Truthful Equilibria in Dynamic Bayesian Games

Johannes Hörner[*], Satoru Takahashi[†] and Nicolas Vieille[‡]

May 12, 2013

Abstract

This paper characterizes an equilibrium payoff subset for Markovian games with private information as discounting vanishes. Monitoring might be imperfect, transitions depend on actions, types correlated or not, values private or interdependent. It focuses on equilibria in which players report their information truthfully. This characterization generalizes those for repeated games, and reduces to a collection of one-shot Bayesian games with transfers. With independent private values, the restriction to truthful equilibria is shown to be without loss, except for individual rationality; in the case of correlated types, results from static mechanism design can be applied, resulting in a folk theorem.

Keywords: Bayesian games, repeated games, folk theorem.

JEL codes: C72, C73

1 Introduction

This paper studies the asymptotic equilibrium payoff set of repeated Bayesian games. In doing so, it generalizes methods that were developed for repeated games (Fudenberg and Levine, 1994; hereafter, FL) and later extended to stochastic games (Hörner, Sugaya, Takahashi and Vieille, 2011, hereafter HSTV).

Serial correlation in the payoff-relevant private information (or *type*) of a player makes the analysis of such repeated games difficult. Therefore, asymptotic results in this literature

[*]Yale University, 30 Hillhouse Ave., New Haven, CT 06520, USA, johannes.horner@yale.edu.

[†]National University of Singapore, ecsst@nus.edu.sg.

[‡]HEC Paris, 78351 Jouy-en-Josas, France, vieille@hec.fr.

have been obtained by means of increasingly elaborate constructions, starting with Athey and Bagwell (2008) and culminating with Escobar and Toikka (2013).[1] These constructions are difficult to extend beyond a certain point, however; instead, our methods allow us to deal with

- moral hazard (imperfect monitoring);
- endogenous serial correlation (actions affecting transitions);
- correlated types (across players) or/and interdependent values.

Allowing for such features is not merely of theoretical interest. There are many applications in which some if not all of them are relevant. In insurance markets, for instance, there is clearly persistent adverse selection (risk types), moral hazard (accidents and claims having a stochastic component), interdependent values, action-dependent transitions (risk-reducing behaviors) and, in the case of systemic risk, correlated types. The same holds true in financial asset management, and in many other applications of such models (taste or endowment shocks, etc.)

We assume that the state profile –each coordinate of which is private information to a player– follows a controlled autonomous irreducible Markov chain. (Irreducibility refers to its behavior under any fixed Markov strategy.) In the stage game, players privately take actions, and then a public signal realizes (whose distribution may depend both on the state and action profile) and the next period state profile is drawn. Cheap-talk communication is allowed, in the form of a public message at the beginning of each round. Our focus is on *truthful* equilibria, in which players truthfully reveal their type at the beginning of each period, after every history.

Our main result characterizes a subset of the limit set of equilibrium payoffs as $\delta \to 1$. While the focus on truth-telling equilibria is restrictive in the absence of any commitment, it nevertheless turns out that this limit set generalizes the payoffs obtained in all known special cases so far –with the exception of the lowest equilibrium payoff in Renault, Solan and Vieille, who also characterize Pareto-inferior "babbling" equilibria. When types are independent (though still possibly affected by one's own action), and payoffs are private,

[1]This not to say that the recursive formulations of Abreu, Pearce and Stacchetti (1990, hereafter APS) cannot be adapted to such games. See, for instance, Cole and Kocherlakota (2001), Fernandes and Phelan (2000), or Doepke and Townsend (2006). These papers provide methods that are extremely useful for numerical purposes for a given discount rate, but provide little guidance regarding qualitative properties of the (asymptotic) equilibrium payoff set.

for instance, all Pareto-optimal payoffs that are individually rational (*i.e.*, dominate the stationary minmax payoff) are limit equilibrium payoffs. In fact, with the exception of individual rationality, which could be further refined, our result is actually a characterization of the limit set of equilibrium payoffs (in case monitoring satisfies the usual identifiability conditions). In this sense, this is a folk theorem. It shows in particular that, when actions do not affect transitions, and leaving aside the value of the best minmax payoff, transitions do not matter for the limit set, just the invariant distribution, as is rather intuitive. When types are correlated, then all feasible and individually rational payoffs can be obtained in the limit.

These findings mirror standard results from static mechanism design, *e.g.* those of Arrow (1979), d'Aspremont and Gérard-Varet (1979) for the independent case, and those of d'Aspremont, Crémer and Gérard-Varet (2003) in the correlated case. This should come as no surprise, as our characterization is a reduction from the repeated game to a (collection of) one-shot Bayesian game with transfers, to which the standard techniques can be adapted. If there is no incomplete information about types, this one-shot game collapses to the algorithm developed by Fudenberg and Levine (1994) to characterize public perfect equilibrium payoffs, used in Fudenberg, Levine and Maskin (1994, hereafter FLM) to establish a folk theorem under public monitoring.

This stands in contrast with the techniques based on review strategies (see Escobar and Toikka for instance) whose adaptation to incomplete information is inspired by the linking mechanism described in Fang and Norman (2006) and Jackson and Sonnenschein (2007). Our results imply that, as for repeated games with public monitoring, transferring continuation payoffs across players is a mechanism that is sufficiently powerful to dispense with explicit statistical tests. Of course, this mechanism requires that deviations in the players' announcements can be statistically distinguished, a property closely related to the budget-balance constraint from static mechanism design. Therefore, our sufficient conditions are reminiscent of conditions in this literature, such as the weak identifiability condition introduced by Kosenok and Severinov (2008).

While the characterization turns out to be a natural generalization of the one from repeated games with public monitoring, it still has several unexpected features, reflecting difficulties in the proof that are not present either in stochastic games with observable states.

Consider the case of independent types for instance. Note that the long-run (or asymptotic) payoff must be independent of the current state of a player, because this state is unobserved and the Markov chain is irreducible. Relative to a stochastic game with observable states, there is a collapse of dimensionality as $\delta \to 1$. Yet the "transient" component of the payoff, which depends on the state, must be taken into account: a player's incentives to take a given action depend on the action's impact on later states. This component must be taken into account, but it cannot be treated as an exogenous transfer. This stands in contrast with the standard technique used in repeated games (without persistent types): there, incentives are provided by continuation payoffs which, as players get patient, become arbitrarily large relative to the per-period rewards, so that, asymptotically, the continuation play can be summarized by a transfer. But with private, irreducible types, the differences in continuation payoffs across types of a given player do not become arbitrarily large relative to the flow payoff (they fade out exponentially fast) and so cannot be replaced by a (arbitrarily large) transfer.

So: the transient component cannot be ignored, although it cannot be exploited as a standard transfer. But, for a given transfer rule and a Markov strategy, this component is easy to compute, using the *average cost optimality equation* (ACOE) from dynamic programming. This equation converts the relative future benefits of taking a particular action, given the current state, into an additional per-period reward. So it can be taken into account, and since it cannot be exploited, incentives will be provided by transfers that are independent of the type (though not of the report). After all, this independence is precisely a feature of transfers in static mechanism design, and our exclusive reliance on this channel illustrates again the lack of linkage in our analysis. What requires considerable work, however, is to show how such type-independent transfers can get implemented, and why we can compute the transient component as if the equilibrium strategies were Markov, which they are not.

Games without commitment but with imperfectly persistent private types were first introduced in Athey and Bagwell (2008) in the context of Bertrand oligopoly with privately observed cost. Athey and Segal (2007, hereafter AS) allow for transfers and prove an efficiency result for ergodic Markov games with independent types. Their team balanced mechanism is closely related to a normalization that is applied to the transfers in one of our proofs in the case of independent private values.

There is also a literature on undiscounted zero-sum games with such a Markovian structure, see Renault (2006), which builds on ideas introduced in Aumann and Maschler (1995). Not surprisingly, the average cost optimality equation plays an important role in this literature as well. Because of the importance of such games for applications in industrial organization and macroeconomics (Green, 1987), there is an extensive literature on recursive formulations for fixed discount factors (Fernandes and Phelan, 1999; Cole and Kocherlakota, 2001; Doepke and Townsend, 2006). In game theory, recent progress has been made in the case in which the state is observed, see Fudenberg and Yamamoto (2012) and Hörner, Sugaya, Takahashi and Vieille (2011) for an asymptotic analysis, and Pęski and Wiseman (2012) for the case in which the time lag between consecutive moves goes to zero. There are some similarities in the techniques used, although incomplete information introduces significant complications.

More related are the papers by Escobar and Toikka (2013), already mentioned, Barron (2012) and Renault, Solan and Vieille (2013). All three papers assume that types are independent across players. Barron (2012) introduces imperfect monitoring in Escobar and Toikka, but restricts attention to the case of one informed player only. This is also the case in Renault, Solan and Vieille. This is the only paper that allows for interdependent values, although in the context of a very particular model, namely, a sender-receiver game with perfect monitoring. In none of these papers do transitions depend on actions.

2 The Model

We consider dynamic games with imperfectly persistent incomplete information. The stage game is as follows. The finite set of players is denoted I. Each player $i \in I$ has a finite set S^i of (private) states, and a finite set A^i of actions. The state $s^i \in S^i$ is private information. We denote by $S := \times_{i\in I} S^i$ and $A := \times_{i \in I} A^i$ the sets of state profiles and action profiles respectively.

In each stage $n \geq 1$, timing is as follows:

1. First, each player privately observes his own state (s^i_n);
2. Players simultaneously make reports $(m^i_n) \in M^i$, where M^i is a finite set to be defined. These reports are publicly observed;
3. The outcome of a public correlation device is observed. For concreteness, it is a draw from the uniform distribution on $[0, 1]$;[2]
4. Players independently choose actions $a^i_n \in A^i$. Actions taken are not observed;
5. A public signal $y_n \in Y$, a finite set, and the next state profile $s_{n+1} = (s^i_{n+1})_{i\in I}$ are drawn according to some joint distribution $p(\cdot \mid s_n, a_n) \in \Delta(S \times Y)$.

[2]We do not know how to dispense with it. But given that public communication is allowed, such a public randomization device is innocuous, as it can be replaced by jointly controlled lotteries.

Throughout, we assume that $p(s, y \mid \bar{s}, \bar{a}) > 0$ whenever $p(y \mid \bar{s}, \bar{a}) > 0$, for all $(\bar{s}, \bar{a}, s)$. This means that (i) the Markov chain (s_n) is irreducible, (ii) public signals, whose probability might depend on $(\bar{s}, \bar{a})$ do not allow players to rule out some type profiles s. This is consistent with perfect monitoring. Note that actions might affect transitions.[3] The irreducibility of the Markov chain is a strong assumption, ruling out among others the case of perfectly persistent types (see Aumann and Maschler, 1995; Athey and Bagwell, 2008).[4] Unfortunately, it is well known that the asymptotic analysis is very delicate without such an assumption (see Bewley and Kohlberg, 1976).

The stage-game payoff function is a function $g : S \times A^i \times Y \to \mathbf{R}^I$ and as usual we define the reward $r : S \times A \to \mathbf{R}^I$ as its expectation, $r(s, a) = \mathbf{E}[g(s, a^i, y) \mid a]$, a function whose domain is extended to mixed action profiles in $\Delta(A)$.

Given the sequence of realized rewards (r_n^i), player i's payoff in the dynamic game is given by

$$\sum (1 - \delta)\delta^{n-1} r_n^i,$$

where $\delta \in [0, 1)$ is common to all players. (Short-run players can be accommodated for, as will be discussed.)

The dynamic game also specifies an initial distribution $p_0 \in \Delta(S)$, which plays no role in the analysis, given the irreducibility assumption and the focus on equilibrium payoffs as $\delta \to 1$.

A special case of interest is *independent private values* (hereafter, IPV). This is the case in which (i) payoffs of a player only depend on his private state, not the others', that is, for all (i, s, a), $r^i(s, a) = r^i(s^i, a)$, (ii) conditional on the public signal y, types are independently distributed. A more precise definition is given in Section 6.

But we do not restrict attention to private values or independent types. In the case of interdependent values, it then matters whether players observe their payoffs or not. It is possible to accommodate privately observed payoffs: simply define a player's private state as including his last realized payoff. As we shall see, the reports of a player's opponents in the

[3] Accommodating observable (public) states, as modeled in stochastic games, requires minor adjustments. One way to model them is to append such states as a component to each player's private state, perfectly correlated across players.

[4] In fact, our results only require that it be unichain, *i.e.* that the Markov chain defined by any Markov strategy has no two disjoint closed sets. This is the standard assumption under which the distributions specified by the rows of the limiting matrix $\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} p(\cdot)^i$ are independent of the initial state; otherwise the average cost optimality equation that is used to analyze, say, the cooperative solution is no longer valid. The full support assumption on states (given a signal) is convenient to avoid specifying out-of-equilibrium beliefs, but note that deterring deviations would become even easier without.

next period are taken into account when evaluating the truthfulness of a player's report, so that one could build on the results of Mezzetti (2004, 2007) in static mechanism design with interdependent valuations. Given this possible interpretation of a private state, we assume that a player's private action, private state and the public signals and reports are all the information that is available to him.[5]

Monetary transfers are not allowed. We view the stage game as capturing all possible interactions among players, and there is no difficulty in interpreting some actions as monetary transfers. In this sense, rather than ruling out monetary transfers, what is assumed is limited liability.

The game defined above allows for public communication among players. In doing so, we follow most of the literature on such dynamic games, Athey and Bagwell (2001, 2008), Escobar and Toikka (2013), Renault, Solan and Vieille (2013), etc.[6] As in static Bayesian mechanism design, communication is necessary for coordination, and makes it possible to characterize what restrictions on behavior are driven by incentives.

But there is no commitment in the dynamic game. As a result, the revelation principle does not apply. As is well known (see Bester and Strausz, 2000, 2001), in the absence of commitment, it is not possible *a priori* to restrict attention to direct mechanisms, corresponding to the choice $M^i = S^i$ (or $M^i = (S^i)^2$, as explained below), let alone obedient or truthful behavior.

Yet this is precisely what we will do. The next section illustrates some of the issues that this raises.

3 Some Examples

Example 1—***A Silent Game.*** This game follows Renault (2006). This is a zero-sum two-player game in which player 1 has two private states, s^1 and $\hat{s}^1$, and player 2 has a single state, omitted. Player 1 has actions $A^1 = \{T, B\}$ and player 2 has actions $A^2 = \{L, R\}$. Player 1's reward is given by Figure 1. Recall that rewards are not observed. Both states s^1

[5]However, our notion of equilibrium is sensitive to what goes into a state: by enriching it, one weakly increases the equilibrium payoff set. For instance, one could also include in a player's state his previous realized action, which following Kandori (2003) is useful even when incomplete information is trivial and the game is simply a repeated game with public monitoring; this is peripheral to our objective and will not be pursued here.

[6]This is not to say that introducing a mediator would be without interest, to the contrary. Following Myerson (1986), we could then appeal to a revelation principle, though without commitment this would simply shift the inferential problem to the stage of recommendations.

$$\begin{array}{c|c|c|}
 & L & R \\ \hline
T & 1 & 0 \\ \hline
B & 0 & 0 \\ \hline
\end{array}
\qquad\qquad
\begin{array}{c|c|c|}
 & L & R \\ \hline
T & 0 & 0 \\ \hline
B & 0 & 1 \\ \hline
\end{array}$$

$$s^1 \qquad\qquad\qquad\qquad \hat{s}^1$$

Figure 1: Player 1's reward in Example 1

and $\hat{s}^1$ are equally likely in the initial period, and the transition is action-independent, with $p \in [1/2, 1)$ denoting the probability that the state remains the same from one stage to the next.

Let $M^1 = \{s^1, \hat{s}^1\}$, so that player 1 can disclose his state if he wishes to. Will he? By revealing the state, player 2 can secure a payoff of 0 by playing R or L depending on player 1's report. Yet player 1 can secure a payoff of 1/4 by choosing messages and actions at random. In fact, this is the (uniform) value of this game for $p = 1$ (Aumann and Maschler, 1995). When $p < 1$, player 1 can actually get more than this by trading off the higher expected reward from a given action with the information that it gives away. He has no interest in giving this information away for free through informative reports. Silence is called for.

Just because we may focus on the silent game does not mean that it is easy to solve. Its value for $p > 2/3$ is still unknown.[7] Because the optimal strategies depend on player 2's belief about player 1's state, the problem of solving for them is infinite-dimensional, and all that can be done is characterize its solution via some functional equation (see Hörner, Rosenberg, Solan and Vieille, 2010).

Non-existence of truthful equilibria in some games is no surprise. The ratchet effect that arises in bargaining and contracting is another manifestation of the tension between truth-telling and lack of commitment (see Freixas, Guesnerie and Tirole, 1985). What Example 1 illustrates is that small message sets are just as difficult to deal with as larger ones. When players hide their information, behavior reflect their private beliefs, which calls for a state space as large as it gets.

The surprise, then, is that the literature on Markovian games (Athey and Bagwell (2001, 2008), Escobar and Toikka (2013), Renault, Solan and Vieille (2013)) manages to get positive results at all: in most games, efficiency requires coordination, and thus disclosure of private information. As will be clear from Section 6, existence is much easier to obtain in the IPV environment, the focus of most of these papers.

[7] It is known for $p \in [1/2, 2/3]$ and some specific values. Pęski and Toikka (private communication) have recently shown that this value is decreasing in p.

Example 2—*A Game that Leaves No Player Indifferent.* Player 1 has two private states, s^1 and $\hat{s}^1$, and player 2 has a single state, omitted. Player 1 has actions $A^1 = \{T, B\}$ and player 2 has actions $A^2 = \{L, R\}$. Rewards are given by Figure 2 (values are private).

	L	R
T	$1, 1$	$1, -1$
B	$0, -1$	$0, 1$

s^1

	L	R
T	$0, 1$	$0, -1$
B	$1, -1$	$1, 1$

$\hat{s}^1$

Figure 2: A two-player game in which the mixed minimax payoff cannot be achieved.

The two types s^1 and $\hat{s}^1$ are i.i.d. over time and equally likely. Monitoring is perfect. To minmax player 2, player 1 must randomize uniformly, independently of his type. Yet in any equilibrium in which player 1 always reports his type truthfully, there is no history after which he is indifferent between both actions, for both types simultaneously. To play B when his type is s^1, or T when his type is $\hat{s}^1$, he must be compensated by \$1 in continuation utility. But then he has an incentive to report his type incorrectly, to pocket this promised utility while playing his favorite action.[8]

This still leaves open the possibility of a player randomizing for *one* of his types. This is very useful when each player has only one type, like in a standard repeated game,as it then delivers the usual mixed minimax payoff. But except in this particular case, it will only introduce some intermediate notion of minimax payoff that has little intrinsic interest. Our focus, then, will be on strict equilibria.

