



# Restricted feedback in long term relationships <sup>☆</sup>

Ulrich Doraszelski <sup>a</sup>, Juan F. Escobar <sup>b,\*</sup>

<sup>a</sup> Wharton School, Business and Public Policy Department, University of Pennsylvania, Philadelphia, PA 19104, United States

<sup>b</sup> Center for Applied Economics, Department of Industrial Engineering, University of Chile, Republica 701, Santiago, Chile

Received 7 May 2011; final version received 5 October 2011; accepted 12 October 2011

Available online 9 November 2011

---

## Abstract

This paper studies long term relationships, modeled as repeated games, with restricted feedback. Players condition current play on summary statistics of past play rather than the entire history, as may be the case in online markets. Our state strategy equilibrium framework allows for arbitrary restrictions on strategies. We derive a recursive characterization for the set of equilibrium payoffs similar to that of Abreu, Pearce, and Stacchetti (1986, 1990) [2,3] for perfect public equilibria and show that the set of equilibrium payoffs is the largest fixed point of a monotone operator. We use our characterization to derive necessary and sufficient conditions for efficient trade in a repeated product choice game where costumers condition their purchase decisions only on the last performance signal.

© 2011 Elsevier Inc. All rights reserved.

*JEL classification:* C72; C73

*Keywords:* Repeated games; Restricted feedback; Product choice game

---

## 1. Introduction

Consider a market in which a sequence of short lived costumers faces a long lived seller. The seller is tempted to provide a low quality good, but each transaction generates a signal about

---

<sup>☆</sup> The authors thank Johannes Hörner, Andy Skrzypacz, an Associate Editor, and two anonymous referees for useful comments. Escobar gratefully acknowledges financial support by Basal project Centro de Modelamiento Matemático.

\* Corresponding author.

*E-mail addresses:* [doraszelski@wharton.upenn.edu](mailto:doraszelski@wharton.upenn.edu) (U. Doraszelski), [jescobar@dii.uchile.cl](mailto:jescobar@dii.uchile.cl) (J.F. Escobar).

her performance. If costumers have access to the entire sequence of past performance signals, then the theory of repeated games (Abreu, Pearce, and Stacchetti [2,3]) allows us to characterize the equilibrium set and understand the conditions under which the seller's temptation to provide a low quality good can be moderated. However, assuming that a short lived costumer has access to all past signals seems demanding. For example, each party to an online transaction may acquire some, but not all, information about its counterparty's past behavior. Another example arises when costumers are part of a social network of information transmission and the costumer buying in the previous round can meaningfully convey his trading experience to the costumer buying in the current round, but the experiences of costumers further back cannot be communicated.

This paper introduces a state strategy equilibrium framework where players condition current play on summary statistics of past play rather than the entire history. We provide a recursive characterization for the set of equilibrium payoffs in repeated games with limited feedback in the form of arbitrary restrictions on strategies. The tools we develop can be useful for deriving comparative statics results and for solving for the set of equilibrium payoffs in applications.

Our main contribution is to extend the machinery developed by Abreu, Pearce, and Stacchetti [2,3] to an alternative equilibrium concept for repeated games, namely state strategy equilibrium, whereas several other papers have adapted it to richer dynamic settings, including games with a payoff relevant state variable (Atkeson [7], Phelan and Stacchetti [31]), games with private information (Cole and Kocherlakota [13], Fernandes and Phelan [20]), repeated games with private monitoring (Ely, Hörner, and Olszewski [18], Cherry and Smith [12]), and games with hyperbolic discounting (Chade, Prokopovych, and Smith [11]). Our state strategy equilibrium framework builds on the small literature on repeated games with restricted feedback, including the OLG model in Bhaskar [9], the repeated prisoners dilemma in Cole and Kocherlakota [14], and the repeated minority game in Renault, Scarsini, and Tomala [32], by providing a recursive characterization of the set of equilibrium payoffs for a fairly general class of games. More recently, Barlo, Carmona, and Sabourian [8] provide a folk theorem in one period memory strategies for repeated games with perfect monitoring and rich action sets, Mailath and Olszewski [28] provide a folk theorem in finite memory strategies for perfect monitoring games, and Hörner and Olszewski [23] also allow for imperfect monitoring. We complement this literature by characterizing the set of equilibrium payoffs for fixed discount factor and memory restrictions (as encoded in the state space).

In Section 2 we present an infinitely repeated game and introduce a state space  $S$  such that the state  $s^t \in S$  in period  $t \geq 1$  is drawn from a distribution  $Q(\cdot; a^{t-1}, s^{t-1})$ , where  $a^{t-1} \in A$  is the action profile in period  $t - 1$ . A state is simply a summary statistic of past play. A *state strategy* for player  $i \in I$  is a sequence of functions  $(\sigma_i^t)_{t \geq 0}$  such that  $\sigma_i^t$  maps states  $s^t \in S$  into actions  $a_i^t \in A_i$ . A *state strategy equilibrium*  $\sigma$  is a perfect equilibrium in state strategies. Our state strategy equilibrium framework is general enough to encompass repeated games with memory restrictions (as in Mailath and Morris [27], Cole and Kocherlakota [14], Liu and Skrzypacz [26]), as well as more general repeated game strategies in which the history of play is summarized by a publicly observable state variable (as in Doraszelski and Escobar [15], Ekmekci [16]).

In Section 3 we show that state strategy equilibrium payoffs can be analyzed using recursive techniques similar to those introduced by Abreu, Pearce, and Stacchetti [3]. To this end, we introduce the set  $E$  of all functions  $v$  that map states into payoff vectors such that there exists a state strategy equilibrium  $\sigma$  for which  $v_i(s)$  is player  $i$ 's continuation value when play transpires according to  $\sigma$  and the initial state is  $s$ . Given an arbitrary set  $W$  of functions that map states into payoff vectors, we say that a function  $v$  is *decomposed* on  $W$  if there exists a function  $\alpha$  that maps states into pure actions and a continuation value function  $w$  selected from  $W$  such that, in

each state  $s$ ,  $\alpha_i(s)$  is a best response for player  $i$  and results in a payoff of  $v_i(s)$ . We also say that  $v$  is decomposed by  $\alpha$  and  $w \in W$ . It is therefore natural to define the set  $\mathcal{B}(W)$  of all functions decomposed on  $W$ . We say that  $W$  is *self-generating* if  $W \subseteq \mathcal{B}(W)$  and prove that self-generating sets are contained in  $E$ . Moreover,  $E = \mathcal{B}(E)$  and therefore  $E$  is the largest self-generating set. We also show that iterative application of the operator  $\mathcal{B}$  results in a decreasing sequence of sets that converge to  $E$ .

Our main point of departure from the existing literature is that our objects of interest are functions that map states into vectors of continuation values (one value per player) and not simply vectors of continuation values. Our operator thus characterizes the set of all functions that can be decomposed using continuation value functions in a given set. This construction of continuation value functions allows us to properly eliminate the dependence of current play on past states.

The tools we develop can also be applied to solve for the set of nonstationary Markov perfect equilibrium payoffs in dynamic stochastic games as usually studied in applied work (Ericson and Pakes [19]). In those games, the current state affects not only the transition probabilities but also the current payoffs. Our recursive characterization in Section 3 directly extends to this more general setting.

In Section 4 we specialize the model and consider strategies with one period memory. These strategies condition on a state drawn from a distribution that is parameterized by the actions in the previous period. Under an absolute continuity restriction on the monitoring technology, we establish a bang bang result implying that a function  $v$  decomposed by  $\alpha$  and  $w \in W$  can also be decomposed by  $\alpha$  and  $\hat{w} \in W$ , with  $\hat{w}$  taking values in the extreme points of the convex hull of the range of  $w$ . As an application we deduce that an improved monitoring technology unambiguously expands the set of equilibrium payoffs and thus provide a result similar to that of Kandori [25] for perfect public equilibria.

In Section 5 we apply our methods to solve for the set of equilibrium payoffs of a repeated product choice game in which players use strategies with one period memory. In our game, a short lived seller is tempted to produce low quality goods when facing each of the members of a sequence of short lived costumers. As in the existing literature (Abreu, Milgrom, and Pearce [1], Fudenberg and Levine [22]), introducing a public randomization device facilitates the analysis. The set of equilibrium payoffs turns out to be surprisingly simple: For discount factors above a certain threshold, the set of equilibrium payoffs with one period memory coincides with the payoff set in perfect public equilibria, while below the threshold the unique equilibrium is to repeat the static Nash equilibrium. Our application shows when and how the dynamics of incentive provision lead to cooperative behavior when the most severe nontrivial memory restriction on strategies is in place. By fully characterizing the conditions under which this restriction does not bind, we are able to sharpen a finding of Cole and Kocherlakota [14]. These authors consider a repeated prisoners dilemma and show when the set of equilibrium payoffs with finite memory strategies approaches the payoff set in strongly symmetric perfect public equilibria. We further demonstrate that a memory length of one is enough to sustain efficient trade provided the discount factor is above a given threshold.

## 2. Model

### 2.1. Set up

We consider an infinitely repeated game with long and short lived players. Time is discrete  $t = 0, 1, \dots$ . The stage game is  $(I, (A_i)_{i \in I}, (u_i)_{i \in I})$ , where  $I$  is the set of players,  $A_i$  is a finite set

of actions for player  $i$ , and  $u_i : A = \prod_{i \in I} A_i \rightarrow \mathbb{R}$  is the payoff of player  $i$ . Players  $1, \dots, n$  are long lived and discount period payoffs geometrically at a rate  $\delta \in ]0, 1[$ .<sup>1</sup> Thus, given a sequence of action profiles  $(a^t)_{t \geq 0}$ , the discounted payoff of long lived player  $i$  is

$$(1 - \delta) \sum_{t \geq 0} \delta^t u_i(a^t),$$

where  $a^t \in A$  is the action profile in period  $t$ . Players  $\{n + 1, \dots, |I|\}$  are short lived and maximize their current payoffs. Following Fudenberg, Kreps, and Maskin [21], a short lived player is active for one period and a new generation of short lived players enters the game in each period. We do not exclude the case  $n = |I|$  in which all players are long lived. At the beginning of period  $t \geq 1$ , a signal  $y^t \in Y$  is drawn from a distribution  $G(dy; a^{t-1})$ , where  $Y \subseteq \mathbb{R}^N$  is endowed with the Borel  $\sigma$  field. This setting corresponds to a standard repeated game of imperfect public monitoring.

