Dynamic Sender-Receiver Games[*]

Jérôme Renault[†], Eilon Solan[‡] and Nicolas Vieille[§]

April 5, 2010

Abstract

We consider a dynamic version of sender-receiver games, where the sequence of states follows a Markov chain observed by the sender. Under mild assumptions, we characterize the limit set of equilibrium payoffs. We obtain a strong dichotomy property: either only uninformative "babbling" equilibria exist, or we can perturb the game so that all equilibrium payoffs can be achieved with strategies where, in most of the stages, the sender reveals the true state to the receiver.

1 Introduction

Since Crawford and Sobel (1982), sender-receiver games, or cheap-talk games, have become a natural framework for studying issues of information transmission between a privately informed 'expert' and an uninformed decision maker, where the two parties have conflicting interests.

When the decision maker acts only *once*, the extent to which information can be shared at equilibrium has been studied extensively, when 'talk' takes place prior to the decision stage. While Crawford and Sobel (1982), see also Green and Stokey (2007), have focused on the case where communication is limited to a single costless and non-verifiable message from the

sender to the receiver, more recent papers have shown that this restriction is not innocuous, and have characterized the equilibrium outcomes for general cheap-talk games, see Krishna and Morgan (2001), Aumann and Hart (2003).[1] This work has been motivated by numerous concrete situations. We refer to Krishna and Morgan (???), Farrell and Rabin (1996), and Sobel (???) for a discussion of these applications.

The present work is motivated by the following observation. When the sender is a financial advisor who provides advice to a client, an expert who is consulted on a project, or a referee on a project/person, the situation often calls for a dynamic approach. Indeed, the financial advisor provides advice on a series of investments, the expert may be consulted on successive, related projects, etc.

Golosov, Skreta, Tsyvinsky and Wilson (2009) consider such a situation. They assume that the sender repeatedly sends messages, the receiver repeatedly makes decisions, while the state of the world remains fixed throughout. Within the Crawford and Sobel framework (continuum of states/messages), they show that (necessarily complex) equilibria exist, that achieve full revelation of the state of the world in finite time.

We here deal with situations in which the state of the world may change through time. Specifically, we assume that the successive states form an irreducible Markov chain over some finite set. In every stage, the sender issues a message/recommendation, and the receiver makes a decision. States are only known to the sender, and payoffs only depend on the current state and on the receiver's decision, but not on the message sent by the sender.

Since states are autocorrelated, any information disclosed in stage n provides valuable information in later stages as well, as in Golosov et al. (2009). Yet, since the Markov Chain is irreducible, this information becomes eventually valueless.

Intuitively, the inter-temporal situation puts some restrictions on the players' behavior. As an illustration, the opinion of an expert who systematically provides laudative reports will eventually come to be discounted, if not ignored, since the decision maker is aware of the fact that the time-average report of the quality of people/projects should reflect the invariant measure of the states of the world. On the other hand, an expert who genuinely provides accurate information to promote efficiency, but sees that the decision maker only acts in his interests, may become wary and may stop to provide valuable information to the decision maker. As is well-known from repeated games, the sender may indeed provide powerful incentives by

---

[1]The case of verifiable messages has also been studied in detail, see Forges and Koessler (2008).

conditioning his future communication policy on the behavior of the decision maker.

Our paper belongs to the recent and growing literature on incomplete information games, in which the uncertainty evolves, see, e.g., Athey and Bagwell (2008), Mailath and Samuelson (2001), Phelan (2006), Renault (2006), Wiseman (2008), and Hörner et al. (2010). Our objective is to study how the various effects of the dynamic environment combine to determine the set of equilibrium payoffs. We provide a characterization of the limit set of sequential equilibrium payoffs, when players become more and more patient.

Our main findings are the following. First, the limit set of equilibrium payoffs, as the players become more and more patient, does not depend on *how* successive states are correlated, nor on fine details of the sequence of states, but only on the invariant measure. In particular, the set of equilibrium payoffs can be computed *as if* successive states were independent.[2]

/// **Eilon: Footnote 2 hides a crucial assumption. Don't we want to elaborate on it in the introduction?** ///

Second, all equilibrium payoffs can be implemented by equilibria in which with high probability, most of the time the sender truthfully reports the current state, while the receiver responds in a stationary manner to the announcements of the sender, and checks that the distribution of these announcements is consistent with the invariant measure. More precisely, a feasible payoff vector is an equilibrium payoff provided two conditions are met. On the one hand, the payoff of the receiver should be at least his babbling equilibrium payoff, since the receiver has the option to ignore the announcements of the sender. On the other hand, the sender may always choose to report *fictitious* states, as long as his reports cannot statistically be proven to be non-truthful. This leads to an incentive constraint, according to which truth-telling should be optimal, given the policy of the decision-maker. As it turns out, this incentive constraint takes the form of finitely many linear inequalities,[3] and therefore can be easily verified. In particular, the equilibrium payoffs of the receiver cannot be lower in the dynamic game than in the one-shot game. Surprisingly, this conclusion is not valid for the sender, as we show by means of an example.

Finally, we show that, unlike in the one-shot case, the limit set of equilibrium payoffs does not depend on how the communication is between the sender and the receiver, as long as messages are costless and non-verifiable.

---

[2]This results hinges on a technical assumption, which is satisfied in all the models quoted above.

[3]These inequalities moreover do not involve the transition function of the Markov chain.

/// Eilon: I do not understand the previous paragraph. ///

The paper is structured as follows. The model is described in Section 2. In Section 3 we provide an example that shows the intuition underlying the results and the proofs. The main results appear in Section 4. Section 5 contains some proofs, and the Appendix contains the more technical ones.

2 Model

We study dynamic sender-receiver games, in which the state of the world changes through time. At each stage $n \geq 1$, the sender (player 1) observes the current state of the world $\mathbf{s}_n \in S$, and chooses a message $\mathbf{a}_n \in A$. Upon observing $\mathbf{a}_n$, the receiver (player 2) chooses an action $\mathbf{b}_n \in B$. The current action $\mathbf{b}_n$, together with the current state $\mathbf{s}_n$, determines the utility vector $u(\mathbf{s}_n, \mathbf{b}_n) = (u^s(\mathbf{s}_n, \mathbf{b}_n), u^r(\mathbf{s}_n, \mathbf{b}_n)) \in \mathbf{R}^2$ at stage n, where the first coordinate is the sender's utility, and the second coordinate is the receiver's utility. Only the action $\mathbf{b}_n$ is then publicly disclosed. As in most of the literature on repeated games with incomplete information, we thus maintain the assumption that payoffs are not observed.[4] The two players share a common discount factor δ.

/// Eilon: Above I denoted payoffs by u^s and u^r. This sounds nice, but it has its problem: $u^s(s, b)$ is the payoff of the sender at state s. Anyone has a novel solution to this? ///

We assume throughout that the set of states S, the set of messages A, and the set of actions B, are finite. We also assume that A coincides with the set of states S. The assumption that there are at least as many messages as states ensures that the only motives for concealing the state are strategic. We thus put aside situations in which, due to capacity constraints, the sender might be forced to choose which feature of the state to reveal. To derive our results, we will not need to assume that there are more messages than states. W.l.o.g., all payoffs are in $[0, 1]$.

We assume that the states $\mathbf{s}_n$ follow a Markov chain over S, with transition function $p(\cdot \mid \cdot)$, which is irreducible and aperiodic.[5] The Markov chain therefore admits a unique invariant

---

[4]In Golosov et al. (2009), this assumption is viewed as an idealization of a situation with random, observed, payoffs. In such a situation, the receiver might perform some Bayesian updating based on the payoffs.

[5]That is, for any two states $s, t \in S$, and for every $N \in \mathbf{N}$ large enough, the probability of moving from s to

measure, $m \in \Delta(S)$. For convenience, we assume that the first state, $\mathbf{s}_1$, is drawn according to m. This ensures that the law of $\mathbf{s}_n$ is equal to m, for every $n \geq 1$. W.l.o.g., the invariant measure m assigns positive probability to each state $s \in S$.

Our goal is to study to what extent the dynamic structure of the game affects the equilibrium outcomes. Formally, we aim at providing a characterization of the limit set of sequential equilibrium payoffs, and at understanding equilibrium behavior, when players become more and more patient.[6]

For some of our results, we will make one substantive assumption on the behavior of the state.

Assumption A. There exist nonnegative numbers α_s, $s \in S$, with $\sum_{s \in S \backslash \bar{s}} \alpha_s \leq 1$ (for every $\bar{s} \in S$), such that $p(s' \mid s) = \alpha_{s'}$ whenever $s' \neq s$.

To motivate this assumption, consider the following story. Changes in the state are due to extraneous shocks, which occur at random times. That is, once drawn, the state remains constant until an exogenous shock occurs. When such a shock occurs, the state is drawn anew, according to m. The inter-arrival times of the successive shocks are i.i.d., and follow a geometric distribution. In that case, **Assumption A** is met. Indeed, it suffices to set $\alpha_t = \pi \times m(t)$ for every $t \in S$, where π is the per-stage probability of a shock. The parameter π is here a measure of the state persistence. When π increases from 0 to 1, the situation evolves from one in which the state remains constant through time, to a situation in which successive states are independent.

When $\pi = 1$, the successive states are independent, and identically distributed according to m. Thus, **Assumption A** holds in the case of i.i.d. states.

Note that **Assumption A** also holds in the benchmark case where there are only two possible states. Indeed, denoting the two states by s_1 and s_2, it suffices to set $\alpha_1 = p(s_1 \mid s_2)$ and $\alpha_2 = p(s_2 \mid s_1)$.[7]

---

t in exactly N stages is positive.

[6]Sequential equilibria are defined only for finite games, but the concept has a natural extension to our setup, see Section 5.1.2.

[7]This case need not fit into the first case, as shown by the Markov chain with 2 states where the probability of moving from s to s' is $2/3$ whenever $s \neq s'$.

As a further simple illustration, consider a symmetric random walk on three states. That is, whenever in a state, the chain moves to each of the two other states with probability $\frac{1}{2}$. Again, **Assumption A** is met, with $\alpha_s = \frac{1}{2}$ for each $s \in S$.

Albeit restrictive (as we will see later), **Assumption A** allows for a variety of Markov chains, and it is met in many existing papers, including Athey and Bagwell (2008), Phelan (2006) //**To be completed**.//

For most of the paper, a strategy of the sender maps past and current realized states, and past play, into a mixed message, and is thus a function $\sigma : \cup_{n \geq 0}(S \times A \times B)^n \times S \to \Delta(A)$, while a strategy of the receiver is a function $\tau : \cup_{n \geq 0}(A \times B)^n \to \Delta(B)$. A *stationary* strategy of the receiver is a map $y : A \to \Delta(B)$, with the interpretation that the receiver chooses his action according to $y(\cdot \mid a) \in \Delta(B)$ whenever told $a \in A$.

3   An Example

We here illustrate the main results using a simple example. There are two states, L and R, and two actions for the receiver, l and r. The set A is irrelevant, and we will identify it with S. Successive states are independent and equally likely. Payoffs are given by the two tables in Figure 1, where c is a fixed parameter, with $c \in (1,2)$.

	l	r
	$c, 2$	$2, 1$

State L

	l	r
	$1, -1$	$2, 1$

State R

Figure 1: The payoffs of the two players.

The one-shot information transmission game has a unique equilibrium, in which the receiver plays r with probability 1. To see this, assume by way of contradiction that there is a message $\bar{a}$ that is sent with positive probability (in some state) and following which the receiver plays l with positive probability. By sequential rationality, the belief held by the receiver following $\bar{a}$ must assign a probability of at least $\frac{2}{3}$ to the state being L. Since both states are *ex ante* equally likely, this implies that there is some message $\tilde{a}$, following which the receiver assigns probability higher than $\frac{1}{2}$ to the state being R. By sequential rationality, the receiver plays r following $\tilde{a}$. As a consequence, in both states, the expected payoff of the sender is equal to 2

when sending $\tilde{a}$, and is less than 2 when sending the message $\bar{a}$. Therefore, it cannot be that $\bar{a}$ is sent with positive probability – a contradiction.

All equilibria in the one-shot game are therefore babbling equilibria.[8] Plainly, the dynamic game admits a babbling equilibrium, in which the receiver believes the announcements of the sender to be uninformative, and plays r in every stage (and the sender always sends the same message). In addition, the receiver can always choose to ignore the announcements of the sender, and to play r in every stage, thereby getting 1. As a result, the babbling equilibrium is the *worst* equilibrium for the receiver, in both the one-shot and in the dynamic game.

We claim that the dynamic game has equilibrium payoffs that are arbitrarily close to $(\frac{2+c}{2}, \frac{3}{2})$. This has the surprising implication that the lowest equilibrium payoff of the sender may be lower in the dynamic game than in the static game. Here is the intuition. Player 1 announces the true state at every stage. Player 2 listens to the announcements of the sender, and plays l when told L, r when told R. This obviously raises the concern that the sender may choose to announce R in every stage. To prevent this from happening, the receiver monitors the announcements of the sender, and stops listening if there is an obvious bias (to either L or R). Under the constraint that he should announce both states equally often, the expected payoff of the sender is highest when he reports truthfully.

While this intuition is simple, formalizing it into an equilibrium of the discounted game is not straightforward. Indeed, because payoffs are discounted, the sender may have a preference to send at first the message R more frequently.

We start with a simple construction that yields an equilibrium payoff that differs from $(2, 1)$. Assume that the discount factor satisfies $\delta > \dfrac{4 - 2c}{3 - c}$.

- At odd stages, the sender announces truthfully the current stage, and the receiver plays l if told L, and r if told R.

