Impermanent types and permanent reputations [☆]

Mehmet Ekmekci [a,*], Olivier Gossner [b,c], Andrea Wilson [d]

[a] *MEDS, Kellogg School of Management, Northwestern University, United States*
[b] *Paris School of Economics, France*
[c] *London School of Economics, United Kingdom*
[d] *Department of Economics, New York University, United States*

Abstract

We study the impact of unobservable stochastic replacements for the long-run player in the classical reputation model with a long-run player and a series of short-run players. We provide explicit lower bounds on the Nash equilibrium payoffs of a long-run player, both ex-ante and following any positive probability history. Under general conditions on the convergence rates of the discount factor to one and of the rate of replacement to zero, both bounds converge to the Stackelberg payoff if the type space is sufficiently rich. These limiting conditions hold in particular if the game is played very frequently.
© 2011 Elsevier Inc. All rights reserved.

1. Introduction

The notion that commitment is valuable has long been a critical insight of non-cooperative game theory, and has deeply affected a number of social science fields, including

macroeconomics, international finance and industrial organization.[1] Existing reputation litera-ture argues that in dynamic relationships, reputation concerns can substitute for a commitment technology. A patient long-run player who faces a sequence of short-run players who believe their opponent might be committed to a particular stage-game action benefits from such a perception, as shown by [11,12].

However, [6,7] show that if the long-run player's actions are imperfectly observed by short-run players, reputation effects eventually disappear almost surely, at every equilibrium. This is particularly troubling since it shows that the original model cannot explain survival of reputation effects in environments where the agents have a long history.[2] On the other hand, the commitment possibilities of a central bank or the managers of a firm may change over time, and market beliefs about the long run player's commitment possibilities are progressively renewed. So the question is whether reputation effects are maintained perpetually if reputations are sown in the market only occasionally.

We model a long-run relationship as a repeated game between a long-run player and an infinite sequence of myopic opponents. The long-run player is either a normal type who takes actions optimally by considering the current and future consequences of his actions, or a commitment type who is committed to using a particular stage-game strategy in every interaction. The actions of the long-run player are imperfectly observed. At the beginning of every period, there is a positive probability that the long-run player is replaced by a new long-run player. The new player may also be a normal type, or he may be a commitment type. Neither the replacements nor the types of the long-run players are observed by the myopic players; hence, there is perpetual uncertainty about the long-run player's type. However, in the course of the game the myopic players receive information regarding the type of their opponent through informative signals they observe about the long-run player's actions.

Our main result is a pair of lower bounds on the Nash equilibrium payoffs of a normal type of long-run player as a function of the discount factor, the replacement rate, and the commitment type probability. The first bound is an ex-ante bound that is calculated at the beginning of the game. The second bound is on the long-run player's equilibrium continuation payoffs at any positive probability history on the equilibrium path.

If replacements are arbitrarily infrequent and the long-run player is arbitrarily patient, our bound on the ex-ante payoff converges to the same bound as that established in [11,12]. This shows that the introduction of infrequent replacements constitutes a small departure from the benchmark model.

When continuation payoffs are considered, replacements play both a positive and a negative role in the permanence of reputations. The negative effect is twofold. First, reputations naturally degrade and the short-run player doubts at every stage that he faces the same long-run player who played in previous stages. This makes reputation building less valuable in the long-run. Second, the long-run player anticipates that he might be replaced, and hence discounts the future more. In the extreme case where replacements occur at every stage, the long-run player doesn't care about future interactions and hence no reputation can be built. The positive effect is that, even

if the long-run player's reputation is severely damaged at some point, renewed doubt about his type in the mind of the short-run player offers the opportunity to rebuild a reputation.

We use our second bound to show that along a sequence of games with varying discount factors and replacement rates, if the discount factor goes to 1 at a faster rate than the rate at which the logarithm of the replacement rate goes to infinity,[3] then the long-run player receives his highest possible commitment payoff after *every equilibrium history* of the repeated game. This shows that for a range of replacement rates (as a function of the discount factor), player 1 benefits from the positive effect of replacements without suffering significantly from the negative effects.

This result has a particularly natural interpretation in the study of frequently repeated games. Increasing the discount factor is sometimes interpreted as increasing the frequency with which a stage game is repeated.[4] The conditions that our result requires are satisfied if the replacement events follow a Poisson distribution in real time with a constant hazard rate, and if the game is played in stages that become arbitrarily more frequent.[5] No matter how rarely or frequently replacements occur in real time, they restore the persistency of reputation effects in frequently repeated games.

To derive our bounds, we calculate the expected discounted average of the one-period prediction errors of the short-run players, where the expectation is taken using the probability distribution function that is generated by conditioning on (i) player 1's type being the commitment type at the beginning, and (ii) his type not changing. This idea is similar to the one introduced by [11,12]. However, the active supermartingale approach used in their work is not naturally adapted in our model since the process that governs the beliefs of the short-run players has an extra drift due to replacements.

In our model, the probability that player 1 is a particular commitment type at every period is zero. In particular, there is no "grain of truth" in the fact that player 1 is of that particular commitment type at every stage. The "grain of truth" allows one to apply merging techniques such as [3] to models in which players have initial uncertainty about the behavior of other players, and to obtain the conclusion that players eventually predict the behavior of other players accurately (see, for example, [17,13,26]). It plays a central role in reputation models like those in [1,11,12].

We rely on an information theoretic tool called relative entropy (see [5], for an excellent introduction to the topic) to measure signal-prediction errors of player 2 more precisely, thus generalizing the approach in [14]. This allows us to measure the positive effects versus negative effects of replacements on reputation building. In [15], relative entropy is used to conveniently derive merging results.

---

[3] This is a weak condition. It is satisfied, for instance, if the discount factor goes to one at a speed which is any positive exponent of the replacement rate.

[4] For instance, [9] studies reputation games in which the time between the two stages of the game get closer to zero. In his model the information structure of the stage game depends on the time increment, whereas in our model the information structure is fixed. The literature on repeated games in continuous time, such as [10], is also motivated by interpreting the continuous-time game as the limit of discrete-time games where the stage game is played very frequently.

[5] If the period length between two periods, Δ, becomes small, then the effective discount factor between the two adjacent periods becomes very close to one (i.e., $\delta(\Delta) = e^{-r\Delta}$ for some r). If the replacement events follow a Poisson distribution with a hazard rate ρ, then the probability of a replacement event between two periods becomes almost $\rho\Delta$. As the period length Δ vanishes, the impatience rate (i.e., $1 - \delta = 1 - e^{-r\Delta}$) vanishes at a rate proportional to Δ.

2. Review of the literature

Reputation models were first introduced in the context of finitely repeated games in [18] and [23]. In infinitely repeated games, [11,12] show that, under very weak assumptions on the monitoring technology, in any Nash equilibrium an arbitrarily patient long-run player obtains a payoff that is at least as much as the payoff he could get by publicly committing to playing any of the commitment type strategies. On the other hand, [6,7,29] show that for a class of stage games, all reputation effects are temporary if the actions of the long-run player are imperfectly monitored.[6]

In the benchmark model in [11,12], the long-run player's type is fixed once and for all at the beginning of the repeated game.[7] Our model is a variant in which the types of the long-run player are impermanent.