### 2.2. State strategy equilibria

We add to the repeated game a measurable space of states  $(S, \mathcal{S})$  and a transition function  $Q(\cdot; s, y) \in \Delta(S)$ , where  $\Delta(S)$  denotes the set of probability measures on  $S$ . The state in period  $t + 1$ ,  $s^{t+1}$ , is drawn from the distribution  $Q(\cdot; s^t, y^t)$  where  $s^t$  and  $y^t$  are the state and the signal in period  $t$ . Given the current state  $s^t$  and the current action profile  $a^t$ , the distribution over next period's state  $s^{t+1}$  takes the form

$$q(M; a^t, s^t) = \int Q(M; s^t, y)G(dy; a^t),$$

where  $M \subseteq S$  is a measurable set. The state in period 0,  $s^0$ , is drawn from a distribution  $q_0 \in \Delta(S)$ . We assume that for all measurable sets  $M \subseteq S$ , the function  $(s, y) \in S \times Y \mapsto Q(M; s, y)$  is measurable so that for all  $a \in A$ ,  $s \in S \mapsto q(M; a, s)$  is measurable and therefore for any measurable function  $w : S \rightarrow \mathbb{R}$ ,  $\int w(s^{t+1})q(ds^{t+1}; a, s^t)$  is measurable as a function of  $s^t \in S$  (Stokey and Lucas [35, Theorems 8.1 and 8.2]).

A *state strategy* for player  $i$  is a collection of measurable functions  $\sigma_i = (\sigma_i^t)_{t \geq 0}$ , with  $\sigma_i^t : S \rightarrow A_i$ , such that in period  $t$ , after observing state  $s^t$ , player  $i$  selects action  $\sigma_i^t(s^t) \in A_i$ . The set of state strategies for player  $i$  is  $\Sigma_i$ . A state strategy profile  $\sigma = (\sigma_i)_{i \in I}$  is a *state strategy equilibrium* if for all periods  $t$  and all states  $s^t$ , the continuation strategy  $(\sigma_i^t)_{i \in I, t' \geq t}$  is a Nash equilibrium of the continuation game. Let EQUIL be the set of state strategy equilibria.

A state strategy equilibrium may not exist. As usually done (Abreu, Pearce, and Stacchetti [3]), we assume the stage game possesses a pure strategy Nash equilibrium  $a^* \in A$ . It is not hard to see that repetition of  $a^*$  is a state strategy equilibrium of the infinitely repeated game. Thus, EQUIL is nonempty.

It is important to point out that whether players know (or recall) the history of states  $(s^0, \dots, s^{t-1})$  at the beginning of period  $t$  is immaterial because, when using state strategies, players condition on the current state  $s^t$  so that  $s^t$  fully determines current play and the distribution over continuation strategies. A state strategy equilibrium can thus be seen as a robust prediction in the sense that it applies even when players' recalls of past states are heterogenous and, in the limit, totally imperfect. In this sense, state strategy equilibria are robust to forgetting.

<sup>1</sup> We use the following notation:  $]a, b[ = \{r \in \mathbb{R} \mid a < r < b\}$  denotes the open interval from  $a$  to  $b$  and  $]a, b] = \{r \in \mathbb{R} \mid a < r \leq b\}$  denotes the interval open at  $a$  but closed at  $b$ .

### 2.3. Examples

Several models fit into our state strategy equilibrium framework.

**Example 1 (Perfect Public Equilibria).** If  $S = \bigcup_{t \geq 0} Y^t$ , where  $Y$  is the set of signals, then the set of state strategy equilibria coincides with the set of perfect public equilibria (Abreu, Pearce, and Stacchetti [3]). In a perfect public equilibrium, each player can condition arbitrarily on the history of public signals but neglects her own private actions.

**Example 2 (Finite Memory Equilibria).** Consider a model in which players use finite memory strategies and condition on the last  $\kappa \geq 1$  signals  $y^t$  as in Cole and Kocherlakota [14]. In our model, define the state space  $S = \bigcup_{k=1}^{\kappa} Y^k$ . The state  $s = (y^1, \dots, y^k) \in S$  is composed of the last  $k$  signals, where  $y^1$  is the most recent signal. The transition is deterministic and given by  $Q(\cdot; s, y) = \mathbb{1}(\cdot; (y, y^1, \dots, y^k))$  if  $s = (y^1, \dots, y^k) \in Y^k$  and  $k \leq \kappa - 1$  and  $Q(\cdot; s, y) = \mathbb{1}(\cdot; (y, y^1, \dots, y^{\kappa-1}))$  if  $s = (y^1, \dots, y^{\kappa}) \in Y^{\kappa}$ .<sup>2</sup> The initial state  $s^0$  is an arbitrary signal  $y^0 \in Y$ . State strategy equilibria of this model are perfect public equilibria with finite memory as studied by Mailath and Morris [27] and Cole and Kocherlakota [14].

**Example 3 (Markov Perfect Equilibria).** Extend our model by assuming that the payoff to player  $i \in I$ ,  $u_i$ , depends not only on the current action profile  $a^t$  but also on the current state  $s^t$ . This model is a dynamic game with payoff relevant public states as studied by Atkeson [7]. A state strategy equilibrium of this model is a (possibly nonstationary) Markov perfect equilibrium as typically considered in applied work (Ericson and Pakes [19], Acemoglu and Robinson [4]). The recursive methods we develop in Section 3 (Theorem 1 and Proposition 1) carry over and can be used to analyze and compute the set of Markov perfect equilibrium payoffs. In contrast, Atkeson [7] characterizes the larger set of subgame perfect equilibrium payoffs.

### 2.4. Equilibrium payoffs

The key aspect of the definition of state strategy equilibria is the irrelevance of past states for continuation play. We are interested in characterizing the set of equilibrium payoffs. To obtain our recursive characterization we must therefore consider richer objects than payoff vectors, namely functions that represent attainable equilibrium payoffs across different states. Working with such functions allows us to avoid any dependence of continuation play on past states.

For each state  $s \in S$  and state strategy  $\sigma \in \Sigma$  define the expected discounted payoff of long lived player  $i$  as

$$V_i(s | \sigma) = (1 - \delta) \mathbb{E}_{\sigma} \left[ \sum_{t \geq 0} \delta^t u_i(a^t) \mid s^0 = s \right],$$

where the probability measure over the set of histories is induced by  $\sigma$  and the initial state is  $s^0 = s$ . The number  $V_i(\cdot | \sigma)$  is the continuation value function and the set of equilibrium payoffs is the set of all such functions obtained from equilibrium strategies:

$$E = \{v : S \rightarrow \mathbb{R}^n \mid \exists \sigma \in \text{EQUIL such that } v(s) = V(s | \sigma) \forall s \in S\}.$$

<sup>2</sup> Here,  $\mathbb{1}(\cdot; \cdot)$  denotes the indicator function so that  $\mathbb{1}(a; b) = 1$  if  $a = b$  and  $\mathbb{1}(a; b) = 0$  otherwise.

Defining  $v^* : S \rightarrow \mathbb{R}^n$  by  $v^*(s) = (u_i(a^*))_{i=1}^n$  for all  $s \in S$  with  $a^*$  being the Nash equilibrium of the stage game, it follows that  $v^* \in E$ . Standard arguments (Stokey and Lucas [35, Theorem 9.2]) can be used to check that functions in  $E$  are measurable. Thus  $E$  is a nonempty set of measurable functions.

Because the set  $E$  plays a key role in the subsequent analysis, we illustrate its construction with an example.

**Example 4.** Consider a prisoners dilemma with payoff matrix

	<i>C</i>	<i>D</i>
<i>C</i>	1, 1	− <i>l</i> , 1 + <i>g</i>
<i>D</i>	1 + <i>g</i> , − <i>l</i>	0, 0

Both players are long lived and monitoring is perfect. To represent trigger strategies as state strategies, suppose that the state space is {On, Off} and

$$s^{t+1} = \begin{cases} \text{On} & \text{if } a^t = (C, C) \text{ and } s^t = \text{On,} \\ \text{Off} & \text{if not.} \end{cases}$$

The initial state is  $s^0 = \text{On}$ . Assume  $\delta \geq \frac{g}{1+g}$ . It is easy to see that there are two state strategy equilibria. In the first of them, players always defect. In the second equilibrium, players cooperate when the state is On and defect otherwise. The set of equilibrium payoffs is therefore  $E = \{(0, 0, 0, 0), (1, 1, 0, 0)\}$  where the first two components of a vector are payoffs in state On and the last two components are payoffs in state Off.