- At even stages, the sender announces a constant message, and the receiver plays the action that he did *not* play in the previous stage.

- If the receiver deviates, both players switch to the babbling equilibrium forever.

---

[8]In the sense that the action of the receiver is independent of the message sent by the sender.

We now verify that this construction is indeed an equilibrium. If the players follow this construction, their expected payoffs are equal to $\frac{1}{1+\delta}\left(\frac{2+c}{2}+\delta\frac{5+c}{4}\right)$ and $\frac{1}{1+\delta}\left(\frac{3}{2}+\frac{3}{4}\delta\right)$ respectively. Because a deviation of the receiver is followed by the babbling equilibrium, which yields 1 to the receiver, while the expected payoff of the receiver at every stage is at least 1, no other strategy of the receiver yields a higher payoff. Regarding the sender, it is sufficient to show that he cannot profit by deviating in any block of two stages. In such a block the sender has two possible deviations: to announce L in the first stage of the block when the true state is R, and to announce R in the first stage of the block when the true state is L. In the former case, he gets 1 at the first stage and 2 in the second (instead of 2 at the first stage and $\frac{1+c}{2}$ at the second stage if he announces truthfully). In the latter case, he gets 2 at the first stage and $\frac{1+c}{2}$ at the second stage (instead of c at the first stage and 2 at the second stage if he announces truthfully). The choice of δ ensures that none of these deviations is profitable.

To get payoffs closer to $(\frac{2+c}{2},\frac{3}{2})$, we will be relying on a slightly more complex construction. We let the size $2N$ of a block be large enough, so that a law of large numbers will apply. Given N is fixed, we let the discount factor δ be high enough, so that the contribution of any individual block to the overall discounted payoff is very small.

We first describe a pure strategy τ of the receiver. In each block (unless if the receiver has deviated earlier, see below), the receiver listens to the sender's announcements, plays l if told L, r if told R, *until* the number of announcements of either L or R exceeds N. When this happens, the receiver stops listening to the sender's announcements, and plays the least frequent action until the end of the current block.[9] If indeed the sender reports truthfully the current state, with high probability, the receiver will stop listening to the sender only 'shortly' before the end of the block, and the expected payoff is therefore close to $(\frac{2+c}{2},\frac{3}{2})$.

In contrast with the situation examined above, it need not be optimal for the sender to report truthfully when facing τ. However, a crucial insight is that *any* best reply of the sender to τ must be reporting truthfully most of the time, with high probability. To see why, observe that any best reply achieves a payoff of at least, say, $\frac{2+c}{2}-\varepsilon$. But since the receiver plays both actions l and r equally likely on each block, this implies that with high probability the action of the receiver matches the state, most of the time.

---

[9]An alternative construction, that we adopt in the general case, is for the receiver to generate a specific sequence of fictitious announcements, and continue as if the sender's announcements were equal to the fictitious ones.

We let σ be any pure best-reply of the sender to τ. On the equilibrium path, we let players play according to σ and τ. By construction, the equilibrium property holds for the sender. To deter the receiver from deviating, both players switch forever to the babbling equilibrium once a deviation of the receiver is detected. Since blocks are short, the expected continuation payoff of the receiver is close to $\frac{3}{2}$ following any history, while the receiver gets a payoff of 1 (or close to 1) if he deviates.

This construction hinges on few simple properties. On the one hand, the receiver follows most of the time a stationary strategy, $y : S \to \Delta(B)$, and his action choices are monitored by the sender. To ensure that the threat of switching to the babbling equilibrium provides the receiver with the appropriate incentives, it is enough that the expected payoff of the receiver be higher than v^r, the babbling equilibrium payoff. Provided that the sender truthfully reports the state most of the time, this individual rationality condition writes as

$$\sum_{s \in S} m(s) u^r(s, y(\cdot \mid s)) \geq v^r.$$

Meanwhile, the receiver performs basic consistency checks on the successive announcements of the sender. Here, the receiver simply makes sure that the different states are announced according to their theoretical frequency (possibly by substituting actual announcements by fictitious announcements). When facing the stationary strategy $y : S \to \Delta(B)$, the expected payoff of the sender is a function of the average joint distribution of the true state, and of the sender's announcement. Denoting by μ the average (expected) distribution of this pair $(\mathbf{s}_n, \mathbf{a}_n)$, the statistical checks by the receiver restrict the sender to distributions μ whose marginal on S and on A are both equal to m. As soon as truth-telling maximizes the payoff of the sender under this restriction, such checks are effective in preventing deviations by the sender. This condition takes the following form:

$$\sum_{s \in S} m(s) u^s(s, y(\cdot \mid s)) \geq \sum_{s \in S, t \in A} \mu(s, t) y(s, y(\cdot \mid t)),$$

for every distribution $\mu \in \Delta(S \times A)$ whose marginals on both S and A are equal to m.

//**Nicolas: the previous version was mentioning threats by the receiver. This does not work**//

4 Results

4.1 Statement of the main results

We provide a characterization of the payoff vectors that can be approximated by sequential equilibrium payoffs, when players are sufficiently patient. We start with some notations. Let T be a copy of the set of states S. Elements of T are interpreted as *fictitious* states, and will represent messages that the sender sends to the receiver.

/// **Eilon: because I assumes that $A = S$, we can replace throughout T by A. I did not do it yet so that no-one changes it later back to T.** ///

Building on the example, we denote by $\mathcal{M} \subset \Delta(S \times T)$ the set of *copulas* based on m; that is, the set of distributions μ over $S \times T$ whose marginals on S and on T are both equal to m. The set $\mathcal{M}$ is defined with a finite number of linear inequalities, hence it is a compact convex polyhedron, so it has finitely many extreme points.

We denote by $\mu_0 \in \mathcal{M}$ the specific distribution defined as $\mu_0(s, s) = m(s)$ for each $s \in S$, and $\mu_0(s, t) = 0$ if $s \neq t$. Under μ_0, the fictitious and the true states coincide *a.s.* Thus, the distribution μ_0 is the long-run average distribution of the sequence $(\mathbf{s}_n, \mathbf{t}_n)_n$ when the sender reports truthfully.

Given a copula $\mu \in \mathcal{M}$, and a stationary strategy $y : S \to \Delta(B)$, we set

$$U(\mu, y) := \sum_{s \in S, t \in T} \mu(s, t) u(s, y(t)) \in \mathbf{R}^2.$$

This is the expected payoff vector when the sender's report is chosen by $\mu(\cdot \mid s)$, and the receiver plays y. The babbling equilibrium payoff of the receiver is given by

$$v^r := \max_{b \in B} U^r(\mu_0, b) = \max_{b \in B} \sum_{s \in S} m(s) u^r(s, b). \tag{1}$$

Note that v^r is also equal to the $\min \max$ value of the receiver in the dynamic game.

We let $E(\mathcal{M})$ denote the set of payoff vectors $U(\mu_0, y)$, where $y : S \to \Delta(B)$, that satisfy

C1. $U^s(\mu_0, y) \geq U^s(\mu, y)$ for every $\mu \in \mathcal{M}$.

C2. $U^r(\mu_0, y) \geq v^r$;

Condition **C2** is the individual rationality condition of the receiver: the receiver's payoff is at least his min-max value. Condition **C1** reads as an incentive compatibility condition of the

sender. It may be interpreted as saying that truth-telling is optimal for the sender, when the successive announcements of the sender are constrained to be statistically indistinguishable from truth-telling. Note however that the version of 'statistical indistinguishability' that is implicit in **C2** is a very *weak* one. Indeed, it is only required that the frequencies of the different announcements be equal to those of truth-telling, yet the frequency of patterns of states longer than 1 under truth-telling may differ from those under μ.

For a given stationary strategy y, the map $\mu \mapsto U^s(\mu, y)$ is linear. Hence, condition **C1** holds as soon as the inequality holds for each of the finitely many extreme points of $\mathcal{M}$, one of which is μ_0.

For a given μ, the map $y \mapsto U(\mu, y)$ is linear as well. Hence, the set of y such that both **C1** and **C2** hold is a compact convex polyhedron in $\Delta(B)^S$. It follows that the set $E(\mathcal{M})$ is also a compact convex polyhedron.

We define $\widehat{E}(\mathcal{M})$ as the set of payoff vectors $U(\mu_0, y) \in E(\mathcal{M})$ where the inequalities in **C1** and **C2** are strict. That is, $\widehat{E}(\mathcal{M})$ is the set of vectors $U(\mu_0, y)$, $y : S \to \Delta(B))$, such that

D1. $U^s(\mu_0, y) > U^s(\mu, y)$ for every $\mu \in \mathcal{M}$, $\mu \neq \mu_0$.

D2. $U^r(\mu_0, y) > v^r$;

Note that condition **D1** holds as soon as the inequality $U^s(\mu_0, y) > U^s(\mu, y)$ is satisfied for each of the finitely many extreme points $\mu \neq \mu_0$ of $\mathcal{M}$.

The set $\widehat{E}(\mathcal{M})$ is an open subset of $E(\mathcal{M})$, but it need not be equal to the relative interior of $E(\mathcal{M})$. When we want to emphasize the dependency of the set $\widehat{E}(\mathcal{M})$ on the game G, we write $\widehat{E}_G(\mathcal{M})$.

We denote by NE_δ and SE_δ the set of Nash equilibrium payoffs and of sequential equilibrium payoffs. Plainly, $SE_\delta \subseteq NE_\delta$ for every δ,

Our first main result, Theorem 1 below, assumes the existence of a public randomizing device, that sends extraneous signals in every stage.

Theorem 1 *Suppose that there exists a public randomizing device, which outputs a (uniformly distributed) number in $[0, 1]$ in every stage, after the announcement of the sender. If $\widehat{E}(\mathcal{M}) \neq \emptyset$ then*

$$E(\mathcal{M}) \subseteq \liminf_{\delta \to 1} SE_\delta.$$

/// **An inaccuracy:** SE_δ **is the set of sequential equilibrium payoffs in the game with correlation device, but it was defined before the presentation of the correlation device...** ///

Theorem 1 means that for every $\gamma \in E(\mathcal{M})$ and for every $\varepsilon > 0$ there is $\delta_0 < 1$ such that, for every $\delta \geq \delta_0$, the δ-discounted game has a sequential equibrium payoff (in pure strategies) within ε of γ. We will actually prove the stronger statement that δ_0 can be chosen to be independent of γ: as $\delta \to 1$,

$$\sup_{\gamma \in E(\mathcal{M})} d(\gamma, SE_\delta) \to 0.$$

Our second result is valid, whether or not there is such a public randomizing device, but it requires Assumption A.

Theorem 2 *Suppose that **Assumption A** holds. Then, for every $\delta < 1$, one has*

$$NE_\delta \subseteq E(\mathcal{M}).$$

Provided that $\widehat{E}(\mathcal{M}) \neq \emptyset$ and that **Assumption A** is met, Theorems 1 and 2 thus imply that the set of sequential equilibrium payoffs converges to the set $E(\mathcal{M})$.

Checking whether a given payoff vector $U(\mu, y)$ belongs to $E(\mathcal{M})$ seems to require the computation of the extreme points of the set $\mathcal{M}$, which is not an easy task. Fortunately, it turns out that condition **C1** is equivalent to a much simpler one.

Lemma 3 *Let $y : S \to \Delta(B)$ be given. Conditions **C1** and **D1** are respectively equivalent to conditions **Q2** and **Q'2** below.*

Q2. $\displaystyle\sum_{s \in S} u^s(s, y(\cdot \mid s)) \geq \sum_{s \in S} u^s(s, y(\cdot \mid \phi(s)))$, *for every permutation ϕ over S.*

Q'2. $\displaystyle\sum_{s \in S} u^s(s, y(\cdot \mid s)) > \sum_{s \in S} u^s(s, y(\cdot \mid \phi(s)))$, *for every permutation $\phi \neq id$ over S.*

Interestingly, conditions **Q2** and **Q'2** do *not* involve the distribution m.

The statement of Theorem 1 is unsatisfactory in one important respect: while non-empty interior requirements in existing Folk Theorems are generically satisfied, the condition $\widehat{E}(\mathcal{M}) \neq \emptyset$ is *not*, as the next example shows.

Example 4 *Consider the game depicted in Figure 2, where there are two states, and the receiver has two actions.*

	l	*r*
	1, 1	0, 0

State L

	l	*r*
	1, 1	0, 0

State R

Figure 2: The game in Example 4.

Here, $v^r = 1$, and the stationary strategy y_* which plays l irrespective of the announcement is the only stationary strategy that satisfies **C2**. Hence, $E(\mathcal{M})$ contains a single payoff vector, $(1, 1)$, and $\widehat{E}(\mathcal{M})$ is empty. When payoffs are slightly perturbed, the strategy y_* remains the only strategy satisfying **C2**, therefore $\widehat{E}(\mathcal{M}) = \emptyset$ for any such perturbation. ♦

Example 4 suggests that if all strategies y for which $U(\mu_0, y)$ is in $E(\mathcal{M})$ are constant strategies, then the set $\widehat{E}(\mathcal{M})$ is empty, even when payoffs in the game are slightly perturbed. We build on this intuition, and introduce a new condition.

Condition B. There is a *non-constant* map $y : S \to \Delta(B)$ such that $U(\mu_0, y) \in E(\mathcal{M})$.

In Theorem 5 below, we fix the transition function of the Markov chain p, and identify a game to a point in the space $\mathbf{R}^{2 \times S \times B}$ of payoff functions.

