Truthful Equilibria in Dynamic Bayesian Games
PRELIMINARY AND INCOMPLETE

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

March 24, 2013

Abstract

This paper characterizes a subset of equilibrium payoffs for Markovian games with private information as the discount factor vanishes. Monitoring might be imperfect, transitions may depend on the action profile, types might be correlated or not, values can be private or interdependent. It focuses on equilibria in which players report their information truthfully in every period. This characterization generalizes those obtained for repeated (and stochastic) games with public monitoring, and reduces to a collection of Bayesian games with transfers. These Bayesian games can be analyzed using standard techniques from static mechanism design: in the case of independent private values, Pareto-efficient payoffs are obtained by means of a version of the AGV mechanism; in the case of correlated types, the results of Crémer and McLean (1988) can be brought to bear, 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) and later extended to stochastic games (Hörner, Sugaya, Takahashi and Vieille).

[*]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.

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 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.)

More precisely, 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

[1] This not to say that the recursive formulations of Abreu, Pearce and Stacchetti (1990) 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.

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, for instance, all Pareto-optimal payoffs that are individually rational (i.e., dominate the stationary minmax payoff) are limit equilibrium payoffs. The subset that is characterized is larger than this, but just as in the static case with independent types, admits no known simple description. When types are correlated, then all feasible and individually rational payoffs can be obtained in the limit.

Those findings mirror those obtained in static mechanism design, *e.g.* those of Arrow (1979), d’Aspremont and Gérard-Varet (1979) for the independent case, and those of Crémer and McLean (1988) 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.

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: the action in a round having possibly an impact on the state in later rounds affects a player’s incentives. It must be taken into account, but it cannot be treated as a transfer that can be designed to convey incentives. 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 (possibly unbounded) transfer.

So: the transient component cannot be ignored, but cannot be exploited as a transfer either. 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) 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. Related contributed in dynamic mechanism design include Battaglini (2005), Bergemann and Välimäki (2010) and Pavan, Toikka and Segal (2012). Ultimately, though, given that our results rely on a reduction to a one-shot Bayesian game, the mechanisms that are used in the proof bear a stronger resemblance to those in static mechanism design, as discussed. 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 set of players is I, finite. 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. players first privately observe their own state (s_n^i);

2. players simultaneously make reports $(m_n^i) \in M^i$, where M^i is finite. 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 $[0, 1]$;

4. players choose actions (a_n^i);

5. a public signal $y_n \in Y$, a finite set, and the next state profile $s_{n+1} = (s_{n+1}^i)_{i\in I}$ are drawn according to a distribution $p(\cdot \mid s_n, a_n) \in \Delta(S \times Y)$.

Throughout, we assume that $p(s, y \mid \bar{s}, \bar{a}) > 0$ whenever $p(y \mid \bar{s}, \bar{a}) > 0$, for all $(s, \bar{s}, \bar{a})$. 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.[2] 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). 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 $g : S \times Y \times A^i \to \mathbf{R}^I$ and as usual we define the reward $r : S \times A \to \mathbf{R}^I$ as its expectation, $r(s,a) = \mathbf{E}[g(s,y,a^i) \mid a]$, a function whose domain is extended to mixed action profiles in $\Delta(A)$.

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

$$\sum (1-\delta)\delta^{n-1} r^i_n,$$

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*. This is defined as the case in which (i) payoffs of a player only depend on his private state, not the others', i.e. 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 (and p_0 is a product distribution), i.e., for all $(s, y, \bar{s}, \bar{a})$,

$$p(s \mid y, \bar{s}, \bar{a}) = \times_i p(s^i \mid y, \bar{s}, a^i).$$

But we do not restrict attention to private values or independent types. In the case of interdependent values, this raises the question whether players observe their payoffs or not. It is possible to deal with the case in which payoffs are privately observed: 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 next period is taken into account when evaluating a player's report, so that we can build on the results of Mezzetti (2004, 2007) in static mechanism design with interdependent valuations. Given our interpretation of a player's private state, we assume that his private actions, his private states and the public signals and reports are

[2] 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.

all the information that is available to a player.[3]

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.

We consider a subset of perfect Bayesian equilibria of this dynamic game, to be defined, corresponding to a particular choice of M^i, namely $M^i = (S^i)^2$, that we discuss next.

3 Equilibrium

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.[4] 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. Unlike in static mechanism design, however, 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), both the focus on direct mechanisms and on obedient behavior are restrictive. It is not known what the "right" message set is.

At the very least, one would like players to be able to announce their private states, motivating the choice of M^i as (a copy of) S^i that is usually made in this literature. Making this (or any other) choice completes the description of the dynamic game. Following standard arguments, given discounting, such a game admits at least one perfect Bayesian equilibrium. There is no reason to expect that such an equilibrium displays truthful reporting, as Example 1 illustrates makes clear.

Example 1 *(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. Both states s^1 and $\hat{s}^1$ are equally likely in the initial period, and the transition is action-*

[3] One could also include a player's realized action in his next state, to take advantage of the insights of Kandori (2003), but again this is peripheral to our objective.

[4] 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.

	L	R
T	1	0
B	0	0

s^1

	L	R
T	0	0
B	0	1

$\hat{s}^1$

Figure 1: The payoff of player 1 in Example 1

independent, with $p \in [1/2, 1)$ *denoting the probability that the state remains the same from one stage to the next. There is a single message (i.e., no meaningful communication).*

The value of the game described in Example 1 is unknown for $p > 2/3$.[5] It can be (implicitly) characterized as the solution of an average cost optimality equation, or ACOE (see Hörner, Rosenberg, Solan and Vieille, 2010). However, this is a functional equation that involves as unknown a function of the belief assigned by the uninformed player, player 2, to the state being s^1. Clearly, in this game, player 1 cannot be induced to give away any information regarding his private state.

Non-existence of truthful equilibria is not really new. The ratchet effect that arises in bargaining and contracting is another manifestation of the difficulties in obtaining truth-telling when there is lack of commitment. See, for instance, Freixas, Guesnerie and Tirole (1985).

Example 1 illustrates that small message sets are just as difficult to deal with as very large ones. In general, one cannot hope for equilibrium existence without allowing players to hide their private information, which in turn requires their opponents to entertain beliefs that are complicated to keep track of.

In light of this observation, it is somewhat surprising that the aforementioned papers (Athey and Bagwell (2001, 2008), Escobar and Toikka (2013), Renault, Solan and Vieille (2013)) manage to get positive results while insisting on equilibria that insist on truthful reporting (at least, after sufficiently many histories for efficiency to be achieved). Given that our purpose is to obtain a tractable representation of equilibrium payoffs, we will insist on it as well. The price to pay is not only that existence is lost in some classes of games (such as the zero-sum game with one-sided incomplete information), but also that one cannot hope for more than pure-strategy equilibria in general. This is illustrated in Example 2.

[5]It is known for $p \in [1/2, 2/3]$ and some specific values. Pęski and Toikka (private communication) have recently established the intuitive but nontrivial property that this value is decreasing in p.