### 3. A characterization of the set of equilibrium payoffs

Since the short lived players behave myopically, it is useful to define the set  $B$  of all actions that are consistent with their static best responses:

$$B = \left\{ a \in A \mid a_i \in \arg \max_{a'_i \in A_i} u_i(a'_i, a_{-i}) \ i = n + 1, \dots, |I| \right\}.$$

Let  $\mathcal{W} = \{w \mid w : S \rightarrow \mathbb{R}^n \text{ is measurable}\}$  be the set of all possible continuation value functions. We also consider the set  $\mathcal{A}_i = \{\alpha_i \mid \alpha_i : S \rightarrow A_i \text{ is measurable}\}$  of all functions that map states into actions for player  $i$ . We define the operator  $\mathcal{B}$ , mapping a subset of continuation value functions  $W \subseteq \mathcal{W}$  to a subset of continuation value functions  $\mathcal{B}(W) \subseteq \mathcal{W}$ , by

$$\begin{aligned} \mathcal{B}(W) = \left\{ v \in \mathcal{W} \mid \exists \alpha \in \mathcal{A} \text{ and } w \in W \text{ such that} \right. \\ \text{(i) } \alpha(s) \in B \ \forall s \in S, \\ \text{(ii) } v_i(s) = (1 - \delta)u_i(\alpha(s)) + \delta \int w_i(s')q(ds'; \alpha(s), s) \\ = \max_{a_i \in A_i} (1 - \delta)u_i(a_i, \alpha_{-i}(s)) + \delta \int w_i(s')q(ds'; a_i, \alpha_{-i}(s), s) \\ \left. i = 1, \dots, n \ \forall s \in S \right\}. \end{aligned}$$

The set  $\mathcal{B}(W)$  is the set of payoff functions that can be enforced in different states given that arbitrary continuation value functions  $w$  can be chosen from  $W$ . Constraint (i) ensures that the

actions prescribed to short lived players are consistent with their myopic behavior. Constraint (ii) ensures that, given the continuation value function  $w_i$ , long lived player  $i$  is willing to choose the prescribed action  $\alpha_i(s)$  and achieves the target payoff  $v_i(s)$ . When  $v$ ,  $\alpha$ , and  $w$  satisfy (ii) we say that  $v$  can be *decomposed* by  $\alpha$  and  $w$  and when  $v \in \mathcal{B}(W)$ , we say that  $v$  can be *decomposed* on  $W$ .

A key difference between our operator  $\mathcal{B}$  and those previously proposed in the literature to characterize the set of subgame perfect equilibrium payoffs (Abreu, Pearce, and Stacchetti [2,3], Atkeson [7], Phelan and Stacchetti [31]) is that the definition of  $\mathcal{B}$  imposes a continuation value function  $w \in W$  that applies uniformly on current states  $s \in S$ . The fact that the continuation value function  $w$  does not depend on the current state  $s$  means, to put it somewhat crudely, that the way in which incentives are provided in the continuation game does not depend on the current state  $s$ , although, of course, the current state  $s$  determines the distribution over next period's continuation payoffs  $w(s')$ . This aspect of the construction allows us to ensure that payoffs and strategies depend solely on the current state, as they must in a state strategy equilibrium.

To see this point more clearly, consider the recursive characterization of subgame perfect equilibrium payoffs for dynamic games with payoff relevant states in Atkeson [7, pp. 1078–1079]. Adapted to our setting with payoff irrelevant states, his operator  $\tilde{\mathcal{B}}$ , mapping a correspondence  $\tilde{W} : S \rightrightarrows \mathbb{R}^n$  to a correspondence  $\tilde{\mathcal{B}}(\tilde{W}) : S \rightrightarrows \mathbb{R}^n$ , is defined by

$$\tilde{\mathcal{B}}(\tilde{W})(s) = \left\{ v \in \mathbb{R}^n \mid \exists a \in A \text{ and } w : S \rightarrow \mathbb{R}^n, \text{ with } w(s) \in \tilde{W}(s), \text{ such that} \right.$$

$$(i) \ a \in B,$$

$$(ii) \ v_i = (1 - \delta)u_i(a) + \delta \int w_i(s')q(ds'; a, s)$$

$$= \max_{\tilde{a}_i \in A_i} (1 - \delta)u_i(\tilde{a}_i, a_{-i}) + \delta \int w_i(s')q(ds'; \tilde{a}_i, a_{-i}, s)$$

$$\left. \begin{matrix} i = 1, \dots, n \end{matrix} \right\}.$$

Note first that the operator  $\tilde{\mathcal{B}}$  is defined on correspondences  $\tilde{W} : S \rightrightarrows \mathbb{R}^n$ , whereas our operator  $\mathcal{B}$  is defined on a subset of continuation value functions  $W \subseteq \mathcal{W}$ . Further note that in contrast to our operator  $\mathcal{B}$  the operator  $\tilde{\mathcal{B}}$  has a product structure. Inspection of  $\tilde{\mathcal{B}}(\tilde{W})(s)$  shows that there can be a different continuation value function  $w : S \rightarrow \mathbb{R}^n$  depending on the current state  $s \in S$ . Because equilibrium strategies are constructed inductively, the fact that the continuation value function depends on the current state implies that the equilibrium strategies condition on the entire history of states  $(s^0, \dots, s^t)$ . Hence, while the operator  $\tilde{\mathcal{B}}$  is useful to characterize the set of subgame perfect equilibrium payoffs,<sup>3</sup> it cannot be used to study equilibrium payoffs and strategies with restricted feedback.

We proceed to establish the main properties of our operator  $\mathcal{B}$ . Let  $W \subseteq \mathcal{W}$  be an arbitrary set of functions that map states into payoff vectors. We say that  $W$  is *bounded* if there exists  $\kappa > 0$  such that  $|v(s)| \leq \kappa$  for all  $v \in W$  and all  $s \in S$ . We say that  $W \subseteq \mathcal{W}$  is *self-generating* if  $W \subseteq \mathcal{B}(W)$ .

<sup>3</sup> Strictly speaking, the operator  $\tilde{\mathcal{B}}$  characterizes the set of subgame perfect equilibrium payoffs only when this period's state  $s^t$  encodes last period's action profile  $a^{t-1}$ . More generally, the operator  $\tilde{\mathcal{B}}$  characterizes subgame perfect equilibrium payoffs that condition on the history of states  $(s^0, \dots, s^t)$ .

**Theorem 1.** *The following hold:*

- (i) *Let  $W$  be self-generating and bounded. Then  $W \subseteq E$ ;*
- (ii)  *$E$  is the largest bounded fixed point of  $\mathcal{B}$ .*

This and all other results in the paper are proven in [Appendix A](#). The first part of the theorem is the state strategy version of Theorem 1 in Abreu, Pearce, and Stacchetti [3]. It shows that whenever a set  $W$  is contained in the set of all payoffs enforced by continuation values in  $W$ , then  $W$  is contained in the set of equilibrium payoffs. The idea behind the second part of the theorem is that, in equilibrium, continuation payoffs are also equilibrium payoffs. The innovation in the proof comes from the observation that as our operator avoids any dependence of continuation play on past and current states, we can construct equilibrium payoffs and strategies that depend solely on the current state.

Computing  $E$  by enumeration is typically infeasible as strategies may be nonstationary. Because the operator  $\mathcal{B}$  is monotone (in the sense of inclusion), it readily provides us with an algorithm to compute its largest fixed point  $E$ . Given any bounded set  $W_0 \subseteq \mathcal{W}$  such that  $E \subseteq \mathcal{B}(W_0) \subseteq W_0$ , define the sequence  $(W_\nu)_{\nu \in \mathbb{N}}$  recursively by  $W_\nu = \mathcal{B}(W_\nu)$ . The following result implies that the sequence  $(W_\nu)_{\nu \in \mathbb{N}}$  monotonically converges to  $E$ .

**Proposition 1.** *Assume that  $S$  is countable. Then  $W_{\nu+1} \subseteq W_\nu$  and  $E = \bigcap_{\nu \in \mathbb{N}} W_\nu$ .*

The proposition shows that by iteratively applying  $\mathcal{B}$  to a properly chosen initial set, one can approximate the set of equilibrium payoffs arbitrarily closely. When  $S$  is finite, one way to operationalize the algorithm is by dividing each  $W_\nu$  into a grid and then checking whether  $\mathcal{B}(W_\nu)$  is close to  $W_\nu$ . This approach is straightforward but slow.<sup>4</sup>

Alternatively, one can add a randomization device to the model and consider the operator  $\bar{\mathcal{B}}(W) = \text{co}(\mathcal{B}(W))$ , where  $\text{co}$  denotes the convex hull of a set. In period  $t$  strategies condition on the current state  $s^t$  and on the entire history of randomizations  $(\omega^0, \dots, \omega^t)$ .<sup>5</sup> The operator  $\bar{\mathcal{B}}$  is monotone and convex valued and one can use methods similar to Judd, Yeltekin, and Conklin [24] to compute its largest fixed point  $\bar{E}$ .<sup>6</sup> More precisely, the algorithm fixes a number of directions  $(\lambda_m)_{m=1}^M$ , where  $\lambda_m \in \mathbb{R}^{n|S|}$ , and iteratively computes an inner (respectively outer) approximation of  $\bar{\mathcal{B}}(W_\nu)$  by finding its extreme points  $(v_m)_{m=1}^M$ , where  $v_m \in \mathbb{R}^{n|S|}$ , in all directions and then taking their convex hull (respectively intersecting the corresponding supporting hyperplanes). This algorithm effectively keeps track of  $M$  real valued vectors of length  $n|S|$  and updates them by solving  $M$  linear programs with  $n|S|$  variables subject to incentive compatibility constraints.

Sleet and Yeltekin [34] extend Judd, Yeltekin, and Conklin’s [24] algorithm to compute the set of subgame perfect equilibrium payoffs of dynamic games with payoff relevant states. Our algorithm is more burdensome than theirs because, as discussed above, the subgame perfect equilibrium problem has a product structure that allows to update the extreme point  $v_m$  for direction  $\lambda_m$  state by state. That is, the linear program decomposes into  $|S|$  smaller programs with  $n$  variables.

<sup>4</sup> See Wang [36] and Athey and Bagwell [6] for examples in the context of perfect public equilibria.

<sup>5</sup> Later on in the paper, we compute the set of equilibrium payoffs when players condition only on the current realization of the randomization device.

<sup>6</sup> It is immediate that  $E \subseteq \bar{E}$ , but the inclusion is strict in general.