Previous papers, such as [16,4,21,24,27,28] have already shown that reputations can be permanent if long-run types are impermanent. Most of these papers focus on a particular equilibrium or a class of equilibria with interesting equilibrium dynamics.

Other interesting variations of the benchmark model which allow for permanent reputation effects include the study [19] of a model where short-run players must pay a cost to observe past signals, the study [2] of reputation effects in markets, the analysis [20] of reputation dynamics when short-run players have bounded recall and the study [8] of the sustainability of reputation effects when the short run players observe a summary statistic about the history of the play, as in online markets.

Our paper offers a systematic analysis of reputation effects in terms of payoff bounds in repeated games with replacements. In particular, we obtain explicit bounds on equilibrium payoffs without making assumptions about the stage game or the monitoring structure, for any value of the replacement rate and of the long-run player's discount factor. The long-run player does not need to know how good or bad his reputation is.

3. Model

There is an infinite sequence of long-run and short-run players. At every period $t \in \mathbb{N}$, a long-run player (player 1) plays a fixed, finite stage game G with a short-run player (player 2). The set of actions available to player i in G is A_i. Given any finite set X, $\Delta(X)$ represents the set of probability distributions over X, so the set of mixed stage-game strategies for player i is $S_i := \Delta(A_i)$.

The set of types available to player 1 is $\Omega = \{\tilde{\omega}\} \cup \hat{\Omega}$. The type $\tilde{\omega}$ is called player 1's normal type. The remaining types belong to $\hat{\Omega} \subseteq S_1$ which is a finite or countable set of *simple commitment types*. Each type $\hat{\omega} \in \hat{\Omega}$ is committed to playing the corresponding strategy $\hat{\omega} \in S_1$ at each stage of the interaction. The initial distribution over player 1's types is $\mu \in \Delta(\Omega)$.

A short-run player lives only one period. The life span of each long-run player is stochastic. The first long-run player starts the game at time $t = 1$, and plays the stage game against a new

---

[6] In particular, they show that, for fixed parameters of the game, there exists a period T such that the long-run player's equilibrium-continuation payoff starting at T is close to some equilibrium payoff of the complete-information game, with probability close to one. Moreover, the equilibrium play of the game after period T resembles some equilibrium of the complete-information repeated game. Therefore, all reputation effects (both payoff and behavior effects) eventually disappear.

[7] For an excellent presentation of reputation models, we refer the reader to part 4 of the book [22].

short-run player every period until he is replaced by a new long-run player. Once a player is replaced, he does not re-enter the game. The index $i_t \in \mathbb{N}$ indicates the identity of player 1 playing at stage t, and ω_t is the type of player 1 at stage t. The first player 1 has identity $i_1 = 1$. Between any two stages $t-1$ and t, with probability ρ, player 1 is replaced (i.e., $i_t = i_{t-1} + 1$), in which case a new type ω_t is drawn according to μ and independent of the past play. With the remaining probability $1 - \rho$, player 1's identity (and hence his type) continues from stage $t-1$ to t, $i_t = i_{t-1}$, and $\omega_t = \omega_{t-1}$. There is thus a succession of players 1, and when it is necessary to make a distinction between them, we refer to player $(1, i)$ as the i-th instance of player 1. Player $(1, i)$'s lifespan is the set of stages t such that $i_t = i$.[8]

Actions in the stage game are imperfectly observed. Player i's set of signals is a finite set Z_i. When an action profile $a = (a_1, a_2) \in A_1 \times A_2$ is chosen in the stage game, the profile of private signals $z = (z_1, z_2) \in Z := Z_1 \times Z_2$ is drawn according to the distribution $q(\cdot|a) \in \Delta(Z)$. Each player i is privately informed of the component z_i of z. For $\alpha = (\alpha_1, \alpha_2) \in S_1 \times S_2$, we let $q(\cdot|\alpha) = \mathbf{E}_{\alpha_1, \alpha_2} q(\cdot|a)$. Particular cases of the model include public monitoring (where $z_1 = z_2$ almost surely for every a), and perfect monitoring (where $z_1 = z_2 = (a_1, a_2)$ with probability 1).

3.1. Histories and strategies

The stage-game payoff functions are u_1 for player 1's normal type, and u_2 for player 2, where $u_i : A_1 \times A_2 \to \mathbb{R}$.

The set of private histories prior to stage t for player 1 is $H_{1,t} = (\mathbb{N} \times \Omega \times Z_1)^{t-1}$, with $H_{1,1} = \{\emptyset\}$. Such a private history contains the identities of player 1's, their types and their signals[9] up to stage $t-1$. A generic element of this set is $h_{1,t} = (i_1, \omega_1, z_{1,1}, \ldots, i_{t-1}, \omega_{t-1}, z_{1,t-1})$. In addition to $h_{1,t}$, player 1's behavior at stage t is allowed to depend on his identity and type at stage t. A strategy for player 1 describes the behavior of all instances of player 1, i.e., it is a mapping

$$\sigma_1 : \bigcup_{t \geqslant 1} H_{1,t} \times \mathbb{N} \times \Omega \to S_1$$

with the restriction that $\sigma_1(h_{1,t}, i_t, \omega_t) = \omega_t$ whenever $\omega_t \in \hat{\Omega}$, since commitment types are required to play the corresponding strategy in the stage game. The set of all strategies for player 1 is denoted as Σ_1.

The set of private histories prior to stage t for player 2 is $H_{2,t} = Z_2^{t-1}$, with $H_{2,1} = \{\emptyset\}$. A strategy for player 2 at stage t is

$$\sigma_{2,t} : H_{2,t} \to S_2.$$

We let $\Sigma_{2,t}$ be the set of all such strategies, and $\Sigma_2 := \prod_{t \geqslant 1} \Sigma_{2,t}$ denotes the set of all sequences of such strategies.

A history h_t at stage t is an element of $(\mathbb{N} \times \Omega \times Z_1 \times Z_2)^{t-1}$ which describes all signals observed by both players, as well as the types, and identities of player 1 up to stage t. Since a history h_t corresponds to a pair of private histories $h_{1,t}$ and $h_{2,t}$ for players 1 and 2, respectively,

---

[8]   Another way to think of our model is to assume that there is a pool of inactive long-run players. At every period only one player is active. The active player continues to play until he is replaced by a new player from the pool. Once an active player is replaced, he never plays the game again.

[9]   We do not need to assume that histories reveal past actions, in particular, a player 1 could be ignorant of the actions of past instances of player 1

we write $h_t = (h_{1,t}, h_{2,t})$. Given two histories $h_t \in H_t$ and $h_{t'} \in H'_t$, we let $h_t \cdot h_{t'} \in H_{t+t'}$ be the history obtained by the concatenation of h_t and $h_{t'}$, and we use similar notations for the concatenation of two private histories of either player 1 or 2.