Example 3—*Waiting for Evidence.* There are two players. Player 1 has $K+1$ types, $S^1 = \{0, 1, \ldots, K\}$, while player 2 has only two types, $S^2 = \{0, 1\}$. Transitions do not depend

[8] To see this formally, fix such a history, and consider the continuation payoff of player 1, V^1, which we index by the announcement and action played. Note that this continuation payoff, for a given pair of announcement and action, must be independent of player 1's current type. Suppose that player 1 is indifferent between both actions whether his type is s^1 or $\hat{s}^1$. If his type is s^1, we must then have

$$(1-\delta) + \delta V^1(s^1, T) = \delta V^1(s^1, B) \geq \max\{(1-\delta) + \delta V^1({}^1\hat{s}^1, T), \delta V^1(\hat{s}^1, B)\},$$

which implies that $(1-\delta) + \delta V^1(s^1, T) + \delta V^1(s^1, B) \geq (1-\delta) + \delta V^1(\hat{s}^1, T) + \delta V^1(\hat{s}^1, B)$, or $V^1(s^1, T) + V^1(s^1, B) \geq V(\hat{s}^1, T) + V(\hat{s}^1, B)$. The constraints for type $\hat{s}^1$ imply the opposite inequality, so that $V^1(s^1, T) + V^1(s^1, B) = V(\hat{s}^1, T) + V(\hat{s}^1, B)$. Revisiting the constraints for type s^1, it follows that the inequality must hold with equality, and that $V^1(T) := V^1(s^1, T) = V^1(\hat{s}^1, T)$, and $V^1(B) := V(s^1, B) = V^1(\hat{s}^1, B)$. The two indifference conditions then give $\frac{1-\delta}{\delta} = V^1(B) - V^1(S) = -\frac{1-\delta}{\delta}$, a contradiction.

on actions (ignored), and are as follows. If $s_n^1 = k > 0$, then $s_n^2 = 0$ and $s_{n+1}^1 = s_n^1 - 1$. If $s_n^1 = 0$, then $s_n^2 = 1$ and s_{n+1}^1 is drawn randomly (and uniformly) from S^1. In words, s_n^1 stands for the number of stages until the next occurrence of $s^2 = 1$. By waiting no more than K periods, all reports by player 1 can be verified.

This example makes two closely related point. First, in order for player $-i$ to statistically discriminate between player i's states, it is not necessary that his set of signals (here, states) be as rich as player i's, unlike in static mechanism design with correlated types (the familiar "spanning condition" of Crémer and McLean, generically satisfied only if $|S^{-i}| \geq |S^i|$). Two states for one player can be enough to cross-check the reports of an opponent with many more states, provided that states in later rounds are informative enough.

Second, the long-term dependence of the stochastic process implies that one player's report should not always be evaluated on the fly. It is better to hold off until more evidence is collected. Note that this is not the same kind of delay as the one that makes review strategies effective, which take advantage of the central limit theorem to devise powerful tests even when signals are independent over time (see Radner, 1986; Fang and Norman, 2006, and Jackson and Sonnenschein, 2007). Here, it is precisely because of the dependence that waiting is useful.

This raises an interesting statistical question: does the tail of the sequence of private states of player $-i$ contain indispensable information in evaluating the truthfulness of player i's report in a given round, or is the distribution of this sequence, conditional on (s_n^i, s_{n-1}), summarized by the distribution of an initial segment? This question appears to be open in general. In the case of transitions that do not depend on actions, it has been raised by Blackwell and Koopmans (1957) and answered by Gilbert (1959): it is enough to consider the next $2|S^i| + 1$ values of the sequence $(s_{n'}^{-i})_{n' \geq n}$.[9]

At the very least, when types are correlated and the Markov chain exhibits time dependence, it is useful to condition player i's continuation payoff given his report s_n^i on $-i$'s next private state, s_{n+1}^{-i}. Because this turns out to suffice to obtain sufficient conditions analogous to those in the static case, we will limit ourselves to this conditioning.[10]

[9]The reporting strategy defines a hidden Markov chain on pairs of states, messages and signals that induces a stationary process over messages and signals; Gilbert assumes that the hidden Markov chain is irreducible and aperiodic, which here need not be (with truthful reporting, the message is equal to the state), but his result continues to hold when these assumptions are dropped, see for instance Dharmadhikari (1963).

[10]See Obara (2008) for some of the difficulties encountered in dynamic settings while attempting to extend results from static mechanism design with correlated types.

4 Truthful Equilibria

A public history at the start of round n is a sequence $h_n = (m_1, y_1, \ldots, m_{n-1}, y_{n-1}) \in H_n := (M \times Y)^{n-1}$. Player i's private history at the start of round n is a sequence $h_n^i = (s_1^i, m_1, a_1^i, y_1, \ldots, s_{n-1}^i, m_{n-1}, a_{n-1}^i, y_{n-1}) \in H_n^i := (S^i \times M \times A^i \times Y)^{n-1}$. A (behavior) strategy for player i is a pair of sequences $(\mathfrak{m}^i, \mathfrak{a}^i) = (\mathfrak{m}_n^i, \mathfrak{a}_n^i)_{n \in \mathbb{N}}$ with $\mathfrak{m}_n^i : H_n^i \times S^i \to \Delta(M^i)$, and $\mathfrak{a}_n^i : H_n^i \times S^i \times M \to \Delta(A^i)$ that specify i's message and action as a function of his private information, his current state and the message profile in the current period.[11] A strategy profile $(\mathfrak{m}, \mathfrak{a})$ defines a distribution over histories in the usual way, and we consider the perfect Bayesian equilibria of this game.

A special class of games are "standard" repeated games with public monitoring, in which S^i is a singleton set for each player i and we can ignore the $\mathfrak{m}$-component of players' strategies. For such games, Fudenberg and Levine provide a convenient algorithm to describe and study a subset of equilibrium payoff–public perfect equilibrium payoffs. A public perfect equilibrium (PPE) is an equilibrium in which players' strategies are public; that is, $\mathfrak{a}$ is measurable with respect to H_n, so that player i ignores any additional private information (his past actions). Their characterization of the asymptotic set of PPE payoffs as $\delta \to 1$ relies on the notion of a *score* defined as follows.

Definition 1 *Fix $\lambda \in \mathbf{R}^I$. Let*

$$k(\lambda) = \sup_{v,x,\alpha} \lambda \cdot v,$$

where the supremum is taken over all $v \in \mathbf{R}^I$, $x : Y \to \mathbf{R}^I$ and $\alpha \in \times_{i \in I} \Delta(A^i)$ such that

(i) *α is a Nash equilibrium with payoff v of the game with payoff $r(a) + \sum_y p(y \mid a) x(y)$;*

(ii) *For all $y \in Y$, it holds that $\lambda \cdot x(y) \leq 0$.*

Let $\mathcal{H} = \bigcap_{\lambda \in \mathbf{R}^I} \{v \in \mathbf{R}^I \mid \lambda \cdot v \leq k(\lambda)\}$. FL prove the following.

Theorem 1 (FL) *It holds that $E(\delta) \subseteq \mathcal{H}$ for any $\delta < 1$; moreover, if $\mathcal{H}$ has non-empty interior, then $\lim_{\delta \to 1} E(\delta) = \mathcal{H}$.*

Our purpose is to obtain a similar characterization for the broader class of games considered here. To do so while preserving the recursive nature of the equilibrium payoff set that will be described leads us to focus on a particular class of equilibria in which players report truthfully their private state in every round, on and off path.

[11] Recall however that a public correlation device is assumed, although it is omitted from the notations.

The complete information game with transfers x that appears in the definition of the score must be replaced with a two-stage Bayesian game with communication, formally defined in the next section. To describe a Bayesian game, one needs a type space and a prior. Clearly, in the dynamic game, player i's beliefs about his opponents' private states depend on the previous reports s_{n-1}. Hence, we are led to consider a family of Bayesian games, parametrized by such a report. In addition, as Example 3 made clear, player i's transfer x not only depends on this parameter, but also on s_{n+1}^{-i}, the reports of the other players in the following period.

What is a player's type, *i.e.*, what is player's information that is payoff-relevant in round n? Certainly this includes his private state s_n^i. Because player i does not observe s_n^{-i}, his conditional belief about these states is also payoff-relevant, to predict $-i$'s behavior and because values need not be private. This is what creates the difficulty in Example 1: because player 1 does not want to disclose his state, player 2 must use all available information to make the best prediction, which is the entire history of play.

However, when players report truthfully their information, player i knows s_{n-1}^{-i}; to predict s_n^{-i}, this is a sufficient statistic for the entire history of player i, given (s_{n-1}^i, s_n^i). Note that s_{n-1}^i matters, because the Markov chains (s_n^i) and (s_{n-1}^i) need not be independent across players, and s_n need not be independent of s_{n-1} either.

Off path, these conditional beliefs about s_n^{-i} are "private," as they depend on his previous type s_{n-1}^i (as well as on a_{n-1}^i). The natural choice is then $M^i = (S^i)^2$ (or even $(S^i)^2 \times A^i$), so that player i be able to report all his private information. Along the equilibrium path, this involves repetitions. But it matters when $m_{n-1}^i \neq s_{n-1}^i$. Players $-i$ cannot detect such a deviation, which is "on-schedule" according to Athey and Bagwell (2008). For truthful reporting off path, the choice of M^i makes a difference: by setting $M^i = S^i$, player i is asked to tell the truth regarding his payoff-type, but to lie about his belief-type (which will be incorrectly believed to be determined by his report of s_{n-1}^i, along with his current report).

In the IPV case, however, there is no need for this, as the past deviation does not affect i's conditional beliefs, and we will then proceed as if $M^i = S^i$.

A strategy $(\mathfrak{m}^i, \mathfrak{a}^i)$ is public and truthful if $\mathfrak{m}_n^i(h_n^i, s_n^i) = s_n^i$ (or (s_{n-1}^i, s_n^i)) for all histories h_n^i, and $\mathfrak{a}^i(h_n^i, s_n^i, m_n)$ depends on (h_n, s_n^i, m_n) only. The solution concept is perfect Bayesian equilibrium in public and truthful strategies. Because type, action and signal sets are finite, and given our "full-support" assumption on S, there is no difficulty in adapting Fudenberg and Tirole (1991a and b)'s definition to our set-up –the only issue that could arise is due to the fact that we have not imposed full support on the public signals. In our proofs, actions do not lead to further updating on beliefs, conditional on the reports.

The next section describes formally the family of Bayesian games.

5 Characterization

5.1 The main theorem

In this section, $M = S \times S$. Messages are written $m = (m_p, m_c)$, where m_p (resp. m_c) are interpreted as reports on previous (resp. current) states.

We set $\Omega_{\text{pub}} := M \times Y$, and we refer to the pair (m_n, y_n) as the *public outcome* of stage n. This is the additional public information available at the end of stage n. We also refer to (s_n, m_n, a_n, y_n) as the outcome of stage n, and denote by $\Omega := \Omega_{\text{pub}} \times S \times A$ the set of possible outcomes in any given stage.

5.1.1 The ACOE

Our analysis makes use of the so-called ACOE, which plays an important role in dynamic programming. For completeness, we provide here an elementary and incomplete statement, that is sufficient for our purpose and we refer to Puterman (1994) for details.

Let be given an irreducible (or unichain) transition function q over the finite set S with invariant measure μ, and a payoff function $u : S \to \mathbf{R}$. Assume that the states (s_n) follow a Markov chain with transition function q and that a decision maker receives the payoff $u(s_n)$ in stage n. The long-run payoff of the decision maker is $v = \mathbf{E}_\mu[u(s)]$. While this long-run payoff is independent of the initial state, discounted payoffs are not. Lemma 1 below provides a normalized measure of the differences in discounted payoffs, for different initial states.

Lemma 1 *There is $\theta : S \to \mathbf{R}$ such that*

$$v + \theta(s) = u(s) + \mathbf{E}_{t \sim p(\cdot|s)}\theta(t).$$

The map θ is unique, up to an additive constant. It admits an intuitive interpretation in terms of discounted payoffs. Given $\delta < 1$, denote by $\gamma_\delta(s)$ the discounted payoff when starting for s. Then the difference $\theta(s) - \theta(s')$ is equal to $\lim_{\delta \to 1} \dfrac{\gamma_\delta(s) - \gamma_\delta(s')}{1-\delta}$. For this reason, we call θ the (vector of) *relative rents*.

5.1.2 Admissible contracts

The characterization of FLM for repeated games involves a family of optimization problems. One optimizes over pairs (α, x) where α is an equilibrium in the underlying stage game augmented with transfers.

Because we insist on truthful equilibria, and because we need to incorporate the dynamic effects of actions upon types, we need to consider instead action plans and transfers (ρ, x), such that reporting truthfully and playing $\rho : \Omega_{\text{pub}} \times S \to A$ constitutes a *stationary* equilibrium of the dynamic game augmented with transfers.

When players report truthfully and choose actions according to ρ, the sequence (ω_n) of outcomes is a Markov chain, and so does the sequence $(\tilde{\omega}_n)$, where $\tilde{\omega}_n = (\omega_{\text{pub},n-1}, s_n)$, with transition function denoted π_ρ. By the irreducibility assumption on p, the Markov chain has a unique invariant measure $\mu[\rho] \in \Delta(\Omega_{\text{pub}} \times S)$.

Given transfers $x : \Omega_{\text{pub}} \times S \times (Y \times S) \to \mathbf{R}^I$, we denote by $\theta[\rho, r + x] : \Omega_{\text{pub}} \times S \to \mathbf{R}^I$ the associated relative rents, which are obtained when applying Lemma 1 to the latter chain (and to all players).

Thanks to the ACOE, the above condition that reporting truthfully and playing ρ be a stationary equilibrium of the dynamic game with stage payoffs $r+x$ is equivalent to requiring that, for each $\bar{\omega}_{\text{pub}} \in \Omega_{\text{pub}}$, reporting truthfully and playing ρ is an equilibrium in a one-shot Bayesian game. In this Bayesian game, types s are drawn according to p (given $\bar{\omega}_{\text{pub}}$), players submit reports m, then choose actions a, and obtain the (random) payoff

$$r(s, a) + x(\bar{\omega}_{\text{pub}}, m) + \theta[\rho, r + x](\omega_{\text{pub}}, t),$$

where (y, t) are chosen according to $p(\cdot \mid s, a)$ and the public outcome ω_{pub} is the pair (m, y).

Thus, the map θ provides a "one-shot" measure of the relative value of being in a given state; with persistent and possibly action-dependent transitions, this measure is essential in converting the dynamic game into a one-shot game, just as the invariant measure $\mu[\rho]$. Both μ and θ are defined by a finite system of equations, as it is the most natural way of introducing them. But in the ergodic case that we are concerned with explicit formulas exist for both of them (see, for instance, Iosifescu, 1980, p.123, for the invariant distribution; and Puterman, 1994, Appendix A for the relative rents).

Because we insist on off-path truth-telling, we need to consider arbitrary private histories of the players, and the formal definition of *admissible contracts* (ρ, x) is therefore more involved. Fix a player i. Given any private history of player i, the belief of player i over future plays only depends on the previous public outcome $\bar{\omega}_{\text{pub}}$, and on the report $\bar{m}^i$ and action $\bar{a}^i$ of player i in the previous stage. Given such a $(\bar{\omega}_{\text{pub}}, \bar{m}^i, \bar{a}^i)$, let $D^i(\bar{\omega}_{\text{pub}}, \bar{m}^i, \bar{a}^i)$ denote the two-step decision problem in which

Step 1 $s \in S$ is drawn according to the belief held by player i;[12] player i is informed of s^i, then submits a report $m^i \in M^i$;

[12]Given $\bar{\omega}_{\text{pub}}$, player i assigns probability 1 to $\bar{s}^{-i} = \bar{m}^i_c$, and to previous actions being $\bar{a}^{-i} = \rho^{-i}(\bar{m}^{-i}, \bar{m}^i)$;

Step 2 player i learns s^{-i} and then chooses an action $a^i \in A^i$. The payoff to player i is given by

$$r^i(s,a) + x^i(\bar{\omega}_{\text{pub}}, \omega_{\text{pub}}, t^{-i}) + \theta^i[\rho, r+x](\omega_{\text{pub}}, t), \tag{1}$$

where $m^{-i} = (\bar{m}_c^{-i}, s^{-i})$, $a^{-i} = \rho^{-i}(\bar{\omega}_{\text{pub}}, m)$, the pair (y,t) is drawn according to $p(\cdot \mid s, a)$, and $\omega_{\text{pub}} := (m, y)$.

We denote by $\mathcal{D}^i_{\rho,x}$ the collection of decision problems $D^i_{\rho,x}(\bar{\omega}_{\text{pub}}, \bar{m}^i, \bar{a}^i)$.

Definition 2 *The pair (ρ, x) is* admissible *if all optimal strategies of player i in $\mathcal{D}^i$ report truthfully $m^i = (\bar{s}^i, s^i)$ in Step 1, and then (after reporting truthfully) choose the action $\rho^i(\bar{\omega}_{\text{pub}}, m)$ prescribed by ρ in Step 2.*

Some comments are in place. The condition that ρ be played once types have been reported truthfully simply means that, for each $\bar{\omega}_{pub}$ and $m = (\bar{s}, s)$ such that $\bar{m}_c = \bar{s}$, the action profile $\rho(\bar{\omega}_{\text{pub}}, m)$ is a strict equilibrium in the complete information one-shot game with payoff function $r(s,a) + x(\bar{\omega}_{\text{pub}}, (m,y)) + \theta((m,y),t)$.

The truth-telling condition is slightly more delicate to interpret. Consider first an outcome $\bar{\omega} \in \Omega$ such that $\bar{s}^i = \bar{m}^i_c$ for each i – no played lied in the previous stage. Given such an outcome, all players share the same belief over next types, $\dfrac{p(\cdot, \bar{y} \mid \bar{s}, \bar{a})}{p(\bar{y} \mid \bar{s}, \bar{a})}$. Consider the Bayesian game in which $s \in S$ is drawn according to the latter distribution, players make public reports m then choose actions a, and get the payoff $r(s,a) + x(\bar{\omega}_{\text{pub}}, (m,y)) + \theta((m,y),t)$. The admissibility condition for such an outcome $\bar{\omega}$ is equivalent to requiring that truth-telling followed by ρ is a "strict" equilibrium of this Bayesian game.[13] The admissibility requirement is however more demanding, in that it requires in addition truth-telling to be optimal for player i at any outcome $\bar{\omega}$ such that $\bar{s}^j = \bar{m}^j_c$ for $j \neq i$, but $\bar{s}^i \neq \bar{m}^i_c$. At such an $\bar{\omega}$, players do not share the same belief over the next types, and it is inconvenient to state the admissibility requirement by means of an auxiliary, subjective, Bayesian game.

In loose terms, truth-telling followed by ρ^i is the *unique* best-reply of player i to truth-telling and ρ^{-i}. Note that we require truth-telling to be optimal ($m^i = (\bar{s}^i, s^i)$) even if player i has lied in the previous stage ($\bar{m}^i_c \neq \bar{s}^i$ on his current state). On the other hand, Definition 2 puts no restriction on player i's behavior if he lies in step 1 ($m^i \neq (\bar{s}^i, s^i)$). The

---

hence this belief assigns to $s \in S$ the probability $\dfrac{p(s, \bar{y} \mid \bar{s}, \bar{a})}{p(\bar{y} \mid \bar{s}, \bar{a})}$.

[13]Quotation marks are needed, since we have not defined off-path behavior. What we mean is that any on-path deviation leads to a lower payoff.

second part of Definition 2 is equivalent to saying that $\rho^i(\bar{\omega}_{\text{pub}}, m)$ is the unique best-reply to $\rho^{-i}(\bar{\omega}_{\text{pub}}, m)$ in the complete information game with payoff function given by (26) when $m = (\bar{s}, s)$.

We denote by $\mathcal{C}_0$ the set of admissible pairs (ρ, x).

5.1.3 The characterization

Let S_1 denote the unit sphere of $\mathbf{R}^I$. For a given system of weights $\lambda \in S_1$, we denote by $\mathcal{P}_0(\lambda)$ the optimization program $\sup \lambda \cdot v$, where the supremum is taken over (v, ρ, x) such that

- $(\rho, x) \in \mathcal{C}_0$;
- $\lambda \cdot x(\cdot) \leq 0$;
- $v = \mathbf{E}_{\mu[\rho]}\left[r(s,a) + x(\bar{\omega}_{\text{pub}}, \omega_{\text{pub}}, t)\right]$ is the long-run payoff induced by (ρ, x).

We denote by $k_0(\lambda)$ the value of $\mathcal{P}_0(\lambda)$ and set $\mathcal{H}_0 := \{v \in \mathbf{R}^I, \lambda{\cdot}v \leq k_0(\lambda) \text{ for all } \lambda \in S_1\}$.