Example 2 *This is a 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 $A^2 = \{L, R\}$. Rewards are given by Figure 2. The two types s^1 and $\hat{s}^1$ are i.i.d.*

	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 minmax payoff cannot be achieved.

over time and equally likely. Monitoring is perfect. Note that values are private. To minmax player 2, player 1 must randomize between both his actions, independently of his type. Yet in any equilibrium in which player 1 always reports his type truthfully, there is no history after which player 1 is indifferent between both actions, for both types simultaneously.[6]

Note that this example still leaves open the possibility of a player randomizing for one of his types. This does not suffice to achieve player 2's mixed minmax payoff, but it still improves on the pure-strategy minmax payoff. Such refinements introduce complications that are mainly notational. We will not get into them, but there is no particular difficulty in adapting the proofs to cover special cases in which mixing is desirable, as when a player has only one type, the standard assumption in repeated games. Instead, our focus will be on strict equilibria.

So we restrict attention to equilibria in which players truthfully (or obediently) report their private information. This raises two related questions: what is their private information, and should the requirement of obedience be imposed after all histories?

[6] To see this, 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.

As in repeated games with public monitoring, players have private information that is not directly payoff-relevant, yet potentially useful: namely, their history of past private actions. Private strategies are powerful but still poorly understood in repeated games, and we will follow most of the literature in focusing on equilibria in which players do not condition their continuation strategies on this information.[7] What is payoff-relevant, however, are the players' private states s_n. Because player i does not directly observe s_n^{-i}, his beliefs about these states become relevant (both to predict $-i$'s behavior and because values need not be private). This is what creates the difficulty in analyzing games such as Example 1: because player 1 does not want to disclose his state, player 2 must use all available information to make the best prediction. Player 2's belief relies on the entire history of play.[8] Here, however, we restrict attention to equilibria in which players report truthfully their information, so that 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.

Given the public information available to player i, which –if players $-i$ do not lie– includes s_{n-1}^{-i}, his conditional beliefs about the future evolution of the Markov chain $(s_{n'})_{n' \geq n}$ are then determined by the *pair* (s_{n-1}^i, s_n^i). This is his "belief-type," which pins down his payoff-type. This is why we fix $M^i = (S^i)^2$.

Truthful reporting in stage n, then, implies reporting both player i's private state in stage n and his private state in stage $n-1$. Along the equilibrium path, this involves a lot of repetition. It makes a difference, however, when a player has deviated in round $n-1$, reporting incorrectly s_{n-1}^i (alongside s_{n-2}^i). Note that players $-i$ cannot detect such a lie, so this deviation is "on-schedule," using the terminology of Athey and Bagwell (2008). If we insist on truthful reporting after such deviations, the choice of M^i makes a difference: by setting $M^i = S^i$ (with the interpretation of the current private state), 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 announcement of s_{n-1}^i, along with his current report.) Setting $M^i = (S^i)^2$ allows him to report both truthfully.

It is easy to see that the choice is irrelevant if obedience is only required on the equilibrium path. However, we will impose obedience after all histories. The choice $M^i = (S^i)^2$ is not only conceptually more appropriate, but also allows for a very simple adaptation of the AGV

[7]See Kandori and Obara (2006) for an illuminating analysis.

[8]It is well-known that existence cannot be expected within the class of strategies that would rely on finite memory or that could be implemented via finite automata. There is a vast literature on this topic, starting with Ben-Porath (1993).

mechanism (Arrow, 1979; d'Aspremont and Gérard-Varet, 1979).

The solution concept then, is truthful-obedient perfect Bayesian equilibrium, where we insist on obedience after all histories. 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. This is automatically the case when signals have full support, and hopefully uncontroversial in the general case.[9]

4 Characterization

4.1 Dimension

Before stating the program, there are two "modeling" choices to be made.

First, regarding the payoff space. There are two possibilities. Either we associate to each player i a unique equilibrium payoff, corresponding to the limit as $\delta \to 1$ of his equilibrium payoff given the initial type profile, focusing on equilibria in which this limit does not depend on the initial type profile. Or we consider payoff vectors 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, the case usually considered in the literature, this makes no difference, as the players' limit equilibrium payoff *must* be independent of the initial type profile. This is an immediate consequence of low discounting, the irreducibility of the Markov chain on states, and incentive-compatibility. On the other hand, when types are correlated, it is possible to assign different (to be clear, long-run) payoffs to a given player, depending on the initial state, using ideas along the lines of Crémer and McLean (1988). To provide a unified analysis, we will focus on payoffs in $\mathbf{R}^I$, though our analysis will be tightly linked to Crémer and McLean when types are correlated, as will be clear.[10]

Indeed, to elicit obedience when types are correlated across players, it is natural to "cross-check" player i's report about his type, (s^i_{n-1}, s^i_n), with the reports of others, regarding the

[9]Note however that our choice $M^i = (S^i)^2$ means that, even when signals have full support, there are histories off the equilibrium path, when a player's reported type conflict with his previous announcement. Beliefs of players $-i$ assign probability 1 to the latest report as being correct.

[10]That is to say, incentives will depend on "current" types, and this will be achieved by exploiting the possible correlation across players. When types are correlated across players, the long-run payoffs could further depend on the *initial* types, despite the chain being ergodic.

realized state s_n^{-i}. In fact, one should use all statistical information that is or will be available and correlates with his report to test it.

One such source of information correlated with his current report is player i's future reports. This information is also present in the case of independent Markov chains, and this "linkage" is the main behind some of the constructions in this literature (see Escobar and Toikka, 2013, building on the static insight of Fang and Norman, 2006, and Jackson and Sonnenschein, 2007). Such statistical tests are remarkably ingenious, but, as we show, unnecessary to generalize all the existing results. One difficulty with using this information is that player i can manipulate it. Yet the AGV mechanism in the static setting suggests that no such statistical test is necessary to achieve efficiency when types are independently distributed. We will see that the same holds true in this dynamic setting. And when types are correlated, we can use the *others*' players reports to test player i's signal –just as is done in static Bayesian mechanism design with correlated types, i.e., in Crémer and McLean.

One difficulty is that the size of the type space that we have adopted is $|S^i|^2$, yet the state profile of the other players in round n is only of size $|S^{-i}|$; of course, players $-i$ also report their previous private state in round $n-1$, but this is of no use, as player i can condition his report on it. Whereas we cannot take advantage of i's previous report regarding his state in period $n-1$, because it is information that he can manipulate.

This looks unpromising, as imposing as an assumption that $|S^i|^2 \leq |S^{-i}|$ for all i (and requiring the state distribution to be "generic" in a sense to be made precise) is stronger than what is obtained in the static framework; in particular, it rules out the two-player case. This should not be too surprising: Obara (2008) provides an instructive analysis of the difficulties encountered when attempting to generalize the results of Crémer and McLean to dynamic environments.

Fortunately, this ignores that player $-i$'s *future* reports are in general correlated with player i's current type. Generically, s_{n+1}^{-i} depend on s_n^i, and there is nothing that player i can do about these future reports. Taking into account these reports when deciding how player i should be punished or rewarded for his stage n-report is the obvious solution, and restores the familiar "necessary and generically sufficient" condition $|S^i| \leq |S^{-i}|$.

But as the next example makes clear, it is possible to get even much weaker conditions, unlike in the static set-up.

Example 3 *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 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^1_{n+1} is drawn randomly (and uniformly) from S^1. In words, s^1_n 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.