A strategy profile $\sigma = (\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2$ induces a probability distribution P_σ over the set of infinite histories $(\mathbb{N} \times \Omega \times A_1 \times A_2 \times Z_1 \times Z_2)^\infty$. A player 1, when evaluating the stream of future outcomes, cares only about the payoffs at the periods when he is active. This is equivalent to assuming that an inactive player receives a payoff of zero. The intertemporal discount factor of player 1 is $0 < \delta_0 < 1$. After history $h_{1,t}$ with $P_\sigma(h_{1,t}) > 0$, the expected discounted continuation payoff from strategy profile σ for an active player 1 at stage $t - 1$ of a normal type is

$$U_{1,\sigma}[h_{1,t}] = \sum_{\tau=1}^{\infty} \delta_0^{\tau-1} \mathbf{E}_{P_\sigma(h_t \cdot h_{\tau+1}|h_{1,t})} \mathbf{1}_{i_{t+\tau-1}=i_{t-1}} u_1(a_\tau)$$

$$= \sum_{\tau=1}^{\infty} \delta_0^{\tau-1} \mathbf{E}_{P_\sigma(h_{1,t} \cdot h_{\tau+1}, i_{t+\tau-1}=i_{t-1}|h_{1,t})} u_1(a_\tau)$$

where $\mathbf{1}_{i_{t+\tau}=i_{t-1}}$ is the indicator function, which equals 1 when player 1's identity between stages $t - 1$ and $t + \tau - 1$ remains the same. We have found, however, that it is more convenient to express player 1's payoff using the probability distribution that is conditional on his not being replaced. Let $\hat{P}_\sigma$ be the probability distribution on future histories conditional on history $h_{1,t} \in H_{1,t}$ and conditional on player 1's identity at stage $t + \tau - 1$ being the same as $t - 1$ (i.e., $i_{t+\tau-1} = i_{t-1}$). This probability distribution function is given by

$$\hat{P}_\sigma[h_{1,t}](h_{\tau+1}) = P_\sigma(h_t \cdot h_{\tau+1}|h_{1,t}, i_{t+\tau-1}=i_{t-1}) = \frac{P_\sigma(h_t \cdot h_{\tau+1}, i_{t+\tau-1} = i_{t-1}|h_{1,t})}{(1-\rho)^\tau}$$

Then, by defining δ as $\delta_0(1 - \rho)$ we express player 1's payoffs as follows:

$$U_{1,\sigma}[h_{1,t}] = \sum_{\tau=1}^{\infty} \delta_0^{\tau-1}(1-\rho)^\tau \mathbf{E}_{\hat{P}_\sigma[h_{1,t}](h_{\tau+1})} u_1(a_\tau)$$

$$= (1-\rho) \sum_{\tau=1}^{\infty} \delta^{\tau-1} \mathbf{E}_{\hat{P}_\sigma[h_{1,t}](h_{\tau+1})} u_1(a_\tau).$$

The *normalized* discounted continuation payoff to player 1 at history $h_{1,t}$ is thus

$$\pi_{1,\sigma}[h_{1,t}] = (1-\delta) \sum_{\tau=1}^{\infty} \delta^{\tau-1} \mathbf{E}_{\hat{P}_\sigma[h_{1,t}](h_{\tau+1})} u_1(a_\tau).$$

We also let $\hat{P}_\sigma[h_t]$ and $\hat{P}_\sigma[h_{2,t}]$ be the probabilities over histories h_τ following h_t and $h_{2,t}$ respectively, conditional on player 1 not being replaced between stages $t - 1$ and $t + \tau - 1$. Similar expressions are used for the expected payoffs $\pi_{1,\sigma}[h_{2,t}]$ following $h_{2,t} \in H_{2,t}$.

3.2. Equilibrium

Nash equilibria are defined as follows:

Definition 1. A Nash equilibrium is a strategy profile $(\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2$ such that:

1. For every $h_{1,t}$ such that $P_\sigma(h_{1,t}) > 0$, the strategy σ_1 maximizes $\pi_{1,(\sigma_1',\sigma_2)}[h_{1,t}]$ over all $\sigma_1' \in \Sigma_1$.
2. For every $h_{2,t}$ with $P_\sigma(h_{2,t}) > 0$, the strategy $\sigma_{2,t}$ maximizes $\pi_{2,(\sigma_1,\sigma_2')}[h_{2,t}]$ over all $\sigma_2' \in \Sigma_2$.

Condition (1) requires that at any history $h_{1,t}$, the strategy σ_1 maximizes the continuation payoff of the player 1 who is active at that history. Nash equilibria require that what σ_1 prescribes for a player who becomes active at history $h_{1,t}$ is optimal for that player. In a model without replacements, requiring σ_1 to be optimal at period zero is enough, but not in our model.

Condition (2) requires that every player 2 plays a myopic best-response to her expectations of the opponent's current-period play.

3.3. ε-entropy confirming best responses

The relative entropy between two probability measures P and Q defined over the same finite set X is:

$$d(P\|Q) = \mathbf{E}_P \ln \frac{P(x)}{Q(x)} = \sum_{x\in X} P(x) \ln \frac{P(x)}{Q(x)},$$

with the convention that $0 \ln 0 = 0$. The relative entropy between P and Q is always nonnegative, it is finite if and only if the support of Q contains that of P, and it equals zero if and only if $P = Q$.

Assume that player 1's strategy in the stage game is α_1, and player 2 plays a best-response α_2 to her belief α_1' about player 1's strategy. Under (α_1, α_2), the distribution of signals to player 2 is the marginal $q^2(\cdot|\alpha_1, \alpha_2)$ on Z^2 of $q(\cdot|\alpha_1, \alpha_2)$, while it is $q^2(\cdot|\alpha_1', \alpha_2)$ under (α_1', α_2). Thus, player 2's signal-prediction error, measured by the relative entropy, is $d(q^2(\cdot|\alpha_1, \alpha_2)\|q^2(\cdot|\alpha_1', \alpha_2))$.

We say that α_2 belongs to the set $B_\varepsilon(\alpha_1)$ of ε-entropy confirming best responses to $\alpha_1 \in S_1$ whenever this prediction error is bounded by ε.[10] More precisely, $\alpha_2 \in B_\varepsilon(\alpha_1)$ when there exists α_1' such that:

- α_2 is a best response to α_1',
- $d(q^2(\cdot|\alpha_1, \alpha_2)\|q^2(\cdot|\alpha_1', \alpha_2)) \leqslant \varepsilon$.

When player 1's strategy is α_1 and player 2 plays an ε-entropy confirming best response to it, the payoff to player 1 is bounded below by the quantity $\underline{v}_{\alpha_1}(\varepsilon) = \inf_{\alpha_2 \in B_\varepsilon(\alpha_1)} u_1(\alpha_1, \alpha_2)$. We define the *lower Stackelberg payoff* as $\underline{v}^* = \sup_{\alpha_1} \underline{v}_{\alpha_1}^*(0)$. This is the smallest payoff player 1 obtains when (i) he chooses α_1, and (ii) player 2 plays a best response while accurately predicting the distribution of her own signals. The supremum of all convex functions below $\underline{v}_{\alpha_1}$ is denoted as $\tilde{v}_{\alpha_1}$.

---

[10] It is a consequence of Pinsker's inequality (see [25], or Lemma 12.6.1 in [5]) that every ε-entropy confirming best response that is not weakly dominated is an $\sqrt{\varepsilon/2}$-confirming best response in the sense of [12].