#### 4. Strategies with one period memory

We specialize our model by equating states and signals and assume  $S = Y \subseteq \mathbb{R}^N$ ,  $s^t = y^t$ , and the transition takes the form  $q(dy; a) = Q(dy; a) = G(dy; a)$ . Therefore, a state strategy is actually a finite memory strategy with memory one (henceforth, a one period memory strategy). One period memory strategies are attractive as they are the most severe memory restriction on strategies that makes the problem of long term relationships plausible. Bhaskar [9], Renault, Scarsini, and Tomala [32], and Barlo, Carmona, and Sabourian [8] study alternative properties of one period memory equilibria.

The key property of one period memory is stated in the following lemma:

**Lemma 1.** *Let  $W \subseteq \mathcal{W}$  and  $v \in \mathcal{B}(W)$ . Let  $\hat{v} \in \mathcal{W}$  be such that  $\text{range}(\hat{v}) \subseteq \text{range}(v)$ . Then  $\hat{v} \in \mathcal{B}(W)$ .*

This lemma shows that in order to characterize  $\mathcal{B}(W)$  it suffices to characterize the maximal (in the sense of inclusion) range of the members of  $\mathcal{B}(W)$ . As Example 4 shows, the lemma does not extend to more general state strategies.<sup>7</sup>

The bang bang result in the following proposition shows that the operator is fully characterized by the extremal points of the range of its members:

**Proposition 2.** *Assume that  $S = Y$  is a set of positive Lebesgue measure in  $\mathbb{R}^N$  and that the distribution of signals  $G(dy; a) = q(dy; a)$  is absolutely continuous with respect to the Lebesgue measure  $dy$  in  $\mathbb{R}^N$ . Let  $v : S \rightarrow \mathbb{R}^n$  be decomposed by  $\alpha \in A$  and  $\tilde{w} : S \rightarrow \text{range}(\tilde{w})$ . Suppose there exists a bounded function  $\hat{w} \in \mathcal{W}$  such that  $\text{range}(\tilde{w}) \subseteq \text{co}(\text{range}(\hat{w}))$  and  $\hat{w} \in \mathcal{B}(W)$  for some  $W \subseteq \mathcal{W}$ . Then there exists  $\bar{w} \in \mathcal{B}(W)$  such that  $v$  is decomposed by  $\alpha$  and  $\bar{w}$ ,  $\int \bar{w}(s')q(ds'; a) = \int \tilde{w}(s')q(ds'; a)$  for all  $a \in A$ , and  $\bar{w}(s') \in \text{ext}(\text{range}(\hat{w}))$  for almost all  $s' \in S$ .*

This bang bang result is weaker than that for perfect public equilibria (Abreu, Pearce, and Stacchetti [3]). The recursive characterization of the set of one period memory equilibrium payoffs applies to sets of functions. Because in general no function in  $W$  has a range containing the range of other members of  $W$ , it may not be possible to decompose all possible continuation value functions by using continuation values that are extreme points of a set in  $\mathbb{R}^n$ ; instead, one may have to use continuation value functions with disjoint ranges. This observation is further illustrated in Section 5. Before moving on to solve for the set of equilibrium payoffs in an application, we demonstrate the usefulness of our bang bang result for comparative statics.

##### 4.1. Comparative statics: Improving the monitoring technology

How does an increase in the precision of the signal impact the equilibrium set? Consider two different monitoring technologies,  $q(\cdot; a)$  and  $q'(\cdot; a)$ . We say that  $q'$  is a *quasi-garbling* of  $q$  if for all  $a \in A$

$$q'(M; a) = \int \Phi(x; M)q(dx; a),$$

<sup>7</sup> The function represented by  $v = (1, 1, 0, 0)$  belongs to  $\mathcal{B}(E)$ . Yet, while the range of  $\hat{v} = (1, 1, 1, 1)$  is contained in that of  $v$ ,  $\hat{v}$  is not contained in  $\mathcal{B}(E)$  because players cannot cooperate in state Off.

where  $\Phi : S \times S \rightarrow [0, 1]$  is such that for all  $s \in S$ ,  $\Phi(s; \cdot)$  is a probability measure defined on  $S$  (the Borel  $\sigma$  field of  $S = Y$ ) and for each measurable set  $M \subseteq S$ , the function  $s \in S \mapsto \Phi(s; M)$  is measurable (additional details can be found in Section 8.1 in Stokey and Lucas [35]). This definition corresponds to the natural notion of informativeness introduced by Blackwell [10], in which a signal distributed according to  $q$  provides more “precise” information about the actions than a signal distributed according to  $q'$ .

**Proposition 3.** *Assume that  $q(dy; a)$  is absolutely continuous with respect to the Lebesgue measure  $dy$  in  $\mathbb{R}^N$  and that  $q'$  is a quasi-garbling of  $q$ . Then  $\mathcal{B}(W'; q') \subseteq \mathcal{B}(W'; q)$  for all  $W' \subseteq \mathcal{W}$ .*

The idea behind the proposition is that if  $v$  can be decomposed on  $W'$  with a monitoring technology  $q'$ , then it is also possible to decompose  $v$  on  $W'$  with an improved monitoring technology  $q$ . Denoting by  $E'$  the set of equilibrium payoffs with monitoring technology  $q'$ , the following comparative statics result follows by noting that  $E' \subseteq \mathcal{B}(E'; q)$ :

**Corollary 1.** *Under the conditions of Proposition 3,  $E' \subseteq E$ .*

### 5. A product choice game

We use our results to solve for the set of equilibrium payoffs of a repeated game with one period memory strategies. While we study a product choice game between a long lived seller and a sequence of short lived costumers, our methods and results extend to other settings such as the strongly symmetric public perfect equilibria of a repeated prisoners dilemma (Abreu, Milgrom, and Pearce [1], Cole and Kocherlakota [14]).

Our product choice game has the following payoffs:

	$h$	$l$
$H$	$\bar{u}, 1$	$0, -1$
$L$	$\bar{u} + g, -1$	$\underline{u}, 1$

where  $0 < \underline{u} < \bar{u}$  and  $g > 0$ . The seller (row player) is long lived and customers (column player) are short lived. The seller can exert high ( $H$ ) or low ( $L$ ) effort and the customer can buy a high ( $h$ ) or a low ( $l$ ) quality product. Effort is costly for the seller. The customer prefers to buy a high quality product if the seller exerts effort, otherwise the customer prefers the low quality product. The unique Nash equilibrium of the one shot game is  $(L, l)$  and attains the minimax value for both players.

We study an infinitely repeated version of the game. Once the action profile  $a^t$  in period  $t$  has been selected,  $x^{t+1} \in \{0, 1\}$  is drawn from the distribution

$$\hat{q}(1; a^t) = \begin{cases} p & \text{if } a^t = (H, h), \\ q & \text{if } a^t = (H, l), \\ r & \text{if } a^t = (L, l) \text{ or } a^t = (L, h), \end{cases}$$

where  $1 \geq p \geq q > r \geq 0$  and  $\hat{q}(0; a^t) = 1 - \hat{q}(1; a^t)$ . We think of  $x^t$  as being a signal about the performance of the seller. As in the received literature (Abreu, Milgrom, and Pearce [1], Fudenberg and Levine [22]), we simplify the analysis by allowing players to use a public randomization device  $\omega^t \in [0, 1]$  drawn from a uniform distribution. We define the signal  $y^t = x^t + \omega^t \in [0, 2]$ , with  $x^0 = 0$ . Denote by  $G(dy; a^{t-1})$  the distribution function of the random variable  $y^t$  conditional on  $a^{t-1}$ . Observe that knowing  $y^t$  is equivalent to knowing its components  $x^t$  and  $\omega^t$ , thus,

from the perspective of equilibrium behavior, whether players condition on  $y^t$  or on  $(x^t, \omega^t)$  is immaterial. We write  $Y = [0, 2]$ .

*5.1. Perfect public equilibria*

As **Example 1** shows, a perfect public equilibrium can be seen as a state strategy equilibrium in which the state encodes the entire history of past signals. The tools introduced by Abreu, Pearce, and Stacchetti [3] can be applied to characterize the set of equilibrium payoffs in this repeated game. Let  $\Pi_\infty \subseteq \mathbb{R}$  be the set of perfect public equilibrium payoffs for the seller.

**Proposition 4.** *Let  $\delta_\infty = \frac{g}{(\bar{u}-\underline{u})(p-r)+pg}$  and  $v^* = \bar{u} - (1 - p)\frac{g}{p-r}$ . Then*

$$\Pi_\infty = \begin{cases} [\underline{u}, v^*] & \text{if } \delta \geq \delta_\infty, \\ \{\underline{u}\} & \text{if not.} \end{cases}$$

When  $\delta < \delta_\infty$ , the unique equilibrium is to repeat the static Nash equilibrium  $(L, l)$ . When  $\delta \geq \delta_\infty$ , the optimal equilibrium is in trigger strategies. In the first period, players play  $(H, h)$ . In period  $t \geq 1$ , players randomize, permanently playing  $(L, l)$  with positive probability after  $x^t = 0$ .

*5.2. One period memory equilibria*

It is natural to assume that costumers do not have access to the entire history of past signals. We thus apply our general results to investigate the equilibrium payoffs when players use strategies that depend solely on the current signal  $y^t$  or, in other words, players use strategies conditioning only on the current realizations of the monitoring signal  $x^t$  and the randomization device  $\omega^t$ . This assumption contrasts with our discussion of Judd, Yeltekin, and Conklin’s [24] algorithm in Section 3 where we allowed players to condition play on the entire history of realizations of the randomization device. In the context of the product choice game, the current assumption ensures analytic tractability and is perhaps also more appealing from a conceptual viewpoint.

We consider the following condition that restricts the informativeness of the monitoring technology:

**Condition 1.**  $\frac{u}{q-r} \geq \frac{g}{p-r}$ .

The relevance of this sort of condition to attain efficient economic transactions with memory restrictions has also been stressed by Cole and Kocherlakota [14] and Liu and Skrzypacz [26]. In contrast to these papers, we do not allow strategies to have arbitrarily long memory, but explore how **Condition 1** allows efficient economic transactions when strategies are restricted to have one period memory.