The logic behind this example is that potentially many future states $s^{-i}_{n'}$ are correlated with s^i_n: the "signal" that can be used is the entire sequence $(s^{-i}_{n'})_{n' \geq n}$. This raises an interesting statistical question: are signals occurring arbitrarily late in this sequence useful, or is the distribution of this sequence, conditional on (s^i_n, s_{n-1}), summarized by the distribution of an initial segment? This question 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 $(s^{-i}_{n'})_{n' \geq n}$.[11]

This leads us to our second modeling choice. We will only include s^{-i}_{n+1} as additional signal, because it already suffices to recover the condition that is familiar from static Bayesian mechanism design. We will return to this issue at the end of the paper.

4.2 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.

4.2.1 Admissible contracts

Let $\rho : \Omega_{\text{pub}} \times M \to A$ be a map which specifies an action profile contingent on the previous public outcome and on the current reports. We refer to such maps as *(profiles) of plans of action.*[12] Assuming reports are always truthful, any such map ρ induces a Markov chain over Ω, with a unique ergodic set. We denote by $\mu[\rho] \in \Delta(\Omega \times \Omega \times S)$ the (unique) stationary distribution of *consecutive* outcomes and states. That is, $\mu[\rho](\bar{\omega}, \omega, t)$ is the (stationary)

[11] 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).

[12] As opposed to strategies.

probability that the outcomes be $\bar{\omega}$ then ω, next the state profiles be t in three consecutive stages. Since reports are truthful, the distribution $\mu[\rho]$ is concentrated on vectors $(\bar{s}, \bar{m}, \bar{y}, \bar{a}, s, m, y, a, t)$ such that $\bar{m}_c = \bar{s}$ and $m = (\bar{s}, s)$.

Let now a map x specify transfers. Incentivizing transfers in stage n will depend on the public outcome in stages $n-1$ and n and on other players' (reported) states in stage $n+1$. Hence, x maps $\Omega_{\text{pub}} \times \Omega_{\text{pub}} \times S \to \mathbf{R}^I$, and is such that $x^i(\bar{\omega}_{pub}, \omega_{pub}, t)$ does not depend on t^i (and will sometimes be written $x^i(\bar{\omega}_{pub}, \omega_{pub}, t^{-i})$).

Given ρ and x, we set

$$v := \mathbf{E}_{\mu[\rho]}\left[r(s,a) + x(\bar{m}, \bar{y}, m, y, t)\right] \in \mathbf{R}^I.$$

Note that v is the long-run payoff vector, assuming (i) types are always truthfully reported, (ii) actions are chosen using ρ and (iii) stage payoffs are augmented by x. While this long-run payoff is independent of the initial types, discounted payoffs are not. This is best formalized when introducing *private rents*, see Lemma 1 below. These private rents θ provide a normalized measure of the difference between the long-run payoff and the actual payoff, given the initial types.

Lemma 1 *There exists $\theta : \Omega_{pub} \times S \to \mathbf{R}^I$ such that*

$$v + \theta(\bar{\omega}_{pub}, s) = r(s, \rho(\bar{\omega}_{pub}, s)) + \mathbf{E}_{p(\cdot|s,\rho(\bar{\omega}_{pub},s))}\left[x(\bar{\omega}_{pub}, \omega_{pub}, t) + \theta(\omega_{pub}, t)\right],$$

where $\omega_{pub} = (\bar{s}, s, y)$, θ being independent of $\bar{s}$, and the expectation is taken over $(y,t) \sim p(\cdot \mid s, \rho(\bar{\omega}_{pub}, s))$.

The map θ is unique, up to an additive (and player-dependent) constant. We denote by $\theta[\rho, x]$ any such map.

The map θ plays an important role in dynamic programming, and is the counterpart in our set-up (given the modified state space and the transfer-augmented reward) of the *relative* (cost or) value function that is pinned down by the average cost optimality equation (see, for instance, Puterman 1994). It admits an intuitive interpretation in terms of discounted payoffs. Having fixed the plan of action ρ, the transfer rule x and some arbitrary state $(\bar{\omega}^*_{\text{pub}}, s^*)$, one can compute, for a fixed discount factor, the difference in the discounted payoff from starting in some state $(\bar{\omega}^0_{\text{pub}}, s^0)$ relative to $(\bar{\omega}^*_{\text{pub}}, s^*)$. Because the plan of action is "Markov" relative to such a state space, and the Markov chain (s_n) is irreducible, this payoff difference $V^\delta(\bar{\omega}^0_{\text{pub}}, s^0) - V^\delta(\bar{\omega}^*_{\text{pub}}, s^*)$ tends to zero as $\delta \to 1$. On the other hand, the normalized difference $(V^\delta(\bar{\omega}^0_{\text{pub}}, s^0) - V^\delta(\bar{\omega}^*_{\text{pub}}, s^*))/(1-\delta)$ converges to a well-defined limit, $\theta(\bar{\omega}^0_{\text{pub}}, s^0)$ –the additive constant reflecting the arbitrariness in picking $(\bar{\omega}^*_{\text{pub}}, s^*)$.

As discussed in introduction, 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 $\mu[\rho]$ that appears in the definition of v. 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 value function).

Fix now a player $i \in I$. We introduce a parametrized family of (two-step) decision problems. These problems correspond to the best-reply problem faced by player i in a typical stage of the game, assuming players $-i$ report truthfully, play according to ρ^{-i}, and stage payoffs are augmented with x.

Let an outcome $\bar{\omega} = (\bar{s}, \bar{m}, \bar{a}, \bar{y})$ be given, such that $\bar{s}^{-i} = \bar{m}_c^{-i}$. We denote by $D^i(\bar{\omega})$ the following two-step decision problem:

Step 1 $s \in S$ is drawn according to the conditional distribution $\dfrac{p(\cdot, \bar{y} \mid \bar{s}, \bar{a})}{p(\bar{y} \mid \bar{s}, \bar{a})}$, player i is informed of s^i, and then makes a report $m^i \in M^i$.

Step 2 player i learns s^{-i} and then chooses an action $a^i \in A^i$. Finally, (y, t) is drawn according to $p(\cdot \mid s, a^i, \rho^{-i}(\bar{\omega}_{\text{pub}}, m))$, where $\bar{\omega}_{\text{pub}} = (\bar{m}, \bar{y})$ and $m = (m^i, (\bar{s}^{-i}, s^{-i}))$.

The payoff to player i is given by

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

Let $\mathcal{R}^{-i}(m) = \{a^{-i} \in A^{-i} : \exists \omega_{\text{pub}} \in \Omega_{\text{pub}} : \rho^{-i}(\omega_{\text{pub}}, m) = a^{-i}\}$. We denote by $\mathcal{D}^i$ the collection of decision problems $D^i(\bar{\omega})$, where $\bar{a}^{-i} \in \mathcal{R}^{-i}(\bar{m})$. A strategy of player i in the collection $\mathcal{D}^i$ consists of (i) an announcement policy, which specifies a report in each decision problem, contingent on s^i and (ii) an action policy, which specifies an action in each decision problem, and in each contingency that may arise.

Definition 1 *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 $\rho^i(\bar{\omega}_{pub}, m)$ prescribed by ρ in Step 2.*

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 ($\bar{m}_c^i \neq \bar{s}^i$ on his current state). On the other hand, Definition 1 puts no restriction on player i's behavior if he lies in step 1 ($m^i \neq (\bar{s}^i, s^i)$). The second part of Definition 1 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 (1) when $m = (\bar{s}, s)$.

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

4.2.2 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 1 *Assume that $\mathcal{H}_0$ has non-empty interior. Then it is included in the limit set of truthful equilibrium payoffs.*

To be clear, there is no reason to expect Theorem 1 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, Pareto-optimal equilibrium payoffs under independent private values, as well as a folk theorem under correlated values.