Theorem 2 *Assume that $\mathcal{H}_0$ has non-empty interior. Then it is included in the limit set of truthful equilibrium payoffs.*

This result is simple enough. Yet, the strict incentive properties required for admissible contracts make it useless in some cases. As an illustration, assume that successive states are independent across stages, so that $p(t, y \mid s, a) = p_1(t)p_2(y \mid s, a)$, and let the current state of i be s^i. Plainly, if player i prefers reporting $(\bar{s}^i, s^i)$ rather than $(\tilde{s}^i, s^i)$ when his previous state was $\bar{s}^i$, then he still prefers reporting $(\bar{s}^i, s^i)$ when his previous state was $\tilde{s}^i$. So there are no admissible contracts! The same issue arises when successive states are independent across players, or in the spurious case where two states of i are identical in all relevant dimensions.

In other words, the previous theorem is well-adapted (as we show later) to setups with correlated and persistent states, but we now need a variant to cover all cases.

This variant is parametrized by report maps $\phi^i : S^i \times S^i \to M^i (= S^i \times S^i)$, with the interpretation that $\phi^i(\bar{s}^i, s^i)$ is the equilibrium report of player i when his previous and current states are $(\bar{s}^i, s^i)$.

Given transfers $x^i : \Omega_{\text{pub}} \times \Omega_{\text{pub}} \times M^{-i} \to \mathbf{R}$, the decision problems $D^i(\bar{\omega})$ are defined as previously. The set $\mathcal{C}_1(\phi)$ of ϕ-admissible contracts is the set of pairs (ρ, x), with $\rho : \Omega_{\text{pub}} \times M \to A$, such that all optimal strategies of player i in $D^i(\bar{\omega})$ report $m^i = \phi^i(\bar{s}^i, s^i)$ in Step 1, and then choose the action $\rho^i(\bar{\omega}_{\text{pub}}, m)$ in Step 2.

We denote by $\mathcal{P}_1(\lambda)$ the optimization problem deduced from $\mathcal{P}_0(\lambda)$ when replacing the constraint $(\rho, x) \in \mathcal{C}_0$ by the constraint $(\rho, x) \in \mathcal{C}_1(\phi)$ for some ϕ.

Set $\mathcal{H}_1 := \{v \in \mathbf{R}^I, \lambda \cdot v \leq k_1(\lambda)\}$ where $k_1(\lambda)$ is the value of $\mathcal{P}_1(\lambda)$.

Theorem 3 generalizes Theorem 2.

Theorem 3 *Assume that $\mathcal{H}_1$ has a non-empty interior. Then it is included in the limit set of truthful and pure perfect Bayesian equilibrium payoffs.*

To be clear, there is no reason to expect Theorem 3 to provide a characterization of the entire limit set of truthful equilibrium payoffs. One might hope to achieve a bigger set of payoffs by employing finer statistical tests (using the serial correlation in states), just as one can achieve a bigger set of equilibrium payoffs in repeated games than the set of PPE payoffs, by considering statistical tests (and private strategies). There is an obvious cost in terms of the simplicity of the characterization. As it turns out, ours is sufficient to obtain all the equilibrium payoffs known in special cases, and more generally, all individually rational Bayes Nash equilibrium payoffs (including the Pareto frontier) under independent private values, as well as a folk theorem under correlated values.[14]

The definition of ϕ-admissible contracts allows for the case where different types are collapsed into a single report. At first sight, this may appear to be inconsistent with the notion of truthful equilibrium. Yet, this is not so. Indeed, observe that, when properly modifying the definition of ρ following messages that are not in the range of ρ, and of x as well, the adjusted pair $(\tilde{\rho}, \tilde{x})$ satisfies the condition of Definition 2, except that truth-telling inequalities may no longer be strict. However, equality may hold only between those types that were merged into the same report, in which case the actual report has no behavioral impact in the current stage.

5.2 Proof overview

We here explain the main ideas of the proof. For simplicity, we assume perfect monitoring, action-independent transitions, and we focus on the proof of Theorem 2. For notational simplicity also, we limit ourselves to admissible contracts (ρ, x) such that the action plan

[14] Besides, an exact characterization would require an analysis in $\mathbf{R}^{I \times S}$, mapping each type profile into a payoff for each player. When the players' types follow independent Markov chains and values are private, this makes no difference, as the players' limit equilibrium payoff *must* be independent of the initial type profile, given irreducibility and incentive-compatibility. But when types are correlated, it is possible to assign different (to be clear, long-run) equilibrium payoffs to a given player, as a function on the initial state.

$\rho : M \to A$ only depends on current reports, and transfers $x : M \times M \times A \to \mathbf{R}^I$ does not depend on previous public signals (which do not affect transitions here). This is not without loss of generality, but going to the general case is mostly a matter of notations.

Our proof is best viewed as an extension of the recursive approach of FLM to the case of persistent, private information. To serve as a benchmark, assume first that types are iid across stages, with law $\mu \in \Delta(S)$. The game is then truly a repeated game and the characterization of FLM applies. In that setup, and according to Definition 2, (ρ, x) is an admissible contract if for each $\bar{m}$, reporting truthfully then playing ρ, is an equilibrium in the Bayesian game with prior distribution μ, and payoff function $r(s,a) + x(\bar{m}, m, a)$ (and if the relevant incentive inequalities are strict).

In order to highlight the innovations of the present paper, we provide a quick reminder of the FLM proof (specialized to the present setup). We let Z be a smooth compact set in the interior of $\mathcal{H}$, and a discount factor $\delta < 1$. Given an initial target payoff vector $v \in Z$, (and $\bar{m} \in M$), one picks an appropriately chosen direction $\lambda \in \mathbf{R}^I$ in the player set[15] and we choose an admissible contract (ρ, x) such that (ρ, x, v) is feasible in $\mathcal{P}_0(\lambda)$. Players are required to report truthfully their type and to play (on path) according to ρ, and we define $w_{\bar{m},m,a} := v + \dfrac{1-\delta}{\delta} x(\bar{m}, m, a)$ for each $(m,a) \in M \times A$. Provided δ is large enough, the vectors $w_{\bar{m},m,a}$ belong to Z, and this construction can thus be iterated,[16] leading to a well-defined strategy profile σ in the repeated game. The expected payoff under σ is v, and the continuation payoff in stage 2, conditional on public history (m,a) is equal to $w_{\bar{m},m,a}$, when computed at the *ex ante* stage, before players learn their stage 2 type. The fact that (ρ, x) is admissible implies that σ yields an equilibrium in the one-shot game with payoff $(1-\delta) r(s,a) + \delta w_{\bar{m},m,a}$. A one-step deviation principle then applies, implying that σ is a Perfect Bayesian Equilibrium of the repeated game, with payoff v.

Assume now that the type profiles (s_n) follow an irreducible Markov chain with invariant measure μ. The proof outlined above fails as soon as types as auto-correlated. Indeed, the initial type of player i now provides information over types in stage 2. Hence, at the interim stage in stage 1, (using the above notations) the expected continuation payoffs of player i are no longer given by $w_{\bar{m},m,a}$. This is the rationale for including continuation private rents into the definition of admissible contracts.

But this presents us with a problem. In any recursive construction such as the one outlined above, continuation private rents (which help define current play) are defined by

[15] If v is a boundary point, λ is an outwards pointing normal to Z at v.

[16] with $w_{\bar{m},m,a}$ serving as the target payoff vector in the next, second, stage.

continuation play, which itself is based on current play, leading to an uninspiring circularity. On the other hand, our definition of an admissible contract (ρ, x) involves the private rents $\theta[\rho, x]$ induced by an indefinite play of (ρ, x). This difficulty is solved by adjusting the recursive construction in such as way that players always expect the current admissible contract (ρ, x) to be used in the foreseeable future. On the technical side, this is achieved by letting players stick to an admissible contract (ρ, x) during a random number of stages, with a geometric distribution of parameter η. The target vector is updated only when switching to a new direction (and to a new admissible contract). The date when to switch is determined by the correlation device. The parameter η is chosen large enough compared to $1-\delta$, ensuring that target payoffs always remain within the set Z. Yet, η is chosen small enough so that the continuation private rents be approximately equal to $\theta[\rho, x]$: in terms of private rents, it is almost as if (ρ, x) were used forever.

Equilibrium properties are derived from the observation that, by Definition 2, the incentive to report truthfully and then to play ρ would be strict if the continuation private rents were truly equal to $\theta[\rho, x]$ and thus, still holds when equality holds only approximately. All the details are provided in the Appendix.

6 Independent private values

This section applies Theorem 2 and Theorem 3 to the case of independent private values. Throughout the section, the following two assumptions (referred to as IPV) are maintained:

- The stage-game payoff function of i only depends only on his own state: for every i and (s, a^i, y), $g^i(s, a^i, y) = g^i(s^i, a^i, y)$.

- Transitions of player i's state only depend on his own state, his own action, conditional on the public signal: for every (s, a, y, t), $p(t, y \mid s, a) = (\times_j p(t^j \mid y, s^j, a^j))p(y \mid s, a)$.[17,18]

For now, we will restrict attention to the case in which

$$p(y \mid s, a) = p(y \mid \tilde{s}, a),$$

[17]Note that this does not imply that other players' private states cannot matter for i's rewards and transitions, but only that they do not matter *conditional* on the public information, namely the public signal. At the cost of more notation, one can also include t^{-i} (for instance, $-i$'s realized stage-game payoff, see Mezzetti 2004, 2007) in this public information.

[18]With some abuse of notation, we use for simplicity the same symbol, $p(\cdot)$, for the relevant marginal probabilities.

for all $(s, \tilde{s}, a, y)$, so that the public signal conveys no information about the state profile. This is the case under perfect monitoring, but rules out interesting cases (Think, for instance, about applications in which the signal is the realization of an accident, whose probability depends both on the agent's risk and his effort.) This simplifies the statements of the results considerably, but is not necessary: We will return to the general case at the end of this section.

There is no gain from having $M^i = (S^i)^2$ here, and so we set $M^i = S^i$, although we use the symbol M^i whenever convenient. Under IPV, one cannot expect all feasible and individually rational payoffs to be equilibrium payoffs under low discounting. Incentive compatibility imposes restrictions on what can be hoped for. A policy is a plan of action that only depends on the last message, *i.e.* a map $\rho : M \to A$. We define the feasible (long-run) payoff set as

$$F = \left\{ v \in \mathbf{R}^I \mid v = \mathbf{E}_{\mu[\rho]}[r(s,a)], \text{ some policy } \rho \right\}.$$

When defining feasible payoffs, the restriction to policies rather than arbitrary strategies (not only plans of actions) is clearly without loss. Recall also that a public randomization device is assumed, so that F is convex.

Not all feasible payoffs can be equilibrium payoffs, however, because types are private information. For all $\lambda \in S_1$, let $I(\lambda) = \{i : \lambda^i \geq 0\}$, and define

$$\bar{k}(\lambda) = \max_{\rho} \mathbf{E}_{\mu[\rho]} \left[\lambda \cdot r(s,a) \right],$$

where the maximum is over all policies $\rho : \times_{i \in I(\lambda)} S^i \to A$, with the convention that $\rho \in A$ for $I(\lambda) = \emptyset$. Furthermore, let

$$V^* = \cap_{\lambda \in S_1} \left\{ v \in \mathbf{R}^I \mid \lambda \cdot v \leq \bar{k}(\lambda) \right\}.$$

Clearly, $V^* \subseteq F$. Furthermore, V^* is an upper bound on the set of *all* equilibrium payoff vectors.

Lemma 2 *The limit set of Bayes Nash equilibrium payoffs is contained in V^*.*

Proof. Fix $\lambda \in S_1$. Fix also $\delta < 1$ (and recall the prior p_0 at time 0). Consider the Bayes Nash equilibrium σ of the game (with discount factor δ) that maximizes $\lambda \cdot v$ among all equilibria (where v^i is the expected payoff of player i given p_0). This equilibrium need not be truthful or in pure strategies. Consider $i \notin I(\lambda)$. Along with σ^{-i} and p_0, player i's equilibrium strategy σ^i defines a distribution over histories. Fixing σ^{-i}, let us consider an alternative strategy $\tilde{\sigma}^i$ where player i's reports are replaced by realizations of the public randomization device with the same distribution (period by period, conditional on the realizations so far), and player i's action is determined by the randomization device as well, with the same conditional distribution (given the simulated reports) as σ^i would specify if this had been i's report. The new profile $(\sigma^{-i}, \tilde{\sigma}^i)$ need no longer be an equilibrium of the game, but it gives players $-i$ the same payoff, and player i a weakly lower payoff (because his choices are no longer conditioned on his true types). Most importantly, the strategy profile $(\sigma^{-i}, \tilde{\sigma}^i)$ no longer depends on the history of types of player i. Clearly, this argument applies to all players $i \notin I(\lambda)$, so that $\lambda \cdot v$ is lower than the maximum inner product achieved over strategies that only depend on the history of types in $I(\lambda)$. Maximizing this inner product over such strategies is a standard Markov decision process, which admits a solution within the class of deterministic policies. Taking $\delta \to 1$ yields that the limit set is in $\left\{v \in \mathbf{R}^I \mid \lambda \cdot \leq \bar{k}(\lambda)\right\}$, and this is true for all $\lambda \in S_1$. ■

It is worth emphasizing that this result does not rely on the choice of any particular message space.[19] We define

$$\rho[\lambda] = \arg \max_{\Pi_{i \in I(\lambda)} S^i \to A} \mathbf{E}_{\mu[\rho]} \left[\lambda \cdot r(s,a)\right] \tag{2}$$

to be the policy that achieves this maximum, and let $\Xi = \{\rho[\lambda] : \lambda \in S_1\}$ denote the set of such policies. We call V^* the set of *incentive-compatible* payoffs.

Because V^* is not equal to F, it is natural to wonder how large this set is. Let Ext^{po} the (weak) Pareto frontier of F. We also write Ext^{pu} for the set of payoff vectors obtained from pure state-independent action profiles, *i.e.* the set of vectors $v = \mathbf{E}_{\mu[\rho]}[r(s,a)]$ for some ρ that takes a constant value in A. Escobar and Toikka (2013) show that, in their environment, all individually rational (as defined below) payoffs in $co(Ext^{pu} \cup Ext^{po})$ are equilibrium payoffs (whenever this set has non-empty interior). Indeed, the following is easy to show.

Lemma 3 *It holds that $co(Ext^{pu} \cup Ext^{po}) \subset V^*$.*

As discussed, V^* contains the Pareto frontier and the constant action profiles, but this might still be a strict subset of F. The following example illustrates the difference.

[19] Incidentally, it appears that the role of V^* is new even in the context of static mechanism design with transfers. There is no known exhaustive description of the allocation rules that can be implemented under IPV, but it is clear that the payoffs in V^* (replace μ with the prior) can be achieved using the AGV mechanism; conversely, no payoff above $\bar{k}(\lambda)$ can be achieved, so that V^* provides a description of the achievable payoff set in that case as well.

	L	R
T	$3 - \frac{c(s^i)}{2}, 3 - \frac{c(s^{-i})}{2}$	$3 - c(s^i), 3$
B	$3, 3 - c(s^{-i})$	$0, 0$

Figure 3: Payoffs of Example 4

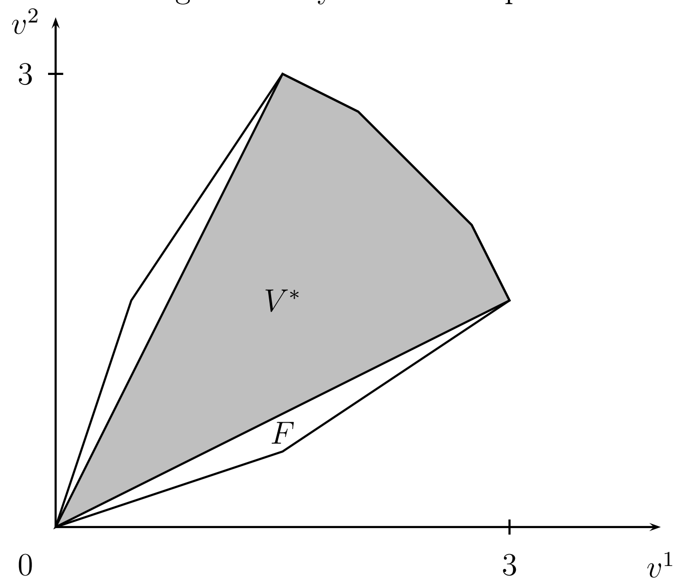

Figure 4: Incentive-compatible and feasible payoff sets for the example

Example 4. Actions do not affect transitions. Rewards are given by Figure 3. Each player has two states $s^i = \underline{s}^i, \bar{s}^i$, with $c(\underline{s}^i) = 2, c(\bar{s}^i) = 1$. (The interpretation is that a pie of size 3 is obtained if at least one agent works; if both choose to work only half the amount of work has to be put in by each worker. Their cost of working is fluctuating.) This an IPV model. From one period the next, the state changes with probability p, constant and independent across players. Given that actions do not affect transitions, we can take it equal to $p = 1/2$ (i.i.d) for the sake of computing V^* and F, shown in Figure 4. Of course, each player can secure at least $3 - \frac{2+1}{2} = \frac{3}{2}$ by always working, so the actual equilibrium payoff set, taking into account the incentives at the action stage is smaller.

6.1 Truth-telling

In this section, we ignore the action stage and focus on the incentives of players to report their type truthfully.

Let us say that (ρ, x) is *truthful* if the pair satisfies Definition 2 if in Step 2 the requirement that ρ^i be optimal is ignored. That is, we take the action choice $\rho(\cdot)$ as given. Furthermore, (ρ, x) is *weakly truthful* if in addition in Step 1 of Definition 2 the requirement that truth-telling is *uniquely* optimal is dropped. That is, we only require truth-telling to be an optimal reporting strategy, although not necessarily the unique one.

Suppose that some extreme point of V is a limit equilibrium payoff vector. Then there is $\lambda \in S^1$ such that this extreme point maximizes $\lambda \cdot v'$ over $v' \in V$. Hence, for every $(\bar{\omega}_{\text{pub}}, \omega_{\text{pub}}) \in \Omega_{\text{pub}} \times \Omega_{\text{pub}}$ and $t \in S$ such that $p(t^{-i}, y \mid m, \rho(\bar{\omega}_{\text{pub}}, m)) > 0$, it must hold that $\lambda \cdot x(\bar{\omega}_{\text{pub}}, \omega_{\text{pub}}, t) = 0$. This motivates the familiar notion of orthogonal enforceability from repeated games (see FLM). The plan of action ρ is *orthogonally enforced* by x in the direction λ (under truth-telling) if for all $(\bar{\omega}_{\text{pub}}, \omega_{\text{pub}}, t)$, all i, $p(t^{-i}, y \mid m, \rho(\bar{\omega}_{\text{pub}}, m)) > 0$ implies $\lambda \cdot x(\bar{\omega}_{\text{pub}}, \omega_{\text{pub}}, t) = 0$. If ρ is orthogonally enforced, then the payoff achieved by (ρ, x) is $\lambda \cdot v = \mathbf{E}_{\mu[\rho]} \lambda \cdot r(s, a)$. We also recall that a direction λ is non-coordinate if $\lambda \neq \pm e^i$; that is, it has at least two nonzero coordinates.

Proposition 1 *Fix a non-coordinate direction λ. Let (ρ, x) be a (weakly) truthful pair. Then there exists $\hat{x}$ such that $(\rho, \hat{x})$ is (weakly) truthful and ρ is orthogonally enforced by $\hat{x}$ in the direction λ.*

Proposition 1 implies that budget-balance ($\lambda \cdot x \leq 0$) comes "for free" in all directions $\lambda \neq \pm e^i$. Proposition 1 is the undiscounted analogue of a result by Athey and Segal (2007), and its proof follows similar steps as theirs. It is worth mentioning that it does not rely on the independence of the signal distribution on the state profile.

