

Ekmekci, Mehmet; Wilson, Andrea

Working Paper

Maintaining a permanent reputation with replacements

Discussion Paper, No. 1511

Provided in Cooperation with:

Kellogg School of Management - Center for Mathematical Studies in Economics and Management Science, Northwestern University

Suggested Citation: Ekmekci, Mehmet; Wilson, Andrea (2010) : Maintaining a permanent reputation with replacements, Discussion Paper, No. 1511, Northwestern University, Kellogg School of Management, Center for Mathematical Studies in Economics and Management Science, Evanston, IL

This Version is available at:

<https://hdl.handle.net/10419/59635>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Maintaining a Permanent Reputation with Replacements*

Mehmet Ekmekci[†], Andrea Wilson[‡]

Northwestern University Kellogg / Meds

NYU Department of Economics

April 2010

First version, November 2008

Abstract

We study the impact of unobservable replacements on the sustainability of reputation effects in frequently repeated games played by a long run player facing a sequence of short run players. At the beginning of every period the long-run player is replaced with a new long run player with probability $(1 - \lambda) \in (0, 1)$. The new long run player is either a commitment type who plays the same strategy in every period when he is in the game, or a normal type. The long run player's choice of stage game strategy is imperfectly observed by the short run players. We show that the long run player's payoff, in any Nash equilibrium, is bounded below by what he could get by committing to his most favorite commitment type strategy after every history of the game, even as his rate of impatience $(1 - \delta)$ vanishes at the same rate as his replacement probability. Hence arbitrarily infrequent replacements are sufficient to prevent reputations and their effects from disappearing when the stage game is played frequently enough.

*We thank Umberto Garfagnini and Nuh Aygun Dalkiran for excellent research assistance. We also thank Alp Atakan for very helpful comments.

[†]Kellogg Meds, Northwestern University, e-mail:m-ekmekci@kellogg.northwestern.edu

[‡]Department of Economics, NYU, e-mail:andrea.wilson@nyu.edu

Introduction

Consider a dynamic relationship between a large firm and a large pool of myopic buyers. At every period, the firm decides how much to invest for a higher quality production technology, and the myopic buyers make a purchase decision before they observe the quality level of the product. The quality of the product is stochastically determined by the production technology, higher investments produce higher quality products more frequently. The buyers would like to incentivize the firm to invest, but since they are myopic, they can't coordinate their behavior in later periods to punish the firm for low quality products. On the other side, the firm may suffer from the lack of efficient punishments, since then he loses his commitment to investment, and the whole trade possibilities may be destroyed.¹

In credit markets, long lived borrowers periodically seek to borrow money. They choose whether to invest in a high risk project or a low risk project. The only information publicly observed is the outcome of the project. If the credit market is sufficiently large, then the borrower at every period may interact with a different lender, hence wriggling out of a potential punishment from the same lender in the future interactions. The ability to escape punishments may harm the borrower, since then the moral hazard problem inherent in the absence of dynamic interactions appears again.²

In both of our opening examples, the long run player would like to commit to a variety of stage game strategies in the one shot game. The dynamic nature of the problem would resolve the commitment problem, even under imperfect monitoring, if the opponents were long lived too, since then efficient punishment schemes could be sustained in equilibrium. However if the opponents are short lived, or if the opponents are a continuum of infinitesimal long-lived players and only their aggregate play is observed publicly, then punishments may be inefficient and the moral hazard problem and commitment issues arise again.

In this paper we consider repeated games played with a long run player and an infinite sequence of myopic opponents. The long run player is either a normal type who takes actions optimally considering the current and future consequences of his actions, or a commitment type who is committed to playing a particular stage game action at every period he is interacting with an opponent. The actions of the long run player are imperfectly and publicly observed. At the beginning of every period the long run player may be replaced by a new long run player, possibly with a small probability. The new player may also be a normal type, or a commitment type. Replacement corresponds to a management change in the firm, or a change of the economic situations that change the commitment possibilities of the borrower. In our model, neither replacements nor the type of the long run player are observed by the myopic players, hence there is perpetual uncertainty on the long run player's type. However in equilibrium the myopic players learn about the type of their opponent from the public signals.

¹This is the classic Klein and Leffler (1981) model.

²See Fudenberg and Levine (1994), Celentani and Pesendorfer (1996) and section 3.6 of Mailath Samuelson (2007) .