4. Main result

We define the infimum over all Nash equilibria of all continuation payoffs of player 1 at any history of the repeated game that is on the equilibrium path, as follows:

$$v(\mu, \delta, \rho) = \inf \left\{ \pi_{1,\sigma}[h_1] \text{ s.t. } h_1 \in \bigcup_t H_{1,t}, \sigma \text{ is a Nash equilibrium and } P_\sigma(h_1) > 0 \right\}.$$

We also consider the infimum of all Nash equilibrium payoffs of player 1 at the start of the game, following the initial history $h_{1,1} = \emptyset$.

$$v_1(\mu, \delta, \rho) = \inf \left\{ \pi_{1,\sigma}[\emptyset] \text{ s.t. } \sigma \text{ is a Nash equilibrium} \right\}.$$

Clearly, $v_1(\mu, \delta, \rho) \geqslant v(\mu, \delta, \rho)$. Our main Theorem below offers bounds on both ex-ante and continuation equilibrium payoffs to player 1 as a function of the distribution of commitment types, the discount factor, and the replacement rate.

Theorem 1. *For every value of the parameters $\mu(\hat{\omega}) > 0$, $\delta < 1$, and $\rho < 1$, it follows that*:

(Ex-Ante): $v_1(\mu, \delta, \rho) \geqslant \sup_{\hat{\omega} \in \hat{\Omega}} \tilde{v}_{\hat{\omega}} \left(-(1 - \delta) \ln \mu(\hat{\omega}) - \ln(1 - \rho) \right).$

(Continuation payoffs): $v(\mu, \delta, \rho) \geqslant \sup_{\hat{\omega} \in \hat{\Omega}} \tilde{v}_{\hat{\omega}} \left(-(1 - \delta) \ln \rho \mu(\hat{\omega}) - \ln(1 - \rho) \right).$

In order to understand the meaning of the bounds of the theorem, it is worthwhile considering their implications when player 1 is arbitrarily patient and replacements are arbitrarily rare. In this case we obtain:

Corollary 1. *Along a sequence of games* $\{G_n^\infty(\mu, \delta_n, \rho_n)\}_{n=1}^\infty$, *such that $\rho_n \to 0$ and $\delta_n \to 1$,*

1. $\liminf_n v_1(\mu, \delta_n, \rho_n) \geqslant \sup_{\hat{\omega} \in \hat{\Omega}} \tilde{v}_{\hat{\omega}}(0).$
2. *If $(1 - \delta_n) \ln \rho_n \to 0$, then $\liminf_n v(\mu, \delta_n, \rho_n) \geqslant \sup_{\hat{\omega} \in \hat{\Omega}} \tilde{v}_{\hat{\omega}}(0).$*
3. *If $(1 - \delta_n) \ln \rho_n \to 0$ and Ω is dense in S_1, then $\liminf_n v(\mu, \delta_n, \rho_n) \geqslant \underline{v}^*.$*

Part (1) of the corollary provides a robustness check for the results in [11,12] when replacements are allowed. If replacements occur infrequently and the long-run player is sufficiently patient, then all of player 1's equilibrium payoffs are bounded below by what can be obtained from his favorite commitment strategy, given that player 2 accurately predicts her own signals.

Part (2) of the corollary says that if the replacements are infrequent, but they disappear at a rate such that $(1 - \delta_n) \ln \rho_n \to 0$, then the lower bound on equilibrium payoffs conditional on any history of player 1 converges to the same lower bound as the ex-ante payoffs.[11] This result shows that if replacements are infrequent, but not too infrequent compared to the discount rate, player 1 enjoys reputation benefits after any history of the game on the equilibrium path. Sufficiently likely replacements are needed to restore reputations after histories when those reputations have been badly damaged, but overly frequent replacements damage reputation. Note that the condition is rather weak because it is satisfied whenever $\rho_n = (1 - \delta_n)^\beta$, for some $\beta > 0$.

---

[11] Note that since $\delta_n = \delta_{0,n}(1 - \rho_n)$, it follows that $(1 - \delta_n) \ln \rho_n \to 0$ if and only if $(1 - \delta_{0,n}) \ln \rho_n \to 0$.

Thus, part (2) establishes the permanency of reputation effects under fairly general conditions. This is in sharp contrast to the results in [6,7] showing that all reputation effects eventually vanish if replacements never occur.

Part (3) of the corollary concludes that if the support of the type distribution is sufficiently rich, then continuation payoffs of player 1 are bounded below his lowest Stackelberg payoff across all Nash equilibria.

A natural economic environment in which the condition on the relative speed of convergence rates assumed in (2) is satisfied is a repeated game played very frequently. Consider a game played in periods $t \in \{1, 2, \ldots\}$ where the time between two adjacent periods is $\Delta > 0$. Suppose the long-run player is replaced according to a Poisson probability distribution with parameter ρ over the real time, and his instantaneous time preference is $r > 0$. Then his effective discount factor between two periods is $\delta(\Delta) = \exp(-(r + \rho)\Delta)$ and the replacement probability at any period is $\rho\Delta$. As the time Δ between periods approaches zero, the replacement probability $\rho\Delta$ approaches zero, the discount factor $\delta(\Delta)$ approaches one, and therefore $(1 - \delta(\Delta)) \ln \rho\Delta$ approaches zero. Therefore part (2) of the corollary applies.

Example. To illustrate the bounds from our main result, we consider a perfect monitoring version of the quality game. Player 1 chooses to produce a good of high (H) or low (L) quality, and player 2 decides whether to buy the good (b) or not (n). Payoffs are given by the matrix:

	n	b
H	0, 0	1, 1
L	0, 0	3, −1

The quality game

We identify player 1's strategies with the probability they assign to H, and consider a commitment type $\hat{\omega} > \frac{1}{2}$ such that $\mu(\hat{\omega}) > 0$. Let $d^* = d(\hat{\omega}\|\frac{1}{2}) = \ln 2 + \hat{\omega}\ln(\hat{\omega}) + (1 - \hat{\omega})\ln(1 - \hat{\omega}) > 0$. There is a unique best-response to any strategy α'_1 of player 1 if $\alpha'_1 > \frac{1}{2}$ and this best response is b. Therefore, $B_{\hat{\omega}}(d) = \{b\}$ for every $d < d^*$. We have $\underline{v}_{\hat{\omega}}(0) = 3 - 2\hat{\omega}$, and for every d,

$$\tilde{v}_{\hat{\omega}}(d) = (3 - 2\hat{\omega}) \max\left\{1 - \frac{d}{d^*}, 0\right\}.$$

We obtain the bounds on ex-ante and continuation payoffs:

$$v(\mu, \delta, \rho) \geqslant \left(1 + \frac{(1 - \delta)\ln \mu(\hat{\omega}) + \ln(1 - \rho)}{d^*}\right)(3 - 2\hat{\omega}),$$

$$v_1(\mu, \delta, \rho) \geqslant \left(1 + \frac{(1 - \delta)\ln \rho\mu(\hat{\omega}) + \ln(1 - \rho)}{d^*}\right)(3 - 2\hat{\omega}).$$

For a numerical application, suppose that the expected time between management changes is 5 years, and the interest rate is 5% per annum. If the period length is a week, then $\delta_0 \simeq (\frac{1}{1.05})^{\frac{1}{52}}$, $\rho \simeq \frac{1}{260}$, and $\delta = \delta_0(1 - \rho) \simeq 0.9952$. For the commitment type $\hat{\omega} = 1$, $d^* = \ln 2 = 0.693$. With $\mu(\hat{\omega}) = 0.01$, we obtain the bounds for ex-ante and continuation equilibrium payoffs of approximately 0.9624 and 0.9236, respectively, which are comparable to the commitment payoff of 1.