Our next goal is to obtain a characterization of all policies ρ for which there exists x such that (ρ, x) is (weakly) truthful. With some abuse, we say then that ρ is (weakly) truthful. Along with ρ and truthful reporting by players $-i$, a reporting strategy by player i, that is, a map $m^i_\rho : \Omega_{\text{pub}} \times S^i \to \Delta(M^i)$ from the past public outcome and the current state into a report, induces a distribution π^i_ρ over $\Omega_{\text{pub}} \times S^i \times M^i$.[20] Conversely, given any such distribution, we can define the corresponding reporting strategy by

$$m^i_\rho(\bar{\omega}_{\text{pub}}, s^i)(m^i) = \frac{\sum_{\bar{\omega}_{\text{pub}}, s^i} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i)}{\sum_{\bar{\omega}_{\text{pub}}, s^i, m^i} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i)},$$

whenever $\sum_{\bar{\omega}_{\text{pub}}, s^i, m^i} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) > 0$, $m^i_\rho(\bar{\omega}_{\text{pub}}, s^i)(m^i)$ defined arbitrarily otherwise. Let us define

$$r^i_\rho(\bar{\omega}_{\text{pub}}, m^i, s^i) = \mathbf{E}_{s^{-i} \mid \bar{\omega}_{\text{pub}}} r^i(\rho(s^{-i}, m^i), s^i)$$

[20]Note that, under IPV, player i's private information contained in $\bar{\omega}$ is not relevant for his incentives in the current period, conditional on $\bar{\omega}_{\text{pub}}$.

as the expected reward of player i given his message, type and the previous public outcome $\bar{\omega}_{\text{pub}}$. Let us also define, for $\omega_{\text{pub}} = (s^{-i}, m^i, y)$ and $\bar{\omega}_{\text{pub}} = (\bar{s}, \bar{y})$,

$$q_\rho(\omega_{\text{pub}}, t^i \mid \bar{\omega}_{\text{pub}}, s^i, m^i) = p(t^i \mid y, s^i, \rho^i(s^{-i}, m^i))p(y \mid \rho(s^{-i}, m^i))p(s^{-i} \mid \bar{y}, \bar{s}^{-i}, \rho^{-i}(\bar{s}^{-i}, \bar{s}^i)),$$

which is simply the probability of $(\omega_{\text{pub}}, t^i)$ given $(\bar{\omega}_{\text{pub}}, s^i, m^i)$. The distribution π^i_ρ must satisfy the balance equation

$$\sum_{\tilde{m}^i} \pi^i_\rho(\omega_{\text{pub}}, t^i, \tilde{m}^i) = \sum_{\bar{\omega}_{\text{pub}}, s^i} q_\rho(\omega_{\text{pub}}, t^i \mid \bar{\omega}_{\text{pub}}, s^i, m^i)\pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i), \tag{3}$$

for all states $(\omega_{\text{pub}}, t^i)$. A distribution π^i_ρ over $\Omega_{\text{pub}} \times S^i \times M^i$ is in Π^i_ρ if (i) it satisfies (3), and (ii) for all $(\bar{\omega}_{\text{pub}}, m^i)$,

$$\sum_{s^i} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) = \mu[\rho](\bar{\omega}_{\text{pub}}, m^i). \tag{4}$$

Equation (4) states that π^i_ρ cannot be statistically distinguished from truth-telling. Its significance comes from the next Lemma.

Lemma 4 *Given a policy ρ, there exists x such (ρ, x) is (weakly) truthful if for all i, truth-telling maximizes*

$$\sum_{(\bar{\omega}_{\text{pub}}, s^i, m^i)} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) r^i_\rho(\bar{\omega}_{\text{pub}}, m^i, s^i) \tag{5}$$

over $\pi^i_\rho \in \Pi^i_\rho$.

Fix $\lambda \in S_1$. We claim that, for all $i \in I(\lambda)$, truth-telling maximizes (5) with $\rho = \rho[\lambda]$ over $\pi^i_{\rho[\lambda]} \in \Pi^i_{\rho[\lambda]}$; suppose not, *i.e.* there exists $i \in I(\lambda)$ such that (5) is maximized by some other reporting strategy $m^i_{\rho[\lambda]}$. Note that players $-i$'s payoffs are the same whether player i uses truth-telling or $m^i_{\rho[\lambda]}$, because of (4). We can then define the policy $\tilde{\rho}$ as, for all s,

$$\tilde{\rho}(s^i, s^{-i}) = \mathbf{E}_{m^i \sim m^i_{\rho[\lambda]}(s^i)} \rho[\lambda](m^i, s^{-i}).$$

By construction, truth telling maximizes (5) (with $\rho = \tilde{\rho}$) over $\pi^i_{\tilde{\rho}} \in \Pi^i_{\tilde{\rho}}$. Player i is strictly better off under truth-telling with $\tilde{\rho}$ than under truth-telling under $\rho[\lambda]$. Furthermore, players $-i$'s payoffs are the same (under truth-telling) under $\tilde{\rho}$ as under $\rho[\lambda]$. This means that $\rho[\lambda]$ could not satisfy (2), a contradiction.

By Proposition 1, given any ρ that satisfies Lemma 4 and any non-coordinate direction λ, we can take the transfer x such that $\lambda \cdot x = 0$. Note that the same arguments apply to the case

$\lambda = \pm e^i$: while Proposition 1 does not hold in that case, the maximum $\mathbf{E}_{\mu[\rho]}\left[\pm e^i \cdot r(s,a)\right]$ over $\rho : \times_{i \in I(\lambda)} S^i \to A$ is orthogonally enforceable in the case $-e^i$ because the maximizer is then a constant action ($I(\lambda) = \emptyset$), and so truth-telling is trivial; in the case $+e^i$ it is also trivial, because $I(e^i) = \{i\}$ and truth-telling maximizes player i's payoff.

It then follows that the maximum score over weakly truthful pairs (ρ, x) is equal to the maximum possible one, $\bar{k}(\lambda)$.

Lemma 5 *Fix a direction $\lambda \in S_1$. Then the maximum score over weakly truthful (ρ, x) such that $\lambda \cdot x \leq 0$ is given by $\bar{k}(\lambda)$.*

For now incentives to report truthfully have been taken to be weak. Our main theorem requires strict incentives, at least (by Theorem 3) for reports that affect the action played. To ensure that strict incentives can be given, a minimal assumption has to be made.

Assumption 1 *For any two (pure) policies $\rho, \tilde{\rho} \in \Xi$, it holds that*

$$\mathbf{E}_{\mu[\rho]}[r(s, \rho(s))] \neq \mathbf{E}_{\mu[\tilde{\rho}]}[r(s, \tilde{\rho}(s))].$$

The meaning of this assumption is clear: there is no redundant pure policy among those achieving payoff vectors on the boundary of V^*. This assumption is stronger than necessary, but at least it holds generically (over r and p). This assumption allows us to strengthen our previous result from weak truthfulness to truthfulness.

Lemma 6 *Fix a direction $\lambda \in S_1$. Under Assumption 1, the maximum score over truthful (ρ, x) such that $\lambda \cdot x \leq 0$ is given by $\bar{k}(\lambda)$.*

The conclusion of this section is somewhat surprising: at least in terms of payoffs, there is no possible gain (in terms of incentive-compatibility) from linking decisions (and restricting attention to truthful strategies) beyond the simple class of policies and transfer functions that we consider. In other words, *ignoring individual rationality and incentives at the action stage*, the set of "equilibrium" payoffs that we obtain is equal to the set of incentive-compatible payoffs V^*.

If transitions are action-independent, note that this means also that the persistence of the Markov chain has no relevance for the set of payoffs that are incentive-compatible. (If actions affect transitions, even the feasible payoff set changes with persistence, as it affects the extreme policies.) Note that this does it rely not on any full support assumption on the transition probabilities, although of course the unichain assumption is used (cf. Example 1 of Renault, Solan and Vieille (2013) that shows that this conclusion –the sufficiency of the invariant distribution– does not hold when values are interdependent).

6.2 Actions

Only one step is missing to go from truthful policies to admissible ones: the lack of commitment and imperfect monitoring of actions. This lack of commitment curtails how low payoffs can be. We define player i's (pure) minimax payoff as

$$\underline{v}^i = \min_{a^{-i} \in A^{-i}} \max_{\rho^i : S^i \to A^i} \mathbf{E}_{\mu[\rho^i, a^{-i}]}[r^i(s^i, a)].$$

This is the state-independent pure-strategy minmax payoff of player i. We let $\underline{\rho}_i$ denote a policy that achieves this minmax payoff, and assume that $\underline{\rho}_i^i$ is unique given $\underline{\rho}_i^{-i}$.

It clearly is a restrictive notion of individual rationality, see Escobar and Toikka (2013) for a discussion. In particular, it is known that to punish player i optimally one should consider strategies outside of the class that we consider –strategies that have a nontrivial dependence on beliefs (after all, to compute the minimax payoff, we are led to consider a zero-sum game, as Example 3 above –private values does not change this).

Denote the set of incentive-compatible, individually rational payoffs as

$$V^{**} = \left\{ v \in V^* \mid v^i \geq \underline{v}^i, \text{ all } i \right\}.$$

We now introduce assumptions on monitoring that are rather standard, see Kandori and Matsushima (1998). Let $Q^i(a) = \{p(\cdot \mid \hat{a}^i, a^{-i}) : \hat{a}^i \neq a^i\}$ be the distribution over signals y induced by a unilateral deviation by i at the action stage. The first assumption involves the minimax policies.

Assumption 2 *For all i, $j \neq i$, all $\underline{\rho}_i^j \in A^j$.*

$$p(\cdot \mid a) \notin co\ Q^j(a).$$

The second involves the extreme points (in policy space) of the incentive-compatible payoff set.

Assumption 3 *For all $\rho \in \Xi$, all s, $a = \rho(s)$,*

1. *For all $i \neq j$, $p(\cdot \mid a) \notin co(Q^i(a) \cup Q^j(a))$;*

2. *For all $i \neq j$,*

$$co\left(p(\cdot \mid a) \cup Q^i(a)\right) \cap co\left(p(\cdot \mid a) \cup Q^j(a)\right) = \{p(\cdot \mid a)\}.$$

We may now state the main result of this section.

Theorem 4 *Suppose that V^{**} has non-empty interior. Under Assumptions **1–3**, the limit set of equilibrium payoffs includes V^{**}.*

6.3 State-dependent Signalling

Let us now return to the case in which signals depend on actions.

Fix $\bar{\omega}_{\text{pub}}$. The states s^i and $\tilde{s}^i$ are *indistinguishable*, denoted $s^i \sim \tilde{s}^i$, if for all s^{-i} such that $p(s^{-i} \mid \bar{\omega}_{\text{pub}}) > 0$ and all (a, y), $p(y \mid s^{-i}, s^i, a) = p(y \mid s^{-i}, \tilde{s}^i, a)$. Indistinguishability defines a partition of S^i, given $\bar{\omega}_{\text{pub}}$, and we denote by $[s^i]$ the partition cell to which s^i belongs. If signals depend on action, this partition is non-trivial for at least one player.

By definition, if $[s^i] \neq [\tilde{s}^i]$ there exists s^{-i} such that $p(\cdot \mid s^{-i}, s^i, a) \neq p(\cdot \mid s^{-i}, \tilde{s}^i, a)$ for some $a \in A$ and $\tilde{s}^i \in [\tilde{s}^i]$. We assume that we can pick this state s^{-i} such that $p(s^{-i} \mid \bar{\omega}_{\text{pub}}) > 0$ for all $\bar{\omega}_{\text{pub}} \in \Omega_{\text{pub}}$ (where $\bar{s}^{-i} = \bar{m}^{-i}$). This avoids dealing with the case in which, for some public outcomes $\bar{\omega}_{\text{pub}}$, player i knows that s^i cannot be statistically distinguished from $\tilde{s}^i$, but not in others. (Note that this assumption is trivial if states do not affect signals, the case considered in the previous section.) This ensures that $S^i/\!\sim$ is independent of $\bar{\omega}_{\text{pub}}$. Let $D^i = \{(s^{-i}, a)\} \subset S^{-i} \times A$ denote a selection of such states, along with the discriminating action profile: for all $[s^i] \neq [\tilde{s}^i]$, there exists $(s^{-i}, a) \in D^i$ such that $p(\cdot \mid s^{-i}, s^i, a) \neq p(\cdot \mid s^{-i}, \tilde{s}^i, a)$.

We must redefine the relevant set of payoffs: Let

$$\bar{k}(\lambda) = \max_{\rho} \mathbf{E}_{\mu[\rho]} \left[\lambda \cdot r(s, a) \right],$$

where the maximum is now over all policies $\rho : S \to A$ such that if $s^i \sim \tilde{s}^i$ and $\lambda^i \leq 0$ then $\rho(s^i, s^{-i}) = \rho(\tilde{s}^i, s^{-i})$. Furthermore, let

$$V^* = \cap_{\lambda \in S_1} \left\{ v \in \mathbf{R}^I \mid \lambda \cdot v \leq \bar{k}(\lambda) \right\}.$$

This set is larger than before, as strategies can depend on some information of the players whose weight is negative, namely information that can be statistically verified. Following the exact same argument as for Lemma 2, V^* is a superset of the set of Bayes Nash equilibrium payoffs. We retain the same notation for $\rho[\lambda]$, the policies that achieve the extreme points of V^*, and Ξ, the set of such policies. We maintain Assumption 1, with this new definition: any two policies in Ξ yield different average payoff vectors $\mathbf{E}_{\mu[\rho]}[r(s, \rho(s))]$.

We may strengthen (4) to, for all $(\bar{\omega}_{\text{pub}}, m^i)$,

$$\sum_{s^i \in [m^i]} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) = \mu[\rho](\bar{\omega}_{\text{pub}}, m^i). \tag{4'}$$

We then redefine Π^i_ρ: A distribution π^i_ρ over $\Omega_{\text{pub}} \times S^i \times M^i$ is in Π^i_ρ if it satisfies (3) as well as (4'). Lemma 4 then holds without modification. We then obtain the analogue of Lemma 6.

Lemma 7 *Fix a direction $\lambda \in S_1$. Under Assumption 1, the maximum score over truthful (ρ, x) such that $\lambda \cdot x \leq 0$ is given by $\bar{k}(\lambda)$.*

We now turn to actions. The definition of minimax payoff $\underline{v}^i$ does not change, and we maintain the assumption that $\underline{\rho}_i^i$ is unique, given $\underline{\rho}_i^{-i}$. The definition of V^{**} does not change either, given our new definition of V^*.

In what follows $p(\cdot \mid a, s)$ refers to the marginal distribution over signals $y \in Y$ only. (Because types are conditionally independent, players' $-i$ signals in round $n+1$ are uninformative about a^i, conditional on y.) Let $Q^i(a, s) = \{p(\cdot \mid \hat{a}^i, a^{-i}, \hat{s}^i, s^{-i}) : \hat{a}^i \neq a^i, \hat{s}^i \in S^i\}$ be the distribution over signals y induced by a unilateral deviation by i at the action stage, whether or not the reported state s^i corresponds to the true state $\hat{s}^i$ or not. The first assumption involves the minimax policies.

Assumption 4 *For all i, for all s, $a = \underline{\rho}_i(s)$, $j \neq i$,*

$$p(\cdot \mid a, s) \notin co\ Q^j(a, s).$$

The second involves the extreme points (in policy space) of the relevant payoff set.

Assumption 5 *For all $\rho \in \Xi$, all s, $a = \rho(s)$; also, for all (s, a) where $(s^{-i}, a) \in D^i$ for some i:*

1. *For all $i \neq j$, $p(\cdot \mid a, s) \notin co(Q^i(a, s) \cup Q^j(a, s))$;*
2. *For all $i \neq j$,*

$$co\left(p(\cdot \mid a, s) \cup Q^i(a, s)\right) \cap co\left(p(\cdot \mid a, s) \cup Q^j(a, s)\right) = \{p(\cdot \mid a, s)\}.$$

This assumption states that deviations of players can be detected, as well as identified, even if player i has "coordinated" his deviation at the reporting and action stage.

We then get the natural generalization from Theorem 4.

Theorem 5 *Suppose that V^{**} has non-empty interior. Under Assumptions **1**, **4**–**5**, the limit set of equilibrium payoffs includes V^{**}.*

7 Correlated Types

We drop the assumption of independent types and extend here the static insights from Crémer and McLean (1988). We must redefine the minimax payoff of player i as

$$\underline{v}^i = \min_{\rho^{-i}: S^{-i} \to A^{-i}} \max_{\rho^i: S^i \to A^i} \mathbf{E}_{\mu[\rho]}[r^i(s, a)],$$

As before, we let $\underline{\rho}_i$ denote a policy that achieves this minimax payoff. We maintain Assumption **4** with this change of definition in the minimax policies. Similarly, throughout this section we maintain Assumption **5** *with the set of relevant policies* ρ *being those* $\rho \in Ex(F)$, namely those policies achieving extreme points of the feasible payoff set. This ensures that, following the same steps as above, arbitrarily strong incentives can be provided to players to follow any plan of action $\rho : \Omega_{\text{pub}} \times M \to A$, whether or not they deviate in their reports.

We ignore the i.i.d. case, covered by similar arguments that invokes Theorem 3 rather than Theorem 2.

Given $\bar{m}$, $\bar{y}$, $\bar{a}$, and a map $\rho : M \to A$, given i and any pair $\zeta^i = (\bar{s}^i, s^i)$, we use Bayes' rule to compute the distribution over (t^{-i}, s^{-i}, y), conditional on the past messages being $\bar{m}$, the past action and signal being $\bar{y}$, $\bar{a}$, player i's true past and current state being $\bar{s}^i$ and s^i, and the policy $\rho : S \to A$. This distribution is denoted

$$q_{-i}^{\bar{m},\bar{y},\bar{a},\rho}(t^{-i}, s^{-i}, y \mid \zeta^i).$$

The tuple $\bar{m}$, $\bar{y}$, $\bar{a}$ also defines a joint distribution over profiles s, y and t, denoted

$$q^{\bar{m},\bar{y},\bar{a},\rho}(t, s, y),$$

which can be extended to a prior over $\zeta = (\bar{s}, s)$, y and t that assigns probability 0 to types $\bar{s}^i$ such that $\bar{s}^i \neq \bar{m}^i_c$.

We exploit the correlation in types to induce players to report truth-fully. As always, we must distinguish between directions $\lambda = -e^i$ (minimaxing) and other directions. First, we assume

Assumption 6 *For all* i, $\rho = \underline{\rho}_i$, *all* $(\bar{m}, \bar{y}, \bar{a})$, *for any* i, $\hat{\zeta}^i \in (S^i)^2$, *it holds that*

$$q_{-i}^{\bar{m},\bar{y},\bar{a},\rho}(t^{-i}, s^{-i}, y \mid \hat{\zeta}^i) \neq co\left(q_{-i}^{\bar{m},\bar{y},\bar{a},\rho}(t^{-i}, s^{-i}, y \mid \zeta^i) : \zeta^i \neq \hat{\zeta}^i\right).$$

If types are independent over time, and signals y do not depend on states (as is the case with perfect monitoring, for instance), this reduces to the requirement that the matrix with entries $p(s^{-i} \mid s^i)$ have full row rank, a standard condition in mechanism design (see d'Aspremont, Crémer and Gérard-Varet (2003) and d'Aspremont and Gérard-Varet (1982)'s condition B). Here, beliefs can also depend on player i's previous state, $\bar{s}^i$, but fortunately, we can also use player $-i$'s future state profile, t^{-i}, to statistically distinguish player i's types.