Our main result shows that, along a sequence of games with varying discount factors and replacement rates, if the replacement rates are vanishing, but at a rate not faster than that by which the discount factor is approaching to one, then the long run player receives his highest possible commitment payoff after every history of the repeated game.

Our method of proof delivers a new insight in the study of repeated games and reputations as a by product. We show that for every small replacement rate, there is a T such that for every block of T periods, the long run player can guarantee that the ratio of the periods at which his payoffs fall short of his commitment payoff is small. This ratio decreases (and eventually vanishes to zero) as the replacement rate vanishes. Hence, if replacements are non negligible with respect to impatience, yet still very small, then an arbitrarily patient long run player gets his commitment payoff after every history of the repeated game.

Our result fits very nicely in applications where the same stage game is played very frequently. Sometimes the exercise of increasing the discount factor to one corresponds to increasing the frequency with which the stage game is repeated. 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). Now consider a situation where the replacement events follow a Poisson distribution with a hazard rate ρ . Then the expected time between two replacement events is $1/\rho$. In this case, as the period length Δ becomes small, the probability of a replacement event between two periods also becomes very small (to be exact, $\rho\Delta$). As Δ vanishes, the impatience rate (i.e., $1 - \delta = 1 - e^{-r\Delta}$) vanishes at a rate proportional to Δ . Therefore the conditions that our theorem require are satisfied. Hence our result says that no matter how infrequent replacements are (or how small ρ is), if the stage game is played frequently enough, then the long run player receives his commitment payoff after every history of these repeated games.

We view replacements as a perturbation of the complete information game. Fudenberg and Levine (1989) and (1992) (FL from hereon) studied a similar problem. Unlike FL's model where the long run player's identity is fixed once and for all at the beginning of the game, we assume that at the beginning of every period, the long run player is replaced by a new long run player with some small probability, $1 - \lambda$. If the long run player is replaced, the new player may be one of several commitment types, or a normal type. The replacement event is not observed by the short run players.

There are couple of subtleties we solve when we prove our result. The first one is that, we know from FL's result that for any initial prior μ , there is a critical discount factor such that for discount factors higher than this critical level, long run player's equilibrium payoffs are not lower than his commitment payoffs. However the result is silent on how fast this critical number should increase if the prior probability of the commitment types gets smaller. In our model, the replacement probability (and hence the lowest possible reputation level after all possible histories) is allowed to vanish at the same rate or at a slower rate than the rate at which the discount factor goes to one. Hence, we extend FL's result in a model with replacements and where the prior uncertainty

on types vanishes as the discount factor goes to one.

Second, replacement possibility may have an adverse effect on the long run player's ability to build a reputation. When long run player almost convinces his opponent that he is the Stackelberg type, his reputation for being this type decreases next period since he might be replaced by a normal type, or some other commitment type that he wouldn't like to mimic. We show that the adverse effect of replacements disappears rapidly as the replacement probability vanishes.

We elaborate on both of these issues, and why the replacement rate is not allowed to vanish much faster than the impatience rate in the section where we sketch the argument for a perfect monitoring stage game.³

Next we discuss related literature. After the introduction, we present our model, and then we present the main result of the paper. Later we provide a sketch of the proof with a perfect monitoring stage game example, and discuss how perfect monitoring case is extended to stage games with imperfect monitoring. We discuss some of the assumptions of our model, and present a complete proof in the Appendix.

Discussion of Related Literature The first papers that introduced the idea that adverse selection approach to repeated games may create reputation building dynamics are Kreps, Milgrom, Roberts, and Wilson (1982), Kreps and Wilson (1982) and Milgrom and Roberts (1982). Their results show that reputation effects may explain cooperation in early rounds of the finitely repeated prisoners' dilemma and entry-deterrence in early rounds of the chain store game. In infinitely repeated games, the multiplicity of equilibria provided by the folk theorem contrasts with the intuitive attraction of equilibria that provide relatively high payoffs. Reputation effects can again rescue such intuition by imposing lower bounds on equilibrium payoffs.

Our paper is most closely related to FL, where they show that an arbitrarily patient long run player can guarantee himself a payoff that he could get by publicly committing to playing any of the commitment type strategies, in any Nash equilibrium of the infinitely repeated game. In particular if the commitment types have full support, long run player's equilibrium payoffs get arbitrarily close to his Stackelberg payoff, as his discount factor approaches one.

