposition is the key to the proof.

Proposition 5.1. *Let λ_1 assign probability one to some $\theta_1 \in \Theta_1$. For all $q > 0$, there exists $T^* < \infty$ such that, for all $T \geq T^*$, if player 1 plays the honest strategy*

[33]Given the above derivation, we assume in the sequel that λ_1 is degenerate, and that player 2 plays a pure strategy $\rho_2 \in \bar{R}_2^{\alpha,T}$. While these restrictions simplify the proofs somewhat, Lemma 5.3 and Proposition 5.1 can be extended to general λ_1 and $\rho_2 \in \Delta(R_2^{\alpha,T})$.

$\rho_1^* \in R_1^{\frac{q}{2|\Theta|},T}$ and player 2 plays a pure strategy $\rho_2 \in \bar{R}_2^{\frac{q}{2|\Theta|},T}$, then the empirical distribution of messages, π^T, satisfies

$$\mathbb{P}(\|\pi^T - \pi\| < q) \geq 1 - q,$$

where π is the invariant distribution of the joint type process.

Proof. Let λ_1 be degenerate. Fix $q > 0$ and put $\alpha = \frac{q}{2|\Theta|}$. We argue first that it suffices to consider an honest player 1 who is not subject to the message spaces (i.e., set $M_1^\alpha(h_m^t) = \Theta_1$ for all h_m^t). To this end, fix T and $\rho_2 \in \bar{R}_2^{\frac{q}{2|\Theta|},T}$, and suppose that player 1 is not subject to the message spaces. Note that, by construction, player 2's message spaces are non-empty even at histories that include infeasible histories by player 1. Player 2's strategy ρ_2 can be extended arbitrarily to such histories as they play no role in what follows. Suppose we have found a set $C \subset \Theta_1^T$ of probability $1 - \frac{q}{2}$ such that $\|\pi^T - \pi\| < q$ if player 1's realized type sequence $(\theta_1^t)_{t=1}^T$ is in C. By Lemma 5.3 there exists a set $D \in \Theta_1^T$ of probability $1 - \frac{q}{2}$ such that for all $(\theta_1^t)_{t=1}^T \in D$ the honest player 1 is truthful even if he is subject to the message spaces. But then for any $(\theta_1^t)_{t=1}^T \in C \cap D$ we have $\|\pi^T - \pi\| < q$ even if player 1 is subject to the message spaces. Moreover, $\mathbb{P}(C \cap D) \geq 1 - q$. So for the rest of the proof we put $M_1^\alpha(h_m^t) = \Theta_1$ for all h_m^t.

It is convenient to generate player 1's types by means of an auxiliary probability space $([0,1], \mathcal{B}, \hat{\mathbb{P}})$. (The construction that follows is adapted from Billingsley, 1961.) On this space, define a countably infinite collection of *independent* random variables

$$\tilde{\psi}^n_{\theta,\theta_2'} \colon [0,1] \to \Theta_1, \quad \theta \in \Theta,\ \theta_2' \in \Theta_2,\ n \in \mathbb{N},$$

where

$$\hat{\mathbb{P}}(\tilde{\psi}^n_{(\theta_1,\theta_2),\theta_2'} = \theta_1') = P_1(\theta_1, \theta_1').$$

That is, for fixed $\theta = (\theta_1, \theta_2)$ the variables $\tilde{\psi}^n_{\theta,\theta_2'}$, $\theta_2' \in \Theta_2$, $n = 1, 2, \ldots$ are independent draws from $P_1(\theta_1, \cdot)$. Imagine the variables $\tilde{\psi}^n_{\theta,\theta_2'}$ set out in the following array:

$$\begin{array}{ccccc}
\tilde{\psi}^1_{1,1} & \tilde{\psi}^2_{1,1} & \cdots & \tilde{\psi}^n_{1,1} & \cdots \\
\tilde{\psi}^1_{1,2} & \tilde{\psi}^2_{1,2} & \cdots & \tilde{\psi}^n_{1,2} & \cdots \\
\vdots & & & & \\
\tilde{\psi}^1_{1,|\Theta_2|} & \tilde{\psi}^2_{1,|\Theta_2|} & \cdots & \tilde{\psi}^n_{1,|\Theta_2|} & \cdots \\
\tilde{\psi}^1_{2,1} & \tilde{\psi}^2_{2,1} & \cdots & \tilde{\psi}^n_{2,1} & \cdots \\
\vdots & & & & \\
\vdots & & & & \\
\tilde{\psi}^1_{|\Theta|,|\Theta_2|} & \tilde{\psi}^2_{|\Theta|,|\Theta_2|} & \cdots & \tilde{\psi}^n_{|\Theta|,|\Theta_2|} & \cdots
\end{array}$$

Then along each of the $K = |\Theta||\Theta_2|$ rows the variables are independent draws from a fixed distribution. We can apply Lemma 5.1 along any fixed row of the array to conclude that with $\hat{\mathbb{P}}$-probability at least $1-\frac{q}{2K}$, for all n the empirical measure for the first n observations along the row is within

$$c_n = \sqrt{\frac{2}{n}\log\frac{\pi^2 n^2 2K}{3q}}$$

of the true distribution in the sup-norm. (Note that $c_n \to 0$.) Thus, if we let $E \in \mathcal{B}$ denote the event where this is true along all K rows, then $\hat{\mathbb{P}}(E) \geq 1-\frac{q}{2}$.

Given any T and any strategy $\rho_2 \in \bar{R}_2^{\frac{q}{2|\Theta|},T}$ for player 2, the array can be used to generate a sequence $(\theta_1^t, m_2^t)_{t\geq 1}$ of player 1's types (which equal his messages) and player 2's messages as follows. Player 2's first period message is some m_2^1. Since λ_1 puts probability one on some θ_1, player 1's period 1 type is simply $\theta_1^1 = \theta_1$. Player 2's message in period 2 is given by $m_2^2 = \rho_2^2(\theta_1^1)$. Player 1's period 2 type θ_1^2 is then drawn by sampling the first variable in the row indexed by $(\theta_1^1, m_2^1), m_2^2$ (i.e., by setting $\theta_1^2 = \psi^1_{(\theta_1^1,m_2^1),m_2^2}$). Player 2's period 3 message is then given by $m_2^3 = \rho_2^3(\theta_1^1, \theta_1^2)$. Player 1's period 3 type is then drawn by sampling the first element of the row indexed by $(\theta_1^2, m_2^2), m_2^3$, unless $(\theta_1^1, m_2^1), m_2^2 = (\theta_1^2, m_2^2), m_2^3$, in which case the second variable in the row indexed by $(\theta_1^1, m_2^1), m_2^2$ is sampled instead. And so forth.

To see that this construction indeed gives rise to the right process for the T-period environment, fix $(\theta_1^t, m_2^t)_{t=1}^T$. Obviously we must have

$$m_2^t = \rho_2^t(\theta_1^1, \ldots, \theta_1^{t-1})$$

for all $t = 1, \ldots, T$ since otherwise the probability of this sequence is trivially zero. So suppose this is the case. Then the probability of the sequence according to the original description of the process is simply

$$\lambda_1(\theta_1^1) P_1(\theta_1^1, \theta_1^2) \cdots P_1(\theta_1^{T-1}, \theta_1^T).$$

On the other hand, the above array construction assigns this sequence the probability

$$\lambda_1(\theta_1^1)\hat{\mathbb{P}}(\tilde{\psi}^1_{(\theta_1^1,m_2^1),m_2^2} = \theta_1^2) \cdots \hat{\mathbb{P}}(\tilde{\psi}^k_{(\theta_1^{T-1},m_2^{T-1}),m_2^T} = \theta_1^T),$$

where $k-1$ is the number of times the combination $(\theta_1^{T-1}, m_2^{T-1}), m_2^T$ appears in the sequence, and where we have used independence of the $\tilde{\psi}^n_{\theta,\theta_2}$ to write the joint probability as a product. By construction of the array,

$$\hat{\mathbb{P}}(\tilde{\psi}^1_{(\theta_1^1,m_2^1),m_2^2} = \theta_1^2) = P_1(\theta_1^1, \theta_1^2)$$

and

$$\hat{\mathbb{P}}(\tilde{\psi}^k_{(\theta_1^{T-1},m_2^{T-1}),m_2^T} = \theta_1^T) = P_1(\theta_1^{T-1}, \theta_1^T),$$

(and similarly for the elements we haven't explicitly written out) so both methods assign the sequence the same probability.

It suffices to show that if T is large enough, then conditional on E, given any $\rho_2 \in \bar{R}_2^{\frac{q}{2|\Theta|},T}$, we have $\|\pi^T - \pi\| < q$. Let P^T denote the empirical transition matrix for the sequence $(\theta_1^t, m_2^t)_{t=1}^T$.[34] By Lemma A.1 in Appendix A, there exists $\bar{T}$ and $\nu > 0$ such that if $T \geq \bar{T}$ and $\|P^T - P\| < \nu$, then $\|\pi^T - \pi\| < q$. (Recall that P is the transition matrix for the joint type process.) Thus, in order to show that the distribution of messages π^T converges to π on E, it is enough to show that the transitions P^T converge to P on E.