Let  $\Pi^1 \subseteq \mathbb{R}$  be the set of equilibrium payoffs for the seller that can be attained with one period memory strategies. The main result of this section is the following:

**Proposition 5.** *Let  $\delta_1 = \frac{g}{(p-r)(\bar{u}-\underline{u}+g)}$ . Then*

$$\Pi_1 = \begin{cases} [\underline{u}, v^*] & \text{if Condition 1 holds and } \delta \geq \delta_1, \\ \{\underline{u}\} & \text{if not.} \end{cases}$$

By stating necessary and sufficient conditions for  $v \in \mathbb{R}$  to be an equilibrium payoff, this proposition fully characterizes the set of equilibrium payoffs when strategies are restricted to have one period memory. It shows that this restriction is not binding in terms of what payoffs can be achieved if  $\delta \geq \delta_1$ . In doing so it refines Cole and Kocherlakota’s [14] finding by showing that in this case arbitrarily long memory is not needed to obtain the full set of perfect public equilibrium payoffs. When  $r > 0$ ,  $\delta_1 > \delta_\infty$  and there is a range of discount factors for which the set of perfect public equilibrium payoffs strictly contains the set of one period memory equilibrium payoffs, even when Condition 1 holds.

Proposition 5 illustrates the usefulness of our recursive characterization. Proving that  $v \in [\underline{u}, v^*]$  is an equilibrium payoff when  $\delta$  is sufficiently large can be done by direct albeit tedious calculation without using our tools (as shown by Cole and Kocherlakota [14]). However, showing that a payoff  $v \in [\underline{u}, v^*]$  cannot be attained when  $\delta < \delta_1$  is not obvious and our recursive characterization allows us simplify this task by focusing on the dynamic programming problem of the seller.

To prove Proposition 5, we employ the tools introduced in the previous sections. Note that the state space is  $S = [0, 2]$ , the set  $E$  contains functions of the form  $v : [0, 2] \rightarrow \mathbb{R}$  and the domain of the operator  $\mathcal{B}$  is the set of subsets of measurable functions  $v : [0, 2] \rightarrow \mathbb{R}$ . Since the public randomization device is drawn from a uniform distribution, the realizations of  $y^t$  are absolutely continuous and Proposition 2 allows us to characterize the set of functions  $\mathcal{B}(W)$  by the extreme points of the convex hull of the range of its members. We therefore simplify the problem by representing a set  $W$  containing measurable functions  $w : [0, 2] \rightarrow \mathbb{R}$  by  $\{(\underline{w}, \bar{w}) \in \mathbb{R}^2 \mid (\underline{w}, \bar{w}) = \text{ext}(\text{co}(\text{range}(w))), w \in W\}$ .

Given  $w \in \mathbb{R}^2$  and  $\psi_1, \psi_0 \in [0, 1]$ , define

$$V_1(\psi_1, \psi_0, w) = (1 - \delta)\bar{u} + \delta(\underline{w} + (p\psi_1 + (1 - p)\psi_0)(\bar{w} - \underline{w})),$$

$$V_0(\psi_1, \psi_0, w) = (1 - \delta)\underline{u} + \delta(\underline{w} + (r\psi_1 + (1 - r)\psi_0)(\bar{w} - \underline{w})).$$

$V_0(\psi_1, \psi_0, w)$  is the expected payoff if the current action profile is  $(L, l)$  and the continuation values are given by the vector  $w$ .  $\psi_1$  and  $\psi_0$  are the cutoffs for the randomization in the next period; below these cutoffs the low continuation value  $\underline{w}$  is applied.  $V_1(\psi_1, \psi_0, w)$  is defined analogously if the current action profile is  $(H, h)$ . A vector  $w = (\underline{w}, \bar{w}) \in \mathbb{R}^2$ , with  $\underline{w} < \bar{w}$ , may be used to enforce several profiles  $v = (\underline{v}, \bar{v}) \in \mathbb{R}^2$ . Profiles with  $\underline{v} = \bar{v}$  are not suitable to provide incentives for the seller to choose  $H$ . The set of all enforceable profiles with  $\bar{v} > \underline{v}$  is

$$\Phi(w) = \left\{ v \in \mathbb{R}^2 \mid \bar{v} = V_1(\psi_1, \psi_0, w) \right. \\ \left. \geq (1 - \delta)(\bar{u} + g) + \delta(\underline{w} + (r\psi_1 + (1 - r)\psi_0)(\bar{w} - \underline{w})), \right. \\ \left. \underline{v} = V_0(\psi_1, \psi_0, w) \geq \delta(\underline{w} + (q\psi_1 + (1 - q)\psi_0)(\bar{w} - \underline{w})) \right\}, \psi_1, \psi_0 \in [0, 1].$$

$\Phi(w)$  is nonempty if and only if  $\frac{u}{q-r} \geq \frac{g}{p-r}$  and  $w \in \mathcal{C}$ , where

$$\mathcal{C} = \left\{ w \in \mathbb{R}^2 \mid \bar{w} - \underline{w} \geq \frac{1 - \delta}{\delta} \frac{g}{p - r} \right\}.$$

We deduce that  $E = \{(0, 0)\}$  when Condition 1 does not hold.

We characterize  $\Pi_1$  when Condition 1 holds. This characterization is derived in two steps.

**Lemma 2.** *Suppose that  $\delta < \delta_1$ . Then the unique equilibrium is to repeat the static Nash equilibrium  $(L, l)$ .*

The idea behind this lemma is that to provide incentives, continuation values after  $x^t = 1$  must be sufficiently large compared to continuation values after  $x^0 = 0$ . But the distribution over continuation values cannot depend on the current signal (or more generally on the current state) and this puts an upper bound on how much variation we can impose on continuation values. The proof of the lemma shows that these two bounds are not compatible when  $\delta < \delta_1$ .

**Lemma 3.** *Let  $\eta = (1 - \delta)(\bar{u} - \underline{u} + g)$ ,  $v = (v^* - \eta, v^*)$ , and  $W = \{v\}$ . If  $\delta \geq \delta^1$  and Condition 1 holds, then  $W$  is self-generating.*

The lemma implies that  $\{v^*, \underline{u}\} \in \Pi_1$  when  $\delta \geq \delta_1$ . Proposition 5 follows by showing that elements in between  $\underline{u}$  and  $v^*$  can also be attained in an equilibrium with one period memory. Details are given in Appendix A.

From Lemma 3 we can construct the equilibrium strategies sustaining  $v^*$  as

$$\sigma^t(x^t, \omega^t) = \begin{cases} (H, h) & \text{if } t = 0 \text{ or } x^t = 1 \text{ or } [x^t = 0, \omega^t \geq \frac{g}{\delta(p-r)(\bar{u}-\underline{u}+g)}], \\ (L, l) & \text{if not.} \end{cases}$$

This strategy profile is stationary. When  $x^t = 1$ , players play  $(H, h)$  with probability 1 as in the infinite memory case. When  $x^t = 0$ , players choose  $(L, l)$  with probability  $\frac{g}{\delta(p-r)(\bar{u}-\underline{u}+g)}$  which is strictly greater than the probability with which permanent play of  $(L, l)$  is triggered in the infinite memory case. This is so because in the former case the punishment consists of only one period of defection and therefore it must be carried out more often to provide incentives to produce high quality products. In other words, with infinite memory the continuation value in the punishment phase is harsher, but with finite memory punishment is triggered more often.

The public randomization device not only simplifies the analysis of the model but also plays a substantive role. Without it, the incentive constraint enforcing  $(L, l)$  imposes an upper bound on the discount factor  $\delta$ , so that  $(H, h)$  can be enforced only for intermediate values of the discount factor. Moreover, in our model, unless  $\underline{u}$  is sufficiently large or  $q$  is sufficiently larger than  $r$ , efficient transactions cannot be attained with stationary strategies and no randomization device regardless of the discount factor (Mailath and Samuelson [29, Section 7.2.2]).<sup>8</sup> As in Ellison’s [17] community enforcement model, the public randomization device allows us to fine-tune the severity of the punishment so that the incentive constraint enforcing the punishment profile  $(L, l)$  does not bind and payoffs in  $[\underline{u}, v^*]$  can be attained.

### 6. Concluding remarks

We ask how restrictions on strategies shape the extent to which players can use long term relationships to align private and public incentives. We show that the methods introduced by Abreu, Pearce, and Stacchetti [2,3] can be adapted to characterize equilibrium payoffs in a state strategy equilibrium framework. Our recursive characterization can be useful for deriving comparative statics results and for solving for the set of equilibrium payoffs in applications.

---

<sup>8</sup> It can also be shown the set of equilibrium payoffs with stationary strategies and a randomization device is strictly contained in  $\Pi_1$  and contains  $\underline{u}$  as an isolated point. Observe that if strategies can condition on the entire history of realizations of the randomization device, this result does not hold as the initial randomization can be used to convexify the set of equilibrium payoffs.

**Private monitoring.** While our results apply to repeated games of public monitoring, they also have implications for repeated games of private monitoring. Mailath and Morris [27] show that strict perfect public equilibria in finite memory strategies exhaust the set of equilibria that are robust to private monitoring. The set of equilibrium payoffs in one period memory strategies characterized in Proposition 5 is a lower bound for the set of perfect public equilibrium payoffs that survives the introduction of a tiny amount of private monitoring. Phelan and Skrzypacz [30] provide necessary and sufficient conditions for a state strategy profile to be a sequential equilibrium of a private monitoring game and use their methods to check whether tit-for-tat is a sequential equilibrium in a private monitoring repeated prisoners dilemma. The results presented in Section 5 suggest that by introducing a randomization device, tit-for-tat may be an equilibrium without imposing an upper bound on the discount factor.