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 $(\bar{s}^i, s^i)$. To capture our insistence on truthful equilibria, we focus on report maps ϕ^i, such that transitions $p(\cdot \mid s, a)$ only depend on reports. Formally, we require that $p(\cdot \mid (s^i, s^{-i}), a) = p(\cdot \mid (t^i, s^{-i}), a)$ for all i, s^{-i}, a, whenever $\phi^i(\bar{s}^i, s^i) = \phi^i(\bar{t}^i, t^i)$.

Given transfers $x^i : \Omega_{pub} \times \Omega_{pub} \times M^{-i} \to \mathbf{R}$, the decision problems $D^i(\bar{\omega})$ are defined as previously. The set $\mathcal{C}_0(\phi)$ of ϕ-admissible contracts is the set of pairs (ρ, x), with $\rho : \Omega_{pub} \times M \to A$, 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 $\rho^i(\bar{\omega}_{pub}, m)$ in Step 2.

We denote by $\mathcal{P}_0^\phi(\lambda)$ the optimization problem deduced from $\mathcal{P}_0(\lambda)$ when substituting the constraint $(\rho, x) \in \mathcal{C}_0(\phi)$ to the constraint $(\rho, x) \in \mathcal{C}_0$. Set $\mathcal{H}_0(\phi) := \{v \in \mathbf{R}^I, \lambda \cdot v \leq k_0^\phi(\lambda)\}$ where $k_0^\phi(\lambda)$ is the value of $\mathcal{P}_0^\phi(\lambda)$.

Theorem 2 generalizes Theorem 1.

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

5 Independent private values

5.1 A "folk" theorem

We assume independent private values. Recall that this means that:

- The reward of i only depends on his own state: $r^i(s, a) = r^i(s^i, a)$.
- Transitions of player i's state only depend on his own state, his own action, conditional on the public signal: $p(t, y \mid s, a) = (\times_j p(t^j \mid y, s^j, a^j))p(y \mid s, a)$.

First, we define the feasible (long-run) payoff set as

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

(The restriction to plans of actions rather than arbitrary strategies is clearly without loss.)

We define the state-independent pure-strategy minmax payoff of player i as

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

where $\mu[\rho^i, a^{-i}]$ is the invariant distribution under the plan of action in which players $-i$ play a^{-i} always and i uses ρ^i. We denote the set of feasible, individually rational payoffs as

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

The "folk" theorem requires two assumptions. First, we need to make some minimal assumption to ensure that truth-telling is uniquely optimal for some objective. Given some vector of weights $\lambda \in S^1$, we let

$$r^\lambda(s,a) := \sum_i \lambda^i \cdot r^i(s^i, a).$$

This defines a Markov decision process (M.D.P.) with state space S, action space A, transitions $p(t \mid s, a)$ and reward $r^\lambda(s,a)$. Such a M.D.P. admits average optimal policies $\{\rho^\lambda\}$ that are deterministic and Markov stationary. (In what follows, a *policy* is a plan of action that only depends on the last report m_c, i.e. a map $\rho : S \to A$). We assume:

Assumption 1 *There exists $\lambda > 0$ such that the M.D.P. admits a unique average optimal (stationary) policy $\rho^\lambda : S \to A$. In addition, this policy is such that, for all i, $s^i, \hat{s}^i \in S^i$, $s^i \neq \hat{s}^i$, there exists $s^{-i} \in S^{-i}$ such that $\rho^\lambda(s^i, s^{-i}) \neq \rho^\lambda(\hat{s}^i, s^{-i})$.*

This implies that ρ^λ is the unique solution to the average cost optimality equation, and that a player who misreports his types takes the chance (given that s^{-i} has positive probability of occurring) of changing what is believed to be the optimal action.

Second, we must make some assumptions on the monitoring structure. 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 following assumptions generalize those in Kandori and Matsushima (1998), to which they reduce when states do not affect the signal distribution.

Assumption 2 *For all $a \in A$, $s \in S$,*

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 $i \neq j$,*

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

For each $\lambda \in S^1$, let $I^\lambda = \{i \in I : \lambda^i > 0\}$, and

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

where $\rho : S^{I^\lambda} \to A$. The value $k(\lambda)$ can be interpreted as the efficient payoff when private states of players whose weight is negative are unobserved. Note that, for $\lambda > 0$, this reduces to Pareto-efficient payoffs. Let

$$V^{**} = V^* \cap \left(\cap_{\lambda \in S^1}\{v : \lambda \cdot v \leq k(\lambda)\}\right).$$

We may now state:

Theorem 3 *Under Assumptions 1 and 2, the limit set of truthful equilibrium payoffs includes V^{**}.*

First, note that Assumption 2 ensures that for each state s, pure action profile a and $d > 0$,

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 $\hat{x}^h : S \times Y \to \mathbf{R}$, $h = i, j$,

   $$\lambda^i \hat{x}^i(s,y) + \lambda^j \hat{x}^j(s,y) = 0, \tag{2}$$

   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 (2), since (2) 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.

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 $m^i = s^i$.

This transfer addresses deviations at the action stage. There exists another kind of deviations, namely, those that consist in setting $m^i \neq s^i$ for some s^i (but play $\rho^i(m^i, s^{-i})$).

Consider $\lambda > 0$ from Assumption 1 and the corresponding policy ρ^λ. Let $(v^\lambda, \theta^\lambda) \in \mathbf{R}^I \times \mathbf{R}^{S \times I}$ denote the value (per player) and private rents that arises from the play of ρ^λ. We construct a (modified) AGV mechanism to implement ρ^λ. We fix the mapping from reported messages to implemented action profiles to ρ^λ. Define, for all i,

$$\breve{x}^{i,\lambda}(\bar{\omega}_{\text{pub}}, m) := \mathbf{E}_{p(\cdot|\bar{\omega}_{\text{pub}})}\left[\sum_{j \neq i} \frac{\lambda^j}{\lambda^i}\left(r^j(s^j, \rho^\lambda(s^{-i}, m_c^i)) + \mathbf{E}_{p(\cdot|s,\rho^\lambda(s^{-i}, m_c^i))}[\theta^{j,\lambda}(t)]\right)\right],$$

where the outer expectation is with respect to $m_c^{-i} = s^{-i}$. By construction, given $\rho^{-i,\lambda}$ and the transfers $\breve{x}^{i,\lambda}$, $\rho^{i,\lambda}$ satisfies the ACOE for player i, with value and private rents $\frac{1}{\lambda^i}(\sum_j v^{j,\lambda}, \sum_j \theta^{j,\lambda})$. Next, we normalize $\breve{x}^{i,\lambda}$ to get budget-balance, and define $\breve{x}^{i,\lambda}$ as

$$\breve{x}^{i,\lambda}(\bar{\omega}_{\text{pub}}, m) = \breve{x}^{i,\lambda}(\bar{\omega}_{\text{pub}}, m) - \frac{1}{I-1}\sum_{j \neq i} \frac{\lambda^j}{\lambda^i} \breve{x}^{j,\lambda}(\bar{\omega}_{\text{pub}}, m).$$

Finally, we set

$$\bar{x}^{i,\lambda}(\bar{\omega}_{\text{pub}}, m) = \breve{x}^{i,\lambda}(\bar{\omega}_{\text{pub}}, m) - \mathbf{E}_{p(\cdot|\bar{\omega}_{\text{pub}})}[\breve{x}^{i,\lambda}(\bar{\omega}_{\text{pub}}, m)],$$

so that the (unconditional) expected transfer of player i is independent of his announcement. This last normalization ensures that the transfer does not affect the private rent $\theta^{i,\lambda}$.