Let $p = \min_{i\in I} \min\{P_i(\theta_i, \theta_i') : P_i(\theta_i, \theta_i') > 0\}$ denote the smallest positive transition probability. For any $x \in \mathbb{R}_+$, let $\lfloor x \rfloor = \max\{n \in \mathbb{N}_0 : n \leq x\}$.

Claim 5.1. *Suppose that conditional on E, the message profile $\bar{\theta} \in \Theta$ is sent at least $n+1$ times during T periods. Then*

(1) $\|P^T(\bar{\theta}, \cdot) - P(\bar{\theta}, \cdot)\| \leq c_{\lfloor (p - b_n^\alpha)n \rfloor} + b_n^\alpha$, *and*
(2) *the number of times each θ in the support of $P(\bar{\theta}, \cdot)$ is sent is at least*

$$\lfloor (p - c_{\lfloor (p-b_n^\alpha)n\rfloor}) \lfloor (p - b_n^\alpha)n \rfloor \rfloor.$$

Proof. Define the sets

$$\tau(\bar{\theta}) = \{t \in \{2, \ldots, T\} : m^{t-1} = \bar{\theta}\}$$

and

$$\tau(\bar{\theta}, \theta_2) = \{t \in \{2, \ldots, T\} : (m^{t-1}, m_2^t) = (\bar{\theta}, \theta_2)\},\ \theta_2 \in \Theta_2.$$

Let $P_2^T(\bar{\theta}, \cdot)$ denote the empirical distribution of player 2's messages over $\tau(\bar{\theta})$. By assumption $|\tau(\bar{\theta})| \geq n$ so that by construction of the message spaces we have

$$\|P_2^T(\bar{\theta}, \cdot) - P_2(\bar{\theta}, \cdot)\| \leq b^\alpha_{|\tau(\bar{\theta})|} \leq b_n^\alpha,$$

where we have included $\bar{\theta}_1$ as an argument in player 2's type transition P_2 for convenience. Thus for all θ_2 in the support of $P_2(\bar{\theta}, \cdot)$ we have

$$|\tau(\bar{\theta}, \theta_2)| \geq \lfloor (P_2(\bar{\theta}, \theta_2) - b_n^\alpha)n \rfloor \geq \lfloor (p - b_n^\alpha)n \rfloor.$$

Let $P_1^T((\bar{\theta}, \theta_2), \cdot)$ denote the empirical distribution of player 1's types over $\tau(\bar{\theta}, \theta_2)$. By the above construction $P_1^T((\bar{\theta}, \theta_2), \cdot)$ can be taken to be the empirical distribution for the first $|\tau(\bar{\theta}, \theta_2)|$ elements of the $\bar{\theta}, \theta_2$-row in our array. Since we are conditioning on the event E, we thus have

$$\|P_1^T((\bar{\theta}, \theta_2), \cdot) - P_1(\bar{\theta}, \cdot)\| \leq c_{|\tau(\bar{\theta}, \theta_2)|} \leq c_{\lfloor (p-b_n^\alpha)n \rfloor}.$$

[34] That is, put

$$P^T((\theta_1, m_2), (\theta_1', m_2')) = \frac{|\{s < T : ((\theta_1^s, m_2^s), (\theta_1^{s+1}, m_2^{s+1})) = ((\theta_1, m_2), (\theta_1', m_2'))\}|}{|\{s < T : (\theta_1^s, m_2^s) = (\theta_1, m_2)\}|}.$$

But then the joint message distribution $P^T(\bar\theta,\cdot)$ over $\tau(\bar\theta)$ satisfies

$$\begin{aligned}|P^T(\bar\theta,\theta)-P(\bar\theta,\theta)| &= |P_2^T(\bar\theta,\theta_2)P_1^T((\bar\theta,\theta_2),\theta_1)-P_2(\bar\theta,\theta_2)P_1(\bar\theta,\theta_1)|\\ &\le P_2^T(\bar\theta,\theta_2)|P_1^T((\bar\theta,\theta_2),\theta_1)-P_1(\bar\theta,\theta_1)|\\ &\quad + P_1(\bar\theta,\theta_1)|P_2^T(\bar\theta,\theta_2)-P_2(\bar\theta,\theta_2)|\\ &\le c_{\lfloor (p-b_n^\alpha)n\rfloor}+b_n^\alpha,\end{aligned}$$

where the equality is simply by definition, the first inequality is by triangle inequality, and the second inequality follows by the above results. This establishes (1). Furthermore, the number of times each θ in the support of $P(\bar\theta,\cdot)$ is sent over $\tau(\bar\theta)$ is bounded from below by

$$\lfloor (p-b_n^\alpha)n\rfloor (P_1(\bar\theta,\theta_1)-c_{\lfloor (p-b_n^\alpha)n\rfloor})\ge \lfloor (p-b_n^\alpha)n\rfloor (p-c_{\lfloor (p-b_n^\alpha)n\rfloor}),$$

where $\lfloor (p-b_n^\alpha)n\rfloor$ is the lower bound on $|\tau(\bar\theta,\theta_2)|$ from above, and $P_1(\bar\theta,\theta_1)-c_{\lfloor (p-b_n^\alpha)n\rfloor}$ is a lower bound on $P_1^T((\bar\theta,\theta_2),\theta_1)$. This establishes (2). □

Since P is irreducible, there exists $L<\infty$ such that it is possible to go from any $\bar\theta\in\Theta$ to any other $\theta\in\Theta$ in at most L steps. Thus iterating the claim at most L times we obtain bounds for $\|P^T(\theta,\cdot)-P(\theta,\cdot)\|$ for all $\theta\in\Theta$. (Indeed, in the special case where P has full support only one iteration is needed.) By inspection the bounds in (1) and (2) are independent of $\bar\theta$, so this procedure yields a bound on $\|P^T-P\|$ which is independent of $\bar\theta$. It is straightforward to check that this bound converges to zero as $n\to\infty$ since only finitely many iterations are needed.

We will now use Claim 5.1 and Lemma A.1 to finish the proof of the proposition. For any T and any ρ_2 there exists some $\bar\theta\in\Theta$ that is sent at least $\frac{T}{|\Theta|}$ times during the T periods. We may thus put $n+1=\frac{T}{|\Theta|}$ in Claim 5.1 and iterate it at most L times to get a bound on $\|P^T-P\|$ conditional on E. The bound so obtained is independent of $\bar\theta$ and hence independent of ρ_2. Moreover, since n grows linearly in T, the bound is arbitrarily small if T is large enough. So for T large enough we can apply Lemma A.1 to conclude that $\|\pi^T-\pi\|<q$ conditional on E. □

We can now establish condition $(*)$. Since A and Θ are finite, there exists $B<\infty$ such that $|u_1(a,\theta_1)|\le B$. Put

$$q=\frac{\varepsilon}{4B|\Theta|}\quad\text{and}\quad \alpha=\frac{q}{2|\Theta|}.$$

Since Θ_1 is finite, Proposition 5.1 implies that if player 1 is honest, then there exists T^* such that for all initial types θ_1, all $\rho_2\in\bar R_2^{\alpha,T}$, and all $T\ge T^*$, we have $\mathbb{P}(\|\pi^T-\pi\|<q)>1-q$. Now fix θ_1, $T\ge T^*$, and some $\rho_2\in\bar R_2^{\alpha,T}$ that achieves $w(\theta_1,1,T,\alpha)$. Since $0\le\|\pi^T-\pi\|\le 1$, we then have

$$\mathbb{E}_{(\rho_1^*,\rho_2)}[\|\pi^T-\pi\|]<(1-q)q+q\le 2q=\frac{\varepsilon}{2B|\Theta|}.$$

This implies that

$$\begin{aligned}\Big|w(\theta_1,1,T,\alpha)-v_1\Big| &= \Big|\mathbb{E}_{(\rho_1^*,\rho_2)}\Big[\frac{1}{T}\sum_{t=1}^{T}u_1(f(\theta_1^t,m_2^t),\theta_1^t)\Big]-\sum_{\theta\in\Theta}\pi_\theta u_1(f(\theta),\theta_1)\Big]\Big| \\ &= \Big|\mathbb{E}_{(\rho_1^*,\rho_2)}\Big[\sum_{\theta\in\Theta}(\pi_\theta^T-\pi_\theta)u_1(f(\theta),\theta_1)\Big]\Big| \\ &\leq B|\Theta|\mathbb{E}_{(\rho_1^*,\rho_2)}[\|\pi^T-\pi\|]\leq\frac{\varepsilon}{2},\end{aligned}$$

where the equalities follow by simply writing out the definitions and rearranging terms, the first inequality follows by passing the absolute value through the expectation and the sum, and the last inequality is by the bound on $\mathbb{E}_{(\rho_1^*,\rho_2)}[\|\pi^T-\pi\|]$. This implies condition $(*)$.

To complete the proof of Theorem 5.1, note that the choice of α above is independent of the identity of the players. Hence, reversing the roles of the players in the above argument and taking the maximum over the cutoff times and discount factors across players implies the result.

6. Proving Theorem 4.1

Under the conditions of Theorem 4.1, we can find $\bar{v},w^1,w^2\in\mathbf{Fr}(V)$ such that for all i

$$\underline{v}_i<\min\{\bar{v}_i,w_i^i\}$$

and

$$w_i^i<\min\{\bar{v}_i,w_i^{-i}\},$$

with $\bar{v}$ arbitrarily close to v. In the sequel, without loss we assume that $\bar{v}=v$ keeping in mind that otherwise the subsequent arguments can be applied to $\bar{v}$ close enough to the target payoff profile v.