The incomplete information game can be seen as a perturbation of the complete information game. In particular the complete information game is the limit of incomplete information games as the ex ante commitment type probabilities approach to zero. From this point of view FL's result highlights two of the reputation effects: i) Many equilibria of the complete information game don't remain equilibria in the nearby incomplete information games ii) There are equilibrium payoffs in the incomplete information game that are not sustained by any equilibrium of the complete

³It is very intuitive that for a fixed discount factor δ , if the replacement rate is arbitrarily small, then there may be some histories after which the long run player's equilibrium continuation payoffs are very close to his equilibrium payoffs in the complete information repeated game. We don't have a proof of this, however we offer this "continuity argument" only to clarify why our result requires the replacement rate to be not "much" smaller than the impatience rate.

information game.

However Cripps, Mailath and Samuelson (2004) (CMS from hereon) showed that this reputation phenomena is only temporary if monitoring is imperfect. In particular, they show that, fixing the parameters of the game, there exists a period T such that in the continuation game starting at T , long run player's payoff 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.

There are examples in the literature that sustain persistent reputation effects by assuming replacements as in our model, that is, the type of the player is governed by a stochastic process and changes through time, rather than being determined once and for all at the beginning of the game. Holmstrom (1999), Cole, Dow and English (1995), Mailath and Samuelson (2001), Phelan (2006) and Vial (2008) maintain permanent reputations by assuming particular types of replacement in their models. Liu (2009) studies a model where short-run players must pay a cost to observe past signals. Bar-Isaac and Tadelis (2008) studies reputation effects in markets. Liu and Skrzypacz (2009) study reputation dynamics when short run players have bounded recall and Ekmekci (2010) studies sustainability of reputation effects when the short run players observe a summary statistic about the history of the play, such as in online markets.

Most of these papers construct a particular equilibrium (or a class of equilibria) that has interesting equilibrium dynamics. However none of these papers find bounds on long run player's equilibrium payoffs as in FL.

Model

The setup of the model closely resembles the setup of FL (92) and CMS (04). A long run player plays a fixed stage game against a new short run player in every period of an infinitely repeated game. Player 1 is the long run player and player 2 is the short run player in the stage game G . Player 1 selects an action a_1 from a finite set A_1 while player 2 selects from a finite set A_2 . An action profile is denoted $a \in A_1 \times A_2$. The stage game is a simultaneous move game. However our results easily generalize to sequential move games as in FL. At the end of each period a public signal y is observed, and y is drawn from a finite set Y according to the probability distribution $\rho(\cdot|a)$.⁴ Note that perfect monitoring is a special case of this setup. For any set X , let $\Delta(X)$ be the set of all probability distributions on X . The set of mixed actions for player i is $\Delta(A_i)$, and α_i represents a typical element of this set. A mixed action profile is $\alpha = (\alpha_1, \alpha_2) \in \Delta(A_1) \times \Delta(A_2)$. We denote $\rho(y|\alpha) = \sum_{a \in A} \rho(y|a)\alpha_1(a_1)\alpha_2(a_2)$. We assume an identification assumption as in CMS.

⁴We make this assumption for notational simplicity. Our results would still be true if monitoring technology were private.

Assumption: (Identification) $\rho(\cdot|\alpha'_1, a_2) \neq \rho(\cdot|\alpha_1, a_2)$ for any $\alpha_1 \neq \alpha'_1 \in \Delta(A_1)$ and $a_2 \in A_2$.

Identification assumption ensures that player 2 learns any stage game strategy of player 1 after observing sufficiently many signals.

Histories, Strategies and Payoffs

Let $R = \{r, nr\}$ denote the set of events "replaced" or "not replaced". We assume that a long run player at time t observes all replacement events that occurred before period t .

Both the long-run and short-run players can observe and condition their play at time t on the entire history of the realized outcomes. Let H_i^t be the set of player i 's histories at period t . In particular $H_1^t \equiv \{A_1 \times A_2 \times Y \times R\}^t$ denotes a t -period history for player 1, and $H_2^t \equiv \{A_2 \times Y\}^t$ denotes a t -period history for player 2. Let $\{\tilde{H}_1^t\}_{t=0}^\infty$ denote the filtration on $\{A_1 \times A_2 \times Y \times R\}^\infty$ induced by player 1's histories, and $\{\tilde{H}_2^t\}_{t=0}^\infty$ denote the filtration on $\{A_2 \times Y\}^\infty$ induced by player 2's histories. A pure strategy for a player 2 is a map $s_2^t : H_2^t \rightarrow A_2$, while the set of all such strategies is denoted S_2^t . A pure strategy s_1 for player 1 is a sequence of maps $s_1^t : H_1^t \rightarrow A_1$, and the set of all pure strategies is denoted S_1 .