Theorem 5 *Let a game G be given.*

*If condition **B** holds for G, then any neighborhood of G contains a game G' with $\widehat{E}_{G'}(\mathcal{M}) \neq \emptyset$.*

*If condition **B** does not hold for G, there is a neighborhood $\mathcal{N}$ of G such that, for every game in $\mathcal{N}$, condition **B** does not hold.*

Theorem 5 allows us to complete the picture provided by Theorems 1 and 2. Indeed, let G be a game. If Condition B holds for G, Theorems 1 and 2 provide a characterization of the limit set of equilibrium payoffs for games arbitrarily close to G. If condition B does not hold for G, then all games close enough to G have only babbling equilibrium payoffs.

4.2   Comments

The relation between transitions and equilibria.

Note that the set $\mathcal{M}$ of copulas only depends on the invariant measure m, and not on finer details of the transition function. A striking implication of the characterization is that, under **Assumption A**, the limit set of equilibrium payoffs therefore only depends on the invariant measure m. In particular, the limit set of equilibrium payoffs is the same as when the states are drawn independently across stages. That is, the amount of state persistence is irrelevant for the determination of the limit set of equilibrium payoffs.

This property hinges on **Assumption A**, and may fail to hold more generally, as we will show through an example.

//**Nicolas: I left the comment below, but it sounds weird to have uniform equilibria popping up**//

/// **Eilon: We can delete it.** ///

If the initial state s were to remain fixed throughout the play, the game would fall into the class of repeated games with incomplete information introduced by Aumann and Maschler (1995). (This is the setup studied in Golosov et al. (2009)). In this case, the set of undiscounted equilibrium payoffs, which correspond to the case $\delta = 1$, is typically not equal to $E(\mathcal{M})$. Hence, there is a discontinuity in the set of equilibrium payoffs when δ goes to 0, where successive states become perfectly autocorrelated.[10]

By contrast, for a fixed discount factor, the set of equilibrium payoffs is upper hemi-continuous with respect to the transition function. The source of this apparent paradox can be traced back to the fact that, in loose terms, the convergence of the set SE_δ to $E(\mathcal{M})$ is slowlier, the more successive states are correlated.

 Equilibrium behavior.

Our construction (to be made precise in later sections), has the somewhat surprising feature that the sender reports truthfully, at least most of the time and with high probability. A direct intuition can be provided, that is reminiscent of the revelation principle in mechanism design. Let an equilibrium (σ, τ) be given. Consider the strategy profile where the sender reports truthfully, and the receiver first computes the message that the strategy σ would have sent,

---

[10]Our Theorems 1 nd 2 extend to cover the case of uniform equilibrium payoffs.

and next plays what τ would have played given this message. We argue loosely that this new profile (when supplemented with threats) is an equilibrium. The key to the argument is twofold. On the one hand, the sender can check that the receiver does indeed play as prescribed, and does not use the additional information provided by the knowledge of the true state. On the other hand, the threat of switching to the babbling equilibrium is effective because the knowledge of the state at a given stage becomes eventually valueless in predicting distant stages, because of the irreducibility property of the sequence of states.

//**Nicolas: I deleted the comment on the Folk Theorem.**//

The impact of repetition on the players.

The characterization implies that every equilibrium payoff of the one-shot game remains an equilibrium payoff in the dynamic game, provided players are patient enough. We stress that this property is not obvious here, since the game is not a repeated game. In particular, it would typically fail to hold if the state were constant throughout the play.

Let (σ, τ) be an equilibrium of the one-shot game. Let $y : S \to \Delta(B)$ be the stationary strategy defined as

$$y(b \mid a) = \sum_{s \in S} \sigma(a \mid s)\tau(a)[b].$$

Note that the expected payoff under (σ, τ) is $U(\mu_0, y)$. We claim that $U(\mu_0, y) \in E(\mathcal{M})$, so that by Theorem 1 it is a sequential equilibrium payoff in the repeated game. Indeed, because the receiver can guarantee v^r in the one-shot game, condition **C2** holds. Because σ is a best reply to τ in the one-shot game, the inequality in **C1** holds for every μ, and in particular for every $\mu \in \mathcal{M}$.

This result has the implication that the lowest equilibrium payoff of the sender in the repeated game cannot be higher than his lowest equilibrium payoff in the one shot game. As the example in Section 3 shows, it can be in fact strictly lower.

On the other hand, the lowest equilibrium payoff of the receiver in both the one-shot game and the repeated game is equal to his babbling equilibrium payoff v^r.

The information of the sender.

Theorems 1 and 2 hold as soon as the sender knows the current state. As will be clear from the proof, they continue to hold if the sender knows *more*. In particular, they hold in the extreme case where the sender learns the entire sequence of realized states in stage 1, or in any intermediate setup.

The assumption $\widehat{E}(\mathcal{M}) \neq \emptyset$.

In the light of the existing results for repeated games, it is not surprising that some non-empty interiority type of assumption is needed. We refer to Mailath and Samuelson (???)for a survey on similar results and a discussion of the role of this assumption.

As the next example illustrates, the conclusion of Theorem 1 fails to hold if $\widehat{E}(\mathcal{M}) = \emptyset$.

Example 6 *Let there be two states and two actions for the receiver. The payoffs in the two states are given by the tables in Figure 3. We assume that the successive states are independent and that the two states are equally likely.*

	l	r			l	r
	$0.5, 1$	$1, 1$			$0, 0$	$1, 1$

State L   *State R*

Figure 3: The game in Example 6.

The strategy which plays r irrespective of the announcement is weakly dominant in the one-shot game, and thus, $v^r = 1$. Consider now the stationary strategy y defined by $y(l \mid L) = y(r \mid R) = 1$. The payoff vector $U(\mu_0, y) = (\frac{3}{4}, 1)$ is in $E(\mathcal{M})$. However, we claim that $(1, 1)$ is the unique equilibrium payoff, irrespective of δ. Here is why. Consider any equilibrium (σ, τ). Plainly, the equilibrium payoff of the receiver is equal to 1. In particular, with probability 1 the receiver plays r whenever the current state is R. This implies that in every stage, and for a.e. past history, there is one (possibly history-dependent) message following which the receiver plays r, and which is assigned positive probability by σ. But then, the sender gets a payoff of one by assigning probability 1 to this specific message in every stage. ♦

//**Nicolas: the previous conclusion was too weak for our purposes, and I was not convinced by the loose argument; I rewrote it**//

On the role of Assumption A.

As the next example shows, the conclusion of Theorem 2 fails to hold if **Assumption A** is not satisfied. Thus, in general, the limit set of sequential equilibrium payoffs does not only depend on the invariant measure, but also on fine details of the transition function.

Example 7 *Consider a game with 5 states $S := \{0, 1, 2, 3, 4\}$. The sequence of states follows a random walk on S. When in s, the chain moves either to $s+1 \pmod 5$ or to $s-1 \pmod 5$ with*

equal probabilities. The action set B of the receiver coincides with S, and the payoff function is described in Figure 4, where $c > 1$.

	$b = 0$	$b = 1$	$b = 2$	$b = 3$	$b = 4$
$s = 0$	$1, 1$	$c, 0$	$0, 0$	$0, 0$	$0, 0$
$s = 1$	$c, 0$	$1, 1$	$0, 0$	$0, 0$	$0, 0$
$s = 2$	$0, 0$	$0, 0$	$1, 1$	$0, 0$	$0, 0$
$s = 3$	$0, 0$	$0, 0$	$0, 0$	$1, 1$	$0, 0$
$s = 4$	$0, 0$	$0, 0$	$0, 0$	$0, 0$	$1, 1$

Figure 4: The game in Example 7.

Thus, both players receive a payoff of one if the action matches the current state, and of zero otherwise, except when the receiver chooses action 1 in state 0, or action 0 in state 1.

The payoff vector $(1, 1)$ is not in $E(\mathcal{M})$ as soon as $c > 1$. Indeed, the stationary strategy $y : S \to \Delta(B)$ defined by $y(s \mid s) = 1$ is the only strategy such that $U(\mu_0, y) = (1, 1)$. But then, the sender profits by reporting $a = 1$ whenever $s = 0$, and $a = 0$ whenever $s = 1$. On the other hand, $(1, 1)$ is an equilibrium payoffs, as soon as $c < \frac{3}{2}$, provided the players are patient enough. Indeed, consider the strategy of the receiver in which he matches the announcement of the sender, as long as $|\mathbf{a}_{n+1} - \mathbf{a}_n| = 1$ modulo 5, and switches forever to the babbling equilibrium (e.g., playing always $b = 4$) if $|\mathbf{a}_{n+1} - \mathbf{a}_n| \neq 1$ modulo 5 for some stage n. Provided c is not too large, the best response of the sender is to report the true state. If instead, say, the sender chooses to report $a = 1$ when in fact $s = 0$ in a given stage, he gains $c - 1$, but then in the next period, with probability $\frac{1}{2}$ the new state will be $s = 4$, and then he will either report $a \in \{0, 2\}$ and receive 0, or report $a \in \{1, 3, 4\}$ and be punished with the babbling equilibrium payoff $\frac{1}{5}$ forever. Provided the players are patient enough, such a deviation is not profitable.

In this equilibrium, the receiver checks that the one-step transitions between *successive* announcements are consistent with the transitions of the Markov chain. As it turns out, under **Assumption A**, such a sophisticated statistical analysis of the announcements is not more powerful than a statistical analysis which is based only on the empirical frequencies of the different announcements. ♦

Imperfect monitoring.

Let us assume here that successive states are independent. Results continue to hold if the receiver only observes a noisy, *public* version of the sender's message (provided the definition

of $U(\mu, y)$ is modified in an appropriate way). They still hold if the receiver observes a noisy, *public* signal of the current state, provided the individual rationality level v^r is modified in the proper way. They also hold, without changes, if the sender only observes a noisy, public signal of the receiver's action. What happens in any of these variants when signals are private is beyond the scope of the paper.

We briefly conclude this section by discussing the case where the sender fails to receive any information relative to the receiver's choices. In spite of this feature, the game does not reduce to a sequence of successive, independent, one-shot games, because of the ability of the receiver to monitor the sender. In particular, it is easy to construct examples with equilibrium payoffs that lie outside of the convex hull of the set of equilibrium payoffs in the one-shot game.

We refer to the game where the sender does not observe the actions of the receiver as to the *blind* game. Denote by NE_δ^b the set of all Nash equilibrium payoffs of the blind game. We prove that the value of monitoring is positive, in the sense that allowing the sender to monitor the receiver has a non-ambiguous effect on the equilibrium set.

Proposition 8 *The set NE_δ^b is a subset of NE_δ.*

The inclusion is strict in general, as Example 9 below shows.

Example 9 *There are two states $\{L, R\}$, and three actions for the receiver, $B = \{l, m, r\}$. The payoffs in the two states are given in Figure 5.*

l	m	r
$2, 2$	$0, 0$	$0, 0$

State L

l	m	r
$0, 0$	$2, 2$	$0, 3$

State R

Figure 5: The game in Example 9.

We claim that $(2, 2)$ is an equilibrium payoff when the sender observes the actions of the receiver, but it is no longer an equilibrium payoff when the sender does not observe the receiver's actions.

Note first that $v^r = \frac{3}{2}$, and that $\widehat{E}(\mathcal{M}) \neq \emptyset$. By Theorem 1, $(2, 2) \in E(\mathcal{M})$, so that $(2, 2) \in \liminf_{\delta \to 1} SE_\delta$.

We now argue that $(2, 2)$ is bounded away from the set NE_δ^b. Indeed, assume to the contrary that there is some equilibrium profile (σ, τ) of the blind game with a payoff close to $(2, 2)$. In

particular, with a probability close to one, there is a positive fraction of the stages in which the current state is R and the receiver plays m. Consider the strategy τ' which plays as τ, except that τ' plays r whenever τ would play m. Because the sender does not observe the receiver's actions, he cannot tell whether the receiver uses τ or τ', and therefore τ' is a profitable deviation of the receiver: it yields the receiver payoff close to $2\frac{1}{2}$. ♦

4.3   On the role of the randomizing device

The randomizing device is not needed in the proof of Theorem 2 to implement payoffs $U(\mu_0, y)$, whenever $y(\cdot \mid s)$ is a pure strategy: it assigns probability one to some action $b(s)$, for each $s \in S$. However, as soon as $y(\cdot \mid s)$ is a truly mixed distribution for some state s, it may be impossible to dispense with the randomizing device, as we now argue by means of an example.

Let there be two states, L and R. The successive states are drawn independently in every period, and each of the two states is equally likely. The receiver has three actions, denoted $B = \{l, m, r\}$. The payoffs are given in Figure 6.

l	m	r
$3, 0$	$0, 4$	$2, 1$

State L

l	m	r
$1, -5$	$4, -4$	$2, 1$

State R

Figure 6: The payoffs of the players.

Plainly, $v^r = 1$. Define y_* to be the stationary strategy such that $y_*(\cdot \mid R)$ assigns probability one to r, and $y_*(\cdot \mid L)$ assigns probabilities $\frac{2}{3}$ and $\frac{1}{3}$ to l and m, respectively. Then $U(\mu_0, y) = (2, \frac{7}{6})$, and one can verify that $U(\mu_0, y_*) \in E(\mathcal{M})$ while $\widehat{E}(\mathcal{M}) \neq \emptyset$. Thus, using Theorem 1, the vector $(2, \frac{7}{6})$ can be approximated by sequential equilibrium payoffs, when players are sufficiently patient, provided a randomizing device is available.