Let f and f^i be the decision rules giving expected payoffs v and w^i respectively. Before turning into the analysis of the dynamic game strategies resulting in payoffs approximately equal to v, it is useful to introduce and study an auxiliary dynamic mechanism we use as off path punishment.

6.1. **Preliminaries: The Punishment Mechanism.**

For $i\in I$, take a minmaxing action

$$a_{-i}^i\in\arg\min_{a_{-i}\in A_{-i}}\Big\{\max_{a_i\in A_i}\mathbb{E}_{\pi_i}[u_i(a,\theta_i)]\Big\}.$$

For each L^i, T^i, and α, we consider an infinite-horizon dynamic mechanism, characterized by the tuple $(i,L^i,(f^i,M^{\alpha^i,T^i})^\infty)$, running through $t=1,2,\ldots$. At each date $t=1,\ldots,L^i$, player i picks an action $a_i^t\in A_i$; player $-i$ has no choice but to pick $a_{-i}^t=a_{-i}^i$. At $t=L^i+1$, the block mechanism $(f^i,M^{\alpha^i,T^i})^\infty$ starts. Note that the construction of the block mechanism starting at L^i+1 does not depend on how play transpires during the first L^i rounds of the mechanism. We think of this mechanism as embedded in our main dynamic game model. In particular, the evolution of private types is characterized by the transition matrix P and players' payoffs are the discounted sum of period payoffs. We allow the initial beliefs to be different from λ and equal to some μ.

The mechanism described above is what we call the "punishment mechanism against i." It starts with a stick phase in which player $-i$ is restricted to minmax player i during L^i periods while player i can choose arbitrary actions $a_i^t \in A_i$. The mechanism then moves on to a carrot phase in which players simultaneously report their types, subject to the restrictions imposed by the block mechanism $(f^i, M^{\alpha^i,T^i})^\infty$.

The punishment mechanism inherits many of the properties of the block mechanism already discussed in Section 5.3, important among them is the following corollary.

Corollary 6.1. *Fix $i \in I$, let w^i be on the Pareto-frontier of V, and let $\varepsilon > 0$. There exists a block mechanism $(f^i, M^{\alpha^i,T^i})^\infty$ and a discount factor δ^* such that, for all $\delta \geq \delta^*$, all L^i, all μ, and all sequential equilibria of the punishment mechanism $(i, L^i, (f^i, M^{\alpha^i,T^i})^\infty)$, the expected continuation payoff profile is within distance ε of w^i at all feasible histories of length at least $L^i + 1$.*

The punishment mechanism possesses a sequential equilibrium (Fudenberg and Levine, 1983).

6.2. **Strategies and Beliefs.** Equilibrium strategies can be informally described as follows.[35] Players start in the *cooperative phase* by reporting as in a sequential equilibrium of the block mechanism $(f, M^{\alpha,T})^\infty$, where $v = \mathbb{E}_\pi[u(f(\theta), \theta)]$, and given a message profile $m^t \in \Theta$, by taking actions according to $f(m^t)$. Any observable deviation by player i (i.e., reporting a message that would have been infeasible in the mechanism, or, given messages m^t, choosing an action other than $f_i(m^t)$) triggers the players to mimic the play of a sequential equilibrium of the punishment mechanism against i. As described above, this consists of an L^i-period *stick phase* followed by a *carrot phase*. An observable deviation by any player j from the equilibrium of the punishment mechanism against i triggers the players to mimic the play of a sequential equilibrium of the punishment mechanism against j, unless the deviation is by player i during the stick phase against himself, in which case play continues to mimic the equilibrium of the punishment mechanism against i.

Let us now describe more formally the assessment. To simplify the notation, in the sequel we assume that $f(\theta), f^i(\theta) \in A$ for all $\theta \in \Theta$ and all $i \in I$. (We note that when f is a randomized rule, players coordinate their actions by conditioning

[35] How we set the free parameters describing the strategies will be discussed in the next subsection.

on the realization of the public randomization device.) Take $H^{\alpha,T}$ as the set of all public histories of feasible messages of the block mechanism $(f, M^{\alpha,T})^\infty$. These histories are not proper dynamic game public histories as they do not specify actions ensuing reports. It is therefore useful to define the set of cooperative histories $H_f^{\alpha,T}$ as the set of histories in which the reports belong to $H^{\alpha,T}$ and each report m^t is followed by an action profile $f(m^t)$. It is also useful to define $H_{f^j}^{\alpha^j,L^j,T^j}$ as the set of dynamic game public histories in which players play as in arbitrary feasible histories of the punishment mechanism $(j, L^j, (f^j, M^{\alpha^j,T^j})^\infty)$ with arbitrary reports during the cheap talk stages of the first L^j rounds, j picking arbitrary actions $a_j \in A_j$ but $-j$ minmaxing j by choosing a_{-j}^j during the first L^j rounds, whereas from round L^j+1 on messages are restricted by the block mechanism $(f^j, M^{\alpha^j,T^j})^\infty$ and actions coincide with $f^j(m^t)$.[36]

The assessment (σ, μ) is constructed as follows. Pick first a sequential equilibrium (ρ^0, μ^0) of the block mechanism $(f, M^{\alpha,T})^\infty$, given initial beliefs $\lambda \in \Delta(\Theta)$. For cooperative histories $h \in H_f^{\alpha,T}$, σ_i mandates player i to report as ρ_i^0, and to pick actions according to $f_i(m^t)$. Beliefs are as given by the belief system μ^0.

Take now a history $h \notin H_f^{\alpha,T}$ such that all histories preceding it belong to $H_f^{\alpha,T}$. If it is not enough to change the action of only one player to transform h into a history in $H_f^{\alpha,T}$, then restart the mechanism with some given beliefs $\bar{\mu}$. Suppose then that it is enough to change the play of only one of the players, say player j, to transform h into a history belonging to $H_f^{\alpha,T}$. As discussed informally above, play now mimics the behavior in a punishment mechanism against j. Take the beliefs players have at the beginning of the punishment mechanism about the rivals' current type as $\bar{\mu}_{-j}$ and $\tilde{\mu}_j$, where $\bar{\mu}_{-j}$ is fixed and $\tilde{\mu}_j$ is the belief player j can form about $-j$'s type by using Bayes rule after observing the preceding on-path behavior by $-j$. Let (ρ^j, μ^j) be the sequential equilibria of the punishment mechanism $(j, L^j, (f^j, M^{\alpha^j,T^j})^\infty)$, given the beliefs $\bar{\mu}_{-j}$ and $\tilde{\mu}_j$. For histories (h, h') with $h' \in H_{f^j}^{\alpha^j,L^j,T^j}$ of length less than L^j (i.e., stick-phase histories) pick reports uniformly in Θ_i; play the minmaxing action a_i^j if $j \neq i$, or pick actions as prescribed by the equilibrium ρ_i^j of the punishment mechanism against j if $j = i$. For histories of length greater than L^j (i.e., carrot-phase histories) the reports are as prescribed by ρ^j and actions are taken as mandated by f^j. Beliefs are as given by the belief system μ^j associated to the punishment mechanism equilibrium.

[36]Formally, $H_{f^j}^{\alpha^j,L^j,T^j}$ is composed of two types of histories. Histories of length less than L^j, say $t \leq L^j$, belonging to $\Theta^t \times A_j^{t'} \times \{a_{-j}^j\}^{t'}$, with $t' = t$ or $t' = t-1$; and histories of length greater than L^j obtained by concatenating the aforementioned histories, for $t' = t = L^j$, with histories in $H_{f^j}^{\alpha,T^j}$.

If after some history player k deviates while $-k$ conformed to the above punishment phase, then start mimicking the punishment mechanism against k as explained computing beliefs about current types by Bayes rule when possible. The exception is that after deviations by player i in the stick phase against himself the play continues as specified by ρ^i.

6.3. **Proof of Theorem 4.1.** Suppose, without loss, that $\varepsilon > 0$ is small enough such that there exists $\gamma \in]0,1[$ satisfying

$$\underline{v}_i + \varepsilon < \min\{v_i, w_i^i\}$$
$$w_i^i + \varepsilon < \min\{v_i, w_i^{-i}\}$$
$$\gamma > \frac{\varepsilon}{w_i^i - \underline{v}_i}$$

and

$$\gamma\big(\underline{v}_{-i} + \frac{\varepsilon}{2} - b\big) + (1-\gamma)\big(w_{-i}^{-i} - w_{-i}^i + \varepsilon\big) < 0.$$

for all $i \in I$, where $b = \min\{u_i(a, \theta_i) \mid i \in I, a \in A, \theta_i \in \Theta_i\}$. Such γ can always be found when ε is small enough, given the construction of w^i.

Take now the block mechanism $(f, M^{\alpha,T})^\infty$ yielding payoffs within distance $\varepsilon/2$ of v for all sequential equilibria (Corollary 5.2) and the punishment mechanism $(i, L^i, (f^i, M^{\alpha^i,T^i})^\infty)$ yielding payoffs within distance $\varepsilon/2$ of w^i during the carrot phase (Corollary 6.1), for all $\delta \geq \delta^0$ and all L^i. We will prove that when δ is large enough we can pick L^i, for each $i \in I$, such that the assessment described in the previous subsection forms a PBE.