(This is analogous to the construction of Athey and Segal (2007)'s team mechanism.) By construction, this transfer ensures that player i has no incentive to report $m_c^i \neq s^i$.

Finally, combining the two (action-complying and truth-telling inducing) transfers, we set

$$x^{i,\lambda}(\bar{\omega}_{\text{pub}}, (m, y), t^{-i}) = \bar{x}^{i,\lambda}(\bar{\omega}_{\text{pub}}, m) + \hat{x}^i(m_c, y).$$

(Note that this transfer does not take advantage of t^{-i}, which is only useful with correlated types.) By construction, the transfer is balanced: for all $\bar{\omega}_{\text{pub}}, (m, y), t^i$, $\lambda \cdot x^\lambda(\bar{\omega}_{\text{pub}}, (m, y), t) = 0$. By Assumption 1 (along with the AGV component of the transfer), incentives to tell the truth are strict. By the choice of $\hat{x}^i$, the incentives to take the action $\rho^\lambda(s)$ is strict as well.

The same construction applies to all non-negative directions $\lambda \in S^1$. To make the truth-telling incentives strict, it suffices to use the public randomization to randomize between the desired plan of action ρ^λ and the policy from Assumption 1, with arbitrarily high probability on the former plan of action. Similarly, we can (with arbitrarily high probability) enforce any constant plan of action $\rho : \Omega_{\text{pub}} \times M \to A$ in any non-coordinate direction.

In the directions $\lambda = -e^i$, an analogous argument gives that the score is (at least)

$$k_0(-e^i) = - \min_{a^{-i} \in A^{-i}} \max_{\rho^i : S \to A^i} \mathbf{E}_{\mu[\rho^i, a^{-i}]}[r(s, a)] = -\underline{v}^i.$$

6 Correlated Types

We drop the assumption of independent types and extend here the static insights from Crémer and McLean (1988). We maintain Assumption 2. We ignore the i.i.d. case, handled by Theorem 2. 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.

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 $\bar{y}$, $\bar{a}$, player i's true past and current state being $\bar{s}^i$ and s^i, and the action mapping current announcement into action profiles ρ. 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 $\bar{s}^i \neq \bar{m}_c^i$. The Crémer and McLean condition states the following.

Assumption 3 *For each* $(\bar{m}, \bar{y}, \bar{a}, \rho)$*, 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, the standard Crémer and McLean condition (see also 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 3 ensures that for any fixed plan of action ρ, 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 4 *For all* $(\bar{m}, \bar{y}, \bar{a}, \rho)$*,*

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 4.1 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 4.2 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.[13]

It is then routine to show:

Theorem 4 *Assume that V^* has non-empty interior. Under Assumptions 2–4, the limit set of truthful equilibrium payoffs includes V^*.*

Assumptions 3–4 are generically satisfied if $|S^{-i}| \geq |S^i|$ for all i.

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

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

[13] 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 $-p(\cdot \mid c^j)$.

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

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

[9] Baron, D. and D. Besanko (1984). “Regulation and Information in a Continuing Relationship,” *Information Economics and Policy*, **1**, 447–470.

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

[11] Battaglini, M. (2005). “Long-Term Contracting with Markovian Consumers,” *American Economic Review*, **95**, 637–658.

[12] Ben-Porath, E. (1993). “Repeated Games with Finite Automata,” *Journal of Economic Theory*, **59**, 17–32.

[13] Bergemann, D. and J. Välimäki (2010). “The Dynamic Pivot Mechanism,” *Econometrica*, **78**, 771–789.

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

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

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

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

[18] Crémer, J. and R. McLean (1988). “Full Extraction of Surplus In Bayesian and Dominant Strategy Auctions,” *Econometrica*, **56**, 1247–1257.

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

[20] 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.

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

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

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

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

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

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

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

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

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

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

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

[32] 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.

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

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

[35] 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.

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

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

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

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

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

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

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

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

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

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

[46] Pavan, A., I. Segal and J. Toikka (2012). "Dynamic Mechanism Design: Incentive Compatibility, Profit Maximization and Information Disclosure," working paper, Stanford University.

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

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

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

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

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

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

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

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

A Proof of Theorem 1

A.1 Preliminaries

The proof is inspired by FLM but there are a number of 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 quote without proof the following classical result, which relies on the smoothness of Z_η.

Lemma 2 *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 $\lambda \in S_1$ 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 a direction $\lambda \in S_1$, 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 S_1, this proves Lemma 3 below.

Lemma 3 *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 $z \in \mathbf{R}^I$, and every direction $\lambda \in S_1$, 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 choose $\kappa_0 \in \mathbf{R}$ 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 2 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 $\delta \geq \bar{\delta}$.

For $(v, x, \rho) \in \mathcal{S}_0$, we denote by $\theta_{\rho,x} : \Omega_{\text{pub}} \times S \to \mathbf{R}^I$ the payoff bias under (ρ, x). The condition $(\rho, x) \in \mathcal{C}_0$ is equivalent to a finite number of strict inequalities, which we now proceed to list.

Fix a player $i \in I$, a private outcome $\bar{\omega}^i \in \Omega^i$ with public part $\bar{\omega}$, and a (private) state $s^i \in S^i$. Assume that player i reports $m^i \in M^i$. If players $-i$'s states are s^{-i}, and player i then plays $a^i \in A^i$, player i's expected payoff in the decision problem $D^i(\bar{\omega}^i)$ is then

$$\pi^i(a^i \mid \bar{\omega}^i, s, m^i) := r^i(s, (a^i, \rho^{-i}(\bar{\omega}_{\text{pub}}, m))) + \mathbf{E}^{y,t}[x^i(\bar{\omega}_{\text{pub}}, \omega_{\text{pub}}, t^{-i}) + \theta^i(\omega_{\text{pub}}, t)],$$

m is (m^i, m^{-i}), with $m^{-i} := (\bar{m}_c^{-i}, s^{-i})$, and the public outcome is $\omega_{\text{pub}} := (m, y)$. Thus, when states are s and when reporting m^i, the highest payoff in $D^i(\bar{\omega}^i)$ is equal to

$$\pi^i_{\rho,x}(\bar{\omega}^i, s, m^i) := \max_{a^i \in A^i} \pi^i(a^i \mid \bar{\omega}^i, s, m^i).$$

Since $(\rho, x) \in \mathcal{C}_0$, any optimal policy in $D^i(\bar{\omega}^i)$ first reports truthfully $m^i_* := (\bar{s}^i, s^i)$ then plays the action $\rho^i(\bar{\omega}_{\text{pub}}, (m^i_*, m^{-i})$ dictated by the action plan ρ^i.

The truth-telling condition writes

$$\mathbf{E}^{s^{-i}}[\pi^i_{\rho,x}(\bar{\omega}^i, (s^i, s^{-i}, m^i_*) \mid \bar{\omega}^i, s^i] > \mathbf{E}^{s^{-i}}[\pi^i_{\rho,x}(\bar{\omega}^i, (s^i, s^{-i}), m^i) \mid \bar{\omega}^i, s^i] \text{ for all } \bar{\omega}^i, s^i \text{ and } m^i \neq (\bar{s}^i, s^i), \tag{3}$$

where the expectation is taken over s^{-i}, and is computed under the conditional distribution of s^{-i}, given the previous outcome $\bar{\omega}^i$, and s^i.