Let Σ_1 and Σ_2^t be the sets of probability distributions over S_1 and S_2^t (equivalently the set of all behavioral strategies), let $\Sigma_2 = \prod_{t=0}^\infty \Sigma_2^t$. Each strategy profile $\sigma = (\sigma_1, \sigma_2) \equiv (\sigma_1, \sigma_2^0, \sigma_2^1, \dots) \in \Sigma_1 \times \Sigma_2$ gives rise to a probability distribution over infinite sequences of actions and signals. Let E_σ denote the expectation with respect to this distribution (and we denote the conditional expectations in the standard way), and the expected average discounted payoff to player 1 from the strategy profile σ is:

$$U_1(\sigma) = (1 - \delta)E_\sigma \left[\sum_{t=0}^\infty \delta^t u_1(a_1^t, y^t) \right], \quad (1)$$

where a^t is the action profile at period t , and a_1^t is player 1's action at period t and $u_1 : A_1 \times Y \rightarrow R$ is the stage game payoff function of player 1. Let $v_1(\alpha) = \sum_{a=(a_1, a_2) \in A, y \in Y} u_1(a_1, y) \rho(y|a) \alpha_1(a_1) \alpha_2(a_2)$ denote the expected stage game payoff to player 1 of a mixed action profile α . For every $h_1^t \in H_1^t$, let σ_{1, h_1^t} denote the continuation strategy of player 1 at history $h_1^t \in H_1^t$ for any $t \geq 0$.⁵ Similarly, let

$$U_1(\sigma|h_1^t) = (1 - \delta)E_\sigma \left[\sum_{s=t}^\infty \delta^{s-t} u_1(a_1^s, y^s) | h_1^t \right]$$

be the expected average discounted payoff of player 1 at history h_1^t .

All short run players have the same expected utility function $u_2 : Y \times A_2 \rightarrow R$, and each maximizes her expected payoff $v_2 : \Delta(A_1) \times \Delta(A_2) \rightarrow R$, where

$$v_2(\alpha) = \sum_{a=(a_1, a_2) \in A, y \in Y} u_2(y, a_2) \rho(y|a) \alpha_1(a_1) \alpha_2(a_2)$$

⁵For a definition of continuation strategy, see page 20 in Mailath and Samuelson (2007).

denotes the expected payoff from the mixed action profile α .

The payoff function of the short run players is common knowledge. On the other side, only player 1 knows his own type. Player 1 is either a simple commitment type⁶ or a normal type with an expected utility function specified as in equation 1. Let $W = \{w_0, w_1, \dots, w_N\}$ be a finite set of types. Each $w_i \in \Delta(A_1)$ with $i \geq 1$ corresponds to a commitment type who is committed to playing the strategy w_i at every period, and w_0 corresponds to the normal type with the payoff function U_1 as in equation 1.

Player 2's Beliefs

A reputation function $\mu : \cup_{t=0}^{\infty} H_2^t \rightarrow \Delta(W)$ describes the beliefs of short-run players about the type of their opponent after every history. A belief function $\sigma^{b,t} : H_2^t \rightarrow \Delta(A_1)$ describes the beliefs of the short run players about player 1's play after every history. Each pair of reputation function and strategy profile induces a belief function, in particular

$$\sigma^{b,t}(h_2^t)(a) = \sum_{i \geq 0} \mu(h_2^t)(w_i) w_i(a) + \mu(h_2^t)(w_0) E(\sigma_1(h_1^t)(a) | h_2^t) \quad \forall a \in A_1. \quad (2)$$

The set of all sequences of all such belief functions is $\Sigma_b = \{\sigma^b : \cup_{t=0}^{\infty} H_2^t \rightarrow \Delta(A_1)\}$.