Note that in the punishment mechanism against i, player i's total expected payoff during the first L^i rounds in which he is being minmaxed is at most

$$\max_{\theta_i \in \Theta_i} \sum_{t=1}^{L^i} \delta^t \mathbb{E}[\max_{a_i \in A_i} u_i(a_i, a_{-i}^i, \theta_i^t) \mid \theta_i^1 = \theta_i].$$

The following lemma uses the irreducibility of each player's type process to provide an upper bound for this term for large L^i. It will be convenient to consider $d^i \in \mathbb{N}$ as the period of the Markov chain with transition matrix P_i (see Norris 1997 for details) and define

$$L^i(\delta) = \max\left\{nd^i \mid n \in \mathbb{N}, \quad nd^i \leq \frac{\ln(1-\gamma)}{\ln(\delta)}\right\}$$

We observe that, as $\delta \to 1$, $L^i(\delta) \to \infty$ and $\delta^{L^i(\delta)} \to 1 - \gamma$.

Lemma 6.1. *There exists $\delta^1 \geq \delta^0$ such that for all $\delta > \delta^1$, all $i \in I$, and all $\theta_i \in \Theta_i$*

$$\frac{1-\delta}{1-\delta^{L^i(\delta)}} \sum_{t=1}^{L^i(\delta)} \delta^{t-1} \mathbb{E}\Big[\max_{a_i \in A_i} u_i(a_i, a_{-i}^i, \theta_i^t) \mid \theta_i^1 = \theta_i\Big] \leq \underline{v}_i + \frac{\varepsilon}{2}.$$

The proof of this result is presented in the Appendix B.

Let us now show that for all δ sufficiently large, when $L^i = L^i(\delta)$, the prescribed assessment forms an equilibrium. Since in all equilibria of the block and punishment mechanisms beliefs are consistent, it is enough to show that (σ, μ) is sequentially rational. Now, deviations that do not trigger a change in phase cannot be optimal as the prescribed behavior corresponds to a sequentially rational behavior in a game having, at each round, the same expected continuation payoffs as in our dynamic game. (In other words, such a deviation would be a profitable deviation in the block mechanism or in the punishment mechanism, which is impossible given that the play within each phase mimics a sequential equilibrium of the mechanism.) Thus, it is enough to show that deviations triggering a change in phase cannot be optimal.[37]

Consider first the incentives at the cheap-talk stage (i.e., at $t.2$). At any stick-phase history each player i randomizes uniformly over Θ_i and hence all messages are on the equilibrium path. At any cooperative or carrot phase history, conforming to the equilibrium strategy player i is getting at least $w_i^i - \frac{\varepsilon}{2}$, while a deviation will result in a payoff of at most

$$(1-\delta^{L^i})(\underline{v}_i + \frac{\varepsilon}{2}) + \delta^{L^i}(w_i^i + \frac{\epsilon}{2})$$

where we use Lemma 6.1 to bound the stick payoffs. Taking the limit as $\delta \to 1$, the incentive constraint becomes $\gamma \geq \frac{\varepsilon}{w_i^i - \underline{v}_i}$, which holds strictly by definition of γ. Thus we can find $\delta^2 \geq \delta^1$ such that for all $\delta > \delta^2$, the incentive constraint holds.

Consider then the incentives to conform with the prescribed actions at the action stage (i.e., at $t.4$).

Cooperative Histories. At any history $h \in H_f^{\alpha,T}$, player i's payoff is at least $v_i - \frac{\varepsilon}{2}$. A deviation will trigger the punishment mechanism and, from Lemma 6.1, will yield expected payoffs of at most

$$\begin{aligned}(1-\delta)B + &(\delta - \delta^{L^i+1})(\underline{v}_i + \frac{\varepsilon}{2}) + \delta^{L^i+1}\Big(w_i^i + \frac{\varepsilon}{2}\Big) \\ &\leq (1-\delta)B + (\delta - \delta^2)(\underline{v}_i + \frac{\varepsilon}{2}) + \delta^2\Big(w_i^i + \frac{\varepsilon}{2}\Big).\end{aligned}$$

At $\delta = 1$, the right side is strictly less than the $v_i - \frac{\varepsilon}{2}$ and thus we can find $\delta^3 \geq \delta^2$, such that for all $\delta > \delta^3$, the on-path incentives hold.

Stick-Phase Histories. During the first L^i rounds of a punishment mechanism against i, by construction of the strategies, player i has no incentive to deviate from his prescribed equilibrium strategy. It is therefore enough to show that it is

[37]Note that such deviations include "double deviations," where a player first deviates unobservably and only then deviates in a way that triggers the punishment.

in player $-i$'s interest to choose a^i_{-i}. Indeed, by conforming, $-i$'s payoff is at least

$$(1-\delta^{L^i})b + \delta^{L^i}\Big(w^i_{-i} - \frac{\varepsilon}{2}\Big),$$

while a deviation will result in a current payoff of at most B and will trigger a punishment mechanism against $-i$. From Lemma 6.1, this will result in a payoff of at most

$$(1-\delta)B + (\delta - \delta^{L^i+1})(\underline{v}_i + \frac{\varepsilon}{2}) + \delta^{L^i+1}\Big(w^{-i}_{-i} + \frac{\varepsilon}{2}\Big).$$

The incentive constraint can be written as

$$(1-\delta)B + (1-\delta^{L^i})\Big\{\delta\big(\underline{v}_{-i} + \frac{\varepsilon}{2}\big) - b\Big\} + \delta^{L^i}\Big\{\delta\big(w^{-i}_{-i} + \frac{\varepsilon}{2}\big) - \big(w^i_{-i} - \frac{\varepsilon}{2}\big)\Big\} \leq 0.$$

As $\delta \to 1$, the left side goes to $\gamma\Big(\underline{v}_{-i} + \frac{\varepsilon}{2} - b\Big) + (1-\gamma)\Big(w^{-i}_{-i} - w^i_{-i} + \varepsilon\Big)$ which is strictly less than 0. Therefore, there exists $\delta^4 \geq \delta^3$ such that for all $\delta > \delta^4$, the incentive constraint holds.

Carrot-Phase Histories. Consider now the incentives each of the players has during the carrot phase following the stick phase triggered after a deviation by i. It is enough to show that, after each history of reports, it is in each of the players' interest to choose actions as prescribed by $f^i(m^t)$. Conforming to the equilibrium strategy, player j gets a payoff of at least $w^i_j - \frac{\varepsilon}{2}$. A deviation by j will trigger the punishment phase against j resulting in an expected discounted payoff of at most

$$(1-\delta)B + (\delta - \delta^{L^i+1})\Big(\underline{v}_j + \frac{\varepsilon}{2}\Big) + \delta^{L^i+1}\Big(w^j_j + \frac{\varepsilon}{2}\Big).$$

Player j will not deviate provided

$$(1-\delta)B + (\delta - \delta^{L^i+1})\Big(\underline{v}_j + \frac{\varepsilon}{2}\Big) + \delta^{L^i+1}\Big(w^j_j + \frac{\varepsilon}{2}\Big) \leq w^i_j - \frac{\varepsilon}{2}.$$

As $\delta \to 1$, the inequality becomes

$$\gamma\big(\underline{v}_j + \frac{\varepsilon}{2}\big) + (1-\gamma)\big(w^j_j + \frac{\varepsilon}{2}\big) \leq w^i_j - \frac{\varepsilon}{2},$$

and this inequality will hold strictly provided

$$\gamma > \frac{\varepsilon}{w^j_j - \underline{v}_j}.$$

Hence there exists $\delta^5 \geq \delta^4$ such that for all $\delta > \delta^5$, the corresponding inequality holds.

It then follows that by taking $\bar{\delta} = \delta^5$, the prescribed strategies form a PBE when $\delta > \bar{\delta}$.

7. Concluding Remarks

We restrict attention to two-player games but much of the argument extends as such to the n-player case. Essentially the only step that uses the restriction to two players is the proof of the security payoff result of Theorem 5.1. As noted in footnote 28, the obvious extension of the message spaces to the n-player case does not deliver the result: The conditional distributions of reports by each player $j \neq i$ are still forced to converge to the true marginal distributions $P_j(\theta)$, $\theta \in \Theta$, but the construction does not guarantee the convergence of the joint distribution of reports by players $-i$ to the product $\times_{j \neq i} P_j(\theta)$. Hence an honest player i may face the wrong distribution of messages. An analogous problem arises in the static setting with iid types studied by Jackson and Sonnenschein (2007), and we conjecture that augmenting our mechanism with a statistical test similar to theirs would allow the extension of Theorem 4.1 to n players.

We assume that the process governing the evolution of types is autonomous. Extending the results to decision controlled processes studied in the literature on stochastic games (see, e.g., Dutta, 1995, and the references therein) appears feasible but notationally involved.

The main restrictions on the nature of the private information are the assumptions about private values and independence of transitions across players. Both assumptions are crucial for our argument. The literature on mechanism design tells us that when valuations are interdependent (sometimes referred to as common values), efficiency need not be achievable (see Jehiel and Moldovanu, 2001). Thus extending our results to games with interdependent valuations necessitates additional assumptions about how the information effects the players' payoffs. In contrast, going from independent to correlated types in general expands the set of implementable outcomes in a mechanism design setting (see Cremer and McLean, 1988). This suggests that the results can be potentially extended to the case of correlated transitions.