**From one period to finite memory strategies.** While our recursive characterization in Section 3 covers general state strategy equilibria, our applications in Sections 4 and 5 leave a gap between one period and finite memory strategies. Our proofs of Propositions 3 and 5 make use of the bang bang result in Proposition 2, which itself is an implication of Lemma 1. Extending Proposition 2 to more general strategies with finite memory length  $\kappa$  seems promising (we have not been able to come up with a counterexample) but difficult. First, Lemma 1 need not hold. Second, the bang bang result one could presumably obtain is a conditional result in the sense that given the last  $\kappa$  signals,  $(y^1, \dots, y^\kappa)$ , continuation values, as functions of the next signal,  $y^{\kappa+1}$ , can be taken from extreme points of convex sets.<sup>9</sup> Such a result would at most simplify the problem by allowing us to manipulate continuation value functions that depend arbitrarily on  $(y^2, \dots, y^\kappa)$  but in simpler “bang bang” way on  $y^{\kappa+1}$ . But this simplification is not enough to extend Propositions 3 and 5 because the arbitrary dependence of continuation values on all but the last signal renders the operator intractable. We leave these explorations for future research.

### Appendix A. Proofs

**Proof of Theorem 1.** Part (i). We prove the stronger result that  $\mathcal{B}(W) \subseteq E$ . For each  $v \in \mathcal{B}(W)$ , we can find  $\alpha^v \in \mathcal{A}$  and  $w^v \in W$  such that for all  $s \in S$ ,  $\alpha(s) \in B$  and

$$\begin{aligned} v_i(s) &= (1 - \delta)u_i(\alpha^v(s)) + \delta \int w_i^v(s')q(ds'; \alpha^v(s), s) \\ &= \max_{a_i \in A_i} (1 - \delta)u_i(a_i, \alpha_{-i}^v(s)) + \delta \int w_i^v(s')q(ds'; a_i, \alpha_{-i}^v(s), s). \end{aligned}$$

Consider an arbitrary  $v^0 \in \mathcal{B}(W)$ . Define inductively  $v^{t+1} = w^{v^t}$  for all  $t \geq 0$ . This is well defined since  $v^t \in W \subseteq \mathcal{B}(W)$  and therefore  $v^{t+1} = w^{v^t} \in W$ . Consider the state strategy profile  $\sigma = (\sigma^t)_{t \geq 0}$  defined by  $\sigma^t(s^t) = \alpha^{v^t}(s^t)$ . This profile is a state strategy equilibrium. Indeed,  $\sigma_i \in \Sigma_i$  and given  $s^0, \dots, s^t$ , player  $i$ 's continuation value is given by a function  $v_i^{t+1}(s^{t+1})$ , which does not depend on  $s^0, \dots, s^t$ . Therefore, provided  $i$ 's rivals play  $\sigma_{-i}^t(s^t)$ , player  $i$  has incentives to play  $\sigma_i^t(s^t)$ . Moreover,  $\sigma$  results in a value of  $v^0$ . Indeed, let  $p_{\sigma, s}^t(\cdot)$  be the probability measure induced by the random variable  $s^t$  given the strategy  $\sigma$ , conditional on  $s^0 = s$ . By construction, for all  $T \geq 1$  we can write

<sup>9</sup> When the memory length is  $\kappa$ , our recursive characterization effectively has  $s^\kappa = (y^1, \dots, y^{\kappa-1}, y^\kappa)$  as state variable, see Example 2. Combining it with Lyapunov's theorem requires conditioning on  $(y^1, \dots, y^\kappa)$ .

$$v^0(s) = (1 - \delta) \sum_{t=0}^{T-1} \delta^t \mathbb{E}_{p_{\sigma,s}^t} [u(\sigma^t(s^t))] + \delta^T \mathbb{E}_{p_{\sigma,s}^T} [v^T(s^T)].$$

But  $W$  is bounded and thus  $(v^T)_{T \geq 1}$  is a sequence of uniformly bounded functions. Taking the limit, we deduce that

$$v^0(s) = (1 - \delta) \sum_{t \geq 0} \delta^t \mathbb{E}_{p_{\sigma,s}^t} [u(\sigma^t(s^t))].$$

*Part (ii).* We prove that  $E = \mathcal{B}(E)$ . The fact that  $E$  is the largest bounded fixed point then follows from part (i) and the fact that period payoffs are bounded.

We first show that  $E \subseteq \mathcal{B}(E)$ . Let  $v \in E$  and consider the corresponding equilibrium profile  $\sigma = (\sigma_i)_{i \in I}$ . Define  $w : S \rightarrow \mathbb{R}$  by

$$w(s) = V(s \mid (\sigma^t)_{t \geq 1}) = (1 - \delta) \mathbb{E} \left[ \sum_{t \geq 1} \delta^{t-1} u(a^t) \mid s^1 = s, (\sigma^t)_{t \geq 1} \right],$$

where the expectation is with respect to the unique probability measure induced on the set of histories by  $(\sigma^t)_{t \geq 1}$  conditional on  $s^1 = s$ . By construction, the measurable function  $w$  belongs to  $E$ . Define  $\alpha \in \mathcal{A}$  by  $\alpha(s) = \sigma^0(s)$ . Clearly,

$$\begin{aligned} v_i(s) &= (1 - \delta) u_i(\alpha(s)) + \delta \int w_i(s') q(ds'; \alpha(s), s) \\ &= \max_{a_i \in A_i} (1 - \delta) u_i(a_i, \alpha_{-i}(s)) + \delta \int w_i(s') q(ds'; a_i, \alpha_{-i}(s), s). \end{aligned}$$

This proves that  $v \in \mathcal{B}(E)$ .

Let  $v \in \mathcal{B}(E)$ . By definition, there exists  $\alpha^v \in \mathcal{A}$  and  $w^v \in E$  such that

$$\begin{aligned} v_i(s) &= (1 - \delta) u_i(\alpha^v(s)) + \delta \int w_i^v(s') q(ds'; \alpha^v(s), s) \\ &= \max_{a_i \in A_i} (1 - \delta) u_i(a_i, \alpha_{-i}^v(s)) + \delta \int w_i^v(s') q(ds'; a_i, \alpha_{-i}^v(s), s). \end{aligned}$$

Let  $\tilde{\sigma} = (\tilde{\sigma}_t)_{t \geq 0}$  be the state strategy profile generating the payoff  $w^v \in E$ . Define the following state strategy

$$\sigma^t(s) = \begin{cases} \alpha^v(s) & \text{if } t = 0, \\ \tilde{\sigma}^{t-1}(s) & \text{if } t \geq 1. \end{cases}$$

This defines a state strategy equilibrium and  $v_i(s) = V_i(s \mid \sigma)$  so that  $v \in E$ .  $\square$

**Proof of Proposition 1.** Since  $W_1 = \mathcal{B}(W_0) \subseteq W_0$ , it follows that  $W_{n+1} \subseteq W_n$  for all  $n$ . Moreover,  $E \subseteq W_n$  for all  $n$  and therefore  $E \subseteq \bigcap_{n \in \mathbb{N}} W_n$ . To prove that  $\bigcap_{n \in \mathbb{N}} W_n \subseteq E$ , we prove that  $\bigcap_{n \in \mathbb{N}} W_n$  is self-generating. We observe that  $\bigcap_{n \in \mathbb{N}} W_n$  contains only measurable functions as each  $W_n$ , by definition, contains only measurable functions. To prove that  $\bigcap_{n \in \mathbb{N}} W_n$  is self-generating, let  $v \in \bigcap_{n \in \mathbb{N}} W_n$ . Then, for all  $n \geq 1$  there exists  $\alpha_n \in \mathcal{A}$  and  $w_n \in W_{n-1}$  such that for all  $s \in S$ ,  $\alpha_n(s) \in B$ , and

$$\begin{aligned} v_i(s) &= (1 - \delta) u_i(\alpha_n(s)) + \delta \int w_{n,i}(s') q(ds'; \alpha_n(s), s) \\ &= \max_{a_i \in A_i} (1 - \delta) u_i(a_i, \alpha_{n,-i}(s)) + \delta \int w_{n,i}(s') q(ds'; a_i, \alpha_{n,-i}(s), s). \end{aligned}$$

Note that  $(\alpha_n)_{n \in \mathbb{N}}$  and  $(w_n)_{n \in \mathbb{N}}$  have pointwise converging subsequences as both are contained in the countable product of compact metric spaces (recall that  $A$  is finite,  $S$  is countable, and  $W_0$  is bounded). Without loss of generality, we assume that  $(\alpha_n)_{n \in \mathbb{N}}$  and  $(w_n)_{n \in \mathbb{N}}$  converge and denote by  $\alpha$  and  $w$  the limit functions. By passing to the limit, it is easy to see that

$$\begin{aligned} v_i(s) &= (1 - \delta)u_i(\alpha(s)) + \delta \int w_i(s')q(ds'; \alpha(s), s) \\ &= \max_{a_i \in A_i} (1 - \delta)u_i(a_i, \alpha_{-i}(s)) + \delta \int w_i(s')q(ds'; a_i, \alpha_{-i}(s), s). \end{aligned}$$

Since  $w \in \bigcap_{n \in \mathbb{N}} W_n$ , it follows that  $v \in \mathcal{B}(\bigcap_{n \in \mathbb{N}} W_n)$ . In other words,  $\bigcap_{n \in \mathbb{N}} W_n$  is self-generating and thus contained in  $E$ .  $\square$

**Proof of Lemma 1.** Let  $\alpha \in \mathcal{A}$  and  $w \in W$  decompose  $v$ . Let  $V : A \rightarrow \mathbb{R}$  be defined by  $V(a) = (1 - \delta)u(a) + \delta \int w(s')q(ds'; a)$  for all  $a \in A$ . Define the set valued map  $X : S \rightrightarrows A$  by

$$X(s) = \left\{ a \in B \mid \hat{v}(s) = V(a) \right\} \cap \left( \bigcap_{i=1}^n \left\{ a \in A \mid V_i(a) = \max_{a'_i \in A_i} V_i(a'_i, a_{-i}) \right\} \right).$$

Since  $\text{range}(\hat{v}) \subseteq \text{range}(v)$ , for all  $s \in S$  there exists  $\hat{s} \in S$  such that  $\hat{v}(s) = v(\hat{s}) = V(\alpha(\hat{s}))$ . As  $v$  is decomposed by  $\alpha$  and  $w$ , for all  $i = 1, \dots, n$ , the function

$$a_i \in A_i \mapsto V_i(a_i, \alpha_{-i}(\hat{s})) = (1 - \delta)u_i(a_i, \alpha_{-i}(\hat{s})) + \delta \int w(s')q(ds', a_i, \alpha_{-i}(\hat{s}))$$

is maximized at  $a_i = \alpha_i(\hat{s})$ . It follows that for all  $s \in S$ , there exists  $a = \alpha(\hat{s}) \in B$  such that  $a \in X(s)$  and therefore  $X(s)$  is nonempty.