As is well known, Assumption **6** ensures that for any minimaxing policy $\underline{\rho}_i$, truth-telling is Bayesian incentive compatible: there exists transfers $x^i(\bar{\omega}_{\text{pub}}, (m, y), t^{-i})$ for which truth-telling is strictly optimal.

Ex post budget balance requires further standard assumptions. Following Kosenok and Severinov (2008), let $c^i : (S^i)^2 \to M^i$ denote a reporting strategy, summarized by numbers $c^i_{\zeta^i \hat{\zeta}^i} \geq 0$, with $\sum_{\hat{\zeta}^i} c^i_{\zeta^i \hat{\zeta}^i} = 1$ for all ζ^i, with the interpretation that $c^i_{\zeta^i \hat{\zeta}^i}$ is the probability with which $\hat{\zeta}^i$ is reported when the type is ζ^i. Let $\hat{c}^i$ denote the truth-telling reporting strategy where $c^i_{\zeta^i \zeta^i} = 1$ for all ζ^i. A reporting strategy profile c, along with the prior $q^{\bar{m},\bar{y},\bar{a},\rho}$ defines a distribution $\pi^{\bar{m},\bar{y},\bar{a},\rho}$ over (ζ, y, t), according to

$$\pi^{\bar{m},\bar{y},\bar{a},\rho}(\hat{\zeta}, y, t \mid c) = \sum_{\zeta} q^{\bar{m},\bar{y},\bar{a},\rho}(\zeta, y, t) \prod_j c^j_{\zeta^j \hat{\zeta}^j}.$$

We let

$$\mathcal{R}^i(\bar{m}, \bar{y}, \bar{a}, \rho) = \left\{ \pi^{\bar{m},\bar{y},\bar{a},\rho}(\cdot \mid c^i, \hat{c}^{-i}) : c^i \neq \hat{c}^i \right\}.$$

Again, the following is the adaptation of the assumption of Kandori and Matsushima (1998) to the current context.

Assumption 7 *For all $\rho \in Ex(F)$, all $(\bar{m}, \bar{y}, \bar{a})$,*

1. *For all $i \neq j$, $\pi^{\bar{m},\bar{y},\bar{a},\rho}(\cdot \mid \hat{c}) \notin co\,(\mathcal{R}^i(\bar{m}, \bar{y}, \bar{a}, \rho) \cup \mathcal{R}^j(\bar{m}, \bar{y}, \bar{a}, \rho))$;*

2. *For all $i \neq j$,*

$$co\left(\pi^{\bar{m},\bar{y},\bar{a},\rho}(\cdot \mid \hat{c}) \cup \mathcal{R}^i(\bar{m}, \bar{y}, \bar{a}, \rho)\right) \cap co\left(\pi^{\bar{m},\bar{y},\bar{a},\rho}(\cdot \mid \hat{c}) \cup \mathcal{R}^j(\bar{m}, \bar{y}, \bar{a}, \rho)\right) = \{\pi^{\bar{m},\bar{y},\bar{a},\rho}(\cdot \mid \hat{c})\}.$$

Assumption **7.1** is equivalent to the assumption of weak identifiability in Kosenok and Severinov (2008) for two players (whose Lemma 2 can be directly applied). The reason it is required for any pair of players (unlike in Kosenok and Severinov) is that we must obtain budget-balance also for vectors $\lambda \in S_1$ with only two non-zero coordinates (of the same sign). Assumption **7.2** is required (as in Kandori and Matsushima in their context) because we must also consider directions $\lambda \in S_1$ with only two non-zero coordinates whose signs are opposite.[21]

It is then routine to show:

[21]See also Hörner, Takahashi and Vieille (2012). One easy way to understand the second one is in terms of the cone spanned by the vectors $\pi^{\bar{m},\bar{y},\bar{a},\rho}(\cdot \mid c^i, \hat{c}^{-i})$ and pointed at $\pi^{\bar{m},\bar{y},\bar{a},\rho}(\cdot \mid \hat{c})$. The first assumption is equivalent to any two such cones only intersecting at 0; and the second one states that any cone intersected with the opposite cone (of another player) also only intersect at 0. When $\lambda^i > 0 > \lambda^j$, we can rewrite the constraint $\lambda x^i + \lambda^j x^j = 0$ as $\lambda^i x^i + (-\lambda^j)(-x^j) = 0$ and the expected transfer of a player as $p(\cdot \mid c^j)x^j(\cdot) = (-p(\cdot \mid c^j))(-x^j(\cdot))$, so the condition for (λ^i, λ^j) is equivalent to the condition for $(\lambda^i, -\lambda^j)$ if one "replaces" the vectors $p(\cdot \mid c^j)$ with $-p(\cdot \mid c^j)$.

Theorem 6 *Assume that V has non-empty interior. Under Assumptions **4**–**7**, the limit set of truthful equilibrium payoffs includes V.*

Assumptions **6**–**7** are generically satisfied if $|S^{-i}| \geq |S^i|$ for all i.

8 Conclusion

This paper has considered a class of equilibria in games with private and imperfectly persistent information. While the structure of equilibria has been assumed to be relatively simple, to preserve tractability –in particular, we have focused on truthful equilibria– it has been shown, perhaps surprisingly, that in the case of independent private values this is not restrictive as far as incentives go: all that transfers depend on are the current and the previous report. This confirms a rather natural intuition: in terms of equilibrium payoffs at least (and as far as incentive-compatibility is concerned), there is nothing to gain from aggregating information beyond transition counts. In the case of correlated values, we have shown how the standard insights from static mechanism design with correlated values generalize; in this case as well, the "genericity" conditions in terms of number of states are satisfied with a simple mechanism, provided next period's reports by a player's opponent are used.

This is not to say that more complicated mechanisms cannot help, but that they help in particular ways. Our initial examples have illustrated some of the issues. In light of our results, let us revisit those.

Looking backward. In the case of independent private values, using richer information from the past is useful for the purpose of pushing down the lowest equilibrium payoff. There are two distinct channels through which this operates.

First, by considering realized action profiles of players $-i$ (under perfect monitoring, say), one might be able to ensure that their play is *as if* they were randomizing, using ideas from approachability theory. As we have discussed in Example 2, players are typically not willing to randomize. Statistical tests based on longer stretch of behavior might help remedy this.

Second, there is a high price to pay for getting the punished player to reveal his state truthfully: punishing players should not take advantage of that information to fine-tune their action profile. Yet if we drop truth-telling, the punished player cannot help but either reveal some of his private information through the actions that he plays, or play in a way that does not depend on this information. In both cases, this is costly and should allow the other players to punish him more effectively. Here again, approachability seems to be the right tool to use, but it requires taking into account longer sequences of past outcomes (or equivalently, keeping track of richer beliefs). This is obvious from Example 1.

While both channels allow for more effective punishments, their roles are somewhat different: in the first case, it is simply to check that the empirical distribution of actions is closed to the one that i.i.d. behavior would generate; in the second, to draw better inferences about the current state from past observations.

Looking forward. Under correlated values, if the genericity assumptions fail, it becomes useful to take into account future observations. Inferences are reversed with respect to the previous case: here, we use future signals (reports by other players, in particular) to better infer the current state. Example 3 illustrates how powerful this can be, but unless actions do not affect the Markov chain (as in the example), it is not entirely clear how to best exploit this information.

References

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

[2] Arrow, K. (1979). "The Property Rights Doctrine and Demand Revelation Under Incomplete Information," in M. Boskin, ed., *Economics and human welfare*. New York: Academic Press.

[3] d'Aspremont, C. and L.-A. Gérard-Varet (1979). "Incentives and Incomplete Information," *Journal of Public Economics*, **11**, 25–45.

[4] d'Aspremont, C. and L.-A. Gérard-Varet (1982). "Bayesian incentive compatible beliefs," *Journal of Mathematical Economics*, **10**, 83–103.

[5] d'Aspremont, C., J. Crémer and L.-A. Gérard-Varet (2003). "Correlation, independence, and Bayesian incentives," *Social Choice and Welfare*, **21**, 281–310.

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

[7] Athey, S. and K. Bagwell (2008). "Collusion with Persistent Cost Shocks," *Econometrica*, **76**, 493–540.

[8] Athey, S. and I. Segal (2007). "An Efficient Dynamic Mechanism," working paper, Stanford University.

[9] Aumann, R.J. and M. Maschler (1995). *Repeated Games with Incomplete Information*. Cambridge, MA: MIT Press.

[10] Barron, D. (2012). "Attaining Efficiency with Imperfect Public Monitoring and Markov Adverse Selection," working paper, M.I.T.

[11] Bester, H. and R. Strausz (2000). "Contracting with Imperfect Commitment and the Revelation Principle: The Multi-Agent Case," *Economics Letters*, **69**, 165–171.

[12] Bester, H. and R. Strausz (2001). "Contracting with Imperfect Commitment and the Revelation Principle: The Single Agent Case," *Econometrica*, **69**, 1077–1088.

[13] Bewley, T. and E. Kohlberg (1976). "The asymptotic theory of stochastic games," *Mathematics of Operations Research*, **3**, 104–125.

[14] Blackwell, D. and L. Koopmans (1957). "On the Identifiability Problem for Functions of Finite Markov Chains," *Annals of Mathematical Statistics*, **28**, 1011–1015.

[15] Cole, H.L. and N.R. Kocherlakota (2001). "Dynamic games with hidden actions and hidden states," *Journal of Economic Theory*, **98**, 114–126.

[16] Crémer, J. and R. McLean (1988). "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, **56**, 1247–1257.

[17] Dharmadhikari, S. W. (1963). "Sufficient Conditions for a Stationary Process to be a Function of a Finite Markov Chain," *The Annals of Mathematical Statistics*, **34**, 1033–1041.

[18] Doepke, M. and R.M. Townsend (2006). "Dynamic mechanism design with hidden income and hidden actions," *Journal of Economic Theory*, **126**, 235–285.

[19] Escobar, P. and J. Toikka (2013). "Efficiency in Games with Markovian Private Information," *Econometrica*, forthcoming.

[20] Fang, H. and P. Norman (2006). "To Bundle or Not To Bundle," *RAND Journal of Economics*, **37**, 946–963.

[21] Fernandes, A. and C. Phelan (1999). "A Recursive Formulation for Repeated Agency with History Dependence," *Journal of Economic Theory*, **91**, 223–247.

[22] Freixas, X., R. Guesnerie and J. Tirole (1985). "Planning under Incomplete Information and the Ratchet Effect," *Review of Economic Studies*, 173–191.

[23] Fudenberg, D. and D. Levine (1994). "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, **62**, 103–135.

[24] Fudenberg, D., D. Levine, and E. Maskin (1994). "The Folk Theorem with Imperfect Public Information," *Econometrica*, **62**, 997–1040.

[25] Fudenberg, D. and Y. Yamamoto (2012). "The Folk Theorem for Irreducible Stochastic Games with Imperfect Public Monitoring," *Journal of Economic Theory*, **146**, 1664–1683.

[26] Fudenberg, D. and J. Tirole (1991a). *Game Theory*, MIT Press: Cambridge, MA.

[27] Fudenberg, D. and J. Tirole (1991b). "Perfect Bayesian Equilibrium and Sequential Equilibrium," *Journal of Economic Theory*, **53**, 236–260.

[28] Gilbert, E.J. (1959). "On the Identifiability Problem for Functions of Finite Markov Chains," *Annals of Mathematical Statistics*, **30**, 688–697.

[29] Green, E.J. (1987). "Lending and the Smoothing of Uninsurable Income," in Prescott E. and Wallace N., Eds., *Contractual Agreements for Intertemporal Trade*, University of Minnesota Press.

[30] Jackson, M.O. and H.F. Sonnenschein (2007). "Overcoming Incentive Constraints by Linking Decision," *Econometrica*, **75**, 241–258.

[31] Hörner, J., D. Rosenberg, E. Solan and N. Vieille (2010). "On a Markov Game with One-Sided Incomplete Information," *Operations Research*, **58**, 1107–1115.

[32] Hörner, J., T. Sugaya, S. Takahashi and N. Vieille (2011). "Recursive Methods in Discounted Stochastic Games: An Algorithm for $\delta \to 1$ and a Folk Theorem," *Econometrica*, **79**, 1277–1318.

[33] Hörner, J., S. Takahashi and N. Vieille (2012). "On the Limit Perfect Public Equilibrium Payoff Set in Repeated and Stochastic Games," working paper, Yale University.

[34] Iosifescu, M. (1980). *Finite Markov Processes and Their Applications*, Wiley: Chichester, NY.

[35] Kandori, M. (2003). "Randomization, Communication, and Efficiency in Repeated Games with Imperfect Public Information," *Econometrica*, **74**, 213–233.

[36] Kandori, M. and H. Matsushima (1998). "Private Observation, Communication and Collusion," *Econometrica*, **66**, 627–652.

[37] Kosenok, G. and S. Severinov (2008). "Individually Rational, Budget-Balanced Mechanisms and Allocation of Surplus," *Journal of Economic Theory*, **140**, 126–161.

[38] Mezzetti, C. (2004). "Mechanism Design with Interdependent Valuations: Efficiency," *Econometrica*, **72**, 1617–1626.

[39] Mezzetti, C. (2007). "Mechanism Design with Interdependent Valuations: Surplus Extraction," *Economic Theory*, **31**, 473–488.

[40] Myerson, R. (1986). "Multistage Games with Communication," *Econometrica*, **54**, 323–358.

[41] Obara, I. (2008). "The Full Surplus Extraction Theorem with Hidden Actions," *The B.E. Journal of Theoretical Economics*, **8**, 1–8.

[42] Pęski, M. and T. Wiseman (2012). "A Folk Theorem for Stochastic Games with Infrequent State Changes," working paper, University of Toronto.

[43] Puterman, M.L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley: New York, NY.

[44] Radner, R. (1986). "Repeated Partnership Games with Imperfect Monitoring and No Discounting," *Review of Economic Studies*, **53**, 43–57.

[45] Renault, J. (2006). "The Value of Markov Chain Games with Lack of Information on One Side," *Mathematics of Operations Research*, **31**, 490–512.

[46] Renault, J., E. Solan and N. Vieille (2013). "Dynamic Sender-Receiver Games," *Journal of Economic Theory*, forthcoming.

[47] Shapley, L.S. (1953). "Stochastic Games," *Proceedings of the National Academy of Sciences of the U.S.A.*, **39**, 1095–1100.

[48] Wang, C. (1995). “Dynamic Insurance with Private Information and Balanced Budgets,” *Review of Economic Studies*, **62**, 577–595.

[49] Wiseman, T. (2008). “Reputation and Impermanent Types,” *Games and Economic Behavior*, **62**, 190–210.

[50] Zhang, Y. (2009). “Dynamic Contracting, Persistent Shocks and Optimal Taxation,” *Journal of Economic Theory*, **144**, 635–675.

A Proof of Theorem 2

The proof is inspired by FLM but there are significant complications arising from incomplete information. We let Z be a compact set included in the interior of $\mathcal{H}_0$, and pick $\eta > 0$ small enough so that the η-neighborhood $Z_\eta := \{z \in \mathbf{R}^I, d(z, Z) \leq \eta\}$ is also contained in the interior of $\mathcal{H}_0$. We will prove that Z_η is included in the set of perfect Bayesian equilibrium payoffs, for δ large enough

A.1 Preliminaries

The discount factor δ should be chosen large enough so that a few conditions are met. We discuss some of these conditions here. The other ones will appear later.

We quote without proof the following classical result, which relies on the smoothness of Z_η (see Lemma 6 in HSTV for a related statement).

Lemma 8 *Given $\varepsilon > 0$, there exists $\bar{\zeta} > 0$ such that the following holds. For every $z \in Z_\eta$ and every $\zeta < \bar{\zeta}$, there exists a direction $\lambda \in \mathcal{U}_1$ such that if $w \in \mathbf{R}^I$ is such that $\|w - z\| \leq \zeta$ and $\lambda \cdot w \leq \lambda \cdot z - \varepsilon\zeta$, then $w \in Z_\eta$.*

Given $\lambda \in \mathcal{U}_1$, and since Z_η is contained in the interior of $\mathcal{H}_0$, one has $\max_{z \in Z_\eta} \lambda \cdot z < k(\lambda)$. Thus, one can find $v \in \mathbf{R}^I$, and $(\rho, x) \in \mathcal{C}_0$ such that $\max_{z \in Z_\eta} \lambda \cdot z < \lambda \cdot v$ and $\lambda \cdot x(\cdot) < 0$. Using the compactness of $\mathcal{U}_1$, this proves Lemma 9 below.

Lemma 9 *There exists $\varepsilon_0 > 0$ and a finite set $\mathcal{S}_0$ of triples (v, ρ, x) with $v \in \mathbf{R}^I$ and $(\rho, x) \in \mathcal{C}_0$ such that the following holds. For every target payoff $z \in Z_\eta$, and every direction $\lambda \in \mathcal{U}_1$, there is $(v, \rho, x) \in \mathcal{S}_0$ such that (v, ρ, x) is feasible in $\mathcal{P}_0(\lambda)$ and $\lambda \cdot z + \varepsilon_0 < \lambda \cdot v$.*

We let $\kappa_0 \in \mathbf{R}$ be a uniform bound on $\mathcal{S}_0$. Specifically, we pick κ_0 such that $\|x\|_\infty \leq \kappa_0/2$ and $\|z - v\| \leq \kappa_0/2$ for each $(v, x, \rho) \in \mathcal{S}_0$ and every $z \in Z_\eta$. We apply Lemma 8 with $\varepsilon := \varepsilon_0/\kappa_0$ to get $\bar{\zeta}$, and we let $\bar{\delta} < 1$ be large enough so that $\dfrac{(1-\delta)^{1/4}}{\delta} \leq \dfrac{\bar{\zeta}}{\kappa_0}$ for each $\delta \geq \bar{\delta}$.

Let $(\rho, x) \in \mathcal{C}_0$ be arbitrary, and let $\theta_{\rho,x} : \Omega_{\text{pub}} \times S \to \mathbf{R}^I$ denote the relative rents under (ρ, x). Fix a player $i \in I$, a triple $(\bar{\omega}_{\text{pub}}, \bar{s}^i, \bar{a}^i) \in \Omega_{pub} \times S^i \times A^i$, and consider the decision problem $D^i(\bar{\omega}_{\text{pub}}, \bar{s}^i, \bar{a}^i)$.

Given $\tilde{a}^i \in A^i$, we denote by $\tilde{\gamma}^i(\bar{\omega}_{\text{pub}}, s, \tilde{a}^i)$ the (conditional) expected payoff in $D^i(\bar{\omega}_{pub}, \bar{s}^i, \bar{a}^i)$ when states are s, player i plays $\tilde{a}^i$ after reporting truthfully.[22]

Given $s^i \in S^i$, denote by $\gamma^i(\bar{\omega}_{\text{pub}}, \bar{a}^i, (\bar{s}^i, s^i) \to m^i)$ the highest expected payoff of player i in $D^i(\bar{\omega}_{\text{pub}}, \bar{s}^i, \bar{a}^i)$, when the type of player i is s^i and when reporting $m^i \in M^i$.