Appendix A. Proofs for the Mechanism Section

This appendix contains the omitted proofs from Section 5 and an auxiliary lemma used in the proof of Proposition 5.1. They are presented in the order they appear in the main text.

Proof of Lemma 5.1. Without loss we may label the elements of Θ from 1 to $|\Theta|$. Define the cdf G from g by setting $G(k) = \sum_{j=1}^{k} g(j)$. The empirical cdf's G^n are defined analogously from g^n. For all n, all k,

$$|g^n(k) - g(k)| \leq |G^n(k) - G(k)| + |G^n(k-1) - G(k-1)|,$$

so that $\|g^n - g\| \leq 2\|G^n - G\|$. Defining the events $B_n = \{\|g^n - g\| \leq b_n\}$ we then have $\left\{\|G^n - G\| \leq \frac{b_n}{2}\right\} \subset B_n$. Thus,

$$\mathbb{P}(B_n) \geq \mathbb{P}(\|G^n - G\| \leq \frac{b_n}{2}) \geq 1 - 2e^{-2n(\frac{b_n}{2})^2} = 1 - \frac{6\alpha}{\pi^2 n^2},$$

where the second inequality is by Massart (1990) and the equality is by definition of b_n. The lemma now follows by observing that

$$\mathbb{P}(\bigcap_{n\in\mathbb{N}} B_n) = 1 - \mathbb{P}(\bigcup_{n\in\mathbb{N}} B_n^C) \geq 1 - \sum_{n\in\mathbb{N}} P(B_n^C) \geq 1 - \sum_{n\in\mathbb{N}} \frac{6\alpha}{\pi^2 n^2} = 1 - \alpha,$$

where the last equality follows since $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi}{6}$. □

Proof of Lemma 5.2. The following claim presents a useful characterization of the set $V(\delta, T)$.

Claim A.1. *For all δ and all T,*

$$V(\delta,T) = \Big\{ v \in \mathbb{R}^I \mid \exists f\colon \Theta \to \Delta(A) \ v = \frac{1-\delta}{1-\delta^T} \sum_{t=1}^{T} \delta^{t-1} \mathbb{E}[u(f(\theta^t), \theta^t)] \Big\}.$$

To prove the claim, note first that $V(\delta, T)$ is convex and thus it can be obtained as the convex hull of its extreme points. Moreover, for any such extreme point we can find a vector $p \in \mathbb{R}^2$ such that the stationary rule $f\colon \Theta \to A$ resulting in total payoffs v solves $f(\theta) \in \arg\max_{a\in A} p \cdot u(a, \theta)$ for all $\theta \in \Theta$. Considering arbitrary randomizations over the extreme points, we obtain the whole set $V(\delta, T)$.

We now prove the lemma. For each T, denote by π^T the empirical distribution of types, given a realization $(\theta^1, \ldots, \theta^T)$. Then, there exists T' such that for all $T \geq T'$,

$$\mathbb{P}[\left\|\pi^T - \pi\right\| < \frac{\varepsilon}{4B\,|\Theta|}] > 1 - \frac{\varepsilon}{4B\,|\Theta|},$$

and thus

$$\mathbb{E}[\left\|\pi^T - \pi\right\|] \leq \frac{\varepsilon}{2B\,|\Theta|}.$$

For each $T \geq T'$, take $\delta' > 0$ such that for all $(a^t)_{t=1}^T \in A^T$, all $(\theta^t)_{t=1}^T \in \Theta^T$, and all $\delta \geq \delta'$

$$\left| \frac{1}{T} \sum_{t=1}^{T} u(a^t, \theta^t) - \frac{1-\delta}{1-\delta^T} \sum_{t=1}^{T} \delta^{t-1} u(a^t, \theta^t) \right| < \frac{\varepsilon}{4}.$$

To prove the result, we need to show that for $T \geq T'$ and $\delta \geq \delta' (= \delta'(T))$

$$dist(V, V(\delta,T)) = \max \Big\{ \sup_{v\in V} dist(v, V(\delta,T)), \sup_{v'\in V(\delta,T)} dist(v', V) \Big\} < \varepsilon.$$

Take first $v \in V$ and the rule $f\colon \Theta \to \Delta(A)$ such that $v = \mathbb{E}_\pi[u(f(\theta), \theta)]$. Now, take $v' \in V(\delta, T)$ defined as

$$v' = \frac{1-\delta}{1-\delta^T} \sum_{t=1}^{T} \delta^{t-1} \mathbb{E}[u(f(\theta), \theta)].$$

Note that

$$\begin{aligned}\left|v - v'\right| \leq &\left|\mathbb{E}_\pi[u(f(\theta),\theta)] - \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[u(f(\theta),\theta)]\right| \\ &+ \left|\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[u(f(\theta),\theta)] - \frac{1-\delta}{1-\delta^T}\sum_{t=1}^{T}\delta^{t-1}\mathbb{E}[u(f(\theta),\theta)]\right|.\end{aligned}$$

The first term on the right side is less than or equal to $\frac{\varepsilon}{2}$. Indeed,

$$\begin{aligned}\left|\mathbb{E}_\pi[u(f(\theta),\theta)] - \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[u(f(\theta),\theta)]\right| &= \left|\mathbb{E}\big[\sum_{\theta\in\Theta} u(f(\theta),\theta)\big(\pi(\theta) - \pi^T(\theta)\big)\big]\right| \\ &\leq \mathbb{E}\big[\sum_{\theta\in\Theta} |u(f(\theta),\theta)|\left|\pi(\theta) - \pi^T(\theta)\right|\big] \\ &\leq \mathbb{E}[|\Theta|\, B \left\|\pi - \pi^T\right\|] \\ &\leq \frac{\varepsilon}{2}.\end{aligned}$$

To bound the second term,

$$\begin{aligned}&\left|\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[u(f(\theta),\theta)] - \frac{1-\delta}{1-\delta^T}\sum_{t=1}^{T}\delta^{t-1}\mathbb{E}[u(f(\theta),\theta)]\right| \\ &\qquad\qquad \leq \mathbb{E}[\left|\frac{1}{T}\sum_{t=1}^{T}u(a^t,\theta^t) - \frac{1-\delta}{1-\delta^T}\sum_{t=1}^{T}\delta^{t-1}u(a^t,\theta^t)\right|] < \frac{\varepsilon}{4}.\end{aligned}$$

It then follows that $|v - v'| < \frac{3}{4}\varepsilon$ and thus $\sup_{v\in V} dist(v, V(\delta,T)) < \varepsilon$.

Conversely, take $v' \in V(\delta,T)$ and the associated stationary decision rule $f\colon \Theta \to \Delta(A)$ such that $v' = \frac{1-\delta}{1-\delta^T}\sum_{t=1}^{T}\delta^{t-1}\mathbb{E}[u(f(\theta^t),\theta^t)]$. Defining $v = \mathbb{E}_\pi[u(f(\theta,\theta))] \in V$, the bounds above show that $|v' - v| < \frac{3}{4}\varepsilon$ and thus $\sup_{v'\in V} dist(v', V) < \varepsilon$. We have thus established the lemma. □

The following preliminary result is used in the proofs of Corollaries 5.1 and 5.2. Take $p \in \mathbb{R}^I_+ \setminus 0$ such that for all $w \in V$, $p \cdot w \leq p \cdot v$. Define the set $\mathbf{Tr}(\kappa, v) = \{w \in \mathbb{R}^I \mid w'_i \geq v_i - \kappa, p\cdot w \leq p \cdot v\}$, for $\kappa > 0$.

Claim A.2. *Assume that $p \gg 0$ and $\sum_{i\in I} p_i = 1$. Then, for all $w \in \boldsymbol{Tr}(\kappa, v)$, $\|w - v\| \leq \kappa \max\{\frac{1}{p_i} \mid i \in I\}$.*

Proof. Consider the problem $\max\{\|w - v\| \mid w \in \mathbf{Tr}(\kappa, v)\}$. This is a maximization problem, with a convex objective function, and a convex and compact set of restrictions. Corollary 32.3.2 in Rockafellar (1970) implies that the maximum is attained at extreme points of $\mathbf{Tr}(\kappa, v)$. Let w be an extreme point of $\mathbf{Tr}(\kappa, v)$ such that for some i, $w_i \neq v_i - \kappa$. Then, for $j \neq i$, $w_j = v_j - \kappa$ and $p \cdot w = p \cdot v$. (Otherwise, we would contradict the fact that w is an extreme point by obtaining it as a convex combination of points in $\mathbf{Tr}(\kappa, v)$.) It follows that for any such extreme point w, $|w_i - v_i| = \frac{p_{-i}}{p_i}\kappa \leq \frac{1}{p_i}\kappa$. We deduce that

$$\max\{\|w - v\| \mid w \in \mathbf{Tr}(\kappa, v)\} \leq \max\{\kappa, \max_{i \in I} \frac{\kappa}{p_i}\} \leq \kappa \max\{\frac{1}{p_i} \mid i \in I\},$$

which proves the claim. □

We are now in a position to prove Corollaries 5.1 and 5.2.