Uncertainty and Replacements

At period $t = 0$, the type of the long run player is drawn from a probability distribution $\mu \in \Delta(W)$ with $\mu(w_i) > 0$ for each $w_i \in W$. At the beginning of every period with probability $(1 - \lambda)$ player 1 is replaced by a new player 1 whose type is drawn from a probability distribution $\eta^* \in \Delta(W)$ with $\eta^*(w_i) > 0$ for each $w_i \in W$. With the remaining probability λ player 1 is not replaced, so the same player 1 remains in the game. Player 1's discount factor δ may incorporate both his intertemporal time preference and the likelihood of being replaced, and we allow this, but we don't require it.⁷ Moreover we assume that player 1 observes the occurrence of all past replacement events. This corresponds for example to a situation where the new owner of a firm has access to all information about previous owners and previous transactions about the firm.⁸

⁶A simple commitment type plays a fixed mixed action at every stage of the repeated game until he is replaced.

⁷One could think of player 1's intertemporal discounting to be $\beta < 1$, and his total discount factor $\delta = \beta\lambda$. We don't restrict ourselves to this particular formulation.

⁸Alternatively we could assume that a long run player at time t does not observe any of the replacement events, or some of them before the period he joined the game. The observational assumptions about past replacement events by player 1 are irrelevant for our results.

Equilibrium

A Bayesian Nash equilibrium, (σ, μ) is the combination of a strategy profile $\sigma = (\sigma_1, \sigma_2)$ and a reputation function μ such that:

- i) σ_1 is a best response to σ_2 .
- ii) σ_2^t is a best response to player 2's belief $\sigma^{b,t}$ about player 1's play, as defined in equation 2.
- iii) The reputation function is updated via Bayes' rule. In particular for each $h_2^t \in H_2^t$ for $t \geq 1$:

$$\mu(h_2^t)(w) = \lambda \frac{\mu(h_2^{t-1})(w) \rho(y|\alpha_1(h_2^{t-1}, w), a_2^{t-1})}{\sum_{w' \in W} \mu(h_2^{t-1})(w') \rho(y|\alpha_1(h_2^{t-1}, w'), a_2^{t-1})} + (1 - \lambda) \eta^*(w) \quad (3)$$

$$\text{whenever } \sum_{w' \in W} \mu(h_2^{t-1})(w') \rho(y|\alpha_1(h_2^{t-1}, w'), a_2^{t-1}) > 0$$

where,

$$\begin{aligned} h_2^t &= (h_2^{t-1}, a_2^{t-1}, y), \\ \alpha_1(h_2^{t-1}, w_0) &= E(\sigma_1(h_1^{t-1})|h_2^{t-1}) \\ \alpha_1(h_2^{t-1}, w_i) &= w_i \text{ for } i \geq 1 \\ \mu(\oslash) &= \mu. \end{aligned}$$

The first 2 conditions describe optimality of behavior for all players. We express optimality of player 2's actions by requiring her stage game strategy to be a best response to her beliefs about her opponent's current period play. This belief has two components, the probability with which player 1 is a commitment type (captured by the reputation function μ) and the probability with which he is a normal type and he follows the strategy σ_1 . The third condition states that updating is Bayesian when possible.

Permanent Reputations

Let $G^\infty(\lambda, \delta)$ be the repeated game where every period a long run player 1 plays the stage game G with a short run player who serves as player 2 in the stage game. The replacement probability for player 1 is $(1 - \lambda)$, and his discount factor is $\delta < 1$. In the following, $B(\alpha_1)$ denotes the best reply set of player 2 to player 1's mixed action α_1 .

Fix any $w \in W \setminus \{w_0\}$. Let

$$\begin{aligned} b_2(\alpha_1) &\in B(\alpha_1), \text{ and } b_2(\alpha_1) \equiv \arg \min_{\alpha_2 \in B(\alpha_1)} v_1(\alpha_1, \alpha_2), \\ u_1(w) &= v_1(w, b_2(w)). \end{aligned}$$

Verbally, $b_2(\alpha_1)$ is one of the most upsetting (to player 1) strategy of player 2, among player 2's

best responses to player 1's strategy α_1 . $u_1(w)$ is the payoff that player 1 would guarantee himself if he could almost convince player 2 that he is the commitment type w . Note that the stage game is a finite simultaneous move game, so the expression with $\arg \min$ is well defined. Moreover, by upper hemicontinuity of the best response correspondence, there is an $\bar{\epsilon} > 0$ such that for every $\alpha_2 \in B(\alpha_1)$ for any α_1 with $\|\alpha_1 - w\| < \bar{\epsilon}$, $v_1(w, \alpha_2) \geq u_1(w)$.⁹