We now assume that such a device is not available. Since successive states are independent, the dynamic game can be viewed as a infinite repetition of the one-shot information transmission game. With this interpretation, an *action* of the sender in the one-shot game is a map $x : S \to A$, while an action of the receiver is a map $y : A \to B$. Given an action profile (x, y), payoffs are random, and take the value $u(s, y(x(s)))$ with probability $m(s)$, for $s \in S$. Players then receive the public signal $(x(s), y(x(s)))$.

We will rely on Fudenberg, Levine and Maskin's (1994) characterization of the limit set of perfect public equilibrium (PPE) payoffs in repeated games with public signals. Some care is

needed, as there are two dimensions according to which our repeated game does not fit into their setup. First, they assume that a player's payoff depends deterministically on his own action and on the public signal, while payoffs here depend randomly on the entire action profile (x, y). Second, their result is a characterization of public equilibrium payoffs, while we focus on sequential equilibrium payoffs.

We briefly argue that their result nevertheless applies. On the one hand, their result is still valid for games where payoffs depend on the entire action profile.[11] Next, it can be verified that the auxiliary game in which stage payoffs are defined to be the *expected* stage game payoffs in our game (given the action profile) has the same set of PPE payoffs. Thus, their result provides a characterization of the limit set of PPE payoffs for our game. On the other hand, let (σ, τ) be a sequential equilibrium of our game, and define a public strategy profile $(\bar{\sigma}, \bar{\tau})$ as follows. Let any public history $\bar{h}$ be given. At $\bar{h}$, we let $\bar{\sigma}$ play the expectation of the mixed move played by σ, where the expectation is computed w.r.t. the belief held by the receiver at the information set which contains $\bar{h}$. We define $\bar{\tau}(h)$ by exchanging the roles of the two players. It can be verified that $(\bar{\sigma}, \bar{\tau})$ is a public perfect equilibrium of the repeated game.

/// Eilon: below, is indeed $Z = A \times B$??? I rewrote the following paragraph, hopefully for the good. ///

Fudenberg et al. (1994) showed that $\gamma \in \mathbf{R}^2$ is a PPE payoff if and only if there is $\lambda \in \mathbf{R}^2$ such that the solution $k(\lambda)$ of a certain optimization problem $\mathcal{P}(\lambda)$ satisfies $\lambda \cdot \gamma > k(\lambda)$.

We set $\gamma = (2, \frac{7}{6})$, and we denote by $Z = A \times B$ the set of public signals in our game. We show that the condition of Fudenberg et al. (1994) holds for $\lambda_* = (0, 1)$. For the sake of completeness we recall Fudenberg et al. (1994) characterization, and we then highlight the main ideas of the proof that γ is a PPE payoff.

Denote by $k(\lambda_*)$ the value of the optimization problem $\mathcal{P}$:

$$\sup V^r,$$

where the supremum is taken over all $(V^s, V^r) \in \mathbf{R}^2$, and all $\phi : Z \to \mathbf{R}^2$, such that

- $\phi^r(z) \leq 0$ for every $z \in Z$;

- (V^s, V^r) is a Nash equilibrium payoff of the one-shot game, with payoff function defined

---

[11]This can be seen from their proof or, alternatively, deduced from Hörner et al. (2009).

by:

$$\sum_{s \in S} m(s) \left(u(s, y(x(s))) + \phi(x(s), y(x(s))) \right), \tag{2}$$

for each action pair (x, y).

Let $\phi : Z \to \mathbf{R}^2$ be any map such that $\phi^r(z) \leq 0$ for each $z \in Z$, and let (α, β) be any (possibly mixed) equilibrium of the one-shot game (2), with payoff (V^s, V^r). We will prove that $V^r < \frac{7}{6}$. We argue by contradiction, and assume that $V^r \geq \frac{7}{6}$. We distinguish between two cases.

Assume first that $\alpha : S \to \Delta(A)$ is pooling: the distribution of messages is the same in both states. Then, since $\phi^r(z) \leq 0$, the expected payoff of the receiver is not higher than

$$\max_{b \in B} \sum_{s \in S} m(s) u^r(s, b) = 1.$$

Thus, $V^r \leq 1$, which is the desired contradiction.

Assume next that α is not pooling. Up to a relabelling of the messages, we may then assume w.l.o.g. that the sender always tells the truth with positive probability. That is, $\alpha(s \mid s) > 0$, for each $s \in S$. We denote by $\tilde{\beta}(\cdot \mid s)$ the conditional distribution of the receiver's move under (α, β), conditional on the state being $s \in S$. Denoting by $s \neq t$ the two states, the equilibrium property for the sender in the game (2) then implies that

$$\sum_{b \in B} \beta(b \mid s) \left(u^s(s, b) + w^s(s, b) \right) \geq \sum_{b \in B} \beta(b \mid t) \left(u^s(s, b) + w^s(t, b) \right),$$

with equality if $\alpha(\cdot \mid s)$ assigns positive probability to both messages, and

$$\sum_{b \in B} \beta(b \mid t) \left(u^s(t, b) + w^s(s, t) \right) \geq \sum_{b \in B} \beta(b \mid s) \left(u^s(t, b) + w^s(s, b) \right).$$

Using the two inequalities, one can verify that

$$u^s(s, \tilde{\beta}(\cdot \mid s)) + u^s(t, \tilde{\beta}(\cdot \mid t)) \geq u^s(s, \tilde{\beta}(\cdot \mid t)) + u^s(t, \tilde{\beta}(\cdot \mid s)).$$

By Lemma 3, condition **C2** therefore holds for the stationary strategy $\tilde{\beta}$.

On the other hand, since $\phi^r(z) \leq 0$ for each z, the expected payoff V^r to the receiver does not exceed $U^r(\mu_0, \tilde{\beta})$. Hence, $U^r(\mu_0, \tilde{\beta}) \geq \frac{7}{6}$. This readily implies that $U(\mu_0, \tilde{\beta}) \in E(\mathcal{M})$.

Next, one can verify that the highest payoff $U^r(\mu_0, \tilde{\beta})$ to the receiver, over the whole set $U(\mu_0, \tilde{\beta}) \in E(\mathcal{M})$, is equal to $\frac{7}{6}$. In addition, the unique strategy $\tilde{\beta}$ that achieves such a payoff

21

is the strategy y_*. Since the supports of $y_*(L)$ and $y_*(.\mid R)$ are distinct, it must therefore be that α is truth-telling: $\alpha(s \mid s) = 1$ for each s. Therefore, β is equal to y_*.

Since $V^r \geq \frac{7}{6}$ and $V^r \leq U^r(\mu_0, y_*)$, one also has $V^r = U^r(\mu_0, y_*)$. In particular, the expectation of $\phi^r(z)$ under the equilibrium profile (α, β) must be equal to zero. Since $\phi^r(z) \leq 0$ for each z, this implies that $\phi^r(z) = 0$, for each public signal z that receives positive probability under (α, β).

Using this, we finally claim that the equilibrium condition for the receiver in the game (2) is violated. Indeed, when told L, the strategy $\beta = y_*$ assigns positive probability to both l and m. Hence, $\phi^r(L, l) = \phi^r(L, m) = 0$ by the previous paragraph. On the other hand however, $u^r(L, m) > u^r(L, l)$, hence the receiver is not indifferent between both actions. This is the desired contradiction.

5 Proofs of Theorems 1 and 2

In the present section we provide the main steps in the proof of Theorems 1 and 2. The technical parts of the proofs are relegated to the appendix. The proof of Theorem 5, which is particular to the model we study, also appears in the appendix.

5.1 Proof of Theorem 1

Denote the set of all stationary strategies of the receiver that induce a payoff in $E(\mathcal{M})$ by

$$Y(\mathcal{M}) = \{y : S \to \Delta(B) \text{ such that } U(\mu_0, y) \in E(\mathcal{M})\}.$$

Because $E(\mathcal{M})$ is defined by linear inequalities, $Y(\mathcal{M})$ is convex. Since $\widehat{E}(\mathcal{M}) \neq \emptyset$, there is $y_0 : S \to \Delta(B)$ such that $U^r(\mu_0, y_0) > v^r$ and $U^s(\mu_0, y_0) > U^s(\mu, y_0)$ for every $\mu \in \mathcal{M}$, $\mu \neq \mu_0$. Fix $\varepsilon > 0$ once and for all.

Define $Y_\varepsilon(\mathcal{M}) := \varepsilon y_0 + (1 - \varepsilon)Y(\mathcal{M})$, and $E_\varepsilon(\mathcal{M}) = U(\mu_0, Y_\varepsilon(\mathcal{M}))$. Because $Y(\mathcal{M})$ and $E(\mathcal{M})$ are convex, $Y_\varepsilon(\mathcal{M}) \subseteq Y(\mathcal{M})$, and $E_\varepsilon(\mathcal{M}) = \varepsilon U(\mu_0, y_0) + (1 - \varepsilon)E(\mathcal{M}) \subseteq \widehat{E}(\mathcal{M})$.

We will prove the existence of $\delta_0 < 1$, such that $E_\varepsilon(\mathcal{M}) \subseteq SE_\delta$ for every $\delta \geq \delta_0$. This will imply $\sup_{\gamma \in E(\mathcal{M})} d(\gamma, SE_\delta) \to 0$, as desired.

5.1.1 A lower bound for $U^s(\mu_0, y) - U^s(\mu, y)$

By **C1**, the difference $U^s(\mu_0, y) - U^s(\mu, y)$ is non-negative for every $\mu \in E(\mathcal{M})$ and $y \in Y(\mathcal{M})$. We now provide a lower bound to this difference, in terms of the finitely many extreme points of $\mathcal{M}$.

Denote by $\mathcal{M}_e$ the (finite) set of extreme points of $\mathcal{M}$. Recall that $\mu_0 \in \mathcal{M}_e$. Set

$$c_1 := \min_{\{\mu_e \in \mathcal{M}_e, \mu_e \neq \mu_0\}} (U^s(\mu_0, y_0) - U^s(\mu_e, y_0)), \text{ and } c_2 := \max_{\{\mu_e \in \mathcal{M}_e, \mu_e \neq \mu_0\}} \|\mu_e - \mu_0\|_1.$$

Note that both c_1 and c_2 are positive.

Lemma 10 *For every* $\varepsilon > 0$, $y \in Y_\varepsilon(\mathcal{M})$ *and* $\mu \in \mathcal{M}$, *one has*

$$U^s(\mu_0, y) - U^s(\mu, y) \geq \frac{\varepsilon c_1}{c_2} \|\mu - \mu_0\|_1.$$

Proof. See Appendix. ∎

For each $N \in \mathbf{N}$, let m_N be the distribution over S that best approximates m among all distributions with rational coefficients whose denominator is N. Plainly, $\lim_{N \to \infty} \|m_N - m\|_1 = 0$.

5.1.2 The strategies

Let $y \in Y_\varepsilon(\mathcal{M})$ be given. We construct a sequential equilibrium (σ_*, τ_*) with payoff close to $U(\mu_0, y)$. The definition depends on an integer $N \in \mathbf{N}$ and on the discount factor δ.

We start by defining a periodic strategy profile (σ_0, τ_0) with period N that yields payoff close to $U(\mu_0, y)$. We divide the play into blocks of length N. At the beginning of each block, both players discard all previously held information, and re-start playing as from stage 1. Therefore, we focus on one block, say the first block.

Recall that $\mathbf{a}_n \in S$ stands for the announcement of the sender in stage n. Given $s \in S$, and $n \in \mathbf{N}$, we denote by $\mathbf{N}_n(s) = |\{k \leq n : \mathbf{a}_k = s\}|$ the number of stages where the sender announced the state to be s. We set

$$q := \min\{1 \leq n \leq B : \mathbf{N}_n(\mathbf{a}_n) > N m_N(\mathbf{a}_n)\}$$

$(\min \emptyset = +\infty)$. The interpretation is as follows. Each state $s \in S$ is allotted a quota of announcements in the block, equal to the product $N \, m_N(s)$. The stage q is the first stage in which the announcements of the sender exceed one of the quotas.

According to τ_0, the receiver 'listens' to the announcements of the sender until stage q. At that stage, the receiver starts following a predetermined sequence of *fictitious* announcements. To be formal, we let $(\mathbf{t}_n)$ be a process with the following properties:

F1. $\mathbf{t}_n = \mathbf{a}_n$ for $n < q$;

F2. For each $s \in S$, the equality $|\{n \leq N : \mathbf{t}_n = s\}| = N m_N(s)$ always hold;

F3. Conditional on $(\mathbf{a}_1, \ldots, \mathbf{a}_q)$, the variables $(\mathbf{t}_q, \ldots, \mathbf{t}_N)$ are deterministic.

We will refer to $\mathbf{t}_n$ as the *announcement* at stage n. Condition **F1** ensures that the announcements coincide with the true announcements prior to stage q; condition **F2** ensures that the entire sequence of announcements always satisfies the quotas; condition **F3** ensures in particular that the fictitious announcements are commonly known between the two players. The strategy τ_0 plays the mixed action $y(\mathbf{t}_n)$ in each stage $n = 1, \ldots, N$.

Let σ_0 be a pure best-reply strategy to τ_0 in the δ-discounted game. Several remarks are in order.

- We impose that, following any history that is consistent with τ_0, σ_0 is a best-reply of the sender in the continuation game.[12]

- Since τ_0 starts anew at the beginning of each block, we may and will assume that σ_0 has the same property, as long as the play is consistent with τ_0.

A little care is needed to construct a sequential equilibrium in pure strategies based on (σ_0, τ_0).