Proof of Corollary 5.1. Assume, without loss, that for some vector p normal to V at v, $p \gg 0$ and $\sum_{i \in I} p_i = 1$. (If all normal vectors at v have some zero component, v can be approximated by points in the frontier having strictly positive normal vectors.) Take $e > 0$ such that $e(1 + 2\max\{\frac{1}{p_i} \mid i \in I\}) = \varepsilon$. From Theorem 5.1 and Lemma 5.2, there is T^* such that for all $T \geq T^*$ there exists δ^* such that for all $\delta \geq \delta^*$ (i) the Hausdorff distance between V and $V(\delta, T)$ is at most e, and (ii) for any Nash equilibrium payoff $v^{\delta,T} \in V(\delta, T)$ of the mechanism $(f, M^{\alpha,T})$, $v_i^{\delta,T} \geq v_i - e$. Let $w^{\delta,T} \in V$ be such that $\left\|w^{\delta,T} - v\right\| \leq e$ and thus $w_i^{\delta,T} \geq v_i - 2e$ for all $i \in I$, all $T \geq T^*$ and all $\delta \geq \delta^*(T)$. Since $w^{\delta,T} \in \mathbf{Tr}(2e, v)$, Claim A.2 implies that

$$\left\|w^{\delta,T} - v\right\| \leq 2e \max\{\frac{1}{p_i} \mid i \in I\},$$

and thus

$$\left\|v^{\delta,T} - v\right\| \leq \left\|v^{\delta,T} - w^{\delta,T}\right\| + \left\|w^{\delta,T} - v\right\| \leq e(1 + 2\max\{\frac{1}{p_i} \mid i \in I\}) = \varepsilon.$$

The result follows. □

Proof of Corollary 5.2. We start by establishing the result about sequential equilibria. Take p to be a normal vector to V at v, with $p \gg 0$ and $\sum_{i \in I} p_i = 1$ and let $e > 0$ be such that $e(1 + 2\max\{\frac{1}{p_i} \mid i \in I\}) = \frac{\varepsilon}{2}$. Theorem 5.1 and Lemma 3.1 implies the existence of a block mechanism $(f, M^{\alpha,T})^\infty$ and $\delta' < 1$ such that for all $\delta \geq \delta'$ and all initial beliefs (i) the Hausdorff distance between V and $V(\delta)$ is at most e, and (ii) each player i can secure a payoff $v_i - e$ at the beginning of the block mechanism $(f, M^{\alpha,T})^\infty$. (To see (ii), observe that Theorem 5.1 implies the result for each of the blocks and then note that the total payoffs in the block mechanism can be decomposed as a sum of payoffs over all the blocks.) Since (ii) holds irrespective of the initial beliefs and the beginning of each block is the beginning of a block mechanism, we can strengthen (ii) and say that (iii) each player i can secure a payoff $v_i - e$ at the beginning of each block of the block mechanism $(f, M^{\alpha,T})^\infty$. In particular, we have that (iv) for any sequential equilibrium payoff $v^{\delta,Tn}$ of the block mechanism $(f, M^{\alpha,T})^\infty$ accruing at the beginning of some block (or, equivalently, after a history of length Tn with $n \in \mathbb{N}$), $v_i^{\delta,Tn} \geq v_i - e$ for all $i \in I$. Now, combining (i), (iv), and Claim A.2, we deduce, as we did in the proof of Corollary 5.1 that $\left\|v^{\delta,Tn} - v\right\| \leq \frac{\varepsilon}{2}$. Now, the result follows by taking $\delta^* \in [\delta', 1[$

such that for all $\delta \geq \delta^*$, $(1-\delta^T)\max\{|u_i(a,\theta_i)| \mid i \in I, a \in A, \theta_i \in \Theta_i\} \leq \frac{\varepsilon}{4}$ and $(1-\delta^T)\,\|v\| \leq \frac{\varepsilon}{4}$.

The result about Nash equilibrium payoffs after on-path histories follows by noting that after any on-path history a player can always revert to playing the honest strategy. Details can be filled in as in the paragraph above. □

Proof of Lemma 5.3. Let λ_1 be degenerate. Fix T and $\alpha > 0$. Suppose that the honest player 1 is not subject to the message spaces and hence is truthful. It is straightforward to check that, by construction, player 2's sets of feasible messages remain non-empty at histories that include infeasible messages by player 1. Furthermore, player 2's strategy ρ_2 can be extended to such histories arbitrarily as they play no role in what follows.

It is convenient to introduce an auxiliary probability space.[38] Consider the probability space $([0,1],\mathcal{B},\hat{\mathbb{P}})$ and a countable collection of *independent* random variables

$$\tilde{\psi}^n_\theta : [0,1] \to \Theta_1, \quad \theta \in \Theta,\ n \in \mathbb{N},$$

where

$$\hat{\mathbb{P}}(\tilde{\psi}^n_{(\theta_1,\theta_2)} = \theta_1') = P_1(\theta_1,\theta_1').$$

Imagine the variables $\tilde{\psi}^n_\theta$ set out in the following array:

$$\begin{array}{ccccc} \tilde{\psi}^1_1 & \tilde{\psi}^2_1 & \cdots & \tilde{\psi}^n_1 & \cdots \\ \tilde{\psi}^1_2 & \tilde{\psi}^2_2 & \cdots & \tilde{\psi}^n_2 & \cdots \\ \vdots & & & & \\ \tilde{\psi}^1_{|\Theta|} & \tilde{\psi}^2_{|\Theta|} & \cdots & \tilde{\psi}^n_{|\Theta|} & \cdots \end{array}$$

Given the array, we can think of the sequence of player 1's types and player 2's messages, $(\theta_1^t, m_2^t)_{t=1}^T$, as being generated as follows. Since λ_1 puts probability one on some θ_1, player 1's period 1 type is simply $\theta_1^1 = \theta_1$. Player 2's period 1 message is some constant m_2^1. Player 1's period 2 type θ_1^2 is then drawn by sampling the first variable in the row indexed by the first period messages (θ_1^1, m_2^1). (I.e., we observe $\tilde{\psi}^1_{(\theta_1^1,m_2^1)}$ and put $\theta_1^2 = \psi^1_{(\theta_1^1,m_2^1)}$.) Player 2's period 2 message is given by $m_2^2 = \rho_2^2(\theta_1^1)$. Player 1's period 3 type θ_1^3 is then drawn by sampling the first element of the row indexed by the second period messages (θ_1^2, m_2^2), unless $(\theta_1^1, m_2^1) = (\theta_1^2, m_2^2)$, in which case the second variable in the row indexed by (θ_1^1, m_2^1) is sampled instead. And so forth.

To see that this construction indeed gives rise to the right process over the T periods, fix a finite sequence $(\theta_1^1, m_2^1), (\theta_1^2, m_2^2), \ldots, (\theta_1^T, m_2^T)$. Obviously we must have

$$m_2^t = \rho_2^t(\theta_1^1, \ldots, \theta_1^{t-1})$$

[38] The construction is adapted from Billingsley (1961). It is similar to, yet distinct from, the one used in the proof of Proposition 5.1.

for all $t = 1, \ldots, T$ since otherwise the probability of this sequence is trivially zero. So suppose this is the case. Then the probability of the sequence according to the original description of the process is simply

$$\lambda_1(\theta_1^1)P_1(\theta_1^1, \theta_1^2) \cdots P_1(\theta_1^{T-1}, \theta_1^T).$$

On the other hand, the above construction assigns this sequence the probability

$$\lambda_1(\theta_1^1)\hat{\mathbb{P}}(\tilde{\psi}^1_{(\theta_1^1, m_2^1)} = \theta_1^2) \cdots \hat{\mathbb{P}}(\tilde{\psi}^k_{(\theta_1^{T-1}, m_2^{T-1})} = \theta_1^T),$$

where $k-1$ is the number of times the pair $(\theta_1^{T-1}, m_2^{T-1})$ appears in the sequence, and where we have used independence of the $\tilde{\psi}_\theta^n$ to write the joint probability as a product. By construction of the array,

$$\hat{\mathbb{P}}(\tilde{\psi}^1_{(\theta_1^1, m_2^1)} = \theta_1^2) = P_1(\theta_1^1, \theta_1^2),$$

and

$$\hat{\mathbb{P}}(\tilde{\psi}^k_{(\theta_1^{T-1}, m_2^{T-1})} = \theta_1^T) = P_1(\theta_1^{T-1}, \theta_1^T),$$

(and similarly for the elements we haven't explicitly written out) so both methods assign the sequence the same probability. Hence we may work with the auxiliary probability space and the above array.

We may apply Lemma 5.1 along each row of the array to conclude that, with $\hat{\mathbb{P}}$-probability at least $1-\alpha$, for all $n \in \mathbb{N}$ the empirical measure of the first n observations along the row is within b_n^α of the true distribution. Hence with $\hat{\mathbb{P}}$-probability at least $1 - |\Theta|\alpha$ this is true along all $|\Theta|$ rows. But in this event player 1's truthful reports remain feasible even if he was subject to the message spaces: Regardless of how the realized types θ_1^t and the strategy ρ_2 lead us to sample from the array, player 1's types (and hence his messages) are converging fast enough conditional on any previous period message profile θ because, by construction, player 1's types in periods t where $m^{t-1} = \theta$ are drawn along the row indexed by θ.