We now prove that  $X : S \rightrightarrows A$  is a weakly measurable correspondence (Definition 18.1 in Aliprantis and Border [5]). To see this, define the correspondence

$$\varphi_0(s) = \{ a \in B \mid \hat{v}(s) = V(a) \}$$

and observe that  $X$  can be obtained as the intersection of  $\varphi_0$  and constant correspondences. Since a constant correspondence is weakly measurable, proving the weak measurability of  $X$  amounts to proving the weak measurability of  $\varphi_0$  (Aliprantis and Border [5, Lemma 18.4, part 3]). Let  $T \subseteq A$  and let us prove that  $\varphi_0^l(T) = \{ s \in S \mid \varphi_0(s) \cap T \neq \emptyset \}$  is a measurable set. Writing

$$\varphi_0^l(T) = \bigcup_{a \in T} \{ s \in S \mid a \in \varphi_0(s) \} = \bigcup_{a \in T} \{ s \in S \mid a \in B, \hat{v}(s) = V(a) \} = \bigcup_{a \in B \cap T} \hat{v}^{-1}(V(a))$$

the result is deduced by noting that the last expression, being the finite union of measurable sets, is measurable.

Since  $A$  is a Polish space and the correspondence  $X : S \rightrightarrows A$  is weakly measurable and has nonempty closed values, the Kuratowski–Ryll–Nardzewski Selection theorem (Aliprantis and Border [5, Theorem 18.13]) implies that  $X$  has a measurable selection: there exists a measurable function  $\hat{\alpha} : S \rightarrow A$  such that  $\hat{\alpha}(s) \in X(s)$  for all  $s \in S$ . By construction of  $X$ ,  $\hat{v}$  can be decomposed by  $\hat{\alpha} \in \mathcal{A}$  and  $w \in W$ . Thus,  $\hat{v} \in \mathcal{B}(W)$ .  $\square$



**Proof of Proposition 2.** Define

$$\hat{\Gamma} = \left\{ w \in L^\infty(S, \mathbb{R}^n) \mid \alpha \text{ is enforced by } w \in \mathcal{W}, \right. \\ \left. w : S \rightarrow \text{co}(\text{range}(\hat{w})), \int w(s')q(ds'; a) = \int \tilde{w}(s')q(ds'; a) \forall a \in A \right\},$$

where  $\text{co}(A)$  is the convex hull of a set  $A \subseteq \mathbb{R}^n$ . The set  $\hat{\Gamma}$  is nonempty and convex. Observe that  $\hat{w} \in \mathcal{B}(W)$  is bounded and its range,  $\text{range}(\hat{w})$ , is finite. Indeed, there exists  $\alpha^* \in \mathcal{A}$  and  $w^* \in W$  such that for all  $s \in S$ ,  $\hat{w}(s) = (1 - \delta)u(\alpha^*(s)) + \delta \int w^*(s')q(ds'; \alpha(s))$ . Further, since we are restricting attention to pure strategies and there is a finite number of pure strategies, the set

$$\text{range}(\hat{w}) \subseteq \bigcup_{a \in A} \left\{ (1 - \delta)u(a) + \delta \int w^*(s')q(ds'; a) \right\}$$

is finite. Therefore,  $\text{co}(\text{range}(\hat{w}))$  is a compact set (Rockafellar [33, Corollary 2.3.1]). It then follows that  $\hat{\Gamma}$  is also weak\* compact and the Krein–Milman theorem (Aliprantis and Border [5, Theorem 7.68]) implies the existence of an extreme point  $\bar{w} \in \hat{\Gamma}$ .<sup>10</sup> We prove that  $\bar{w}$  has all the properties stated in the proposition.

**Claim 1.** *For almost all  $s' \in S$ ,  $\bar{w}(s')$  is an extreme point of  $\text{co}(\text{range}(\hat{w}))$ .*

To prove the claim, suppose otherwise. Then there exists a set of positive measure  $K \subset S$  such that for all  $s' \in K$ ,  $\bar{w}(s')$  is not an extreme point of  $\text{co}(\text{range}(\hat{w}))$ . Then, there exist  $w', w'' \in L^\infty(S, \mathbb{R}^n)$  with  $w'(s'), w''(s') \in \text{co}(\text{range}(\hat{w}))$  such that  $\bar{w} = \frac{1}{2}(w' + w'')$  and  $w' \neq w''$  for a positive measure set of states. Define  $w^* = \frac{1}{2}(w' - w'')$ . We define the vector valued measure  $\mu$  as  $\mu(S') = (\int_{S'} w_i^*(s')q(ds'; a'))_{i=1, \dots, n, a' \in A}$ . Since  $q(\cdot; a')$  is absolutely continuous,  $\mu$  is a nonatomic measure. Therefore, Lyapunov’s convexity theorem (Aliprantis and Border [5, Theorem 13.33]) implies that  $\{\mu(S') \mid S' \text{ is measurable}\}$  is convex. Therefore, there exists  $S'$  such that  $\mu(S') = \frac{1}{2}\mu(S)$ .

Define

$$\bar{w}(s') = \begin{cases} w'(s') & \text{if } s' \in S', \\ w''(s') & \text{if not,} \end{cases}$$

and

$$\bar{w}''(s') = \begin{cases} w''(s') & \text{if } s' \in S', \\ w'(s') & \text{if not.} \end{cases}$$

Note that  $\bar{w}(s'), \bar{w}''(s') \in \text{co}(\text{range}(\hat{w}))$ . By rearranging terms, it follows that

$$\int_S \bar{w}(s')q(ds'; a') = \int_{S'} w'(s')q(ds'; a') + \int_{S \setminus S'} w''(s')q(ds'; a) = \int_S \bar{w}(s')q(ds'; a')$$

so that  $\bar{w}' \in \hat{\Gamma}$ . The same calculation shows that  $\bar{w}'' \in \hat{\Gamma}$ . But  $\bar{w} = \frac{1}{2}(\bar{w}' + \bar{w}'')$  with  $\bar{w}' \neq \bar{w}''$  on a set of positive measure. This contradicts the fact that  $\bar{w}$  is an extreme point of  $\hat{\Gamma}$ . This establishes the claim.

<sup>10</sup> The space  $L^\infty(S, \mathbb{R}^n)$  is the dual of  $L^1(S, \mathbb{R}^n)$ . The weak\* topology on  $L^\infty(S, \mathbb{R}^n)$  is the weakest topology such that for all  $g \in L^1(S, \mathbb{R}^n)$ , the linear function  $f \in L^\infty(S, \mathbb{R}^n) \rightarrow \int f(s) \cdot g(s) ds$  is continuous.

To conclude the proof of the proposition, note that  $\text{range}(\bar{w}) \subseteq \text{range}(\hat{w})$ . Lemma 1 then implies that  $\bar{w} \in \mathcal{B}(W)$ .  $\square$

**Proof of Proposition 3.** Suppose that  $v \in \mathcal{B}(W', q')$  and let  $w' \in W'$  and  $\alpha \in \mathcal{A}$  be such that for all  $s \in S$ ,  $\alpha(s) \in B$  and

$$\begin{aligned} v_i(s) &= (1 - \delta)u_i(\alpha(s)) + \delta \int w'_i(s')q'(ds'; \alpha(s)) \\ &= \max_{a_i \in A_i} (1 - \delta)u_i(a_i, \alpha_{-i}(s)) + \delta \int w'_i(s')q'(ds'; a_i, \alpha_{-i}(s)) \end{aligned}$$

for all  $i = 1, \dots, n$ . Define  $w : S \rightarrow \mathbb{R}^n$  by

$$w(x) = \int w'(s')\Phi(x; ds').$$

This function is measurable as a consequence of the unnumbered corollary on p. 215 of Stokey and Lucas [35]. Moreover, Theorem 8.3 in Stokey and Lucas [35] implies that for all  $a \in A$

$$\int w'(s')q'(ds'; a) = \int w(x)q(dx; a).$$

It then follows that  $v$  is decomposed by  $\alpha$  and  $w$  given the absolutely continuous monitoring technology  $q$ . Finally, note that for each  $x \in S$ ,  $w(x) \in \text{co}(\text{range}(w'))$ , with  $w' \in W'$ . From Proposition 2, it follows that there exists  $\bar{w} \in W'$  such that  $v$  is decomposed by  $\alpha$  and  $\bar{w}$  given  $q$ . This proves that  $v \in \mathcal{B}(W', q)$ .  $\square$

**Proof of Lemma 2.** Consider any  $v \in E \cap \mathcal{C}$ . In particular,  $\underline{v} < \bar{v}$  and therefore there exists  $w \in E \cap \mathcal{C}$  and  $\psi_1, \psi_0 \in [0, 1]$  such that  $v = V(\psi_1, \psi_0, w)$ . Therefore

$$\bar{v} - \underline{v} \leq (1 - \delta)(\bar{u} - \underline{u}) + \delta(p - r) \sup_{w \in E \cap \mathcal{C}} (\bar{w} - \underline{w}).$$