The first step is to transform τ_0 into a pure strategy. Here we use the public correlation device that sends a public message at every stage after the announcement of the sender. Denote by X_n the public signal at stage n. The random variables (X_n) are i.i.d. uniformly distributed over $[0, 1]$. We will use this sequence to de-randomize τ_0. Label the receiver's actions from 1 to $|B|$. Instead of playing the mixed action $y(\mathbf{t}_n)$, the strategy τ_0 instructs the receiver to choose the action $b \in B$ whenever $\sum_{i=1}^{b-1} y(i \mid t_n) \leq X_n < \sum_{i=1}^{b} y(i \mid t_n)$. Thus, τ_0 can be viewed as a pure

---

[12]That is, we also impose the best-reply condition following histories that are consistent with τ_0, but not with σ_0.

strategy. Note that because the signals $(X_n)_n$ are public, a deviation from the de-randomized strategy is detected by the sender.

We denote by σ_* the strategy of the sender that coincides with σ_0 as long as the receiver does not deviate from τ_0, and after a deviation repeats the same (babbling) announcement $\bar{a}$.

We next address the issue of designing a system of beliefs for the receiver that is consistent with σ_*, and that satisfies an additional property. Since the game involves randomizing devices with uncountably many outcomes, the standard definition of consistency does not apply. We denote by $\lambda \in \Delta(S)$ a distribution with full support and, for $\eta < 0$, we denote by σ_η the strategy that, following any history h_n, plays $\eta\lambda + (1 - \eta)\sigma_*(h_n)$.

One can check that, for $\eta > 0$, the beliefs of the receiver are uniquely defined by Bayes rule, and have a limit when $\eta = 0$.[13] Note that, following any history that is inconsistent with τ_0, the belief of the receiver in stage n is independent of $\mathbf{a}_n$.[14]

We denote by τ_* a strategy that coincides with τ_0 as long as the sender does not deviate, and that plays in each later stage n an action that (i) maximizes the current expected payoff of the receiver, given the belief held by the receiver in stage n, and (ii) does not depend on the announcements made by the sender since the deviation took place.

By construction, the strategy σ_* is sequentially rational at each node of the sender, while the strategy τ_* is sequentially rational at each node of the receiver that is inconsistent with τ_0.

We emphasize that the pure strategy profile (σ_*, τ_*) depends on N and on δ, even if this does not appear explicitly in the notation.

5.1.3 Equilibrium properties

We here show that for every $\eta > 0$, there is an integer $N_0 \in \mathbf{N}$ such that the following holds. For every $N \geq N_0$, there is $\delta_0 < 1$ such that, for every $\delta \geq \delta_0$, the profile (σ_*, τ_*) is a sequential equilibrium, and induces a payoff within η of $U(\mu_0, y)$.

Denote by μ_{σ_0, τ_0} the expected (undiscounted) joint distribution of the pair $(\mathbf{s}_n, \mathbf{t}_n)$ over a block of N stages. That is, for each $(s, t) \in S \times S$,

$$\mu_{\sigma_0, \tau_0}(s, t) := \mathbf{E}_{\sigma_0, \tau_0}\left[\frac{1}{N}\sum_{n=1}^{N} 1_{\{\mathbf{s}_n = s, \mathbf{t}_n = t\}}\right]$$

---

[13] And the convergence is uniform w.r.t. the receiver's information set.

[14] That is, should the sender fail to play the babbling announcement $\bar{a}$, the receiver sill interprets the sender's announcements as babbling.

is the expected frequency of the pair (s, t) over N stages.

The announcement $\mathbf{t}_n$ might differ from $\mathbf{s}_n$ because σ_0 is a best response to τ_0, and the optimal message of the sender need not coincide with the true state. The following proposition shows that nevertheless $\mathbf{t}_n = \mathbf{s}_n$ with high probability.

Proposition 11 *For every $\eta > 0$, there is $N_0 \in \mathbf{N}$, such that the following holds. For every $N \geq N_0$, there is $\delta_0 < 1$, such that, for every $\delta \geq \delta_0$, one has*

$$\|\mu_{\sigma_0, \tau_0} - \mu_0\| < \eta.$$

We first argue that Proposition 11 implies the desired result. Since the proof is standard, we limit ourselves to a sketch of the proof, and omit the details. Observe first that by the definition of μ_{σ_0, τ_0}, the expected *average*[15] payoff induced by (σ_0, τ_0) over a single block is equal to $U(\mu_{\sigma_0, \tau_0}, y)$. Since the profile (σ_0, τ_0) is periodic, for fixed N, as δ approaches one, the discounted payoff induced by (σ_0, τ_0) converges to $U(\mu_{\sigma_0, \tau_0}, y)$ and is thus arbitrarily close to the target payoff $U(\mu_0, y)$.

The construction in Section 5.1.2 implies that it suffices to check the sequential rationality of τ_0 at any information set that is consistent with τ_0. If the receiver fails to play the action prescribed by τ_0, the sender will switch to a babbling play, announcing $\bar{a}$ forever, and the receiver's continuation payoff therefore does not exceed $(1 - \delta) \sum_{k=n}^{\infty} \delta^{k-n} \max_{b \in B} u^r(p_k, b)$, where p_k is the belief that the receiver will hold at stage $k \geq n$ on the current state $\mathbf{s}_k$. Since the sequence of states forms an irreducible and aperiodic chain, p_k converges to m. As δ approaches one, the receiver's continuation payoff therefore converges to v^r.

Assume instead that the receiver sticks to τ_0. His continuation payoff is then equal to the sum of his payoffs until the end of the current block and of the continuation payoff from the next block on, which is the discounted payoff induced by (σ_0, τ_0). For fixed N, as δ approaches one, this continuation payoff converges to $U^r(\mu_{\sigma_0, \tau_0}, y)$. Since $U^r(\mu_{\sigma_0, \tau_0}, y) > v^r$, this proves the sequential rationality of τ_0, provided first N, and then δ, are chosen large enough.

Proof of Proposition 11. Define by σ_{truth} the strategy of the sender that announces truthfully the current state $\mathbf{s}_n$ in each stage $n \leq N$, and by τ_{truth} the strategy of the receiver that plays $y(\mathbf{a}_n)$ in each stage $n \leq N$. Thus, τ_{truth} coincides with τ_0 until stage q.

---

[15]That is, when payoffs in the different stages are not discounted.

Let η be given, and set $\xi = \frac{\eta}{|S|+2}$. For every state $s \in S$ and every $n \in \mathbf{N}$, denote by $F_n(s)$ the empirical frequency of visits to s up to (and including) stage n. Since the Markov chain is aperiodic, by the ergodic theorem there is N_0 such that with probability at least $1 - \xi$, $F_{(1-\xi)N}(s) \leq m_N(S)$ for every state $s \in S$, provided $N \geq N_0$. It follows that τ_0 coincides with τ_{truth} in the first $(1 - \xi)N$ stages, so that with probability at least $1 - \xi$,

$$\|\mu_{\sigma_{truth},\tau_0} - \mu_0\|_1 \leq |S|\xi.$$

Because payoffs are in the interval $[0, 1]$, this implies that

$$U^s(\mu_{\sigma_{truth},\tau_0}, y) > U^s(\mu_0, y) - (|S| + 1)\xi.$$

For fixed N, as δ converges to 1, the discounted payoff in each block converges to the average payoff in that block, and therefore for δ sufficiently large

$$\gamma_\delta^s(\sigma_{truth}, \tau_0) > U^s(\mu_0, y) - (|S| + 2)\xi.$$

Because σ_0 is a best reply to τ_0, we deduce that

$$\gamma_\delta^s(\sigma_0, \tau_0) \geq \gamma_\delta^s(\sigma_{truth}, \tau_0) > U^s(\mu_0, y) - (|S| + 2)\xi = U^s(\mu_0, y) - \eta.$$

We again use the fact that, for fixed N, as δ goes to 1, the payoff $\gamma_\delta^s(\sigma_0, \tau_0)$ converges to $U(\mu_{\sigma_0,\tau_0}, y)$ to deduce that

$$U^s(\mu_{\sigma_0,\tau_0}, y) > U^s(\mu_0, y) - \eta. \tag{3}$$

For fixed N, and for every δ, the marginal distributions of $\mu_{\sigma_0,\tau_0} \in \Delta(S \times T)$ on S and T are respectively equal to m and to m_N.

Since the approximation m^N converges to m as $N \to +\infty$, the distribution μ_{σ_0,τ_0} converges to the set $\mathcal{M}$ of copulas. Using Lemma 10, Proposition 11 therefore follows from (3). ∎

/// **Eilon: above there seems to be a small inaccuracy: shouldn't the coefficient** $\frac{\varepsilon c_1}{c_2}$ **appear somewhere?** ///

5.2   Proof of Theorem 2

We here discuss the logic behind the proof of Theorem 2. We let $\delta < 1$, and we fix a Nash equilibrium (σ, τ) of the δ-discounted game. For $s \in S$, we define $y(\cdot \mid s) \in \Delta(B)$ as the

expected *discounted* distribution of moves of the receiver in state s. Formally, for $s \in S$, and $b \in B$, we set

$$y(b \mid s) = \frac{1}{m(s)} \mathbf{E}_{\sigma,\tau} \left[\sum_{n=1}^{\infty} (1-\delta)\delta^{n-1} 1_{\{\mathbf{s}_n=s, \mathbf{b}_n=b\}} \right] = \frac{\mathbf{E}_{\sigma,\tau} \left[\sum_{n=1}^{\infty} (1-\delta)\delta^{n-1} 1_{\{\mathbf{s}_n=s, \mathbf{b}_n=b\}} \right]}{\mathbf{E}_{\sigma,\tau} \left[\sum_{n=1}^{\infty} (1-\delta)\delta^{n-1} 1_{\{\mathbf{s}_n=s\}} \right]}.$$

By construction, one has $\gamma_\delta(\sigma,\tau) = U(\mu_0, y)$. Indeed,

$$\begin{aligned}
\gamma_\delta(\sigma,\tau) &= (1-\delta) \sum_{n=1}^{\infty} \delta^{n-1} \mathbf{E}_{\sigma,\tau}[u(\mathbf{s}_n, \mathbf{b}_n)] = (1-\delta) \sum_{n=1}^{\infty} \delta^{n-1} \sum_{s \in S, b \in B} \mathbf{E}_{\sigma,\tau}[1_{\{\mathbf{s}_n=s, \mathbf{b}_n=b\}}] u(s,b) \\
&= \sum_{s \in S, b \in B} u(s,b) m(s) y(b \mid s) = U(\mu_0, y),
\end{aligned}$$

as desired.

Since the distribution of $\mathbf{s}_n$ is equal to m for each stage $n \in \mathbf{N}$, one has $\gamma_\delta^r(\sigma,\tau) \geq v^r$, and thus, $U^r(\mu_0, y) \geq v^r$, so that **C2** holds.

We will prove that for every $\mu \in \mathcal{M}$, there is a strategy σ' of the sender, such that $\gamma(\sigma',\tau) = U(\mu, y)$, so that the inequality $U^s(\mu_0, y) \geq U^s(\mu, y)$ will follow from the equilibrium property, and **C1** holds as well.

As previous discussions have hinted, the construction of σ' relies on the sender simulating a sequence $(\mathbf{t}_n)_n$ of fictitious states, that he will substitute to the true states $(\mathbf{s}_n)_n$ when playing σ. The construction of this sequence requires much technical work.

Define $\mathcal{M}' \subseteq \Delta(S \times T)$ to be the set of distributions $\mu \in \mathcal{M}$ such that the following property **P** holds:

Property P. For every $(s,t) \in S \times T$, one has

$$\sum_{s' \in S} \mu(s' \mid t) p(s \mid s') = \sum_{t' \in T} \mu(s \mid t') p(t' \mid t). \tag{4}$$

We prove in the Appendix the following two lemmas.

Lemma 12 *Under **Assumption A**, the set $\mathcal{M}'$ coincides with the set $\mathcal{M}$.*

Lemma 13 *Let $\mu \in \mathcal{M}'$ be given. There exists an S-valued process[16] $(\mathbf{t}_n)_n$, such that:*

---

[16]The process $(\mathbf{t}_n)_n$ is possibly defined on a probability space which is an enlargement of the one on which $(\mathbf{s}_n)_n$ is defined.

P1 *The law of the sequence $(\mathbf{t}_n)_n$ is the same as the law of the sequence $(\mathbf{s}_n)_n$.*

P2 *The law of the pair $(\mathbf{s}_n, \mathbf{t}_n)$ is μ, for each stage $n \in \mathbf{N}$.*

P3 *The conditional law of $\mathbf{s}_n$, given $\mathbf{t}_1, \ldots, \mathbf{t}_n$ is $\mu(\cdot \mid \mathbf{t}_n)$.*

P4 *Conditional on $\mathbf{s}_n$, the vector $(\mathbf{t}_1, \ldots, \mathbf{t}_n)$ is independent of the future states $(\mathbf{s}_{n+1}, \mathbf{s}_{n+2}, \ldots)$.*

We emphasize that only Lemma 12 makes use of **Assumption A**. This has the following consequence. Given $\mu \in \mathcal{M}'$, using Lemma 13 and the construction that follows, one has $U^s(\mu_0, y) \geq U^s(\mu, y)$. Thus, the conclusion $U^s(\mu_0, y) = \max_{\mu \in \mathcal{M}'} U^s(\mu, y)$ holds, irrespective of whether **Assumption A** is met or not.

We will interpret the states $(\mathbf{t}_n)_n$ as *fictitious* states, which are constructed by the sender using the sequence $(\mathbf{s}_n)_n$ of true states. Condition **P1** implies that the sequence of fictitious states is statistically indistinguishable from $(\mathbf{s}_n)_n$. Condition **P4** ensures that fictitious states can be computed, using only the information available at stage n.