The condition that expresses that ρ^i is optimal after a truthful report writes

$$\pi^i_{\rho,x}(\bar{\omega}^i, s, m^i_*) > \pi^i(a^i \mid \bar{\omega}^i, s, m^i_*) \text{ for each } a^i \neq \rho^i(\bar{\omega}_{\text{pub}}, (m^i_*, m^{-i})). \tag{4}$$

We pick $\eta_{\rho,x} > 0$ small enough to be less than the least difference between the left and the right-hand side in each of these inequalities.

As (v, ρ, x) varies through the finite set $\mathcal{S}_0$, one thus obtains finitely many strict inequalities. We pick $\eta_1 > 0$ to be less than the minimal difference between the left and the right-hand side in each of the inequalities (3) and (4).

A.2 Strategies

We let $z_* \in Z_\eta$, and $\delta \in (0, 1)$ be given. We here define a pure strategy profile σ. We check in the next section that for δ large enough, σ is a PBE and induces a payoff arbitrarily close to z.

Under σ^i, all reports of player i are truthful, and his actions in a given stage n (when reporting truthfully) depend on a target payoff $z_n \in Z_\eta$, on the previous public outcome $\bar{\omega}_{\text{pub},n-1} \in \Omega_{\text{pub}}$ 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, pick a unit vector $\lambda_n \in S_1$ using Lemma 2, and use Lemma 3 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

$$w_{n+1} := \frac{1}{\delta} z_n - \frac{1-\delta}{\delta} v_n + \frac{1-\delta}{\delta} x_n(\omega_{\text{pub},n-1}, \omega_{\text{pub},n}, m_{c,n+1}), \tag{5}$$

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

$$w_{n+1} = \xi \tilde{w}_{n+1} + (1 - \xi) z_n, \tag{6}$$

where $\xi := (1 - \delta)^{3/4}$.[14]

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.

[14]The choice of the exponent $3/4$ is to a large extent arbitrary. It is important that ξ vanishes as $\delta \to 1$, more slowly than $1 - \delta$, but that ξ^2 vanishes faster than $(1 - \delta)$.

Lemma 4 *One has $\tilde{w}_{n+1} \in Z_\eta$.*

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

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

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

If the current report of player i is truthful –that is, $m_n^i = (s_{n-1}^i, s_n^i)$–, then 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[15] 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 $(\omega_{pub,n-1}, m_n)$ 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,

$$(1-\delta)r^i(s_n, (a^i, \rho^{-i}(\omega_{\text{pub},n-1}, m_n))) + \delta\mathbf{E}[\gamma_\sigma^i(m_n, y_n), m_{n+1}; z_{n+1})],$$

where the expectation is taken over y_n and m_{n+1} and z_{n+1}.[16]

Theorem 1 follows from Proposition 1, which is proven in the next section.

Proposition 1 *The following holds.*

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

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

[15] Which occurs if successive reports are inconsistent, and observed public signals are inconsistent with reported states.

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

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 1

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 (3) and (4). The crucial observation is that at any given stage n, expected continuation payoffs under σ are close to z_n, and (continuation) payoff biases are close to θ_{ρ_n,x_n}. This in turn hinges on the irreducibility properties of the state process.

These properties are established in Proposition 2 below. Given an arbitrary target $z \in Z_\eta$, we will denote by $(v, \rho, x_{)} \in \mathcal{S}_0$ the triple associated to z. We set

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

which we interpret as the expected continuation payoff under σ, when the target is z; and

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

which we interpret as the continuation bias under σ, when the target is z.

Proposition 2 *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 *:* $\|\theta_\sigma(\cdot; z) - \theta_{\rho,x}\| \leq c_1(1-\delta)^{3/4}$.

Proof of Proposition 2. : We first compare $\gamma_\sigma(\omega_{\text{pub}}, m; z)$ and $\gamma_\sigma(\tilde{\omega}_{\text{pub}}, \tilde{m}; z)$ for arbitrary (ω_{pub}, m) and $(\tilde{\omega}_{\text{pub}}, \tilde{m})$ in $\Omega_{\text{pub}} \times M$. 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 q_ρ, which start from (ω_{pub}, m) and $(\tilde{\omega}_{\text{pub}}, \tilde{m})$ respectively;[17]

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. 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 $r_n := r(s_n, \rho(\omega_{\text{pub},n-1}, m_n))$ and $\tilde{r}_n := r(\tilde{s}_n, \rho(\tilde{\omega}_{\text{pub},n-1}, \tilde{m}_n))$ the payoffs in stage n along the two plays. We also 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. In the same spirit, we 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, as defined in (5). Observe that $z_n = \tilde{z}_n = z$ for each $n < \tau$, while z_τ and $\tilde{z}_\tau$ are obtained from z by equation (6).

We define $\tau_c := \inf\{n : (\omega_{\text{pub},n-1}, m_n) = (\tilde{\omega}_{\text{pub},n-1}, \tilde{m}_n)\}$ to be the first "coincidence" time of the two processes (ω_n) and $(\tilde{\omega}_n)$. Given our assumptions on transitions p, the two chains (ω_n) and $(\tilde{\omega}_n)$ have a unique ergodic set (which is the same for both chains), and are aperiodic. This implies that τ_c has a finite expectation. We let C_0 be an upper bound for $\mathbf{E}[\tau_c]$, valid for all (ω_{pub}, m) and $(\tilde{\omega}_{pub}, \tilde{m})$. Since the two stopping times are independent, this implies the existence of C_1, such that $\mathbf{P}(\tau \leq \tau_c) \leq C_1 \xi$.[18]

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

$$\gamma_\sigma(\omega_{\text{pub}}, m; z) = \mathbf{E}[(1-\delta) \sum_{n=1}^{\tau_*-1} \delta^{n-1} r_n + \delta^{\tau_*-1} \gamma_\sigma(\omega_{\text{pub},\tau_*-1}, s_{\tau_*}; z_{\tau_*})], \tag{7}$$

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

$$\begin{aligned} & \gamma_\sigma(\omega_{\text{pub}}, m; z) - \gamma_\sigma(\tilde{\omega}_{\text{pub}}, \tilde{m}; z) \qquad (8) \\ = \; & \mathbf{E}[(1-\delta) \sum_{n=1}^{\tau_*-1} \delta^{n-1} (r_n - \tilde{r}_n) + \delta^{\tau_*-1} (\gamma_\sigma(\omega_{\text{pub},\tau_*-1}, m_{\tau_*}; z_{\tau_*}) - \gamma_\sigma(\tilde{\omega}_{\text{pub},\tau_*-1}, \tilde{m}_{\tau_*}; \tilde{z}_{\tau_*}))]. \end{aligned}$$

[17]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 $\tilde{\omega}_1$.

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

Our first claim provides a preliminary estimate, which will be refined later.