5. Proofs

The main idea of the proofs of both parts of Theorem 1 follow the classical argument of [11, 12]. Assume that at every stage player 1 follows the strategy corresponding to some commitment type $\hat{\omega}$. The sequence of players 2 should eventually predict more and more accurately the distribution of signals induced by player 1's actions; hence, each player 2 plays a best-response to a strategy of player 1 which is "not too far" from $\hat{\omega}$. This provides a lower bound on player 1's payoff while he plays $\hat{\omega}$ repeatedly, and hence on his equilibrium payoff.

What makes matters more complicated in our model than in the models of [11,12]'s original papers is the replacement process of player 1. If we assumed that player 1 plays $\hat{\omega}$ repeatedly, then the "learning" force according to which player 2 anticipates that the distribution of signals becomes "close" to one induced by $\hat{\omega}$ is countered by the possibility that player 1 is replaced, which acts as a "drift" in player 2's belief toward the initial distribution. The way in which these effects balance each other needs to be carefully measured.

We measure prediction errors using the relative entropy distance as in [14], rather than the L^2 norm as in [11,12]. The fundamental property of relative entropies on which we rely, called the Chain Rule, allows for precise control over the process of player 2's errors in her own signals, assuming that player 1 plays $\hat{\omega}$. This property is explained in Section 5.1 just below.

The proofs of both parts of our theorem follow similar arguments. The proof of the ex-ante part of our main theorem in Section 5.2 is simpler in notation; therefore, we present it first and then present the proof of the continuation payoffs part in Section 5.3.

5.1. Chain Rule for relative entropies

Consider two abstract sets of signals X and Y, and an agent observing first $x \in X$, then $y \in Y$. The distribution of (x, y) is P, while this agent's belief on (x, y) is Q. Decompose the observation of the pair (x, y) in stages. In the first stage, the agent's error in predicting $x \in X$ is $d(P_X \| Q_X)$, where P_X and Q_X are P and Q's marginals on X, respectively. Once $x \in X$ is observed, the prediction error in $y \in Y$ is $d(P_Y(\cdot|x) \| Q_Y(\cdot|x))$, where $P_Y(\cdot|x)$ and $Q_Y(\cdot|x)$ are P and Q's conditional probabilities on Y, respectively, given x. Hence, the expected error in predicting y is $\mathbf{E}_{P_X} d(P_Y(\cdot|x) \| Q_Y(\cdot|x))$, and the total expected error in predicting x, and then y, is $d(P_X \| Q_X) + \mathbf{E}_{P_X} d(P_Y(\cdot|x) \| Q_Y(\cdot|x))$.

The Chain Rule of relative entropies (see, e.g., [5, Theorem 2.5.3]) shows that prediction errors can be counted either globally, or in stages, with the same results, as stated here:

$$d(P \| Q) = d(P_X \| Q_X) + \mathbf{E}_{P_X} d\big(P_Y(\cdot|x) \| Q_Y(\cdot|x)\big).$$

A useful implication of the Chain Rule is the following bound on the relative entropy between two distributions under some "grain of truth" assumption (see [14] for the proof).

Claim 1. *Assume $Q = \varepsilon P + (1 - \varepsilon) P'$ for some $\varepsilon > 0$ and P'; then*:

$$d(P \| Q) \leqslant -\ln \varepsilon.$$

Consider player 1 who repeatedly plays $\hat{\omega}$, either from the start of the game or after some history. We want to bound the total prediction error of player 2 in her own signals, over a sequence of n stages. Since player 1 is actually of type $\hat{\omega}$ over these n stages with positive probability, there is some "grain of truth" in the possibility that player 2's process of signals is induced by

player 1's playing $\hat{\omega}$ at each of these stages. Hence, we can apply Claim 1 in order to obtain a bound on player 2's signal-prediction errors. The Chain Rule allows us to decompose this total error as the sum of the expected errors at each stage, and hence to control for "how far" player 2 is on average from predicting accurately the distribution of her own signals. These arguments are presented in more detail in the next subsection.

5.2. Proof of the ex-ante part

The main idea of the proof is the following. We aim to bound player 1's δ-discounted expected payoff, assuming that this player plays a commitment strategy $\hat{\omega}$. Player 2 may not be best responding to $\hat{\omega}$ if she is anticipating a different behavior from player 1. Thus, a way of bounding player 1's payoff is to bound the δ-discounted expected sum of player 2's prediction errors about signals that are one stage ahead (see Proposition 3 below).

To achieve this end, we use Claim 1 and the Chain Rule in order to derive a bound on the expected arithmetic average of the prediction errors of player 2 over n periods, using the probability distribution function generated by the strategy $\hat{\omega}$ and by conditioning on the event that no replacement has occurred during the n-stages (see Proposition 1). In Proposition 2, we convert these bounds on the n-stage prediction errors into a bound on the discounted sum of prediction errors about the signals that are one stage ahead.

Fix $\hat{\omega} \in \hat{\Omega}$. Let $\sigma'_{1,\hat{\omega}}$ be the modification of a strategy σ_1 in which the first instance of player 1, if he is the normal type, plays $\hat{\omega}$ at every stage of the interaction: $\sigma'_{1,\omega}(h_{1,t}, i_t, \omega_t) = \hat{\omega}$ if $i_t = 1$, $\omega_t = \tilde{\omega}$; and otherwise $\sigma'_{1,\omega}(h_{1,t}, i_t, \omega_t) = \sigma_1(h_{1,t}, i_t, \omega_t)$. Let $\sigma' = (\sigma'_{1,\omega}, \sigma_2)$.

For $n \geqslant 1$, consider the marginal $P^{2,n}_\sigma$ of P_σ over $H_{2,n+1}$ and the probability distribution $\hat{P}^{2,n}_{\sigma'}$ over $H_{2,n+1}$, given by

$$\hat{P}^{2,n}_{\sigma'}(h_{2,n+1}) = P_{\sigma'}(h_{2,n+1} | i_n = 1, \ \omega_1 = \tilde{\omega}).$$

$\hat{P}^{2,n}_{\sigma'}$ is the relevant probability distribution over the n stage ahead play when player 1 of type $\tilde{\omega}$ considers playing $\hat{\omega}$ from the first stage on, conditional on his not being replaced.