Since $(\rho, x) \in \mathcal{C}_0$, there exists $\eta_{\rho,x}$ such that

$$\eta_{\rho,x} + \tilde{\gamma}^i(\bar{\omega}_{\text{pub}}, s, \tilde{a}^i) < \tilde{\gamma}^i(\bar{\omega}_{\text{pub}}, s, \rho^i(\bar{\omega}_{\text{pub}}, m)) \tag{6}$$

for every i, $\bar{\omega}_{\text{pub}} \in \Omega_{\text{pub}}$, $s \in S$, $\bar{s}^i \in S^i$, $\tilde{a}^i \neq \rho^i(\bar{\omega}_{\text{pub}}, m)$ (where $m = (\bar{s}, s)$), and

$$\eta_{\rho,x} + \gamma^i(\bar{\omega}_{\text{pub}}, \bar{a}^i, (\bar{s}^i, s^i) \to m^i) < \gamma^i(\bar{\omega}_{\text{pub}}, \bar{a}^i, (\bar{s}^i, s^i) \to (\bar{s}^i, s^i)) \tag{7}$$

whenever $m^i \neq (\bar{s}^i, s^i)$.

We set $\eta := \min_{(v,\rho,x) \in \mathcal{S}_0} \eta_{\rho,x} > 0$.

A.2 Strategies

We let $z_* \in Z_\eta$, and $\delta \geq \bar{\delta}$ be given. We here define a pure strategy profile σ.

Under σ^i, all reports of player i are truthful, and his actions in any given stage n (when reporting truthfully) depend on a target payoff $z_n \in Z_\eta$, on the previous public outcome $\omega_{\text{pub},n-1} \in \Omega_{\text{pub}}$ and on current reports $m_n \in M$. The target payoff z_n is updated in stage n *after* reports have been submitted and the outcome of the public device has been observed.

We first explain this updating process. Given z_n, we pick a unit vector $\lambda_n \in S_1$ using Lemma 8, and use Lemma 9 to pick $(v_n, \rho_n, x_n) \in \mathcal{S}_0$ which is feasible in $\mathcal{P}_0(\lambda_n)$ and such that $\lambda_n \cdot z_n + \varepsilon_0 < \lambda_n \cdot v_n$. Given the public outcome $\omega_{\text{pub},n} = (m_n, y_n)$ and the reports m_{n+1}, z_n is updated to z_{n+1} as follows. We first set

$$w_{n+1} := \frac{1}{\delta} z_n - \frac{1-\delta}{\delta} v_n + \frac{1-\delta}{\delta} x_n(\omega_{\text{pub},n-1}, \omega_{\text{pub},n}, m_{c,n+1}), \tag{8}$$

[22] To be formal, this is the expectation of $r^i(s, \tilde{a}) + x^i(\bar{\omega}_{\text{pub}}, \omega_{\text{pub}}, t^{-i}) + \theta^i_{\rho,r+x}(\omega_{\text{pub}}, t)$, where $m = (\bar{s}, s)$, $\tilde{a} = (\rho^{-i}(m), \tilde{a}^i)$, $y, t) \sim p(\cdot \mid s, \tilde{a})$, and $\omega_{\text{pub}} = (m, y)$.

and we define $\tilde{w}_{n+1}$ by the equation

$$w_{n+1} = \xi \tilde{w}_{n+1} + (1-\xi) z_n, \tag{9}$$

where $\xi := (1-\delta)^{3/4}$.[23]

The randomizing device sets z_{n+1} equal to z_n or to $\tilde{w}_{n+1}$ with respective probabilities $1-\xi$ and ξ. Observe that w_{n+1} is then equal to the expectation of z_{n+1} (where the expectation is over the outcome of the public device in stage $n+1$).

That z_{n+1} then belongs to Z_η follows from the choice of δ and of ξ.

Lemma 10 *One has $\tilde{w}_{n+1} \in Z_\eta$.*

Proof. Omitting the arguments $(\omega_{\text{pub},n-1}, \omega_{\text{pub},n})$, one has

$$\xi(\tilde{w}_{n+1} - z_n) = w_{n+1} - z_n = \frac{1-\delta}{\delta}\left(z_n - v_n + x_n\right).$$

Thus, $\|\tilde{w}_{n+1} - z_n\| \leq \frac{(1-\delta)^{1/4}}{\delta}\kappa_0$ and $\lambda_n \cdot \tilde{w}_{n+1} \leq \lambda_n \cdot z_n - \frac{(1-\delta)^{1/4}}{\delta}\varepsilon_0$, and the result follows since $\delta \geq \bar{\delta}$. ■

We next explain how actions are chosen under σ. Fix a player i, and a private history $h_n^i = ((\omega_{\text{pub},k}, s_k^i, a_k^i)_{k=1,\ldots,n-1}, m_n)$ including reports in stage n.

Whenever the current report of player i is truthful –that is, $m_n^i = (s_{n-1}^i, s_n^i)$–, σ^i plays the action prescribed by ρ_n^i:

$$\sigma^i(h_n^i) = \rho_n^i(\omega_{\text{pub},n-1}, m_n).$$

If h_n^i is consistent with σ^{-i}, then Bayes rule leads player i to assign probability one to $(s_{n-1}^{-i}, s_n^{-i}) = m_n^{-i}$. If h_n^i is inconsistent[24] with σ^{-i}, we let the beliefs of player i be still computed under the assumption that the current reports of $-i$ are truthful.

Thus, at any history h_n^i at which m_n^i is truthful, the expected continuation payoff of player i under σ is well-defined, and it only depends on $(\omega_{\text{pub},n-1}, m_n)$ and on the current payoff target z_n. We denote it by $\gamma_\sigma^i(\omega_{\text{pub},n-1}, m_n; z_n)$.

We now complete the description of σ. Let h_n^i be a (private) history at which m_n^i is *not* truthful: $m_n^i \neq (s_{n-1}^i, s_n^i)$. At such an history, we let σ^i play an action which maximizes the discounted sum of current payoff and expected continuation payoffs, that is,

$$(1-\delta) r^i(s_n, (a^i, a_n^{-i}) + \delta \gamma_\sigma^i(\omega_{\text{pub},n}, m_{n+1}; z_{n+1}),$$

[23] The choice of the exponent 3/4 is to a large extent arbitrary. We will use the fact that ξ vanishes when $\delta \to 1$, more slowly than $1-\delta$, and faster than $\sqrt{1-\delta}$.

[24] Which occurs if past reports are inconsistent, or if observed public signals are inconsistent with reported states.

where $a_n^{-i} = \rho^{-i}(\omega_{\text{pub},n-1}, m_n)$, $(y_n, s_{n+1}) \sim p(\cdot \mid s_n, a^i, a_n^{-i})$, $\omega_{\text{pub},n} = (m_n, y_n)$, $m_{n+1} = (s_n, s_{n+1})$ and the expectation is taken over y_n, m_{n+1} and z_{n+1}.[25]

Theorem 2 follows from Proposition 2, which is proven in the next section.

Proposition 2 *The following holds.*

1. *For δ large enough, σ is a perfect Bayesian equilibrium.*

2. *One has $\lim_{\delta\to 1} \gamma_\sigma(\omega_{\text{pub}}, m; z) = z$ for every $(\omega_{\text{pub}}, m) \in \Omega_{\text{pub}} \times M$ and $z \in Z_\eta$.*

In FLM, the target payoff z is updated every stage. In HSTV, it is only updated periodically, to account for changing states. Here instead, the target payoff is updated at random times. The durations of the successive blocks (during which z is kept constant) are independent, and follow geometric distributions with parameter ξ. As we already noted, the fact that ξ is much larger than $1-\delta$ ensures that successive target payoffs lie in Z_η. The fact that ξ vanishes as $\delta \to 1$ ensures that the expected duration of a block increases to $+\infty$ as $\delta \to 1$.

A.3 Proof of Proposition 2

We will check that player i has no profitable one-step deviation, provided δ is large enough. By construction, this holds at any history h_n^i such that the current report m_n^i is not truthful. At other histories, this sequential rationality claim will follow from the incentive conditions (6) and (7).

The crucial observation is that at any given stage n, expected continuation payoffs under σ are close to the current target z_n, and (continuation) relative rents are close to θ_{ρ_n,x_n}. These properties are established in Proposition 3 below, and hinge on the irreducibility properties of (s_n).

Given an arbitrary target $z \in Z_\eta$, we will denote by $(v, \rho, x) \in \mathcal{S}_0$ the triple that is associated to z in the construction of σ. All computations in this section are done under the assumption that players follow σ. For simplicity, we will write $\gamma_\sigma(\omega_{\text{pub}}, s; z)$ instead of the more cumbersome $\gamma_\sigma(\omega_{\text{pub}}, (m_c, s); z)$. We set

$$\gamma_\sigma(z) := \mathbf{E}_{\mu[\rho]}[\gamma_\sigma(\omega_{\text{pub}}, s; z)],$$

[25] Recall that the belief of player i assigns probability one to $s_n^{-i} = m_{c,n}^i$.

which we interpret as the expected continuation payoff under σ, given the target z. We also set

$$\theta_\sigma(\omega_{\text{pub}}, s; z) := \frac{1}{1-\delta}\left(\gamma_\sigma(\omega_{\text{pub}}, s; z) - \gamma_\sigma(z)\right),$$

which we interpret as the relative rents under σ, when the target is z.

Proposition 3 *There exist positive numbers c_1 and c_2 such that for every target $z \in Z_\eta$, and every discount factor $\delta > \bar{\delta}$, the following holds:*

P1 *: $\|\gamma_\sigma(z) - z\| \leq c_1(1-\delta)^{1/2}$*

P2 *: $\|\theta_\sigma(\cdot; z) - \theta_{\rho,x}\| \leq c_2(1-\delta)^{3/4}$.*

Proof of Proposition 3.

This is the technically more delicate part of the proof, and it can be skipped. We first compare $\gamma_\sigma(\omega_{\text{pub}}, s; z)$ and $\gamma_\sigma(\tilde{\omega}_{\text{pub}}, \tilde{s}; z)$ for arbitrary (ω_{pub}, s) and $(\tilde{\omega}_{\text{pub}}, \tilde{s})$ in $\Omega_{\text{pub}} \times S$. We rely on a coupling argument. Accordingly, we let $(\mathcal{U}, \mathbf{P})$ be a rich enough probability space to accommodate the existence of:

1. two independent Markov chains (ω_n) and $(\tilde{\omega}_n)$ with values in Ω and transition function π_ρ, which start from (ω_{pub}, s) and $(\tilde{\omega}_{\text{pub}}, \tilde{s})$ respectively;[26]

2. a random time τ, independent of the two sequences (ω_n) and $(\tilde{\omega}_n)$, which has a geometric distribution with parameter $\xi(= (1-\delta)^{3/4})$.

The random time τ simulates the stage when the public randomizing device instructs players to switch to the next block. All stochastic processes will be stopped prior to τ, hence (ω_n) and $(\tilde{\omega}_n)$ simulate (coupled) random plays induced by σ starting from (ω_{pub}, s) and $(\tilde{\omega}_{\text{pub}}, \tilde{s})$ respectively.

For $n \geq 1$, we abbreviate to $r_n := r(s_n, a_n)$ and $\tilde{r}_n := r(\tilde{s}_n, \tilde{a}_n)$ the payoffs in stage n along the two plays (where $m_n = (s_{n-1}, s_n)$ and $a_n = \rho(\omega_{\text{pub},n-1}, m_n)$). We denote by $h_n := (\omega_1, \ldots, \omega_{n-1})$ and $\tilde{h}_n := (\tilde{\omega}_1, \ldots, \tilde{\omega}_{n-1})$ the histories associated with the two plays. We also write $x_n := x(\omega_{\text{pub},n-1}, \omega_{\text{pub},n}, s_{n+1})$ and $\tilde{x}_n := x(\tilde{\omega}_{\text{pub},n-1}, \tilde{\omega}_{\text{pub},n}, \tilde{s}_{n+1})$. Finally, z_n and $\tilde{z}_n$ stand for the current target payoff in stage n, while w_n and $\tilde{w}_n$ stand for the expected targets along the two plays, as defined in (8). Observe that $z_n = \tilde{z}_n = z$ for each $n < \tau$, and that z_τ and $\tilde{z}_\tau$ are obtained from z by equation (9).

[26] To be precise, we mean that ω_1 is randomly set to $(s, m, \rho(\omega_{\text{pub}}, m), y)$, where $y \sim q(\cdot \mid s, \rho_*(\omega_{\text{pub}}, m))$ and similarly for $\tilde{\omega}_1$.

We define $\tau_c := \inf\{n : (\omega_{\text{pub},n-1}, s_n) = (\tilde{\omega}_{\text{pub},n-1}, \tilde{s}_n)\}$ to be the first "coincidence" time of the two processes (ω_n) and $(\tilde{\omega}_n)$. By the unichain assumption, τ_c has a finite expectation. We let C_0 be an upper bound for $\mathbf{E}[\tau_c]$, valid for all (ω_{pub}, s) and $(\tilde{\omega}_{\text{pub}}, \tilde{s})$. Since the two stopping times τ and τ_c are independent, this implies the existence of C_1, such that $\mathbf{P}(\tau \le \tau_c) \le C_1 \xi$.[27]

We denote by $\tau_* := \min(\tau, \tau_c)$ the minimum of the switching time and of the first coincidence time. Since σ coincides with ρ prior to τ and since the expected payoff is equal to the discounted sum of current payoffs and of continuation payoffs, one has

$$\gamma_\sigma(\omega_{\text{pub}}, s; z) = \mathbf{E}\left[(1-\delta)\sum_{n=1}^{\tau_*-1} \delta^{n-1} r_n + \delta^{\tau_*-1}\gamma_\sigma(\omega_{\text{pub},\tau_*-1}, s_{\tau_*}; z_{\tau_*})\right], \tag{10}$$

and an analogous formula holds for $\gamma_\sigma(\tilde{\omega}_{\text{pub}}, \tilde{s}; z)$. Hence

$$\begin{aligned} & \gamma_\sigma(\omega_{\text{pub}}, s; z) - \gamma_\sigma(\tilde{\omega}_{\text{pub}}, \tilde{s}; z) \qquad (11)\\ = \ & \mathbf{E}\left[(1-\delta)\sum_{n=1}^{\tau_*-1} \delta^{n-1}(r_n - \tilde{r}_n) + \delta^{\tau_*-1}\left(\gamma_\sigma(\omega_{\text{pub},\tau_*-1}, s_{\tau_*}; z_{\tau_*}) - \gamma_\sigma(\tilde{\omega}_{\text{pub},\tau_*-1}, \tilde{s}_{\tau_*}; \tilde{z}_{\tau_*})\right)\right]. \end{aligned}$$

We provide a first estimate, which will be refined later.

Claim 7 *There is $C_2 > 0$ such that for every $z \in Z_\eta$, every $(\omega_{\text{pub}}, s), (\tilde{\omega}_{\text{pub}}, \tilde{s}) \in \Omega_{\text{pub}} \times S$, and every $\delta > \bar{\delta}$, one has*

$$\|\gamma_\sigma(\omega_{\text{pub}}, s; z) - \gamma_\sigma(\tilde{\omega}_{\text{pub}}, \tilde{s}; z)\| \le C_2 \xi.$$

Proof of Claim 7. Note that $(\omega_{\text{pub},n-1}, s_n, z_n) = (\tilde{\omega}_{\text{pub},n-1}, \tilde{s}_n, \tilde{z}_n)$ on the event $n = \tau_c < \tau$, so that $\gamma_\sigma(\omega_{\text{pub},n-1}, s_n; z_n) = \gamma_\sigma(\tilde{\omega}_{\text{pub},n-1}, \tilde{s}_n; \tilde{z}_n)$. Since payoffs lie in $[0,1]$, equality (11) yields

$$\begin{aligned} \|\gamma_\sigma(\omega_{\text{pub}}, s; z) - \gamma_\sigma(\tilde{\omega}_{\text{pub}}, \tilde{s}; z)\| &= \left\|(1-\delta)\mathbf{E}\left[\sum_{n=1}^{\tau_*-1}(r_n - \tilde{r}_n)\delta^{n-1} + \delta^{\tau_*-1} 1_{\tau \le \tau_c}\right]\right\| \\ &\le (1-\delta)\mathbf{E}[\tau_c] + \mathbf{P}(\tau \le \tau_c). \qquad (12) \end{aligned}$$

The result follows, with $C_2 := C_0 + C_1$. ■

Since $\gamma_\sigma^i(z)$ lies between $\min_{(\omega_{\text{pub}},s)} \gamma_\sigma^i(\omega_{\text{pub}}, s; z)$ and $\max_{(\omega_{\text{pub}},s)} \gamma_\sigma^i(\omega_{\text{pub}}, s; z)$, Claim 7 yields Claim 8 below.

[27] Actually, the inequality holds with $C_1 = \mathbf{E}[\tau_c]$.

Claim 8 *For each z, $(\omega_{\rm pub}, m)$ and $\delta > \bar{\delta}$, one has*

$$\|\gamma_\sigma(\omega_{\rm pub}, s; z) - \gamma_\sigma(z)\| \leq C_2\xi.$$

We now prove **P1** in Proposition 3.

Fix $z \in Z_\eta$. We rewrite (9) which relates the current target z and the expected target w' in the next stage, as

$$z = \delta w' + (1-\delta)v - (1-\delta)x(\bar{\omega}_{\rm pub}, \omega_{\rm pub}, t).$$

When taking expectations under the invariant measure $\mu[\rho] \in \Delta(\Omega \times \Omega \times S)$, and since w' is the expectation of the next target z' under the randomizing device, this latter equality yields

$$z = (1-\delta)v + \mathbf{E}_{\mu[\rho]}\left[\delta z' - (1-\delta)x(\bar{\omega}_{\rm pub}, \omega_{\rm pub}, t)\right].$$

Recall next that $v := \mathbf{E}_{\mu[\rho]}\left[r(s,a) + x(\bar{\omega}_{\rm pub}, \omega_{\rm pub}, t)\right]$. Hence

$$z = \mathbf{E}_{\mu[\rho]}[(1-\delta)r(s,a) + \delta z']. \tag{13}$$

On the other hand, since discounted payoffs are equal to the discounted sum of current and of continuation payoffs, one has

$$\gamma_\sigma(\omega_{\rm pub}, s; z) = (1-\delta)r(s, \rho(\omega_{\rm pub}, m)) + \delta\mathbf{E}\left[\gamma_\sigma(\omega'_{\rm pub}, s'; z')\right],$$

for each $(\omega_{\rm pub}, s)$. Taking expectations under the invariant measure, one gets

$$\gamma_\sigma(z) = \mathbf{E}_{\mu[\rho]}\left[(1-\delta)r(s,a) + \delta\gamma_\sigma(\omega'_{\rm pub}, s'; z')\right]. \tag{14}$$

Denote by A the event where the public randomization device tells players to continue with the current block. Since $z' = z$ on A, one has

$$\begin{aligned}
\mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(\omega'_{\rm pub}, s'; z')1_A\right] &= \mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(\omega'_{\rm pub}, s'; z)1_A\right] && (15)\\
&= \mathbf{P}(A)\mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(\omega'_{\rm pub}, s'; z)\right] = \mathbf{P}(A)\gamma_\sigma(z) && (16)\\
&= \mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(z')1_A\right]. && (17)
\end{aligned}$$

where the second equality holds since the event A and the pair $(\omega'_{\rm pub}, m')$ are independent.

The complement event $\bar{A}$ is of probability ξ. Denoting by $(v', \rho', x') \in \mathcal{S}_0$ the triple associated to the new target z', one has using Claim 7,

$$\|\gamma_\sigma(\omega'_{\rm pub}, s', z') - \gamma_\sigma(z')\| \leq C_2\xi,$$

for every realization of (ω', s', z').