Claim 5 *There is $C_2 > 0$ such that for every $z \in Z_\eta$, every $(\omega_{pub}, m), (\tilde{\omega}_{pub}, \tilde{m}) \in \Omega_{pub} \times M$, and every $\delta > \bar{\delta}$, one has*

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

Proof. Note that $(\omega_{\text{pub},n-1}, m_n, z_n) = (\tilde{\omega}_{\text{pub},n-1}, \tilde{m}_n, \tilde{z}_n)$ on the event $n = \tau_c < \tau$, so that $\gamma_\sigma(\omega_{\text{pub},n-1}, m_n; z_n) = \gamma_\sigma(\tilde{\omega}_{\text{pub},n-1}, \tilde{m}_n; \tilde{z}_n)$. Since payoffs lie in $[0, 1]$, equality (8) yields

$$\begin{aligned} \|\gamma_\sigma(\omega_{\text{pub}}, m; z) - \gamma_\sigma(\tilde{\omega}_{\text{pub}}, \tilde{m}; z)\| &= \|(1-\delta)\mathbf{E}\left[\sum_{n=1}^{\tau_*-1}(r_n - \tilde{r}_n)\delta^{n-1} + \delta^{\tau_*-1}1_{\tau \leq \tau_c}\right]\| \\ &\leq (1-\delta)\mathbf{E}[\tau_c] + \mathbf{P}(\tau \leq \tau_c). \end{aligned} \tag{9}$$

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

Since $\gamma_\sigma^i(z)$ lies between $\min_{\omega_{\text{pub}},m} \gamma_\sigma^i(\omega_{\text{pub}}, m; z)$ and $\max_{\omega_{\text{pub}},m} \gamma_\sigma^i(\omega_{\text{pub}}, m; z)$, Claim 5 yields Claim 6 below.

Claim 6 *For each z, (ω_{pub}, m) and $\delta > \bar{\delta}$, one has*

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

We now prove **P1** of Proposition 2.

Fix $z \in Z_\eta$. We reformulate (6) which relates the current target z and the expected target w' in the next stage:

$$z = \delta w' + (1-\delta)v - (1-\delta)x(\bar{\omega}_{pub}, \omega_{pub}, m'_c).$$

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, the latter equation yields

$$z = (1-\delta)v + \mathbf{E}_{\mu[\rho]}\left[\delta z' - (1-\delta)x(\bar{\omega}_{pub}, \omega_{pub}, m'_c)\right].$$

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

$$z = \mathbf{E}_{\mu[\rho]}[(1-\delta)r(s, a) + \delta z']. \tag{10}$$

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

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

for each (ω_{pub}, m). Taking expectations under the invariant measure, one gets

$$\gamma_\sigma(z) = \mathbf{E}_{\mu[\rho]}\left[(1-\delta)r(s,a) + \delta\gamma_\sigma(\omega'_{\text{pub}}, m'; z')\right]. \tag{11}$$

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

$$\begin{aligned}
\mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(\omega'_{\text{pub}}, m'; z')1_A\right] &= \mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(\omega'_{\text{pub}}, m'; z)1_A\right] && (12)\\
&= \mathbf{P}(A)\mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(\omega'_{\text{pub}}, s'; z)\right] = \mathbf{P}(A)\gamma_\sigma(z) && (13)\\
&= \mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(z')1_A\right]. && (14)
\end{aligned}$$

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

Denote now by B the event where the public randomizing device tells players to switch to a new target – an event of probability ξ. Denoting by $(v', \rho', x') \in \mathcal{S}_0$ the triple associated to z', one has using Claim 5,

$$\|\gamma_\sigma(\omega'_{\text{pub}}, m', z') - \gamma_\sigma(z')\| \leq c_1\xi$$

for every realization of (ω', m', z'). Since $\gamma_\sigma(z') := \mathbf{E}_{\mu[\rho']}\left[\gamma_\sigma(\omega'_{\text{pub}}, m', z')\right]$, one obtains

$$\|\mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(\omega'_{\text{pub}}, m'; z')1_B\right] - \mathbf{E}_{\mu[\rho]}\left[\gamma_\sigma(z')1_B\right]\| \leq c_1\xi\mathbf{P}(B) = c_1\xi^2. \tag{15}$$

Plugging (14) and (15) into (11), one gets

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

Combining this inequality with (10), one gets

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

Setting $S := \sup_{z\in Z_\eta}[\|\gamma_\sigma(z) - z\|$, this implies in turn

$$S \leq \delta S + C_2\xi^2,$$

so that $S \leq C_2\dfrac{\xi^2}{1-\delta} = C_2(1-\delta)^{1/2}$. This yields **P1** (with $c_1 := C_2$).

We turn to the proof of **P2**. We let (ω_{pub}, m) and $(\tilde{\omega}_{\text{pub}}, \tilde{m})$ in $\Omega_{\text{pub}} \times M$ be given, and use the coupling introduced earlier. We proceed in two steps, Claims 7 and 8 below.

Set $\Delta_n := (r_n + x_n) - (\tilde{r}_n + \tilde{x}_n)$.

Claim 7 *There is $C_3 > 0$ such that for every $\delta > \bar{\delta}$, one has*

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

Claim 8 *There is $C_4 > 0$ such that for every $\delta > \bar{\delta}$, one has*

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

(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 $\theta_{\rho,x}(\omega_{\text{pub}}, m) - \theta_{\rho,x}(\tilde{\omega}_{\text{pub}}, \tilde{m})$, Statement **P2** follows from Claims 7 and 8, with $c_2 := C_3 + C_4$.

Proof of Claim 7.

If $\tau_c < \tau$, then $(\omega_{\text{pub},\tau_*-1}, m_{\tau_*}) = (\tilde{\omega}_{\text{pub},\tau_*-1}, \tilde{m}_{\tau_*})$, and $z_{\tau_*} = \tilde{z}_{\tau_*} = z$, hence

$$\gamma_\sigma(\omega_{\text{pub},\tau_*-1}, m_{\tau_*}; z_{\tau_*}) - \gamma_\sigma(\tilde{\omega}_{\text{pub},\tau_*-1}, \tilde{s}_{\tau_*}; \tilde{z}_{\tau_*}) = 0 = z_{\tau_*} - \tilde{z}_{\tau_*}.$$

If instead $\tau \leq \tau_c$, then by Claim6, $\gamma_\sigma(\omega_{\text{pub},\tau_*-1}, m_{\tau_*}; z_{\tau_*})$ is within $C_2\xi = C_2(1-\delta)^{\frac{3}{4}}$ of $\gamma_\sigma(z_{\tau_*})$ and, by **P1**, $\gamma_\sigma(z_{\tau_*})$ is within $C_2(1-\delta)^{\frac{1}{2}}$ of z_{τ_*}. Hence, $\gamma_\sigma(\omega_{\text{pub},\tau_*-1}, m_{\tau_*}; z_{\tau_*})$ is within $2C_2(1-\delta)^{1/2}$ of z_{τ_*}, and a similar result holds for $\gamma_\sigma(\tilde{\omega}_{\text{pub},\tau_*-1}, \tilde{m}_{\tau_*}; \tilde{z}_{\tau_*})$. Hence the difference $(\gamma_\sigma(\omega_{\text{pub},\tau_*-1}, m_{\tau_*}; z_{\tau_*}) - \gamma_\sigma(\tilde{\omega}_{\text{pub},\tau_*-1}, \tilde{m}_{\tau_*}; \tilde{z}_{\tau_*}))$ is equal to the difference $(z_{\tau_*} - \tilde{z}_{\tau_*})$, up to $4C_2(1-\delta)^{1/2}$. Since $\mathbf{P}(\tau \leq \tau_c) \leq C_1\xi$, it follows that

$$\begin{aligned}
& \|\mathbf{E}\left[\gamma_\sigma(\omega_{\text{pub},\tau_*-1}, m_{\tau_*}; z_{\tau_*}) - \gamma_\sigma(\tilde{\omega}_{\text{pub},\tau_*-1}, \tilde{m}_{\tau_*}; \tilde{z}_{\tau_*})\right] - \mathbf{E}[z_{\tau_*} - \tilde{z}_{\tau_*}]\| \\
\leq \; & 4C_2(1-\delta)^{1/2} \times C_1\xi = 4C_1C_2(1-\delta)^{5/4}\|. \qquad (16)
\end{aligned}$$