Proposition 1. *For every $\hat{\omega}$,*

$$d\big(\hat{P}^{2,n}_{\sigma'} \| P^{2,n}_\sigma\big) \leqslant -\ln\big(\mu(\hat{\omega})\big) - n\ln(1-\rho).$$

Proof. Define A as the event: "$\omega_1 = \hat{\omega}$, $i_n = 1$". By the definitions of $P^{2,n}_\sigma$ and $\hat{P}^{2,n}_{\sigma'}$, we have $P^{2,n}_\sigma(\cdot | A) = \hat{P}^{2,n}_{\sigma'}(\cdot)$; hence,

$$P^{2,n}_\sigma(\cdot) = P^{2,n}_\sigma(A) P^{2,n}_\sigma(\cdot | A) + P^{2,n}_\sigma\big(A^c\big) P^{2,n}_\sigma\big(\cdot | A^c\big)$$
$$= P^{2,n}_\sigma(A) \hat{P}^{2,n}_{\sigma'}(\cdot) + P^{2,n}_\sigma\big(A^c\big) P^{2,n}_\sigma\big(\cdot | A^c\big).$$

Claim 1 yields

$$d\big(\hat{P}^{2,n}_{\sigma'} \| P^{2,n}_\sigma\big) \leqslant -\ln\big(P^{2,n}_\sigma(A)\big) \leqslant -\ln\big(\mu(\hat{\omega})(1-\rho)^n\big). \qquad \square$$

Proposition 1 decomposes the prediction error of player 2 into the sum of two terms. The first term corresponds to the error made by player 2 in not assuming that player 1 is of the commitment type. This error increases and goes to infinity as $\mu(\hat{\omega})$ decreases and goes to zero. This corresponds to the intuition that reputations are harder to build for commitment types that

have low probability. The second term is a measure of the error made in assuming that player 1 could have been replaced between stage 1 and n, when in fact he wasn't. This second term is linear in n, and the slope goes to 0 when the replacement rate vanishes. This second term comes from the "negative effect" of the replacements in reputation building. This is because player 2 is less likely to learn that player 1 plays the commitment type strategy if replacements are more likely.

Under P_σ, player 2's beliefs about the next stage's signals following $h_{2,t}$ are:

$$p_\sigma^2(h_{2,t})(z_{2,t}) = P_\sigma(z_{2,t}|h_{2,t}).$$

We compare $p_\sigma^2(h_{2,t})$ with player 2's belief had player 2 assumed that player 1 was of type $\hat{\omega}$, given by $q^2(\hat{\omega}, \sigma_2(h_{2,t}))$.

The expected discounted-average relative entropy between the predictions of player 2 over her next signal, when she relies either on P_σ or on player 1 being of type $\hat{\omega}$, is:

$$d_\delta^{\sigma,\hat{\omega}} = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \mathbf{E}_{\hat{P}_{\sigma'}^2} d\big(p_\sigma^2(h_{2,t}) \big\| q^2\big(\hat{\omega}, \sigma_2(h_{2,t})\big)\big).$$

Proposition 2. *For every* $\hat{\omega}$,

$$d_\delta^{\sigma,\hat{\omega}} \leqslant -(1 - \delta) \ln \mu(\hat{\omega}) - \ln(1 - \rho).$$

Proof. From the Chain Rule for relative entropies, it follows that:

$$d\big(\hat{P}_{\sigma'}^{2,n} \big\| P_\sigma^{2,n}\big) = \sum_{t=1}^{n} \mathbf{E}_{\hat{P}_{\sigma'}^2} d(p_\sigma^2(h_{2,t}) \| q^2\big(\hat{\omega}, \sigma_2(h_{2,t})\big).$$

We use the fact that for a bounded sequence $(x_n)_n$ the following identity holds:

$$\sum_{t=1}^{\infty} \delta^{t-1} x_t = (1 - \delta) \sum_{n=1}^{\infty} \delta^{n-1} \sum_{t=1}^{n} x_t.$$

Applying this identity to the sequence $x_t = \mathbf{E}_{\hat{P}_{\sigma'}^2} d(p_\sigma^2(h_{2,t}) \| q^2(\hat{\omega}, \sigma_2(h_{2,t})))$, we obtain:

$$\begin{aligned}
d_\delta^{\sigma,\hat{\omega}} &= (1 - \delta)^2 \sum_{n=1}^{\infty} \delta^{n-1} d\big(\hat{P}_{\sigma'}^{2,n} \big\| P_\sigma^{2,n}\big) \\
&\leqslant (1 - \delta)^2 \sum_{n=1}^{\infty} \delta^{n-1} \big(-\ln \mu(\hat{\omega}) - n \ln(1 - \rho)\big) \\
&= -(1 - \delta) \ln \mu(\hat{\omega}) - \ln(1 - \rho),
\end{aligned}$$

where the inequality comes from Proposition 1. $\quad\square$

Proposition 2 bounds the expected discounted error in player 2's next stage signals. This expected error is the sum of two terms, which correspond to the two terms discussed in Proposition 1. When $\delta \to 1$, the first term, which corresponds to player 2's initial error as a result of not knowing that player 1 is of the commitment type $\hat{\omega}$, vanishes, since the average prediction errors are taken over longer and longer histories. The second term, corresponding to the drift in player 2's beliefs that comes from replacements, is now constant in the discount rate since no matter how long a horizon is considered, the per-stage error remains the same.

Proposition 3. *The expected payoff to player* $(1, 1)$ *of type* $\tilde{\omega}$ *playing* $\sigma'_{1,\hat{\omega}}$ *is at least*:

$$\tilde{v}_{\hat{\omega}}\big(-(1 - \delta) \ln \mu(\hat{\omega}) - \ln(1 - \rho)\big).$$

Proof. Conditional on history $h_{2,t}$, player 1's expected payoff at stage t is bounded below by $v_{\hat{\omega}}(d(p_{\sigma}^2(h_{2,t}) \| q^2(\hat{\omega}, \sigma_2(h_{2,t}))))$. Using the convexity of $\tilde{v}_{\hat{\omega}}$, we obtain:

$$\pi_{1,\sigma'} \geqslant (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \mathbf{E}_{\hat{P}_{\sigma'}^{2,t}} \underline{v}_{\hat{\omega}}\big(d\big(p_{\sigma}^2(h_{2,t}) \big\| q^2(\hat{\omega}, \sigma_2(h_{2,t}))\big)\big)$$

$$\geqslant (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \mathbf{E}_{\hat{P}_{\sigma'}^{2,t}} \tilde{v}_{\hat{\omega}}\big(d\big(p_{\sigma}^2(h_{2,t}) \big\| q^2(\hat{\omega}, \sigma_2(h_{2,t}))\big)\big)$$

$$\geqslant \tilde{v}_{\hat{\omega}}\big((1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \mathbf{E}_{\hat{P}_{\sigma'}^{2,t}} d\big(p_{\sigma}^2(h_{2,t}) \big\| q^2(\hat{\omega}, \sigma_2(h_{2,t}))\big)\big)$$

$$= \tilde{v}_{\hat{\omega}}\big(d_{\delta}^{\sigma,\hat{\omega}}\big)$$

$$\geqslant \tilde{v}_{\hat{\omega}}\big(-(1 - \delta) \ln \mu(\hat{\omega}) - \ln(1 - \rho)\big),$$

where the last inequality comes from Proposition 2 and from the fact that $\tilde{v}_{\hat{\omega}}$ is a nonincreasing function. $\square$

This proves the ex-ante part of our main theorem.