In terms of the original description of the process the above argument implies that with at least probability $1 - |\Theta|\alpha$ we get a sample path $(\theta_1^t)_{t=1}^T$ such that truthful reporting is feasible. But given such a path player 1 is truthful even if he was subject to the messages spaces. The claim follows. □

The following lemma is used in the proof of Proposition 5.1.

Lemma A.1. *Let P be an irreducible stochastic matrix on a finite set Θ, and let π denote the unique invariant distribution for P. Let $(\theta^t)_{t \in N}$ be a sequence in Θ. For all t, define the empirical matrix P^t by setting*

$$P^t(\theta, \theta') = \frac{|\{s \in \{1, \ldots, t-1\} : (\theta^s, \theta^{s+1}) = (\theta, \theta')\}|}{|\{s \in \{1, \ldots, t-1\} : \theta^s = \theta\}|},$$

and define the empirical distribution π^t by setting

$$\pi^t_\theta = \frac{|\{s \in \{1, \dots, t\} : \theta^s = \theta\}|}{t}.$$

For all $\varepsilon > 0$ there exists $T < \infty$ and $\eta > 0$ such that for all $t \geq T$,

$$\|P^t - P\| < \eta \quad \Rightarrow \quad \|\pi^t - \pi\| < \varepsilon.$$

P^t is an empirical transition matrix that records for each state θ the empirical *conditional* frequencies of transitions $\theta \to \theta'$ in $(\theta^s)_{s=1}^t$. Similarly, π^t is an empirical measure that records the frequencies of visits to each state in $(\theta^s)_{s=1}^t$. So in words the lemma states roughly that if the conditional transition frequencies converge to those in P, then the empirical distribution converges to the invariant distribution for P.

Proof. Fix $\theta' \in \Theta$ and $t \in \mathbb{N}$. Note that $t\pi^t_{\theta'}$ is the number of visits to θ' in $(\theta^s)_{s=1}^t$. Since each visit to θ' is either in period 1 or preceded by some state θ, we have

$$t\pi^t_{\theta'} \leq 1 + \sum_{\theta \in \Theta} |\{s < t : \theta^s = \theta\}| P^t(\theta, \theta') \leq |\Theta| + \sum_{\theta \in \Theta} t\pi^t_\theta P^t(\theta, \theta').$$

On the other hand,

$$t\pi^t_{\theta'} \geq \sum_{\theta \in \Theta} |\{s < t : \theta^s = \theta\}| P^t(\theta, \theta') \geq \sum_{\theta \in \Theta} t\pi^t_\theta P^t(\theta, \theta') - |\Theta|,$$

where the second inequality follows, since $|\{s < t : \theta^s = \theta\}| \geq t\pi^t_\theta - 1$ and $\sum_\theta P^t(\theta, \theta') \leq |\Theta|$. Putting together the above inequalities gives

$$-\frac{|\Theta|}{t} \leq \pi^t_{\theta'} - \sum_{\theta \in \Theta} \pi^t_\theta P^t(\theta, \theta') \leq \frac{|\Theta|}{t}.$$

Since θ' was arbitrary, we have in vector notation

$$-\frac{|\Theta|}{t}\mathbf{1} \leq \pi^t(I - P^t) \leq \frac{|\Theta|}{t}\mathbf{1},$$

where I is the identity matrix and $\mathbf{1}$ denotes a $|\Theta|$-vector of ones. This implies that for all t, there exists $e^t \in \mathbb{R}^{|\Theta|}$ such that $\|e^t\| \leq \frac{|\Theta|}{t}$ and $\pi^t(I - P^t) = e^t$. Let E be a $|\Theta| \times |\Theta|$-matrix of ones. Then

$$\pi^t(I - P^t + E) = \mathbf{1} + e^t \quad \text{and} \quad \pi(I - P + E) = \mathbf{1}.$$

It is straightforward to verify that the matrix $I - P + E$ is invertible when P is irreducible (see, e.g., Norris, 1997, Exercise 1.7.5). The set of invertible matrices is open, so there exists $\eta_1 > 0$ such that $I - P^t + E$ is invertible if $\|P^t - P\| < \eta_1$. Furthermore, the mapping $Q \mapsto (I - Q + E)^{-1}$ is continuous at P, so there exists $\eta_2 > 0$ such that $\|(I - P^t + E)^{-1} - (I - P + E)^{-1}\| < \frac{\varepsilon}{4|\Theta|}$ if $\|P^t - P\| < \eta_2$. Put

$\eta = \min\{\eta_1, \eta_2\}$ and put

$$T = \frac{2|\Theta|^2 \|(I - P + E)^{-1}\|}{\varepsilon}.$$

If $t \geq T$ and $\|P^t - P\| < \eta$, then

$$\begin{aligned}
\|\pi^t - \pi\| &= \|(\mathbf{1} + e^t)(I - P^t + E)^{-1} - \mathbf{1}(I - P + E)^{-1}\| \\
&\leq \|(\mathbf{1} + e^t)[(I - P^t + E)^{-1} - (I - P + E)^{-1}]\| + \|e^t(I - P + E)^{-1}\| \\
&\leq 2|\Theta|\|(I - P^t + E)^{-1} - (I - P + E)^{-1}\| + \frac{|\Theta|^2}{t}\|(I - P + E)^{-1}\| \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.
\end{aligned}$$

The lemma follows. □

Appendix B. Proof of Lemma 6.1

Proof. Fix $i \in I$ and the initial state $\theta_i^1 = \theta_i$. Let $P^{(t)}(\theta_i, \theta_i') = \mathbb{P}[\theta_i^t = \theta_i' \mid \theta_i^1 = \theta_i]$. From Theorem 1.8.4 in Norris (1997), for each $i \in I$, there exists a partition $(C_r^i)_{r=1}^{d^i}$ of Θ_i such that $P_i^{(n)}(\theta_i, \theta_i') > 0$ only if $\theta_i \in C_r^i$ and $\theta_i' \in C_{r+n}^i$ for some $r \in \{1, \dots, d^i\}$, where we write $C_{nd^i+r}^i = C_r^i$. Observe that, without loss, we can assume that the initial state is such that $\theta_i \in C_1^i$ for all i.

From Theorem 1.8.5 in Norris (1997), there exists $N = N(\theta_i) \in \mathbb{N}$ such that for all $n \geq N$ and all $\theta_i' \in C_r^i$, $\left|P^{(nd^i+r)}(\theta_i, \theta_i') - d^i \pi_i(\theta_i')\right| \leq \frac{\varepsilon}{8B|\Theta_i|}$. Note that for any such $n \geq N$,

$$\begin{aligned}
&\left|\sum_{r=1}^{d^i} \sum_{\theta_i' \in \Theta_i} \max_{a_i} u_i(a_i, a_{-i}^i, \theta_i')(P^{(nd^i+r)}(\theta_i, \theta_i') - \pi_i(\theta_i'))\right| \\
&\qquad = \left|\sum_{r=1}^{d^i} \sum_{\theta_i' \in C_r^i} \max_{a_i} u_i(a_i, a_{-i}^i, \theta_i')\big(P^{(nd^i+r)}(\theta_i, \theta_i') - d^i \pi_i(\theta_i')\big)\right| \\
&\qquad \leq \sum_{r=1}^{d^i} \sum_{\theta_i' \in C_r^i} B \frac{\varepsilon}{8B\,|\Theta_i|} \leq \frac{\varepsilon}{8}.
\end{aligned}$$

Now, note that for any δ and any $L \geq Nd^i + 1$,

$$\begin{aligned}
&\left|\frac{1-\delta}{1-\delta^L} \sum_{t=1}^{L} \delta^{t-1} \mathbb{E}[\max_{a_i \in A_i} u_i(a_i, a_{-i}^i, \theta_i^t) \mid \theta_i] - \underline{v}_i\right| \\
&\leq \frac{1-\delta^{Nd^i}}{1-\delta^L} 2B + \left|\frac{1-\delta}{1-\delta^L} \sum_{t=Nd^i+1}^{L} \delta^{t-1} \sum_{\theta_i' \in \Theta_i} \max_{a_i \in A_i} u_i(a_i, a_{-i}^i, \theta_i')\big(P^{(t)}(\theta_i') - \pi_i(\theta_i')\big)\right|.
\end{aligned}$$