Intuitively, given the true states $(\mathbf{s}_n)_n$, the strategy σ' generates the fictitious states $(\mathbf{t}_n)_n$, and plays according to σ, as if the realized states were given by $\mathbf{t}_n$. Conditions **P2** and **P3** in Lemma 13 will ensure that the correlation between the true and the fictitious states is such that $\gamma_\delta(\sigma', \tau) = U(\mu, y)$.

The strategy σ' takes as inputs the true states to generate fictitious states, and then mimics σ. Formally, we let $(\mathbf{t}_n)_n$ be a process such that **P1-4** hold. Following any history $(s_1, t_1, a_1, b_1, \ldots, s_{n-1}, t_{n-1}, a_{n-1}, b_{n-1}, s_n, t_n)$, the strategy σ' plays the mixed move $\sigma(t_1, a_1, b_1, \ldots, t_n) \in \Delta(A)$.

We now proceed to show that the expected payoff induced by (σ', τ) is equal to $U(\mu, y)$, as claimed.

Below we will denote by s_k, t_k, b_k generic values of the random variables $\mathbf{s}_k, \mathbf{t}_k$ and $\mathbf{b}_k$,

respectively. For any given stage $n \in \mathbf{N}$, the following sequence of equalities holds:

$$
\begin{aligned}
\mathbf{E}_{\sigma',\tau}[u(\mathbf{s}_n, \mathbf{b}_n)] &= \sum_{s_n, b_n} \mathbf{P}_{\sigma',\tau}(\mathbf{s}_n = s_n, \mathbf{b}_n = b_n) u(s_n, b_n) \\
&= \sum_{s_n} \sum_{t_1,\ldots,t_n} \sum_{b_1,\ldots,b_n} \mathbf{P}_{\sigma',\tau}(s_n, t_1,\ldots,t_n, b_1,\ldots,b_n) u(s_n, b_n) \quad (5) \\
&= \sum_{s_n} \sum_{t_1,\ldots,t_n} \sum_{b_1,\ldots,b_n} \mathbf{P}_{\sigma',\tau}(s_n \mid t_1,\ldots,t_n, b_1,\ldots,b_n) \mathbf{P}_{\sigma',\tau}(t_1,\ldots,t_n, b_1,\ldots,b_n) u(s_n, b_n) \\
&= \sum_{s_n} \sum_{t_1,\ldots,t_n} \sum_{b_1,\ldots,b_n} \mathbf{P}(s_n \mid t_1,\ldots,t_n) \mathbf{P}_{\sigma',\tau}(t_1,\ldots,t_n, b_1,\ldots,b_n) u(s_n, b_n) \quad (6) \\
&= \sum_{s_n} \sum_{t_1,\ldots,t_n} \sum_{b_1,\ldots,b_n} \mu(s_n \mid t_n) \mathbf{P}_{\sigma',\tau}(t_1,\ldots,t_n, b_1,\ldots,b_n) u(s_n, b_n) \quad (7) \\
&= \sum_{s_n, t_n, b_n} \mu(s_n \mid t_n) \mathbf{P}_{\sigma',\tau}(t_n, b_n) u(s_n, b_n), \quad (8) \\
& \quad (9)
\end{aligned}
$$

where (6) holds because (b_k) are independent of s_n, and (7) holds by **P3**.

Using **P1**, and by the definition of σ', the δ-discounted sum of $\mathbf{P}_{\sigma',\tau}(\mathbf{t}_n = t_n, \mathbf{b} = b_n)$ is equal to $\mathbf{P}_{\sigma,\tau}(\mathbf{s}_n = t_n, \mathbf{b} = b_n)$, which is equal to $\mu(t_n) \times y(b_n \mid t_n)$. By (9) we now obtain

$$
\gamma_\delta(\sigma', \tau) = \sum_{s,t,b} \mu(s \mid t) \mu(t) y(b \mid t) u(s, b) = U(\mu, y).
$$

5.3 Proof of Proposition 8

Let (σ, τ) be a Nash equilibrium of the blind game. Define τ' to be the following strategy that depends only on the sender's announcements, and not on the receiver's past actions: after a sequence $(a_1, \ldots, a_n)$ of announcements, τ' plays any action $b \in B$ with the probability that the n-th action of the receiver according to τ is b, conditional on the sender's announcements being $(a_1, \ldots, a_n)$:

$$
\tau'(a_1, \ldots, a_n)[b] = \mathbf{E}\left[\tau(a_1, b_1, a_2, b_2, \ldots, b_{n-1}, a_n)[b] \mid a_1, a_2, \ldots, a_n\right].
$$

In words, τ' gets rid of the possible correlation between successive actions of the receiver, that may exist in the strategy τ.

We claim that the strategy profile (σ, τ') is a Nash equilibrium of the blind game. Indeed, τ' is a best-reply to σ because it induces the same payoff as τ. σ is a best-reply to τ' because any strategy of the sender in the *blind* game induces the same expected payoff against τ or τ'.

We next claim that the strategy profile (σ, τ') is a Nash equilibrium of the non-blind game. Indeed, because under σ, the sender does not condition his play on past actions of the receiver, and because τ' is a best response to σ in the blind game, it follows that τ' is a best response to σ in the non-blind game as well. Because the receiver's actions are conditionally independent, given the sender's announcements, any profitable deviation against τ' in the non-blind game is also profitable in the blind game.

References

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

[2] Aumann R.J. and Hart S. (2003) Long Cheap Talk. *Econometrica*, **71**, 1619-1660.

[3] Bochnak J., Coste M. and Roy M.F. (1998) Real Algebraic Geometry. Springer.

[4] Crawford V.P. and Sobel J. (1982) Strategic Information Exchange. *Econometrica*, **50**, 1431-1451.

[5] Farrell J. and Rabin M. (1996) Cheap talk. *Journal of Economic Perspectives*, **10**, 103-118.

[6] Forges F. and Koessler F. (2008) Long Persuasion Games. *Journal of Economic Theory*, **143**, 1-35.

[7] Golosov M., Skreta V., Tsyvinski A. and Wilson A. (2009) Dynamic Strategic Information Transmission. Preprint.

[8] Green J.R. and Stokey N.L. (2007) A Two-Person Game of Information Transmissionstar. *Journal of Economic Theory*, **135**, 90-104.

[9] H orner J., Rosenberg D., Solan E. and Vieille N. (2010) On a Markov Game with One-Sided Incomplete Information. *Operations Research*, forthcoming.

[10] Krishna V. and Morgan J. (2001) A Model of Expertise. *Quarterly Journal of Economics*, **116**, 747-775.

[11] Phelan C. (2006) Public Trust and Goverment Betrayal. *Journal of Economic Theory*, **130**, 27-43.

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

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

Appendix

6   Proof of Lemma 3

To prove Lemma 3 we need the following description of $\mathcal{M}$, which is of independent interest.

A *permutation matrix* is a (square) matrix with entries in $\{0, 1\}$, such that each row and each column contains exactly one entry equal to 1. We denote by Φ the set of $S \times S$ permutation matrices, and by I the matrix that corresponds to the identity permutation.

Lemma 14 *The set $\mathcal{M}(m)$ is equal to*

$$\mathcal{M}(m) = (\mu_0 - I + co\ \Phi) \cap \mathbf{R}_+^{S \times S}.$$

Proof. The inclusion $\supseteq$ is clear. We prove the reverse inclusion. Take μ in $\mathcal{M}(m)$, and define the matrix $J := \mu + I - \mu_0$ in $\mathbf{R}^{S \times S}$. J is a bistochastic matrix, hence it is a convex combination of permutation matrices. Since $\mu = J - I + \mu_0$, the result follows. ∎

Proof of Lemma 3.   We only prove that **C1** is equivalent to **Q2**. For every permutation ϕ over S denote by $\mu^\phi \in S \times S$ the matrix where the entry (s, t) is equal to 1 if $t = \phi(s)$, and is 0 otherwise. Note that $U^s(I, y) = \sum_{s \in S} u^s(s, y(\cdot \mid s))$, and $U^s(\mu^\phi, y) = \sum_{s \in S} u^s(s, y(\cdot \mid \phi(s)))$.

Assume first that **Q2** holds, and let $\mu \in \mathcal{M}(m)$. By Lemma 14, μ can be written $\mu = \mu_0 - I + \sum_\phi \alpha_\phi P^\phi$, where the α_ϕ are non negative reals and sum to one. Because U^s is linear in μ,

$$U^s(\mu, y) = U^s(\mu_0, y) - U^s(I, y) + \sum_\phi \alpha_\phi U^s(\mu^\phi, y).$$

By **Q2**, $U^s(I, y) \geq U^s(\mu^\phi, y)$ for every permutation ϕ, and therefore $U^s(I, y) \geq \sum_\phi \alpha_\phi U^s(\mu^\phi, y)$. It follows that $U^s(\mu_0, y) \geq U^s(\mu, y)$. Because this inequality holds for every $\mu \in \mathcal{M}(\mu)$, **C1** holds.

Assume now that **C1** holds. Fix a permutation ϕ, and define $\mu_\varepsilon = \mu_0 - \varepsilon I + \varepsilon \mu^\phi$, where $\varepsilon > 0$. Because m has full support, one has $\mu_\varepsilon \in \mathcal{M}(m)$ provided ε is sufficiently small. Now, by **C1**, for each such ε,

$$U^s(\mu_0, y) \geq U^s(\mu_\varepsilon, y) = U^s(\mu_0, y) - \varepsilon U^s(I, y) + \varepsilon U^s(\mu^\phi, y).$$

It follows that $U^s(\mu^\phi, y) \leq \varepsilon U^s(I, y)$. As this inequality holds for every permutation ϕ, **Q2** holds. ∎

7   Proof of Lemma 10

Fix $\varepsilon > 0$, a stationary strategy $y \in Y_\varepsilon(\mathcal{M})$, and a copula $\mu \in \mathcal{M}$. Since $y \in Y_\varepsilon(\mathcal{M})$, one has $y = \varepsilon y_0 + (1 - \varepsilon)y_1$ for some $y_1 \in Y(\mathcal{M})$. Present μ as a convex combination of the extreme points $(\mu_e)_e$ of $\mathcal{M}$: $\mu = \sum_{\mu_e \in \mathcal{M}_e} \alpha_e \mu_e$, with $\alpha_e \geq 0$ and $\sum_{\mu_e \in \mathcal{M}_e} \alpha_e = 1$. Recall that μ_0 is one of the extreme points of $\mathcal{M}$.

On the one hand, since U^s is bi-linear, one has

$$
\begin{aligned}
U^s(\mu_0, y) - U^s(\mu, y) &= \varepsilon U^s(\mu_0, y_0) + (1 - \varepsilon)U^s(\mu_0, y_1) - \varepsilon U^s(\mu, y_0) - (1 - \varepsilon)U^s(\mu, y_1) \\
&\geq \varepsilon \left(U^s(\mu_0, y_0) - U^s(\mu, y_0) \right) & (10) \\
&= \varepsilon \left((1 - \alpha_0)U^s(\mu_0, y_0) - \sum_{\mu_e \neq \mu_0} \alpha_e U^s(\mu_e, y_0) \right) & (11) \\
&\geq \varepsilon(1 - \alpha_0)c_1, & (12)
\end{aligned}
$$

where the inequality (10) holds because $y_1 \in Y(\mathcal{M})$ and by **C1**.

On the other hand, one has $\mu - \mu_0 = \sum_{\mu_e \in \mathcal{M}_e} \alpha_e(\mu_e - \mu_0)$, hence

$$\|\mu - \mu_0\|_1 \leq c_2 \sum_{\mu_e \in \mathcal{M}_e, \mu_e \neq \mu_0} \alpha_e = c_2(1 - \alpha_0). \tag{13}$$

The result follows from (12) and (13).

8   Proof of Lemma 13

Let $\mu \in \mathcal{M}'$ be given, and define $\bar{\mu} \in \Delta(T \times S \times T)$ by

$$\bar{\mu}(t', s, t) = \mu(s, t)p(t \mid t')\frac{m(t')}{m(t)}, \quad (t's, t) \in T \times S \times T. \tag{14}$$

For every two indices $i, j \in \{1, 2, 3\}$ with $i < j$, denote by $\bar{\mu}_{i,j}$ the marginal of $\bar{\mu}$ on the i-th and j-th coordinates.

We will use the following properties of $\bar{\mu}$.

Lemma 15 *One has*

1. $\bar{\mu}_{2,3} = \mu$;

2. $\bar{\mu}_{1,3}(t', t) = m(t')p(t \mid t')$ *for every* $t, t' \in T$;

3. $\bar{\mu}_{1,2}(t', s') = \sum_{s \in S} \bar{\mu}_{2,3}(s, t')p(s' \mid s)$, *for each* $t' \in T, s' \in S$;

4. $\bar{\mu}(s \mid t', t) = \mu(s \mid t)$ *for each* $(t', s, t) \in T \times S \times T$.

Proof. We prove the four claims in turn. Let $s, t \in S \times T$ be given. One has

$$
\begin{aligned}
\bar{\mu}_{2,3}(s, t) &= \sum_{t' \in T} \bar{\mu}(t', s, t) = \sum_{t' \in T} \mu(s, t)p(t \mid t')\frac{m(t')}{m(t)} \\
&= \frac{\mu(s, t)}{m(t)} \sum_{t' \in T} p(t \mid t')m(t') = \mu(s, t),
\end{aligned}
$$

which proves the first claim.