5.3. Proof of the continuation payoffs part

Fix a Nash equilibrium σ, a history $h_{1,t}$ for player 1 such that $P_{\sigma}(h_{1,t}) > 0$, and $\hat{\omega} \in \hat{\Omega}$. Let σ'_1 be the strategy of player 1 that plays $\hat{\omega}$ at all histories after $h_{1,t}$ as long as player 1's identity is the same as at history $h_{1,t}$; at all other histories σ'_1 plays σ_1. Let $\sigma' = (\sigma'_1, \sigma_2)$, and, for $h_{2,t}$ such that $P_{\sigma}(h_{1,t}, h_{2,t}) > 0$, consider the probabilities $P_{\sigma}^{2,n}[h_{2,t}]$ and $\hat{P}_{\sigma'}^{2,n}[h_{1,t}, h_{2,t}]$ given by

$$P_{\sigma}^{2,n}[h_{2,t}](h_{2,n+1}) = P_{\sigma}(h_{2,t} \cdot h_{2,n+1} | h_{2,t})$$

and by

$$\hat{P}_{\sigma'}^{2,n}[h_{1,t}, h_{2,t}](h_{2,n+1}) = P_{\sigma'}(h_{2,t} \cdot h_{2,n+1} | h_{1,t}, h_{2,t}, i_{t+n-1} = i_{t-1}).$$

$P_{\sigma}^{2,n}[h_{2,t}]$ is the probability induced over the signals of player 2 for the n stages following $h_{2,t}$ when σ is followed, while $\hat{P}_{\sigma'}^{2,n}[h_{1,t}, h_{2,t}]$ is the probability over player 2's signals following $(h_{1,t}, h_{2,t})$, assuming that player 1 switches to $\hat{\omega}$ after $h_{1,t}$, and conditional on player 1 surviving during these n stages.

Proposition 4. *For any* $h_{2,t}$ *with* $P_{\sigma}(h_{1,t}, h_{2,t}) > 0$,

$$d\big(\hat{P}_{\sigma'}^{2,n}[h_{1,t}, h_{2,t}] \big\| P_{\sigma}^{2,n}[h_{2,t}]\big) \leqslant -\ln\big(\rho \mu(\hat{\omega})\big) - n \ln(1 - \rho).$$

Proof. Define A as the event: "player 1 is of commitment type $\hat{\omega}$ in stage $t - 1$ ($\omega_t = \hat{\omega}$), and is not replaced from stage t to $t + n$ ($i_{t+n-1} = i_{t-1}$)". We have:

$$P_\sigma^{2,n}[h_{2,t}](\cdot) = P_\sigma^{2,n}[h_{2,t}](A)\,P_\sigma^{2,n}[h_{2,t}](\cdot|A) + P_\sigma^{2,n}[h_{2,t}](A^c)\,P_\sigma^{2,n}[h_{2,t}](\cdot|A^c)$$
$$= P_\sigma^{2,n}[h_{2,t}](A)\,\hat{P}_{\sigma'}^{2,n}[h_{1,t},h_{2,t}](\cdot) + P_\sigma^{2,n}[h_{2,t}](A^c)\,P_\sigma^{2,n}[h_{2,t}](\cdot|A^c).$$

Using Claim 1, we obtain:

$$d\big(\hat{P}_{\sigma'}^{2,n}[h_{1,t},h_{2,t}] \| P_\sigma^{2,n}[h_{2,t}]\big) \leqslant -\ln\big(P_\sigma^{2,n}[h_{2,t}](A)\big) \leqslant -\ln\big(\rho\mu(\hat{\omega})(1-\rho)^n\big). \qquad \square$$

The expected discounted-average relative entropy between the predictions of player 2 over her next signal when she relies either on P_σ or on player 1 being of type $\hat{\omega}$ after history $h_{1,t}, h_{2,t}$, is:

$$d_\delta^{\sigma,\hat{\omega}}[h_{1,t},h_{2,t}] = (1-\delta)\sum_\tau \delta^{\tau-1}\mathbf{E}_{\hat{P}_{\sigma'}^2}\,d\big(p_\sigma^2(h_{2,t}\cdot h_{2,\tau}) \| q^2\big(\hat{\omega},\sigma_2(\cdot h_{2,\tau})\big)\big).$$

Proposition 5. *For every $h_{2,t}$ with $P_\sigma(h_{1,t},h_{2,t}) > 0$,*

$$d_\delta^{\sigma,\hat{\omega}}[h_{1,t},h_{2,t}] \leqslant -(1-\delta)\ln\big(\rho\mu(\hat{\omega})\big) - \ln(1-\rho).$$

Proof. The proof relies on Proposition 4 and follows steps identical to those of the proof of Proposition 2. $\square$

It is interesting to compare Proposition 5, which applies to continuation equilibrium payoffs, with Proposition 2 for the ex-ante equilibrium payoffs. In both cases, the second term, which corresponds to the negative effects of replacements on reputations, is the same. This is due to the fact that this term arises from the uncertainty on replacements of player 1, which is the same in both cases. The first term is linear in δ in both cases; it depends on ρ in Proposition 5, while it is independent of ρ in Proposition 2. In the bound for continuation payoffs (Proposition 5), this first term corresponds to the positive effect of replacements in reputations. If there are no replacements, as in [11,12], there may be histories after which player 2 knows for sure that player 1 is of the normal type, after which it is impossible to restore a reputation. On the other hand, replacements cast a permanent doubt in player 2's mind as to whether player 1 is of the commitment type, which may allow reputation effects to be restored after any history. The higher the replacement rate is, the easier it is to restart a reputation, and the lower this first term is.

Proposition 6. *The expected continuation payoff after history $h_{1,t}$ to player 1 of type $\tilde{\omega}$ playing σ_1' is at least*

$$\tilde{v}_{\hat{\omega}}\big(-(1-\delta)\ln\rho\mu(\hat{\omega}) - \ln(1-\rho)\big).$$

Proof. This follows from Proposition 5, when the same sketch is used as in the proof of Proposition 3. $\square$

This proves the continuation payoffs part of the main theorem.

6. Concluding comments

Although the idea that impermanent types may restore reputation effects permanently is not entirely new, our paper is the first to show that this is true without imposing any assumptions on the stage game or without restricting the class of equilibrium strategies. Our main Theorem provides bounds on the equilibrium payoffs of the long-run player that hold uniformly after any

history on the equilibrium path. We now briefly discuss upper bounds on equilibrium payoffs, continuation payoffs after histories outside of the equilibrium path, and several extensions.

6.1. Upper bounds

Theorem 1 provides lower bounds on equilibrium payoffs. The techniques developed in this paper allow us to derive upper bounds as well.

The supremum over all Nash equilibria of all continuation payoffs of player 1 at any history of the repeated game that is on the equilibrium path is:

$$V(\mu, \delta, \rho) = \sup\left\{\pi_{1,\sigma}[h_1] \text{ s.t. } h_1 \in \bigcup_t H_{1,t}, \sigma \text{ is a Nash equilibrium and } P_\sigma(h_1) > 0\right\},$$

and the supremum of all Nash equilibrium payoffs of player 1 at the start of the game is:

$$V_1(\mu, \delta, \rho) = \sup\left\{\pi_{1,\sigma}[\varnothing] \text{ s.t. } \sigma \text{ is a Nash equilibrium}\right\}.$$

The maximum payoff to player 1 if player 2 plays an ε-entropy confirming best response to player 1's strategy is:

$$\overline{V}(\varepsilon) = \max\left\{u_1(\alpha_1, \alpha_2), \ \alpha_1 \in S_1, \ \alpha_2 \in B_{\alpha_1}(\varepsilon)\right\},$$

and we let $\tilde{V}$ represent the infimum of all concave functions above $\overline{V}$.