To bound the second term, assume $L/d^i \in \mathbb{N}$ and note that

$$\begin{aligned}
&\left|\sum_{t=Nd^i+1}^{L} \delta^{t-1} \sum_{\theta_i' \in \Theta_i} \max_{a_i \in A_i} u_i(a_i, a_{-i}^i, \theta_i')\big(P^{(t)}(\theta_i') - \pi_i(\theta_i')\big)\right| \\
&\le \sum_{n=N}^{L/d^i-1} \delta^{nd^i-1} \left|\sum_{r=1}^{d^i} \delta^r \sum_{\theta_i' \in \Theta_i} \max_{a_i \in A_i} u_i(a_i, a_{-i}^i, \theta_i')\big(P^{(nd^i+r)}(\theta_i') - \pi_i(\theta_i')\big)\right| \\
&\le \sum_{n=N}^{L/d^i-1} \delta^{nd^i-1} \Big\{ \left|\sum_{r=1}^{d^i} (1-\delta^r) \sum_{\theta_i' \in \Theta_i} \max_{a_i \in A_i} u_i(a_i, a_{-i}^i, \theta_i')\big(P^{(nd^i+r)}(\theta_i') - \pi_i(\theta_i')\big)\right| \\
&\qquad + \left|\sum_{r=1}^{d^i} \sum_{\theta_i' \in \Theta_i} \max_{a_i \in A_i} u_i(a_i, a_{-i}^i, \theta_i')\big(P^{(nd^i+r)}(\theta_i') - \pi_i(\theta_i')\big)\right| \Big\} \\
&\le \sum_{n=N}^{L/d^i-1} \delta^{nd^i-1} \Big\{ (1-\delta^{d^i}) 2Bd^i\, |\Theta_i| + \frac{\varepsilon}{8} \Big\} \\
&= \frac{\delta^{d^i N-1} - \delta^{L-1}}{1-\delta^{d^i}} \Big\{ (1-\delta^{d^i}) 2Bd^i\, |\Theta_i| + \frac{\varepsilon}{8} \Big\},
\end{aligned}$$

and thus

$$\begin{aligned}
&\left| \frac{1-\delta}{1-\delta^L} \sum_{t=Nd^i+1}^{L} \delta^{t-1} \sum_{\theta_i' \in \Theta_i} \max_{a_i \in A_i} u_i(a_i, a_{-i}^i, \theta_i')\big(P^{(t)}(\theta_i') - \pi_i(\theta_i')\big)\right| \\
&\qquad\qquad \le \frac{1-\delta}{1-\delta^{d^i}} \frac{\delta^{d^i N-1} - \delta^{L-1}}{1-\delta^L} \Big\{ (1-\delta^{d^i}) 2Bd^i\, |\Theta_i| + \frac{\varepsilon}{8} \Big\} \\
&\qquad\qquad\qquad \le \Big\{ (1-\delta^{d^i}) 2Bd^i\, |\Theta_i| + \frac{\varepsilon}{8} \Big\} \le \frac{\varepsilon}{4},
\end{aligned}$$

if δ is big enough (uniformly in $L \ge Nd^i + 1$). Let $\delta(i) \in]0,1[$ be such that the last inequality holds for all $\delta \ge \delta(i)$.

Now, let δ_{θ_i} be such that for all $\delta \ge \delta_{\theta_i}$, $L^i(\delta) \ge N(\theta_i)d^i + 1$ and

$$\frac{1-\delta^{Nd^i}}{1-\delta^{L^i(\delta)}} 2B < \frac{\varepsilon}{4}.$$

Defining $\delta_{i,\theta_i} = \max\{\delta_{\theta_i}, \delta(i)\}$, it then follows that for all $\delta \ge \delta_{i,\theta_i}$,

$$\left| \frac{1-\delta}{1-\delta^{L^i(\delta)}} \sum_{t=1}^{L^i(\delta)} \delta^{t-1} \mathbb{E}[\max_{a_i \in A_i} u_i(a_i, a_{-i}^i, \theta_i^t) \mid \theta_i] - \underline{v}_i \right| < \frac{\varepsilon}{2}.$$

Finally, taking $\delta^1 = \max\{\delta^0, \max\{\delta_{i,\theta_i} \mid i \in I, \theta_i \in \Theta_i\}\}$ gives the result. □

References

Abdulkadiroglu, A., and K. Bagwell (2007): "Trust, Reciprocity and Favors in Cooperative Relationships," Discussion paper, Stanford University, Department of Economics.

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58(5), 1041–1063.

Athey, S., and K. Bagwell (2001): "Optimal Collusion with Private Information," *RAND Journal of Economics*, 32(3), 428–465.

——— (2008): "Collusion with Persistent Cost Shocks," *Econometrica*, 76(3), 493–540.

Athey, S., K. Bagwell, and C. Sanchirico (2004): "Collusion and price rigidity," *Review of Economic Studies*, 71(2), 317–349.

Athey, S., and I. Segal (2007): "An Efficient Dynamic Mechanism," Discussion paper, Harvard University.

Aumann, R., and M. Maschler (1995): *Repeated Games with Incomplete Information*. The MIT press.

Bergemann, D., and J. Välimäki (2007): "Dynamic Marginal Contribution Mechanism," Discussion paper, Cowles Foundation.

Billingsley, P. (1961): "Statistical Methods in Markov Chains," *The Annals of Mathematical Statistics*, 32, 12–40.

Blackwell, D. (1956): "An Analog of the Minimax Theorem for Vector Payoffs," *Pacific Journal of Mathematics*, 6(1), 1–8.

Cole, H., and N. Kocherlakota (2001): "Dynamic Games with Hidden Actions and Hidden States," *Journal of Economic Theory*, 98(1), 114–126.

Cremer, J., and R. P. McLean (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56(6), 1247–1257.

Dutta, P. (1995): "A Folk Theorem for Stochastic Games," *Journal of Economic Theory*, 66(1), 1–32.

Fernandes, A., and C. Phelan (2000): "A Recursive Formulation for Repeated Agency with History Dependence," *Journal of Economic Theory*, 91(2), 223–247.

Fudenberg, D., and D. Levine (1983): "Subgame Perfect Equilibria of Finite and Infinite Horizon Games," *Journal of Economic Theory*, 31(2), 251–268.

Fudenberg, D., D. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62(5), 997–997.

Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54(3), 533–554.

Fudenberg, D., and Y. Yamamoto (2009a): "Learning from Private Information in Noisy Repeated Games," Discussion paper, Department of Economics, Harvard University.

——— (2009b): "The Folk Theorem for Irreducble Stochastic Games with Imperfect Public Monitoring," Discussion paper, Department of Economics, Harvard University.

Gossner, O. (1995): "The Folk Theorem for Finitely Repeated Games with Mixed Strategies," *International Journal of Game Theory*, 24(1), 95–107.

Hauser, C., and H. A. Hopenhayn (2008): "Trading Favors: Optimal Exchange and Forgiveness," Discussion paper, Department of Economics, UCLA, and Collegio Carlo Alberto.

Hörner, J., and J. Jamison (2007): "Collusion with (Almost) No Information," *The RAND Journal of Economics*, 38(3), 804–822.

Hörner, J., and S. Lovo (2009): "Belief Free Equilibria in Games with Incomplete Information," *Econometrica*, 77(2), 453–487.

Hörner, J., S. Lovo, and T. Tomala (2009): "Belief-Free Equilibria in Games with Incomplete Information: Characterization and Existence," Discussion paper, Cowles Foundation.

Hörner, J., T. Sugaya, S. Takahashi, and N. Vielle (2010): "Recursive Methods in Discounted Stochastic Games: An Algorithm for $\delta \to 1$ and a Folk Theorem," Discussion paper, Cowles Foundation.

Jackson, M. O., and H. F. Sonnenschein (2007): "Overcoming Incentive Constraints by Linking Decisions," *Econometrica*, 75(1), 241–258.

Jehiel, P., and B. Moldovanu (2001): "Efficient Design with Interdependent Valuations," *Econometrica*, 69(5), 1237–59.

Kreps, D., and R. Wilson (1982): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253–279.

Levin, J. (2003): "Relational Incentive Contracts," *American Economic Review*, 93(3), 835–857.

Maggi, G., and M. Morelli (2006): "Self-Enforcing Voting in International Organizations," *American Economic Review*, 96(4), 1137–1158.

Mailath, G., and L. Samuelson (2006): *Repeated Games and Reputations*. Oxford University Press, Oxford.

Massart, P. (1990): "The Tight Constant in The Dvoretzky-Kiefer-Wolfowitz Inequality," *The Annals of Probability*, 18, 1269–1283.

Milgrom, P., and J. Roberts (1982): "Predation, Reputation and Entry Deterrence," *Journal of Economic Theory*, 27.

Mobius, M. (2001): "Trading Favors," Discussion paper, Harvard University.

Myerson, R. B. (1986): "Multistage Games with Communication," *Econometrica*, 54(2), 323–58.

Myerson, R. B., and M. A. Satterthwaite (1983): "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29(2), 265–281.

Norris, J. (1997): *Markov Chains*. Cambridge University Press.

Peski, M. (2008): "Repeated Games with Incomplete Information on One Side," *Theoretical Economics*, 3(1), 29–84.

Radner, R. (1981): "Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship," *Econometrica*, 49(5), 1127–1148.

Radner, R. (1985): "Repeated Principal-Agent Games with Discounting," *Econometrica*, 53(5), 1173–1198.

Rockafellar, R. T. (1970): *Convex Analysis*. Princeton University Press.

Skrzypacz, A., and H. Hopenhayn (2004): "Tacit Collusion in Repeated Auctions," *Journal of Economic Theory*, 114(1), 153–169.

Townsend, R. (1982): "Optimal Multi-Period Contracts and the Gain from Enduring Relationships under Private Information," *Journal of Political Economy*, 90(6), 1166–1186.

Watson, J. (2002): "Starting Small and Commitment," *Games and Economic Behavior*, 38(1), 176–199.

Center for Applied Economics, Department of Industrial Engineering, University of Chile, Santiago, Chile

Department of Economics, Massachusetts Institute of Technology, Cambridge, MA, USA