Since $\gamma_\sigma(z') := \mathbf{E}_{\mu[\rho']}\left[\gamma_\sigma(\omega'_{\rm pub}, s'; z')\right]$, one obtains

$$\|\mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(\omega'_{\rm pub}, s'; z')1_B\right] - \mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(z')1_B\right]\| \le C_2\xi\mathbf{P}(B) = C_2\xi^2. \tag{18}$$

Plugging (17) and (18) into (14), this yields

$$\|\gamma_\sigma(z) - \left(\mathbf{E}_{\mu[\rho]}\left[(1-\delta)r(s,a) + \delta\gamma_\sigma(z')\right]\right)\| \le C_2\xi^2.$$

Combining this inequality with (13), one gets

$$\|\gamma_\sigma(z) - z\| \le \delta\mathbf{E}_{\mu[\rho]}\left[\|\gamma_\sigma(z') - z'\|\right] + C_2\xi^2.$$

Setting $S := \sup_{z\in Z_\eta}[\|\gamma_\sigma(z) - z\|$, this implies in turn $S \le \delta S + C_2\xi^2$, so that $S \le C_2\dfrac{\xi^2}{1-\delta} = C_2(1-\delta)^{1/2}$. This yields **P1** (with $c_1 := C_2$).

We turn to the proof of **P2**. We let $(\omega_{\rm pub}, s)$ and $(\tilde\omega_{\rm pub}, \tilde s)$ in $\Omega_{\rm pub}\times S$ be given, and use the coupling introduced earlier. We proceed in two steps, Claims 9 and 10 below.

Set $\Delta_n := (r_n + x_n) - (\tilde r_n + \tilde x_n)$.

Claim 9 *There is $C_3 > 0$ such that for every $\delta > \bar\delta$, one has*

$$\left\|\frac{\gamma_\sigma(\omega_{\rm pub}, s; z) - \gamma_\sigma(\tilde\omega_{\rm pub}, \tilde s; z)}{1-\delta} - \mathbf{E}\left[\sum_{n=1}^{\tau_*-1}\delta^{n-1}\Delta_n\right]\right\| \le C_3(1-\delta)^{1/4}.$$

Claim 10 *There is $C_4 > 0$ such that for every $\delta \ge \bar\delta$, one has*

$$\left\|\mathbf{E}\left[\sum_{n=1}^{\tau_*-1}\delta^{n-1}\Delta_n\right] - \mathbf{E}\left[\sum_{n=1}^{\tau_c-1}\Delta_n\right]\right\| \le C_4\xi.$$

(Observe that the range of the two sums is not the same.) Since $\mathbf{E}\left[\sum_{n=1}^{\tau_c-1}\Delta_n\right]$ is equal to $\theta_{\rho,x}(\omega_{\rm pub}, s) - \theta_{\rho,x}(\tilde\omega_{\rm pub}, \tilde s)$, Statement **P2** follows from Claims 9 and 10, with $c_2 := C_3 + C_4$.

Proof of Claim 9. If $\tau_c < \tau$, then $(\omega_{{\rm pub},\tau_*-1}, s_{\tau_*}) = (\tilde\omega_{{\rm pub},\tau_*-1}, \tilde s_{\tau_*})$, and $z_{\tau_*} = \tilde z_{\tau_*} = z$, hence

$$\gamma_\sigma(\omega_{{\rm pub},\tau_*-1}, s_{\tau_*}; z_{\tau_*}) - \gamma_\sigma(\tilde\omega_{{\rm pub},\tau_*-1}, \tilde s_{\tau_*}; \tilde z_{\tau_*}) = 0 = z_{\tau_*} - \tilde z_{\tau_*}.$$

If instead $\tau \le \tau_c$, then by Claim 8, $\|\gamma_\sigma(\omega_{{\rm pub},\tau_*-1}, s_{\tau_*}; z_{\tau_*}) - \gamma_\sigma(z_{\tau_*})\| \le C_2\xi = C_2(1-\delta)^{\frac{3}{4}}$ and, by **P1**, $\gamma_\sigma(z_{\tau_*})$ is within $C_2(1-\delta)^{\frac{1}{2}}$ of z_{τ_*}. Hence, $\gamma_\sigma(\omega_{{\rm pub},\tau_*-1}, m_{\tau_*}; z_{\tau_*})$ differs from

z_{τ_*} by at most $2C_2(1-\delta)^{1/2}$, and a similar result holds for $\gamma_\sigma(\tilde\omega_{\mathrm{pub},\tau_*-1},\tilde s_{\tau_*};\tilde z_{\tau_*})$. Hence the difference $(\gamma_\sigma(\omega_{\mathrm{pub},\tau_*-1},s_{\tau_*};z_{\tau_*})-\gamma_\sigma(\tilde\omega_{\mathrm{pub},\tau_*-1},\tilde s_{\tau_*};\tilde z_{\tau_*}))$ is equal to the difference $(z_{\tau_*}-\tilde z_{\tau_*})$, up to $4C_2(1-\delta)^{1/2}$. Since $\mathbf{P}(\tau\le\tau_c)\le C_1\xi$, it follows that

$$\begin{aligned}&\left\|\mathbf{E}\left[\gamma_\sigma(\omega_{\mathrm{pub},\tau_*-1},s_{\tau_*};z_{\tau_*})-\gamma_\sigma(\tilde\omega_{\mathrm{pub},\tau_*-1},\tilde s_{\tau_*};\tilde z_{\tau_*})\right]-\mathbf{E}[z_{\tau_*}-\tilde z_{\tau_*}]\right\|\\ \le\ & 4C_2(1-\delta)^{1/2}\times C_1\xi=4C_1C_2(1-\delta)^{5/4}.\end{aligned}\tag{19}$$

Observe next that $z_{n+1}=\tilde z_{n+1}(=z)$ for each $n<\tau_*-1$. So that plugging (19) into (11), the difference $\gamma_\sigma(\omega_{\mathrm{pub}},s;z)-\gamma_\sigma(\tilde\omega_{\mathrm{pub}},\tilde s;z)$ is equal to

$$\mathbf{E}\left[\sum_{n=1}^{\tau'-1}\delta^{n-1}\left((1-\delta)(r_n-\tilde r_n)+\delta(z_{n+1}-\tilde z_{n+1})\right)\right],\tag{20}$$

up to $4C_1C_2(1-\delta)^{5/4}$.

Next, we rewrite

$$\sum_{n=1}^{\tau'-1}\delta^n(z_{n+1}-\tilde z_{n+1})=\sum_{n=1}^{\infty}\delta^n(z_{n+1}-\tilde z_{n+1})1_{\tau'\ge n+1}.$$

Since the random time τ is independent of the plays (ω_n) and $(\tilde\omega_n)$, one has for each stage $n\ge 2$,

$$\mathbf{E}\left[(z_n-\tilde z_n)1_{\tau'\ge n}\right]=\mathbf{E}\left[(w_n-\tilde w_n)1_{\tau'\ge n}\right]=\mathbf{E}\left[\frac{1-\delta}{\delta}\left(x_{n-1})-\tilde x_{n-1})\right)\right].$$

Therefore,

$$\mathbf{E}\left[\sum_{n=1}^{\tau'-1}\delta^n(z_{n+1}-\tilde z_{n+1})\right]=\mathbf{E}\left[\sum_{n=1}^{\tau'-1}\delta^{n-1}(1-\delta)\left(x_n-\tilde x_n\right)\right].$$

The result follows when plugging the latter equation into (20), with $C_3:=4C_1C_2$. ■

Proof of Claim 10. Observe first that, since $\|\Delta_n\|\le(1+\kappa_0)$, the difference between $\mathbf{E}\left[\sum_{n=1}^{\min(\tau,\tau_c)-1}\delta^{n-1}\Delta_n\right]$ and $\mathbf{E}\left[\sum_{n=1}^{\tau_c-1}\delta^{n-1}\Delta_n\right]$ is at most $(1+\kappa_0)\mathbf{E}[\tau_c-\min(\tau,\tau_c)]$. Since τ and τ_c are independent, one has

$$\mathbf{E}[\tau_c-\min(\tau,\tau_c)]=\mathbf{E}[\tau_c-\tau\mid\tau<\tau_c]\times\mathbf{P}(\tau<\tau_c)\le C_0\times C_1\xi.\tag{21}$$

Next, note that, since $(1-\delta^{n-1})\le n(1-\delta)$, one has

$$\left\|\mathbf{E}\left[\sum_{n=1}^{\tau_c-1}\delta^{n-1}\Delta_n\right]-\mathbf{E}\left[\sum_{n=1}^{\tau_c-1}\Delta_n\right]\right\|\le(1+\kappa_0)(1-\delta)\mathbf{E}\left[\sum_{n=1}^{\tau_c-1}n\right]\le(1-\delta)\times(1+\kappa_0)\mathbf{E}\left[\tau_c^2\right].\tag{22}$$

Collecting (21) and (22), there exists $c_4 > 0$ such that

$$\left\| \mathbf{E}\left[\sum_{n=1}^{\tau_*-1} \delta^{n-1}\Delta_n\right] - \mathbf{E}\left[\sum_{n=1}^{\tau_c-1} \Delta_n\right] \right\| \leq c_4\xi. \tag{23}$$

■

This concludes the proof of Proposition 3. ■

A.4 Conclusion

We now check no player i has no profitable one-step deviation. For concreteness, we will focus on deviations at the report stage. Apart from notational issues, the proof is similar for a deviation at an action stage. Consider thus a stage n and any private history $(h_{\text{pub},n}, (s^i_k)_{1\leq k\leq n-1}, (a^i_k)_{1\leq k\leq n-1}, s^i_n)$ of player i.

The continuation payoff of player i only depends on this private history through the previous public outcome $\omega_{\text{pub},n-1}$, on $(s^i_{n-1}, a^i_{n-1}, s^i_n)$ and on the target z_n in stage n. We denote by $(v, \rho, x) \in \mathcal{S}_0$ the triple associated to z_n. We will derive the no profitable deviation property from the incentive properties of (ρ, x) in $D^i_{\rho,x}(\omega_{\text{pub},n-1}, s^i_{n-1}, a^i_{n-1})$ and on the estimates **P1** and **P2** obtained earlier.

We will compare the expected payoff obtained when reporting truthfully $m^i_n = (s^i_{n-1}, s^i_n)$ then playing $\rho^i(m_n)$ with the expected payoff obtained when reporting $\tilde{m}^i_n \neq (s^i_{n-1}, s^i_n)$ then playing some arbitrary action $a^i(m^{-i}_n, \tilde{m}^i_n)$.

In the former case, player i's payoff is the expectation of

$$(1-\delta)r^i(s_n, a_n) + \delta\gamma^i_\sigma(\omega_{\text{pub},n}, s_{n+1}; z_{n+1}),$$

where s^{-i}_n is drawn according to i's belief, $m_n = (s_{n-1}, s_n)$, $a_n = \rho(\omega_{\text{pub},n-1}, m_n)$, $(y_n, s_{n+1}) \sim p(\cdot \mid s_n, a_n)$, $\omega_{\text{pub},n} = (m_n, y_n)$, and z_{n+1} is randomly chosen, using equations (8) and (9).

In the latter case, player i's payoff is the expectation of

$$(1-\delta)r^i(s_n, \tilde{a}_n) + \delta\gamma^i_\sigma(\tilde{\omega}_{\text{pub},n}, \tilde{s}_{n+1}; \tilde{z}_{n+1}),$$

with obvious notations.

We compare these two expectations using again a coupling argument. Specifically, we will assume that the "same" randomizing device is used for both computations. Thus, with probability $1-\xi$, one has $z_{n+1} = \tilde{z}_{n+1} = z$ and $\gamma_\sigma(\omega_{\text{pub},n}, s_{n+1}; z_{n+1}) - \gamma_\sigma(\tilde{\omega}_{\text{pub},n}, \tilde{s}_{n+1}; \tilde{z}_{n+1})$ is then, by **P2**, equal to

$$(z_{n+1} + (1-\delta)\theta_{\rho,x}(\omega_{\text{pub},n}, s_{n+1}; z_{n+1})) - (\tilde{z}_{n+1} + (1-\delta)\theta_{\rho,x}(\tilde{\omega}_{\text{pub},n}, \tilde{s}_{n+1}; \tilde{z}_{n+1})) + o(1-\delta).$$

On the other hand, with probability ξ, the public device instructs to switch to a new block, and the difference $\gamma_\sigma(\omega_{\text{pub},n}, s_{n+1}; z_{n+1}) - \gamma_\sigma(\tilde{\omega}_{\text{pub},n}, \tilde{s}_{n+1}; \tilde{z}_{n+1})$ is then, by **P1**, equal to

$$(z_{n+1} + (1-\delta)\theta_{\rho,x}((s_n, y_n), s_{n+1}; z_{n+1})) - \left(\tilde{z}_{n+1} + (1-\delta)\theta_{\rho,x}((m_n^i, s_n^{-i}), \tilde{y}_n), \tilde{s}_{n+1}; \tilde{z}_{n+1})\right) + 0((1-\delta)^{\frac{1}{2}}).$$

Taking expectations, (A.4) and (A.4) yield

$$\begin{aligned}
&\|\mathbf{E}\left[\gamma_\sigma(\omega_{\text{pub},n}, s_{n+1}; z_{n+1})\right] - \mathbf{E}\left[\gamma_\sigma(\tilde{\omega}_{\text{pub},n}, \tilde{s}_{n+1}; \tilde{z}_{n+1})\right]\| + O((1-\delta)^{3/2}) \\
= &\left\|\mathbf{E}\left[z_{n+1} + (1-\delta)\theta_{\rho,x}(\omega_{\text{pub},n}, s_{n+1}; z_{n+1})\right] - \mathbf{E}\left[\tilde{z}_{n+1} + (1-\delta)\theta_{\rho,x}((m_n^i, s_n^{-i}), \tilde{y}_n), \tilde{s}_{n+1}; \tilde{z}_{n+1})\right]\right\|.
\end{aligned}$$

Note now that the following identity holds by construction:

$$\begin{aligned}
\delta\mathbf{E}\left[z_{n+1}\right] - \delta\mathbf{E}\left[\tilde{z}_{n+1}\right] &= \delta\mathbf{E}\left[w_{n+1}\right] - \delta\mathbf{E}\left[\tilde{w}_{n+1}\right] \\
&= (1-\delta)\mathbf{E}[x_n] - (1-\delta)\mathbf{E}[\tilde{x}_n].
\end{aligned}$$

Thus, when divided by $(1-\delta)$, the difference in expected continuation payoffs (when reporting truthfully or not) is equal to

$$\mathbf{E}\left[r_n + x_n + \theta_{\rho,x}(\omega_{\text{pub},n}, s_{n+1})\right] - \mathbf{E}\left[\tilde{r}_n + \tilde{x}_n + \theta_{\rho,x}(\tilde{\omega}_{\text{pub},n}, \tilde{s}_{n+1})\right] + o(1). \tag{24}$$

To conclude, note that the difference of the expectations which appear in (24) is at least η_1 (see Section A.1) hence the continuation payoff is strictly higher when reporting truthfully than when not, provided δ is high enough.

B Proofs for independent private values

Proof of Lemma 3. Fix a non-coordinate direction λ. Suppose first that $\lambda^i \leq 0$, all i. Consider the vector $v \in Ext^{po}$ that maximizes $\lambda \cdot v$, and the corresponding policy ρ. This policy implements a distribution over $\Delta(A)$. Consider the constant (and hence weakly admissible, for $x = 0$) policy that uses the public randomization device to replicate this distribution (independently of the announcements). The IPV assumption ensures that all players are weakly worse off. Hence $k^*(\lambda) \geq \lambda \cdot v$, so that $v \in \mathcal{H}_0^*(\lambda)$, and so $co(Ext^{pu} \cup Ext^{po}) \subset \mathcal{H}_0^*(\lambda)$ (if another constant policy improves the score further, consider it instead). Suppose next that $\lambda^i < 0$ for all $i \in J \subsetneq I$, and there exists i such that $\lambda^i > 0$. Again, consider the vector $v \in Ext^{po}$ that maximizes $\lambda \cdot v$. Because $v \in Ext^{po}$, v also maximizes $\hat{\lambda} \cdot v$ over $v \in Ext^{po}$, for some $\hat{\lambda} \geq 0$, $\hat{\lambda} \in S_1$. Furthermore, we can take $\hat{\lambda}^i = 0$ for all $\lambda^i < 0$. Such a vector is achieved by a policy that only depends on $(s_i)_{i \notin J}$, because of private values. Truth-telling is trivial for these types, and hence this policy is also weakly admissible in the direction

λ. Hence again $k^*(\lambda) \geq \lambda \cdot v$, so that $v \in \mathcal{H}_0^*(\lambda)$ and so $co(Ext^{pu} \cup Ext^{po}) \subset \mathcal{H}_0^*(\lambda)$. Directions $\lambda \geq 0$ are unproblematic, as both efficient and constant policies are weakly admissible for some choice of x.

Coordinate directions $\lambda = e^i$ are also immediate: the vector that maximizes v^i over $co(Ext^{pu} \cup Ext^{po})$ is part of the Pareto-frontier, and the corresponding policy is weakly admissible using AGV. Scores in the directions $\lambda = -e^i$ are (at least) $-\underline{v}^i$, hence $V^* \cap co(Ext^{pu} \cup Ext^{po}) \subset \mathcal{H}_0^*(-e^i)$. ■

B.1 Proof of Proposition 1

Given an action plan $\rho : S \to A$ and $x : \Omega_{\text{pub}} \times M \to \mathbf{R}^I$, we denote the induced relative rents as $\theta[\rho, r + x] : \Omega_{\text{pub}} \times S \to \mathbf{R}^I$ instead of $\theta[\rho, x]$. In other words, $\theta[\rho, x]$ are the relative rents when transfers are x and stage payoffs are identically zero. When ρ is clear from the context, we write θ_x.

(Assuming truth-telling), any action plan $\rho : S \to A$ induces a probability transition over Ω_{pub}, denoted π_ρ. We use the notation π_ρ whenever distributions should be computed under the assumption that states are truthfully reported, actions chosen according to ρ, and transitions determined using p. For instance, $\pi_\rho(s^{-i} \mid \bar{\omega}_{\text{pub}})$ is the (conditional) distribution of s^{-i} under $p(\cdot \mid \bar{s}, \rho(\bar{s}))$, given $\bar{y}$. Given the IPV assumption, it is thus $\prod_{j \neq i} p^j(s^j \mid \bar{s}^i, \rho^j(\bar{s}), \bar{y})$.

Fix a weakly truthful pair (ρ, x), with $\rho : S \to A$ and $x : \Omega_{\text{pub}} \times S \to \mathbf{R}^I$. For $i \in I$, $(\bar{\omega}_{\text{pub}}, s^i) \in \Omega_{\text{pub}} \times S^i$, set

$$\xi^i(\bar{\omega}_{\text{pub}}, s^i) := \mathbf{E}_{s^{-i} \sim \pi_\rho(\cdot \mid \bar{\omega}_{\text{pub}})}[x^i(\bar{\omega}_{\text{pub}}, s^i, s^{-i})].$$

Plainly, the pair (ρ, ξ) is weakly truthful as well.

The crucial lemma is the next one. It is the long-run analog of Claim 1 in Athey and Segal (AS).

Lemma 11 *Define* $\tilde{x} : \Omega_{\text{pub}} \times S \to \mathbf{R}^I$ *by*

$$\tilde{x}^i(\bar{\omega}_{\text{pub}}, s)(= \tilde{x}^i(\bar{\omega}_{\text{pub}}, s^i)) = \theta_{\xi^i}(\bar{\omega}_{\text{pub}}, s^i) - \mathbf{E}_{\tilde{s}^i \sim \pi_\rho(\cdot \mid \bar{\omega}_{\text{pub}})}[\theta_{\xi^i}(\bar{\omega}_{\text{pub}}, \tilde{s}^i)].$$

Then $(\rho, \tilde{x})$ *is weakly truthful.*

Proof.