Observe next that $z_{n+1} = \tilde{z}_{n+1} (= z)$ for each $n < \tau_* - 1$. So that plugging (16) into (**??**), the difference $\gamma_\sigma(\omega_{\text{pub}}, m; z) - \gamma_\sigma(\tilde{\omega}_{\text{pub}}, \tilde{m}; z)$ is equal to

$$\mathbf{E}\left[\sum_{n=1}^{\tau'-1} \delta^{n-1}\left((1-\delta)(r_n - \tilde{r}_n) + \delta(z_{n+1} - \tilde{z}_{n+1})\right)\right], \qquad (17)$$

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

Next, we rewrite

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

Since the random time τ is independent of the plays (ω_n) and $(\tilde{\omega}_n)$, one has for each stage $n \geq 2$,

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

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 (17), with $C_3 := 4C_1C_2$. ■

Proof of Claim 8.

Observe first that, since $\|\Delta_n\| \leq (1+\kappa_0)$, the difference between (**??**) and $\mathbf{E}\left[\sum_{n=1}^{\tau_c-1} \delta^{n-1}\Delta_n\right]$ is at most $(1+\kappa_0)\mathbf{E}[\tau_c - \min(\tau, \tau_c)]$. Using the independence of τ and τ_c, one has

$$\mathbf{E}[\tau_c - \min(\tau, \tau_c)] = \mathbf{E}[\tau_c - \tau \mid \tau < \tau_c] \times \mathbf{P}(\tau < \tau_c) \leq C_0 \times C_1\xi. \tag{18}$$

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

$$\|\mathbf{E}\left[\sum_{n=1}^{\tau_c-1} \delta^{n-1}\Delta_n\right] - \mathbf{E}\left[\sum_{n=1}^{\tau_c-1} \Delta_n\right]\| \leq (1+\kappa_0)(1-\delta)\mathbf{E}[\sum_{n=1}^{\tau_c-1} n] \leq (1-\delta) \times (1+\kappa_0)\mathbf{E}[\tau_c^2]. \tag{19}$$

Collecting (18) and (19), there exists $c_4 > 0$ such that

$$\|\mathbf{E}\left[\sum_{n=1}^{\tau_*-1} \delta^{n-1}\Delta_n\right] - \mathbf{E}\left[\sum_{n=1}^{\tau_c-1} \Delta_n\right]\| \leq c_4\xi. \tag{20}$$

■

This concludes the proof of Proposition 2. ■

A.4 Conclusion

We now check that player i has no profitable one-step deviation. For clarity, we consider a report node – apart from notational issues, the proof is similar for an action node. Consider a stage n and a report node, at which the private history of i is $(h_{\mathrm{pub},n}, (s_k^i)_{1 \leq k \leq n}, (a_k^i)_{1 \leq k \leq n-1})$.

We compare the continuation payoffs obtained with σ^i or when lying in stage n (and assuming that player i reverts to σ^i in stage $n+1$).

When reporting truthfully, the continuation payoff of i is

$$\mathbf{E}\left[(1-\delta)r^i(s_n, \rho(\omega_{\mathrm{pub},n-1}, s_n)) + \delta\gamma_\sigma((s_n, y_n), s_{n+1}; z_{n+1})\right],$$

where the expectation is taken over s_n^{-i}, s_{n+1}, y_n When instead reporting $m_n^i \neq s_n^i$ and next playing some action $a^i(m_n^i, s_n^{-i})$, the continuation payoff of i is given by

$$\mathbf{E}\left[(1-\delta)r^i(s_n,(a^i(m_n^i,s_n^{-i}),\rho^{-i}(\omega_{\text{pub},n-1},m_n^i,s_n^{-i}))+\delta\gamma_\sigma((m_n^i,s_n^{-i},\tilde{y}_n),s_{n+1};\tilde{z}_{n+1})\right].$$

We compare the latter two expectations relying as above on a coupling argument. More specifically, we will view the pairs (y_n, s_{n+1}) and $(\tilde{y}_n, \tilde{s}_{n+1})$ as (conditionally) independent. Yet, the announcements of the randomizing device are perfectly coupled in the two expectations.

With this interpretation in mind, with probability $1-\xi$, one has $z_{n+1} = \tilde{z}_{n+1} = z$ and $\gamma_\sigma((s_n, y_n), s_{n+1}; z_{n+1}) - \gamma_\sigma((m_n^i, s_n^{-i}, \tilde{y}_n), s_{n+1}; \tilde{z}_{n+1})$ is then, using Lemma **??**, equal to

$$\begin{aligned} & (z_{n+1}+(1-\delta)\theta_{\rho,x}((s_n,y_n),s_{n+1};z_{n+1})) \\ - & \left(\tilde{z}_{n+1}+(1-\delta)\theta_{\rho,x}((m_n^i,s_n^{-i}),\tilde{y}_n),\tilde{s}_{n+1};\tilde{z}_{n+1})\right)+o(1-\delta). \end{aligned} \tag{21}$$

On the other hand, with probability ξ, the public device instructs to switch to a new block, and the difference $\gamma_\sigma((s_n, y_n), s_{n+1}; z_{n+1}) - \gamma_\sigma((m_n^i, s_n^{-i}, \tilde{y}_n), s_{n+1}; \tilde{z}_{n+1})$ is then equal, by Lemmas **??** and **??**, to

$$(z_{n+1}+(1-\delta)\theta_{\rho,x}((s_n,y_n),s_{n+1};z_{n+1})) \tag{22}$$
$$- \quad \left(\tilde{z}_{n+1}+(1-\delta)\theta_{\rho,x}((m_n^i,s_n^{-i}),\tilde{y}_n),\tilde{s}_{n+1};\tilde{z}_{n+1})\right)+o((1-\delta)^{\frac{3}{4}}). \tag{23}$$

At a technical level, the difference between (21) and (22) is that the error term is much smaller in the former (that is, if $z_{n+1} = \tilde{z}_{n+1} = z$). The higher error term in (22) is compensated by the fact that it gets multiplied by the probability of switching to a new block, which is small.

Using obvious notations, the difference in expected continuation payoffs is therefore, up to $o(1-\delta)$, equal to

$$\mathbf{E}[(1-\delta)r_n+\delta(1-\delta)\theta_{\rho,x}(\omega_{\text{pub},n},s_{n+1})+\delta z_{n+1}]-\mathbf{E}[(1-\delta)\tilde{r}_n+\delta(1-\delta)\theta_{\rho,x}(\tilde{\omega}_{\text{pub},n},\tilde{s}_{n+1})+\delta\tilde{z}_{n+1}],$$

which is also equal to

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

which is strictly positive for δ close to one, thanks to (**??**).