This implies that  $\sup_{v \in E \cap \mathcal{C}} (\bar{v} - \underline{v}) \leq (1 - \delta)(\bar{u} - \underline{u}) + \delta(p - r) \sup_{w \in E \cap \mathcal{C}} (\bar{w} - \underline{w})$  and consequently,  $\sup_{v \in E \cap \mathcal{C}} (\bar{v} - \underline{v}) \leq \frac{(1-\delta)(\bar{u}-\underline{u})}{1-\delta(p-r)}$ . By definition of  $\mathcal{C}$ ,  $\inf_{v \in E \cap \mathcal{C}} (\bar{v} - \underline{v}) \geq \frac{1-\delta}{\delta} \frac{g}{p-r}$ . We thus deduce that  $\frac{1-\delta}{\delta} \frac{g}{p-r} \leq \frac{(1-\delta)(\bar{u}-\underline{u})}{1-\delta(p-r)}$ , a condition that contradicts  $\delta < \delta_1$ . Thus,  $E \cap \mathcal{C}$  is empty and consequently no element of  $E$  can enforce  $(H, h)$ .  $\square$

**Proof of Lemma 3.** We prove that  $W \subseteq \mathcal{B}(W)$ . Since  $\eta > 0$ , to prove that  $v \in \mathcal{B}(W)$  we need to show that there exist  $\psi_0, \psi_1 \in [0, 1]$  such that

$$\begin{aligned} v^* &= (1 - \delta)\bar{u} + \delta(v^* - \eta + \eta(p\psi_1 + (1 - p)\psi_0)) \\ &\geq (1 - \delta)(\bar{u} + g) + \delta(v^* - \eta + \eta(q\psi_1 + (1 - q)\psi_0)), \\ v^* - \eta &= (1 - \delta)\underline{u} + \delta(v^* - \eta + \eta(r\psi_1 + (1 - r)\psi_0)) \\ &\geq \delta(v^* - \eta + \eta(q\psi_1 + (1 - q)\psi_0)). \end{aligned}$$

Take  $\psi_1 = 1$  and  $\psi_0 = 1 - \frac{1-\delta}{\delta} \frac{g}{\eta(p-r)} \geq 0$  and let us verify that the conditions above are satisfied. Since  $\psi_1 - \psi_0 = \frac{1-\delta}{\delta} \frac{g}{\eta(p-r)}$  it follows that both incentive constraints hold. It is therefore enough to verify that the two equalities hold. To see the first equality note that

$$\begin{aligned} (1 - \delta)\bar{u} + \delta(v^* - \eta + \eta(p\psi_1 + (1 - p)\psi_0)) &= (1 - \delta)\bar{u} + \delta v^* - \delta\eta(1 - p)(1 - \psi_0) \\ &= \delta v^* + (1 - \delta)\left(\bar{u} - (1 - p)\frac{g}{p - r}\right) = v^*. \end{aligned}$$

The second equality follows analogously.  $\square$

**Proof of Proposition 5.** It immediately follows from Lemma 3 that  $\underline{u}$  and  $v^*$  are equilibrium payoffs when  $\delta \geq \delta_1$ . We claim that  $\bigcup_{n \geq 0} I^n = ]\bar{u}, v^*]$  where  $I^n = (1 - \delta^n)\underline{u} + \delta^n[v^* - \eta, v^*]$ . Indeed,  $I^n \cap I^{n-1} \neq \emptyset$ , each  $I^n$  is an interval, and  $\inf\{v \in I_n\} \rightarrow \underline{u}$  as  $n \rightarrow \infty$ . Therefore, for any  $v \in ]\underline{u}, v^*]$ , we can find  $n$  and  $\lambda \in [0, 1]$  such that  $v = (1 - \delta^n)\underline{u} + \delta^n(\lambda(v^* - \eta) + (1 - \lambda)v^*)$ . Consider the following strategies: During the first  $n - 1$  periods play  $(L, l)$ , in period  $n$  play  $(L, l)$  with probability  $\lambda$  and play  $(H, h)$  with probability  $(1 - \lambda)$ , for  $t > n$ , play according to the stationary strategies sustaining  $(v^* - \eta, v^*) \in E$ . This strategy has one period memory and prescribes optimal behavior after each history. Moreover, by construction, the expected payoff for the long lives player equals  $v = (1 - \delta^n)\underline{u} + \delta^n(\lambda(v^* - \eta) + (1 - \lambda)v^*)$ .  $\square$

## References

- [1] D. Abreu, P. Milgrom, D. Pearce, Information and timing in repeated partnerships, *Econometrica* 59 (6) (1991) 1713–1733.
- [2] D. Abreu, D. Pearce, E. Stacchetti, Optimal cartel equilibria with imperfect monitoring, *J. Econ. Theory* 39 (1) (1986) 251–269.
- [3] D. Abreu, D. Pearce, E. Stacchetti, Toward a theory of discounted repeated games with imperfect monitoring, *Econometrica* 58 (5) (1990) 1041–1063.
- [4] D. Acemoglu, J. Robinson, A theory of political transition, *Amer. Econ. Rev.* 91 (4) (2001) 938–963.
- [5] C. Aliprantis, K. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer-Verlag, 2006.
- [6] S. Athey, K. Bagwell, Optimal collusion with private information, *RAND J. Econ.* 32 (3) (2001) 428–465.
- [7] A. Atkeson, International lending with moral hazard and risk of repudiation, *Econometrica* 59 (4) (1991) 1069–1089.
- [8] M. Barlo, G. Carmona, H. Sabourian, Repeated games with one memory, *J. Econ. Theory* 144 (1) (2009) 312–336.
- [9] V. Bhaskar, Informational constraints and the overlapping generations model: Folk and anti-folk theorems, *Rev. Econ. Stud.* 65 (1) (1998) 135–149.
- [10] D. Blackwell, Comparison of experiments, in: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, University of California Press, 1951, pp. 93–102.
- [11] H. Chade, P. Prokopovych, L. Smith, Repeated games with present-biased preferences, *J. Econ. Theory* 139 (1) (2008) 157–175.
- [12] J. Chery, L. Smith, Unattainable payoffs for repeated games of private monitoring, Discussion paper, University of Michigan, 2010.
- [13] H. Cole, N. Kocherlakota, Dynamic games with hidden actions and hidden states, *J. Econ. Theory* 98 (1) (2001) 114–126.
- [14] H. Cole, N. Kocherlakota, Finite memory and imperfect monitoring, *Games Econ. Behav.* 53 (1) (2005) 59–72.
- [15] U. Doraszelski, J. Escobar, A theory of regular Markov perfect equilibria in dynamic stochastic games: Genericity, stability, and purification, *Theoretical Econ.* 5 (3) (2010) 369–402.
- [16] M. Ekmekci, Sustainable reputations with rating systems, *J. Econ. Theory* 146 (2) (2011) 479–503.
- [17] G. Ellison, Cooperation in the Prisoner's dilemma with anonymous random matching, *Rev. Econ. Stud.* 61 (3) (1994) 567–588.
- [18] J. Ely, J. Hörner, W. Olszewski, Belief-free equilibria in repeated games, *Econometrica* 73 (2) (2005) 377–415.
- [19] R. Ericson, A. Pakes, Markov-perfect industry dynamics: A framework for empirical work, *Rev. Econ. Stud.* 62 (1) (1995) 53–82.
- [20] A. Fernandes, C. Phelan, A recursive formulation for repeated agency with history dependence, *J. Econ. Theory* 91 (2) (2000) 223–247.
- [21] D. Fudenberg, D. Kreps, E. Maskin, Repeated games with long-run and short-run players, *Rev. Econ. Stud.* 57 (4) (1990) 555–573.

- [22] D. Fudenberg, D. Levine, Repeated games with frequent signals, *Quart. J. Econ.* 124 (2009) 233–265.
- [23] J. Hörner, W. Olszewski, How robust is the folk theorem with imperfect public monitoring?, *Quart. J. Econ.* 124 (4) (2009) 1773–1814.
- [24] K. Judd, S. Yeltekin, J. Conklin, Computing supergame equilibria, *Econometrica* 71 (4) (2003) 1239–1254.
- [25] M. Kandori, The use of information in repeated games with imperfect monitoring, *Rev. Econ. Stud.* 59 (3) (1992) 627–652.
- [26] Q. Liu, A. Skrzypacz, Limited records and reputation, Discussion paper, Stanford GSB, 2011.
- [27] G. Mailath, S. Morris, Repeated games with almost-public monitoring, *J. Econ. Theory* 102 (1) (2002) 189–228.
- [28] G. Mailath, W. Olszewski, Folk theorems with bounded recall under (almost) perfect monitoring, *Games Econ. Behav.* 71 (1) (2010) 174–192.
- [29] G. Mailath, L. Samuelson, *Repeated Games and Reputations*, Oxford University Press, Oxford, 2006.
- [30] C. Phelan, A. Skrzypacz, Beliefs and private monitoring, Working paper, University of Minnesota, 2009.
- [31] C. Phelan, E. Stacchetti, Sequential equilibria in a Ramsey tax model, *Econometrica* 69 (6) (2001) 1491–1518.
- [32] J. Renault, M. Scarsini, T. Tomala, A minority game with bounded recall, *Math. Oper. Res.* 32 (4) (2007) 873–889.
- [33] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
- [34] C. Sleet, S. Yeltekin, On the computation of value correspondences, Discussion paper, Northwestern University, 2003.
- [35] N. Stokey, R.E. Lucas with E.C. Prescott, *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge, 1989.
- [36] C. Wang, Dynamic insurance with private information and balanced budgets, *Rev. Econ. Stud.* 62 (4) (1995) 577–595.