To prove the second claim, let $t', t \in T$ be given. One has

$$
\bar{\mu}_{1,3}(t', t) = \sum_{s \in S} \bar{\mu}(t', s, t) = \sum_{s \in S} \mu(s, t)p(t \mid t')\frac{m(t')}{m(t)} = p(t \mid t')m(t'),
$$

where the last equality holds since the marginal distribution of μ on S is m.

We turn to the third claim. Let $t' \in T$, $s' \in S$ be given. By the first claim, and since $\mu \in \mathcal{M}'$, one has

$$
\sum_{s \in S} \bar{\mu}_{2,3}(s, t')p(s' \mid s) = \sum_{s \in S} \mu(s, t')p(s' \mid s) = m(t') \sum_{t \in T} \mu(s' \mid t)p(t \mid t'). \tag{15}
$$

On the other hand,

$$
\bar{\mu}_{1,2}(t', s') = \sum_{t \in T} \bar{\mu}(t', s', t) = \sum_{t \in T} \mu(s', t)p(t \mid t')\frac{m(t')}{m(t)}. \tag{16}
$$

The third claim follows from (15) and (16).

Finally, let $(t', s, t) \in T \times S \times T$ be given. By the second claim,

$$
\bar{\mu}(s \mid t', t) = \frac{\bar{\mu}(t', s, t)}{\bar{\mu}_{1,3}(t', t)} = \frac{\mu(s, t)p(t \mid t')}{p(t \mid t')m(t')} \times \frac{m(t')}{m(t)} = \mu(s \mid t),
$$

and the fourth claim follows. ∎

We construct the sequence $(\mathbf{t}_n)_n$ as follows. The initial values $\mathbf{t}_0$ and $\mathbf{t}_1$ are drawn according to the conditional distribution $\bar{\mu}(\cdot|\mathbf{s}_1) \in \Delta(T \times T)$. For $n \neq 2$, $\mathbf{t}_n$ is drawn according to the conditional distribution $\bar{\mu}(\cdot \mid \mathbf{t}_{n-1}, \mathbf{s}_n)$. In this construction, $\mathbf{t}_0$ is used to unify the treatment of $\mathbf{s}_1$ with that of $(\mathbf{s}_n)_{n\geq 2}$. Property **P4** thus holds by construction. Properties **P1** and **P2** follow from the next lemma.

Lemma 16 *The law of $(\mathbf{t}_{n-1}, \mathbf{s}_n, \mathbf{t}_n)$ is equal to $\bar{\mu}$, for each stage $n \geq 1$.*

Proof. We argue by induction. Observe that the law of $\mathbf{s}_1$ is equal to m. Therefore,

$$\mathbf{P}((\mathbf{t}_0, \mathbf{s}_1, \mathbf{t}_1) = (t', s, t)) = m(s)\bar{\mu}(t', t \mid s) = \bar{\mu}(t', s, t).$$

Assume that the claim holds for some $n \in \mathbf{N}$. We will prove that the law of $(\mathbf{t}_n, \mathbf{s}_{n+1})$ is then equal to $\bar{\mu}_{1,2}$. This follows from the following sequence of equalities, which holds for every $t' \in T, s \in S$:

$$
\begin{aligned}
\mathbf{P}((\mathbf{t}_n, \mathbf{s}_{n+1}) = (t', s)) &= \sum_{s' \in S} \mathbf{P}((\mathbf{s}_n, \mathbf{t}_n, \mathbf{s}_{n+1}) = (s', t', s)) \\
&= \sum_{s' \in S} \mathbf{P}((\mathbf{s}_n, \mathbf{t}_n) = (s', t')) \times \mathbf{P}(\mathbf{s}_{n+1} = s|(\mathbf{s}_n, \mathbf{t}_n) = (s', t)) \\
&= \sum_{s' \in S} \bar{\mu}_{2,3}(s', t')p(s \mid s') = \bar{\mu}_{1,2}(t', s),
\end{aligned}
$$

where the last equality follows from Lemma 15(3) and **P4**. Since the conditional law of $\mathbf{t}_{n+1}$ given $(\mathbf{t}_n, \mathbf{s}_{n+1})$ is equal to $\bar{\mu}(\cdot \mid \mathbf{t}_n, \mathbf{s}_{n+1})$, this yields the claim for $n + 1$. ∎

Finally, property **P3** follows from the second part of the next lemma. The first part of the lemma is needed to the proof of the second part.

Lemma 17 *(1) The conditional law of $\mathbf{t}_n$ given $(\mathbf{t}_0, \ldots, \mathbf{t}_{n-1})$ coincides with the conditional law of $\mathbf{t}_n$ given $\mathbf{t}_{n-1}$.*
(2) The conditional law of $\mathbf{s}_n$ given $(\mathbf{t}_0, \ldots, \mathbf{t}_{n-1}, \mathbf{t}_n)$ coincides with the conditional law of $\mathbf{s}_n$ given $\mathbf{t}_n$.

Proof. The proof is by induction. For $n = 1$, the first statement trivially holds, while the second statement holds by Lemma 15(1). Assume that the claim holds for some $n \in \mathbf{N}$. For

brevity, we denote by $t_n, s_n, \cdots$ generic values of $\mathbf{t}_n, \mathbf{s}_n \cdots$, and we write $\mathbf{P}(t_n, s_n)$ instead of $\mathbf{P}((\mathbf{t}_n, \mathbf{s}_n) = (t_n, s_n))$.

Observe first that by the definition of $(\mathbf{t}_n)$,

$$
\begin{aligned}
\mathbf{P}(t_{n+1} \mid t_0, \ldots, t_n) &= \sum_{s_{n+1} \in S} \mathbf{P}(s_{n+1} \mid t_0, \ldots, t_n) \mathbf{P}(t_{n+1} \mid t_0, \ldots, t_n, s_{n+1}) \\
&= \sum_{s_{n+1} \in S} \mathbf{P}(s_{n+1} \mid t_0, \ldots, t_n) \times \bar{\mu}(t_{n+1} \mid t_n, s_{n+1}).
\end{aligned}
\tag{17}
$$

Moreover,

$$
\begin{aligned}
\mathbf{P}(s_{n+1} \mid t_0, \ldots, t_n) &= \sum_{s_n \in S} \mathbf{P}(s_n \mid t_0, \ldots, t_n) \mathbf{P}(s_{n+1} \mid s_n, t_0, \ldots, t_n) \\
&= \sum_{s_n \in S} \mathbf{P}(s_n \mid t_0, \ldots, t_n) \times p(s_{n+1} \mid s_n) \\
&= \sum_{s_n \in S} \mathbf{P}(s_n \mid t_n) \times p(s_{n+1} \mid s_n),
\end{aligned}
\tag{18}
$$

where the last equality holds by the induction hypothesis. Note that the right-hand side of (18) is independent of $(t_0, t_1, \ldots, t_{n-1})$, and therefore

$$
\mathbf{P}(s_{n+1} \mid t_0, \ldots, t_n) = \mathbf{P}(s_{n+1} \mid t_n).
\tag{19}
$$

Plugging (18) in (17), one obtains

$$
\mathbf{P}(t_{n+1} \mid t_0, \ldots, t_n) = \sum_{s_{n+1} \in S} \sum_{s_n \in S} \mathbf{P}(s_n \mid t_n) \times p(s_{n+1} \mid s_n) \times \bar{\mu}(t_{n+1} \mid t_n, s_{n+1}).
$$

The right hand side is independent of $t_1, \ldots, t_{n-1}$, hence it is equal to $\mathbf{P}(t_{n+1} \mid t_n)$, and the first part of the lemma follows.

We turn to the second statement. One has

$$
\begin{aligned}
\mathbf{P}(s_{n+1} \mid t_0, \ldots, t_{n+1}) &= \frac{\mathbf{P}(s_{n+1}, t_{n+1} \mid t_0, \ldots, t_n)}{\mathbf{P}(t_{n+1} \mid t_0, \ldots, t_n)} = \frac{\mathbf{P}(s_{n+1} \mid t_0, \ldots, t_n) \times \mathbf{P}(t_{n+1} \mid s_{n+1}, t_0, \ldots, t_n)}{\mathbf{P}(t_{n+1} \mid t_0, \ldots, t_n)} \\
&= \frac{\mathbf{P}(s_{n+1} \mid t_n) \bar{\mu}(t_{n+1} \mid s_{n+1}, t_n)}{\mathbf{P}(t_{n+1} \mid t_n} = \mathbf{P}(s_{n+1} \mid t_n) \frac{\mu(s_{n+1}, t_{n+1})}{\bar{\mu}(t_n, s_{n+1})} \frac{m(t_n)}{m(t_{n+1})} \\
&= \frac{\mu(s_{n+1}, t_{n+1})}{m(t_{n+1})} = \mathbf{P}(s_{n+1} \mid t_{n+1}),
\end{aligned}
$$

where the third equality holds by (19), the construction of $(\mathbf{t}_n)_n$ and the first claim, and the fourth equality holds by (14). This concludes the proof of the induction step. ∎

The proof of Lemma 13 is now completed.

9 Proof of Lemma 12

We here verify that if Assumption **A** holds then $\mathcal{M} = \mathcal{M}'$. Let p be a transition function such that $p(s' \mid s) = \alpha_{s'}$ for every two states $s \neq s'$, and $p(s \mid s) = 1 - \sum_{s' \neq s} \alpha_{s'}$. Set $C = \sum_{s \in S} \alpha_s$. One can verify that the invariant measure of p is given by $m(s) = \frac{\alpha_s}{C}$ for each $s \in S$.

Let $\mu \in \mathcal{M}$. We will prove that for every $(t, s') \in T \times S$, the equality

$$\sum_{s \in S} \mu(s \mid t) p(s' \mid s) = \sum_{t' \in T} \mu(s' \mid t') p(t' \mid t) \tag{20}$$

holds. Fix $t \in T$ and $s' \in S$. Observe that

$$
\begin{aligned}
\sum_{s \in S} \mu(s \mid t) p(s' \mid s) &= \mu(s' \mid t) \left(1 - \sum_{s \neq s'} \alpha_s \right) + \sum_{s \neq s'} \alpha_{s'} \mu(s \mid t) \\
&= \mu(s' \mid t) \left(1 - \sum_{s \neq s'} \alpha_s \right) + \alpha_{s'}(1 - \mu(s' \mid t)) \\
&= \alpha_{s'} + \mu(s' \mid t)(1 - C). \tag{21}
\end{aligned}
$$

On the other hand, one has

$$
\begin{aligned}
\sum_{t' \in T} \mu(s' \mid t') p(t' \mid t) &= \mu(s' \mid t) \left(1 - \sum_{t' \neq t} \alpha_{t'} \right) + \sum_{t' \neq t} \mu(s' \mid t') \alpha_{t'} \tag{22} \\
&= \mu(s' \mid t)(1 - C + \alpha_t) + \sum_{t' \neq t} \mu(s' \mid t') \alpha_{t'}. \tag{23}
\end{aligned}
$$

When subtracting (22) from (21) one obtains

$$
\begin{aligned}
\sum_{s \in S} \mu(s \mid t) p(s' \mid s) - \sum_{t' \in T} \mu(s' \mid t') p(t' \mid t) &= \alpha_{s'} - \mu(s' \mid t) \alpha_t - \sum_{t' \neq t} \mu(s' \mid t') \alpha_{t'} \\
&= \alpha_{s'} - \sum_{t' \in S} \mu(s' \mid t') \alpha_{t'} \\
&= \alpha_{s'} - C \sum_{t' \in S} \mu(s' \mid t') m(t') \tag{24} \\
&= \alpha_{s'} - C m(s') = 0,
\end{aligned}
$$

where (24) and (25) hold since $\alpha(s) = C m(s)$ for every $s \in S$. This proves (20), as desired.

9.1 Proof of Theorem 5

The proof of Theorem 5 consists of two independent parts. We first prove that, if condition **B** does not hold for some game G, then it does not hold throughout some neighborhood of G.

Proposition 18 *Let G be a game that does not satisfy condition **B**. Then there is a neighborhood $\mathcal{N}$ of G such that no game in $\mathcal{N}$ satisfies condition **B**.*

   Proof. The proof relies on the theory of semi-algebraic sets. We refer to Bochnak, Coste and Roy (1998) for the results used below. Recall that the set of extreme points of the polytope $\mathcal{M}$ is denoted by $\mathcal{M}_e$.

   We will use the following two properties, that hold for constant functions $y : S \to \Delta(B)$.

R1. If $y : S \to \Delta(B)$ is constant, then $U^s(\mu_0, y) = U^s(\mu, y)$ for every $\mu \in \mathcal{M}$.

R2. If $y : S \to \Delta(B)$ is constant, then $U^r(\mu_0, y) \leq v^r$.

Property **R1** holds because when y is independent of S, the payoff is independent of the sender's announcements. Property **R2** holds because v^r is the maximum of $U^r(\mu_0, y)$ over all constant functions y.

   Given a payoff function $\tilde{u} : S \times B \to \mathbf{R}^2$, we denote by $\mathcal{S}(\tilde{u})$ the system of inequalities

$$\tilde{U}^r(\mu_0, y) \geq v^r_{\tilde{u}} \text{ and } \tilde{U}^s(\mu_0, y) \geq \tilde{U}^s(\mu, y), \text{for all } \mu \in \mathcal{M}_e,$$

with unknowns $y : S \to \Delta(B)$, where $v^r_{\tilde{u}} = \max_{b \in B} \tilde{U}^r(\mu_0, b)$ is the min-max value of the receiver in the game with payoffs $\tilde{u}$.

   We say that a vector $y \in \mathbf{R}^{S \times B}$ is *constant* if $y(s, b)$ only depends on b.