The following result can be proven along lines similar to the proof of Theorem 1:

Theorem 2. *For every value of the parameters $\mu(\tilde{\omega}) > 0$, $\delta < 1$, $\rho < 1$:*

(Ex-Ante): $V_1(\mu, \delta, \rho) \leqslant \tilde{V}\big(-(1 - \delta) \ln \mu(\tilde{\omega}) - \ln(1 - \rho)\big)$;

(Continuation payoffs): $V(\mu, \delta, \rho) \leqslant \tilde{V}\big(-(1 - \delta) \ln \rho\mu(\tilde{\omega}) - \ln(1 - \rho)\big)$.

When $\rho \to 0$, $\delta \to 1$, and $(1 - \delta) \ln \rho \to 0$ (which is the case, for example, when the game is played in time increments that approach to zero), the bound on continuation equilibrium payoffs converges to $\tilde{V}(0)$, which coincides with the upper Stackelberg payoff $\bar{v}^* = \max\{u_1(\alpha_1, \alpha_2), \ \alpha_1 \in S_1, \ \alpha_2 \in B_{\alpha_1}(0)\}$.

As in [12] when commitment types have full support and monitoring allows for identification of player 1's actions, the upper Stackelberg payoff coincides with the Stackelberg payoff $\underline{v}^*$. For this class of games, when the frequency of play increases, both the lower and upper bounds on all continuation equilibrium payoffs to player 1 on the equilibrium path converge to the same limit.

6.2. Equilibrium payoffs outside of equilibrium path

The continuation payoff bounds of Theorems 1 and 2 hold on the equilibrium path. However, these results do not address what happens outside the equilibrium path, since the Nash equilibrium notion puts no restrictions on players' behavior after such histories. However, tight bounds still obtain uniformly after all histories if one considers appropriate refinements of Nash equilibria.

The refinement that is needed is that, after every private history $h_{2,t}$, player 2 holds a belief on player 1's types, and that these beliefs are updated using Bayes's rule whenever possible. Moreover, player 2 plays a best-response to her belief on player 1's types and each type's strategy. These requirements are stronger than a weak perfect Bayesian equilibrium, but weaker than a

sequential equilibrium. When we restrict attention to such equilibria, the continuation payoff bounds of Theorems 1 and 2 hold after every history, whether on the equilibrium path or not.

6.3. Extensions

Although our model specifies the replacement process, it is fairly straightforward to extend our main result to the following extensions of our model.

Nonstationary replacements: Our model makes the simplifying assumption that the replacement rate is fixed through time. Our approach easily generalizes to the case in which the replacement rate is time-dependent, and players may have incomplete information about it. The extension of our result to this context needs the assumption of a lower bound and an upper bound on replacement rates after any history. Such an extension can be interesting to study in games where periods of institutional stability where replacements are less likely follow periods of higher instability where they are more likely.

Nonidentically, independently distributed replacements: Similarly, our techniques easily extend to the study of cases in which the replacement process may depend on the current type of the long-run players and may be nonstationary. The only condition needed in such a context for our result to generalize is a uniform lower bound on the probability of a considered commitment type, given any past history.

References

[1] R.J. Aumann, S. Sorin, Cooperation and bounded recall, Games Econ. Behav. 1 (1989) 5–39.
[2] H. Bar-Isaac, S. Tadelis, Seller reputation, Found. Trends Microeconomics 4 (4) (2008) 273–351.
[3] D. Blackwell, L. Dubins, Merging of opinions with increasing information, Ann. Math. Statist. 33 (1962) 882–886.
[4] H. Cole, J. Dow, W.B. English, Default, settlement, and signalling: Lending resumption in a reputational model of sovereign debt, Int. Econ. Rev. 36 (2) (1995) 365–385.
[5] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley Series in Telecommunications, Wiley, 1991.
[6] M. Cripps, G. Mailath, L. Samuelson, Imperfect monitoring and impermanent reputations, Econometrica 72 (2) (2004) 407–432.
[7] M. Cripps, G. Mailath, L. Samuelson, Disappearing private reputations in long-run relationships, J. Econ. Theory 134 (2007) 287–316.
[8] M. Ekmekci, Sustainable reputations with rating systems, J. Econ. Theory 146 (2011) 479–503.
[9] E. Faingold, Building a reputation under frequent decisions, Technical report, Yale University, 2010.
[10] E. Faingold and Y. Sannikov, Reputation in continuous-time games, Econometrica (2011), doi:10.3982/ECTA7377, in press.
[11] D. Fudenberg, D.K. Levine, Reputation and equilibrium selection in games with a patient player, Econometrica 57 (1989) 759–778.
[12] D. Fudenberg, D.K. Levine, Maintaining a reputation when strategies are imperfectly observed, Rev. Econ. Stud. 59 (1992) 561–579.
[13] D. Fudenberg, D.K. Levine, The Theory of Learning in Games, MIT Press, Cambridge, MA, 1998.
[14] O. Gossner, Simple bounds on the value of a reputation, Econometrica (2011), doi:10.3982/ECTA9385, in press.
[15] O. Gossner, T. Tomala, Entropy bounds on Bayesian learning, J. Math. Econ. 44 (2008) 24–32.
[16] B. Holmström, Managerial incentive problems: A dynamic perspective, Rev. Econ. Stud. 66 (1994) 169–182.
[17] E. Kalai, E. Lehrer, Rational learning leads to Nash equilibrium, Econometrica 61 (5) (1993) 1019–1045.
[18] D.M. Kreps, R.B. Wilson, Reputation and imperfect information, J. Econ. Theory 27 (1982) 253–279.
[19] Q. Liu, Information acquisition and reputation dynamics, Review of Economic Studies Online first, 2011.
[20] Q. Liu, A. Skrzypacz, Limited records and reputation, Research Paper 2030, Stanford University, Graduate School of Business, 2009.
[21] G. Mailath, L. Samuelson, Who wants a good reputation? Rev. Econ. Stud. 68 (2) (2001) 415–441.
[22] G.J. Mailath, L. Samuelson, Repeated Games and Reputations: Long-Run Relationships, Oxford University Press, New York, NJ, 2006.

[23] P. Milgrom, J. Roberts, Predation, reputation and entry deterrence, J. Econ. Theory 27 (1982) 280–312.
[24] C. Phelan, Public trust and government betrayal, J. Econ. Theory 130 (2006) 27–43.
[25] M.S. Pinsker, Information and Information Stability of Random Variables and Processes, Holden-Day Series in Time Series Analysis, Holden Day, San Francisco, 1964.
[26] S. Sorin, Merging, reputation, and repeated games with incomplete information, Games Econ. Behav. 29 (1999) 274–308.
[27] B. Vial, Competitive Equilibrium and Reputation under Imperfect Public Monitoring, Documento de Trabajo, vol. 327, Pontificia Universidad Católica de Chile, 2008.
[28] T. Wiseman, Reputation and impermanent types, Games Econ. Behav. 62 (2008) 190–210.
[29] T. Wiseman, Reputation and exogenous private learning, J. Econ. Theory 144 (2009) 1352–1357.