We first argue that $\theta^i_{\tilde{x}}(\cdot) = \tilde{x}^i(\cdot)$ (up to an additive constant, as usual). It suffices to prove that $\tilde{x}^i$ is a solution to the linear system (with unknowns θ)

$$\theta(\bar{\omega}_{\text{pub},1}, s_1) - \theta(\bar{\omega}_{\text{pub},2}, s_2) = \tilde{x}^i(\bar{\omega}_{\text{pub},1}, s_1) + \mathbf{E}_{\pi_\rho(\cdot|\bar{\omega}_{\text{pub},1}, s_1)}[\theta(\omega, t)] - \left(\tilde{x}^i(\bar{\omega}_{\text{pub},2}, s_2) + \mathbf{E}_{\pi_\rho(\cdot|\bar{\omega}_{\text{pub},2}, s_2)}[\theta(\omega, t)]\right),$$

(for $(\bar{\omega}_{\text{pub}_1}, s_1), (\bar{\omega}_{\text{pub},2}, s_2) \in \Omega_{\text{pub}} \times S^i$). But this follows from the fact that for each $(\bar{\omega}_{\text{pub}}, s) \in \Omega_{\text{pub}} \times S$, one has

$$\mathbf{E}_{\pi_\rho(\cdot|s)}[\tilde{x}^i(\omega_{\text{pub}}, t^i)] = \mathbf{E}_{y \sim \pi_\rho(\cdot|s)}\left[\mathbf{E}_{t \sim \pi_\rho(\cdot|\omega_{\text{pub}})} \tilde{x}^i(\omega_{\text{pub}}, t^i)\right] = 0.$$

Fix next a player i, and an outcome $\bar{\omega} = (\bar{s}, \bar{m}, \bar{s}, \bar{y})$ such that $\bar{s}^{-i} = \bar{m}^{-i}$ and $\bar{a}^{-i} = \rho^{-i}(\bar{m})$. Since (ρ, ξ) is weakly truthful, for each such $\bar{\omega}$ and $s^i \in S^i$, the expectation of

$$r^i(s^i, \rho(s^{-i}, m^i)) + \xi^i(\bar{\omega}_{\text{pub}}, m^i) + \theta_{r^i}(\omega_{\text{pub}}, t) + \theta_{\xi^i}(\omega_{\text{pub}}, t) \tag{25}$$

is maximized for $m^i = s^i$. Here, the expectation is to be computed as follows. First, s^{-i} is drawn according to the belief of i which, given the IPV assumption, is equal to $\pi_\rho(\cdot \mid \bar{\omega}_{\text{pub}})$; next, (y, t) is drawn $\sim p(\cdot \mid s, \rho(s^{-i}, m^i))$, and $\omega_{\text{pub}} = (s^{-i}, m^i, y)$.

To prove that $(\rho, \tilde{x})$ is admissible, we need to prove that the similar expectation of

$$r^i(s^i, \rho(s^{-i}, m^i) + \tilde{x}^i(\bar{\omega}_{\text{pub}}, m^i) + \theta_{r^i}(\omega_{\text{pub}}, t) + \theta_{\tilde{x}^i}(\omega_{\text{pub}}, t) \tag{26}$$

is maximized for $m^i = s^i$ as well. Fix $m^i \in M^i$. Using $\theta^i_{\tilde{x}} = \tilde{x}^i$, and the definition of $\tilde{x}^i$, the expectation of the expression in (26) is equal to the expectation of

$$r^i(s^i, \rho(s^{-i}, m^i)) + \theta_{\xi^i}(\bar{\omega}_{\text{pub}}, m^i) + \theta_{r^i}(\omega_{\text{pub}}, t) + \theta_{\xi^i}(\omega_{\text{pub}}, t^i) - \mathbf{E}_{\tilde{s}^i \sim \pi_\rho(\cdot|\omega_{\text{pub}})} \theta_{\xi^i}(\omega_{\text{pub}}, \tilde{s}^i), \tag{27}$$

up to the constant $\mathbf{E}_{\tilde{s}^i \sim \pi_\rho(\cdot|\bar{\omega}_{\text{pub}})} \theta_{\xi^i}(\bar{\omega}_{\text{pub}}, \tilde{s}^i)$, which does not depend on m^i.

Next, observe that by definition of θ_{ξ^i}, one has

$$\theta_{\xi^i}(\bar{\omega}_{\text{pub}}, m^i) = \xi^i(\bar{\omega}_{\text{pub}}, m^i) + \mathbf{E}_{\pi_\rho} \theta_{\xi^i}(\omega_{\text{pub}}, \tilde{s}^i),$$

again up to a constant that does not depend on m^i.

Thus, (and up to a constant), the expression in (26) has the same expectation as the expression in (25), so that the weak truthfulness of $(\rho, \tilde{x})$ follows from that of (ρ, ξ).

■

Corollary 11 *Let $\mu_{ij} \in \mathbf{R}$ be arbitrary. For $i \in I$, set*

$$\hat{x}^i(\bar{\omega}_{\text{pub}}, m) = \tilde{x}^i(\bar{\omega}_{\text{pub}}, m^i) + \sum_{j \neq i} \mu_{ij} \tilde{x}^j(\bar{\omega}_{\text{pub}}, m^j).$$

Then $(\rho, \bar{x})$ is weakly truthful.

Proof. It is enough to check that, at any $\bar{\omega}_{\text{pub}}$, the expectation of $\theta_{\tilde{x}^j}(\omega_{\text{pub}}, \tilde{s}^j) = \tilde{x}^j(\omega_{\text{pub}}, \tilde{s}^j)$ does not depend on m^i. But this expectation is zero (as in Claim 2 in AS). ■

Proposition 4 *Fix a non-coordinate direction λ, that is, a direction such that $I(\lambda) := \{i : \lambda^i \neq 0\}$ is not a singleton. Let (ρ, x) be a weakly truthful pair. Then there exists $\hat{x}$ such that $(\rho, \hat{x})$ is weakly truthful and $\lambda \cdot \hat{x}(\cdot) = 0$.*

Proof. Set $\hat{x}^i = \tilde{x}^i$ for $i \notin I(\lambda)$. For $i \in I(\lambda)$ set

$$\hat{x}^i(\bar{\omega}_{\text{pub}}, s) := \tilde{x}^i(\bar{\omega}_{\text{pub}}, s^i) - \frac{1}{|I(\lambda)| - 1} \sum_{j \neq i} \frac{\lambda^j}{\lambda^i} \tilde{x}^j(\bar{\omega}_{\text{pub}}, s^j)$$

and apply the previous corollary. ■

Proof of Lemma 4, 5. We focus on a fixed player i. Fix a weakly truthful (ρ, x). The optimality of truth-telling given (ρ, x) is equivalent to truth telling solving the following Markov decision process:

- The state space is $(\Omega_{\text{pub}}, S^i)$, with elements $(\bar{\omega}_{\text{pub}}, s^i)$: last period's public outcome and today's private state;

- The action set is $M^i = S^i$: today's announced state;

- The reward is $r^i(\rho(s^{-i}, m^i), s^i) + x^i(\bar{\omega}_{\text{pub}}, \omega_{\text{pub}})$ (we drop t^{-i} from the arguments of x^i as there is no gain to include such an argument under IPV);

- Transitions are given by $p(\omega_{\text{pub}}, t^i \mid \bar{\omega}_{\text{pub}}, s^i, m^i) = 0$ if ω_{pub} does not specify m^i as i's report, or does not specify a signal y in the support of the distribution determined by $\rho(s^{-i}, m^i)$. Otherwise, it is derived from p and ρ in the obvious way.

Let us define, as in Section 6,

$$r^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) = \mathbf{E}_{s^{-i}|\bar{\omega}_{\text{pub}}} r^i(\rho(s^{-i}, m^i), s^i), x^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) = \mathbf{E}_{s^{-i}, y|\bar{\omega}_{\text{pub}}, \rho} x^i(\bar{\omega}_{\text{pub}}, \omega_{\text{pub}}).$$

Under our unichain assumption, there is an equivalent LP formulation (see Puterman, Ch. 8.8). Namely, agent i maximizes

$$\sum_{(\bar{\omega}_{\text{pub}}, s^i, m^i)} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) \left(r^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) + x^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) \right),$$

subject to $\pi^i_\rho \in \hat{\Pi}^i(\rho)$, defined as the set of $\{\pi(\bar{\omega}_{\text{pub}}, s^i, m^i) : (\bar{\omega}_{\text{pub}}, s^i, m^i) \in \Omega_{\text{pub}} \times S^i \times S^i\}$, $\pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) \geq 0$, $\sum_{(\bar{\omega}_{\text{pub}}, s^i, m^i)} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) = 1$, and for all states $(\omega_{\text{pub}}, t^i)$,

$$\sum_{\tilde{m}^i} \pi^i_\rho(\omega_{\text{pub}}, t^i, \tilde{m}^i) = \sum_{\bar{\omega}_{\text{pub}}, s^i} p(\omega_{\text{pub}}, t^i \mid \bar{\omega}_{\text{pub}}, s^i, m^i) \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i).$$

This is equation 3 from Section 6. Our goal is to examine which ρ are weakly truthful for some choice of x^i_ρ.

Consider first Lemma 4, *i.e.* the case in which $p(y \mid s, a)$ is independent of s. Then $x^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i)$ is a function of $\bar{\omega}_{\text{pub}}, m^i$ only, so we write $x^i_\rho(\bar{\omega}_{\text{pub}}, m^i)$. We can consider the zero-sum game in which we pick x^i_ρ in some large but bounded set $[-M, M]$ to minimize

$$\sum_{(\bar{\omega}_{\text{pub}}, s^i, m^i)} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) r^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) + \sum_{(\bar{\omega}_{\text{pub}}, s^i, m^i)} (\pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) - \mu(\bar{\omega}_{\text{pub}}, m^i)) x^i_\rho(\bar{\omega}_{\text{pub}}, m^i).$$

This is a game between player i who chooses $\pi^i_\rho \in \hat{\Pi}^i(\rho)$ and the designer who chooses $x^i_\rho \in [-M, M]^{\Omega_{\text{pub}} \times S^i}$. By the minimax theorem, we can think of i moving first, and it is then clear that any optimal strategy for i must specify, for all $(\bar{\omega}_{\text{pub}}, m^i)$,

$$\sum_{s^i} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) = \mu(\bar{\omega}_{\text{pub}}, m^i),$$

and we can pick x^i_ρ to be the optimal strategy of the designer. Thus, we may restrict further player i to choose from $\Pi^i(\rho)$ where $\pi^i_\rho \in \Pi^i(\rho)$ if and only if $\pi^i_\rho \in \hat{\Pi}^i(\rho)$ and $\sum_{s^i} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) = \mu(\bar{\omega}_{\text{pub}}, m^i)$ for all $(\bar{\omega}_{\text{pub}}, m^i)$. Note that for $\pi^i_\rho \in \Pi^i(\rho)$, the original objective of the LP becomes

$$\sum_{(\bar{\omega}_{\text{pub}}, s^i, m^i)} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) r^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) + \sum_{\bar{\omega}_{\text{pub}}, m^i} \mu(\bar{\omega}_{\text{pub}}, m^i) \hat{x}^i(\bar{\omega}_{\text{pub}}, s^i, m^i),$$

and because the second term does not involve π^i_ρ, it is irrelevant for the maximization. Hence, transfers cannot achieve more than the restriction to $\Pi^i(\rho)$, and the policy ρ is weakly truthful if and only if the solution to the program $\hat{\mathcal{P}}^i(\rho)$: maximize

$$\sum_{(\bar{\omega}_{\text{pub}}, s^i, m^i)} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) r^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i)$$

over $\pi^i_\rho \in \Pi^i(\rho)$ is given by $\pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) = 0$ if $m^i \neq s^i$, all $(\bar{\omega}_{\text{pub}}, s^i)$ (and so $\pi^i_\rho(\bar{\omega}_{\text{pub}}, m^i, m^i) = \mu(\bar{\omega}_{\text{pub}}, m^i)$). Invoking Proposition 1 to balance the budget, this concludes the proof of Lemma 4 for non-coordinate directions. In coordinate directions $\lambda = \pm e^i$, note that setting

$x = 0$ provides the appropriate truth telling incentives (the only player who makes a report that affects the outcome is player i when $\lambda = e^i$, but in this case the policy $\rho[e^i]$ trivially is a solution to the LP for $x^i_\rho = 0$.

Consider next the case of state-dependent signaling, as defined in Section 6.3. Then x^i is a function of $(\bar{\omega}_{\text{pub}}, [s^i], m^i)$ only. Then we consider the zero-sum game with payoff

$$\sum_{(\bar{\omega}_{\text{pub}}, s^i, m^i)} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) r^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i)$$
$$+ \sum_{(\bar{\omega}_{\text{pub}}, s^i, m^i)} (\pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) - \mu(\bar{\omega}_{\text{pub}}, m^i)) x^i_\rho(\bar{\omega}_{\text{pub}}, [s^i], m^i),$$

and it follows from the minimax theorem again that player i's optimal strategy specifies, for all $(\bar{\omega}_{\text{pub}}, m^i, [m^i])$,

$$\sum_{s^i \in [m^i]} \pi^i_\rho(\bar{\omega}_{\text{pub}}, s^i, m^i) = \mu(\bar{\omega}_{\text{pub}}, m^i, [m^i]).$$

This is (4'). The remainder follows as in the first case. ■

Proof of Lemma 6 and 7. We have shown that the policies $\rho[\lambda] \in \Xi$ are weakly truthful (see Lemma 5). Now we must show that we can make the truth-telling incentives strict. This is where we use Assumption **1**.

Consider the MDP $\mathcal{M}(\lambda)$: maximize over ρ $\mathbf{E}_{\mu[\rho]}[\lambda \cdot r(s, \rho(s))]$.

Given two pure stationary policies ρ and ρ', let $\Lambda(\rho, \rho')$ denote the set of λ such that both ρ and ρ' are optimal in $\mathcal{M}(\lambda)$. Thanks to the ACOE, the set $\Lambda(\rho, \rho')$ is closed and convex.

The map $\lambda \mapsto \mathbf{E}_{\mu[\rho]}[\lambda \cdot r(s, \rho(s))] - \mathbf{E}_{\mu[\rho']}[\lambda \cdot r(s, \rho'(s))]$ vanishes over $\Lambda(\rho, \rho')$. Thus, either $\Lambda(\rho, \rho')$ has a empty-interior, or $\mathbf{E}_{\mu[\rho]}[\lambda \cdot r(s, \rho(s))] = \mathbf{E}_{\mu[\rho']}[\lambda \cdot r(s, \rho'(s))]$ for all λ, in which case the two policies ρ and ρ' yield the same long-run payoff vector: $\mathbf{E}_{\mu[\rho]}[r(s, \rho(s))] = \mathbf{E}_{\mu[\rho']}[\cdot r(s, \rho'(s))] \in \mathbf{R}^I$. This is impossible given Assumption **1**.

Hence, the complement O of $\cup_{\rho,\rho'} \Lambda(\rho, \rho')$ is an open and dense set. Fix a positive $\lambda \in O$, and consider the unique optimal ρ in $\mathcal{M}(\lambda)$. We next define a report function $\phi^i : S^i \times S^i \to M^i$. Two types s^i and $\tilde{s}^i$ are equivalent ($s^i \sim \tilde{s}^i$) if $\rho(s^i, s^{-i}) = \rho(\tilde{s}^i, s^{-i})$ for each s^{-i}. We define ϕ^i such that $\phi^i(\bar{s}^i, s^i) = \phi^i(\bar{t}^i, \tilde{s}^i)$ if and only if $s^i \sim \tilde{s}^i$, and we redefine ρ accordingly. Truth-telling is then uniquely optimal for type/equivalence class, and it remains so if we perturb ρ by assigning a small probability to the elements in D^i.

It then follows that, for all $\lambda \in O$, $k(\lambda) = \mathbf{E}_{\mu[\rho]}[\lambda \cdot r(s, \rho(s))]$. For $\lambda \notin O$, the argument follows from the continuity of the function $\mathbf{E}_{\mu[\rho]}[\lambda \cdot r(s, \rho(s))]$ in λ, for fixed ρ. ■

Proof of Proposition 4 and 5. Since the case of state-dependent signaling is more general (and the assumptions **4**–**5** reduce to **2**–**3** when states do not enter signal distributions), it suffices to prove Proposition 5. First, note that Assumption **5** ensures that for all $\rho \in \Xi$, all s, $a = \rho(s)$ (also, for all (s, a) where $(s^{-i}, a) \in D^i$ for some i) and for all $d > 0$,

1. For each i, there exists $\hat{x}^i : S \times Y \to \mathbf{R}$ such that, for all $\hat{a}^i \neq a^i$, all $\hat{s}^i$,

$$\mathbf{E}[\hat{x}^i(s, y) \mid a, s] - \mathbf{E}[\hat{x}^i(s, y) \mid a^{-i}, \hat{a}^i, s^{-i}, \hat{s}^i] > d;$$

   (The expectation is with respect to the signal y.)

2. For every pair i, j, $i \neq j$, $\lambda^i \neq 0, \lambda^j \neq 0$, there exists $\hat{x}^h : S \times Y \to \mathbf{R}$, $h = i, j$,

$$\lambda^i \hat{x}^i(s, y) + \lambda^j \hat{x}^j(s, y) = 0, \tag{28}$$

   and for all $\hat{a}^h \neq a^h$, all $\hat{s}^h$,

$$\mathbf{E}[\hat{x}^h(s, y) \mid a, s] - \mathbf{E}[\hat{x}^h(s, y) \mid a^{-h}, \hat{a}^h, \hat{s}^h, s^{-h}] > d.$$

See Lemma 1 of Kandori and Matsushima (1998). By subtracting the constant $\mathbf{E}[\hat{x}^i(s, y) \mid a, s]$ from all values $\hat{x}^i(s, y)$ (which does not affect (28), since (28) must also hold in expectations), we may assume that, for our fixed choice of a, it holds that, for all s, $\hat{x}^i$ is such that $\mathbf{E}[\hat{x}^i(s, y) \mid a, s] = 0$, all i.

Given this normalization, we have that

$$\mathbf{E}[\hat{x}^i(s, y) \mid a^{-i}, \hat{a}^i, s^{-i}, \hat{s}^i] < -d,$$

for any choice $(\hat{s}^i, \hat{a}^i)$ that does not coincide with (s^i, a^i) (in which case the expected transfer is zero). Intuitively, the transfer $\hat{x}^i$ ensures that, when chosen for high enough d, it never pays to deviate in action, even in combination with a lie, rather than reporting the true state and playing the action profile a that is agreed upon, holding the action profile to be played constant across reports $\hat{s}^i$, given s^{-i}. Deviations in reports might also change the action profile played, but the difference in the payoff from such a change is bounded, while d is arbitrary.

More formally, fix some pure policy $\rho : S \to A$ with long-run payoff v. There exists $\theta : S \to \mathbf{R}^I$ such that, for all s,

$$v + \theta(s) = r(s, \rho(s)) + \mathbf{E}_{p(\cdot|s,\rho(s))}[\theta(t)].$$

Consider the M.D.P. in which player i chooses messages $m^i \in M^i = S^i$ and action $\hat{\rho}^i : M^i \times S^{-i} \to A^i$, and his realized reward is $r^i(s, a^i, \rho^{-i}(m^i, s^{-i})) + \hat{x}^i(m^i, s^{-i}, y)$. Then we may pick $d > 0$ such that, given $\hat{x}^i$, every optimal policy specifies $\hat{\rho}^i(m^i, s^{-i}) = \rho^i(m^i, s^{-i})$. Note also that because of our normalization of $\hat{x}^i$, the private rents in this M.D.P. are equal to θ^i if player i sets $m^i = s^i$.

The argument is similar in the case of coordinate directions. In case $\lambda = -e^i$ (resp. $+e^i$) use Assumption **4** (resp. again **5**) and follow Kandori and Matsushima (1998, Case 1 and 2, Theorem 1).

This transfer addresses deviations at the action stage. ■