   Let u denote the payoff function of G. By assumption, any solution y to $\mathcal{S}(u)$ is constant. We will show that this implies that all solutions to $\mathcal{S}(\tilde{u})$ are constant, for all $\tilde{u}$ in a neighborhood of u.

   Assume to the contrary that for every $\varepsilon > 0$ there is a payoff function $u_\varepsilon \in \mathbf{R}^{2(S \times B)}$ such that (i) $\|u - u_\varepsilon\| < \varepsilon$, and (ii) the system $\mathcal{S}(u_\varepsilon)$ has a non-constant solution $y_\varepsilon \in \mathbf{R}^{S \times B}$.

   This implies that there is a *semi-algebraic* map $\varepsilon \in (0, 1) \mapsto (u_\varepsilon, y_\varepsilon)$ such that (i) $\lim_{\varepsilon \to 0} u_\varepsilon = u$, and (ii) y_ε is a *non-constant* solution to $\mathcal{S}(u_\varepsilon)$ for every $\varepsilon > 0$ small enough.

   In particular, the map $\varepsilon \mapsto y_\varepsilon$ has an expansion to a Puiseux series in a neighborhood of zero: there exist $\varepsilon_0 > 0$, a natural number r and vectors $y_k \in \mathbf{R}^{S \times B}$ for $k \geq 0$ such that

$$y_\varepsilon = \sum_{k=0}^{\infty} \varepsilon^{\frac{k}{r}} y_k,$$

for every $\varepsilon \in (0, \varepsilon_0)$, and a similar expansion exists for the map $\varepsilon \mapsto u_\varepsilon$.

Note that $y_0 = \lim_{\varepsilon \to 0} y_\varepsilon$. This implies in particular that $y_0(\cdot, s) \in \Delta(B)$ for every $s \in S$, and that y_0 is a solution to $\mathcal{S}(u)$. In particular, y_0 is constant.

Because $y_\varepsilon(\cdot, s) \in \Delta(B)$, it follows that $\sum_{b \in B} y_\varepsilon(b \mid s) = 1$ for every $\varepsilon > 0$ and every $s \in S$, so that $\sum_{b \in B} y_k(b \mid s) = 0$ for every $k \geq 1$ and every $s \in S$.

Let $l \geq 0$ be the maximal integer such that $y_0, y_1, \ldots, y_l$ are constant functions. Because y_ε is non constant for every $\varepsilon > 0$, we have $l < \infty$. Define a vector $d \in \mathbf{R}^B$ by

$$d(b) = \min_{s \in S} y_{l+1}(b, s), \quad \forall b \in B.$$

Note that

$$y_\varepsilon = \sum_{k=0}^{\infty} \varepsilon^{\frac{k}{r}} y_k \tag{25}$$

$$= \left(\sum_{k=0}^{l} \varepsilon^{\frac{k}{r}} y_k + \varepsilon^{\frac{l+1}{r}} d \right) + \varepsilon^{\frac{l+1}{r}} (y_{l+1} - d) + \sum_{k=l+2}^{\infty} \varepsilon^{\frac{k}{r}} y_k. \tag{26}$$

The first term $\left(\sum_{k=0}^{l} \varepsilon^{\frac{k}{r}} y_k + \varepsilon^{\frac{l+1}{r}} d \right)$ is independent of s, and all its coordinates are non-negative because y_ε is non-negative for every $\varepsilon > 0$. Set

$$z_\varepsilon = \frac{\sum_{k=0}^{l} \varepsilon^{\frac{k}{r}} y_k + \varepsilon^{\frac{l+1}{r}} d}{1 + \varepsilon^{\frac{l+1}{r}} \sum_{b \in B} d(b)} \in \mathbf{R}^{S \times B}.$$

Then $z_\varepsilon(\cdot, s) \in \Delta(B)$ for every $s \in S$, and z_ε is independent of s. Set

$$w(\cdot, s) = \frac{y_{l+1}(\cdot, s) - d}{-\sum_{b \in B} d(b)} \in \mathbf{R}^{S \times B}, \quad \forall s \in S.$$

Then $w(s) \in \Delta(B)$ and w is non-constant. We will show that w solves $\mathcal{S}(u)$, contradicting the assumption that all solutions of $\mathcal{S}(u)$ are constant.

By **R2**, for every $\varepsilon > 0$ we have $\tilde{U}^r(\mu_0, z_\varepsilon) \leq v_{u_\varepsilon}^r$. But $\tilde{U}^r(\mu_0, y_\varepsilon) \geq v_{u_\varepsilon}^r$, and y_ε is a convex combination of z_ε, w, and a "tail" which is of a lower order of ε; by taking the limit $\varepsilon \to 0$ and using $v_{u_\varepsilon}^r \to v^2$ we obtain $U^r(\mu_0, w) \geq v^r$.

Fix $\mu \in \mathcal{M}_e$. By **R1** it follows that $\tilde{U}^s(\mu_0, z_\varepsilon) = \tilde{U}^s(\mu, z_\varepsilon)$. Because $\tilde{U}^s(\mu_0, y_\varepsilon) \geq \tilde{U}^s(\mu, y_\varepsilon)$, it follows for the same reasoning as above that $U^s(\mu_0, w) \geq U^s(\mu, w)$. ∎

We turn to the second part of the proof.

Proposition 19 *Let G be a game such that condition **B** holds. Then any neighborhood of the game G contains a game G' such that $\widehat{E}_{G'}(\mathcal{M}) \neq \emptyset$.*

Proof. The proof combines three independent lemmas. We first show that there are perturbations of u^r such that the inequality in (i) holds strictly for the perturbed game. Next, we show that the map y may be assumed to be one-to-one. Finally, we construct perturbations of u^s such that the inequalities in (ii) will be strict.

Lemma 20 *Let G be a game with payoff function u, and let $y : S \to \Delta(B)$ be a non-constant function such that $U^r(\mu_0, y) \geq v^r$. Then, any neighborhood of u^r contains payoff functions $\tilde{u}^r$ such that $\tilde{U}^r(\mu_0, y) > \tilde{v}^r$.*

Proof. Define $P \in \Delta(S \times B)$ by $P(s, b) := m(s)y(b \mid s)$, for $s \in S, b \in B$, and let $\varepsilon > 0$ be given. We abuse notations and still denote by P the two marginals of P over S and B. Note that $P(s) = m(s) > 0$ for each $s \in S$. Define $\tilde{u}^r : S \times B \to \mathbf{R}$ by $\tilde{u}^r(s, b) = u^r(s, b)$ if $P(b) = 0$, and

$$\tilde{u}^r(s, b) = u^r(s, b) + \varepsilon \frac{P(s, b)}{P(s)P(b)} \text{ if } P(b) > 0.$$

We claim that $\tilde{U}^r(\mu_0, y) > \tilde{v}^r$. Since ε is arbitrary, the result will follow. Note first that, for $b \in B$ such that $P(b) > 0$, one has

$$\tilde{U}^r(\mu_0, b) = U^r(\mu_0, b) + \varepsilon \sum_{s \in S} m(s) \frac{P(s, b)}{m(s)P(b)} = U^r(b, \mu_0) + \varepsilon.$$

Hence, $\tilde{v}^r = v^r + \varepsilon$ (see Eq. (1)). On the other hand, since $y(b \mid s) = \frac{P(s, b)}{m(s)} = \frac{P(s, b)}{P(s)} = P(b \mid s)$,

$$\begin{aligned}
\tilde{U}^r(\mu_0, y) &= U^r(\mu_0, y) + \varepsilon \sum_{s \in S, b \in B} m(s)y(b \mid s) \frac{P(s, b)}{P(s)P(b)} \\
&= U^r(\mu_0, y) + \varepsilon \sum_{s \in S} m(s) \sum_{b \in B} \frac{P(b \mid s)^2}{P(b)}.
\end{aligned}$$

Viewed as a function of the probability distribution $q \in \Delta(B)$, the expression $\sum_{b \in B} \frac{(q(b))^2}{P(b)}$ is strictly convex, and admits a unique minimum equal to 1, when $q = P$. Thus, for fixed state $s \in S$, one has $\sum_{b \in B} \frac{P(b|s)^2}{P(b)} \geq 1$, with a strict inequality whenever the conditional distribution $P(\cdot \mid s)$ differs from P. Since y is non-constant, there exist one state s such that $P(\cdot \mid s) \neq P$. Therefore,

$$\tilde{U}^r(\mu_0, y) > U^r(\mu_0, y) + \varepsilon \geq v^r + \varepsilon = \tilde{v}^r,$$

40

as desired. ∎

Lemma 21 *Let G be a game with payoff function u, and let $y : S \to \Delta(B)$ be such that $U^s(\mu_0, y) \geq U^s(\mu, y)$ for each $\mu \in \mathcal{M}$. Then, any neighborhood of y in $\mathbf{R}^{S \times B}$ contains a one-to-one function $\tilde{y} : S \to \Delta(B)$ such that $U^s(\mu_0, \tilde{y}) \geq U^s(\mu, \tilde{y})$ for each $\mu \in \mathcal{M}$.*

Proof. It suffices to show the existence of a one-to-one map $\tilde{z} : S \to \Delta(B)$ such that $U^s(\mu_0, \tilde{z}) \geq U^s(\mu, \tilde{z})$ for each $\mu \in \mathcal{M}$. Indeed, the conclusion of the lemma then follows by setting $\tilde{y} = (1 - \varepsilon)y + \varepsilon\tilde{z}$, for $\varepsilon > 0$ small enough.

Let $(z_s)_{s \in S}$ be arbitrary distinct elements of $\Delta(B)$. Let $\tilde{\phi}$ be a permutation over S that maximizes the sum $\sum_{s \in S} u^s(s, z_{\psi(s)})$ over all permutations ψ, and set $\tilde{z}_s = z_{\tilde{\phi}(s)}$. By construction, one has

$$\sum_{s \in S} u^s(s, \tilde{z}_s) \geq \sum_{s \in S} u^s(s, \tilde{z}_{\phi(s)}),$$

for every permutation ϕ over S. By Lemma 3, this implies $U^s(\mu_0, \tilde{z}) \geq U^s(\mu, \tilde{z})$ for every $\mu \in \mathcal{M}$, as desired. ∎

Lemma 22 *Let G be a game with payoff function u, and let $y : S \to \Delta(B)$ be a one-to-one map such that $U^s(\mu_0, y) \geq U^s(\mu, y)$ for each $\mu \in \mathcal{M}$. Then, any neighborhood of u^s contains payoff functions $\tilde{u}^s$ such that $\tilde{U}^s(\mu_0, y) > \tilde{U}^s(\mu, y)$ for each $\mu \in \mathcal{M}$, $\mu \neq \mu_0$.*

Note that the existence of a stationary strategy y that satisfies the requirements follows from Lemma 21.

Proof. Let G, u and y be as stated. Given $\varepsilon > 0$, we define $\tilde{u}^s : S \times B \to \mathbf{R}$ by

$$\tilde{u}^s(s, b) = u^s(s, b) + \varepsilon y(b \mid s).$$

We will prove that for every $\varepsilon > 0$, one has $\tilde{U}^s(\mu_0, y) > \tilde{U}^s(\mu, y)$ for each $\mu \in \mathcal{M} \setminus \{\mu_0\}$.

Given a permutation ϕ over S, we denote by $Y_\phi \in \mathbf{R}^{S \times B}$ the vector whose (s, b)-component is equal to $y(b \mid \phi(s))$. Then,

$$\sum_{s \in S} \tilde{u}^s(s, y(\cdot \mid \phi(s))) = \sum_{s \in S, b \in B} y(b \mid \phi(s))\tilde{u}(s, b) \tag{27}$$

$$= \sum_{s \in S, b \in B} y(b \mid \phi(s))u(s, b) + \varepsilon \sum_{s \in S, b \in B} y(b \mid \phi(s))y(b \mid s) \tag{28}$$

$$= \sum_{s \in S} u^s(s, y(\cdot \mid \phi(s)) + \varepsilon\langle Y_{Id}, Y_\phi\rangle, \tag{29}$$

41

where $\langle Y_{Id}, Y_\phi \rangle = \sum_{s \in S, b \in B} y(b \mid \phi(s)) y(b \mid s)$ is the standard scalar product in $\mathbf{R}^{S \times B}$.

Since y is one-to-one, the vectors Y_ϕ and Y_{Id} are not co-linear as soon as $\phi \neq Id$. By Cauchy-Schwarz inequality, it follows that

$$\langle Y_{Id}, Y_\phi \rangle < \|Y_{Id}\|_2 \|Y_\phi\|_2 = \|Y_{Id}\|^2 = \langle Y_{Id}, Y_{Id} \rangle \tag{30}$$

where the first equality holds since the components of Y_ϕ are obtained by permuting the components of Y_{Id}.

On the other hand, observe that by Lemma 3, one has

$$\sum_{s \in S} u^s(s, y(\cdot \mid \phi(s))) \leq \sum_{s \in S} u^s(s, y(\cdot \mid s)). \tag{31}$$

Plugging (31) into (27), one obtains

$$\sum_{s \in S} \tilde{u}^s(s, y(\cdot \mid \phi(s))) < \sum_{s \in S} u^s(s, y(\cdot \mid s)) + \langle Y_{Id}, Y_{Id} \rangle = \sum_{s \in S} \tilde{u}^s(s, y(\cdot \mid s)).$$

By Lemma 3 this yields $\tilde{U}^s(\mu_0, y) > \tilde{U}^s(\mu, y)$, for every $\mu \neq \mu_0$ in $\mathcal{M}$, as desired. ∎

The proof of Proposition 19 follows from Lemmas 20, 21 and 22. ∎