In our main theorem, we show that, if the replacement rate and the impatience rate are *comparably* small, then player 1's payoffs, calculated after *every* history of the repeated game, in *every* Bayesian Nash equilibrium is bounded below by almost $u_1(w)$. Our main theorem is stated below:

Theorem 1 *Let $v(\lambda, \delta)$ be the infimum of the set of all Bayesian Nash equilibrium payoffs of the long run player in $G^\infty(\lambda, \delta)$ after all histories $h_1^t \in \bigcup_{s=0}^{\infty} H_1^s$. Let $(\lambda_r, \delta_r)_{r=0,1,\dots}$ be a sequence of replacement probabilities and discount factors such that i) $\lim_{r \rightarrow \infty} \lambda_r = 1$ ii) $\liminf_{r \rightarrow \infty} \frac{1-\lambda_r}{1-\delta_r} > 0$. Then $\liminf_{r \rightarrow \infty} v(\lambda_r, \delta_r) \geq u_1(w)$ for all $w \in W \setminus \{w_0\}$.*

Proof. See Appendix B. ■

In this theorem we show that if the replacement probability $(1 - \lambda)$ vanishes at the same rate or at a slower rate than the impatience rate $(1 - \delta)$ (in particular if $\frac{(1-\lambda)}{(1-\delta)}$ does not go to zero as $(1 - \lambda)$ goes to zero), a lower bound on the long run player's Nash equilibrium payoffs calculated after every history of the game approaches $u_1(w)$. Since this is true for any $w \in W$, the long run player guarantees himself a payoff that is almost what he could get by fully mimicking his most favorite commitment type.

The most natural economic environment in which conditions i) and ii) are satisfied is a repeated game played very frequently. Consider a game which is played in periods $t = 1, 2, \dots$ in which the time between the periods is $\Delta > 0$. Suppose the long run player is replaced according to a Poisson probability distribution that has parameter ρ , 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 between periods, Δ , goes to zero, $\rho\Delta \rightarrow 0$ and $\frac{\rho\Delta}{1-\delta(\Delta)} \rightarrow \frac{\rho}{r+\rho}$, hence player 1's impatience and replacement probability go to 0 at the same rate.

As we argue in the sketch of argument for a perfect monitoring example, we may strengthen the theorem by replacing condition ii) by the following weaker condition that $\ln \delta_r \ln(1 - \lambda_r) \rightarrow 0$. This weaker condition is satisfied, for instance if $\liminf \frac{(1-\lambda_r)^\beta}{1-\delta_r} > 0$ for some $\beta > 0$. In other words, we can allow the log of the replacement rate to explode at the same rate as the log of the impatience rate.

In order to gain some intuition for the result, consider a game with perfect monitoring and with only one commitment type. Moreover assume that $\delta = \lambda$. Whenever player 2 plays an action that is not a best response to the commitment type's strategy, mimicking the commitment type increases player 1's reputation level by a multiple that is bounded from below by a positive number.

⁹For any two vectors x and y , $\|x - y\|$ denotes the Euclidean distance between x and y .

If there were no replacement possibility, then along a sequence of periods where player 2 is playing an action different from a best response to the commitment type strategy, player 1 can increase his reputation level from $1 - \delta$ to any number $\pi < 1$ in $O\left(\ln\left(\frac{1}{1-\delta}\right)\right)$ periods.¹⁰ The first observation we make is that as δ goes to one, $\ln\left(\frac{1}{1-\delta}\right)$ increases to ∞ *more slowly* than $\ln \delta$ approaches to zero. In particular $\lim_{\delta \rightarrow 1} \delta^{\ln\left(\frac{1}{1-\delta}\right)} = 1$.¹¹

The possibility of replacements makes the previous argument easier when the reputation level is low, since then replacement event increases the reputation of player 1. However, when the reputation of player 1 is relatively high, the argument is challenged by the fact that replacement possibility may decrease the reputation of player 1. Here we observe that when the replacement probability is low, if player 2 is not best responding to the commitment type, the reputation level will still increase significantly if player 1 mimics the commitment type. However at higher reputation levels, if player 2 best responds to the commitment type strategy, and if player 1 mimics the commitment type, then the reputation will decrease by at most an amount proportional to $1 - \delta$. Combining these observations we show that the fraction of periods in the repeated game where player 2 may not be best responding to the commitment type decreases to zero as δ goes to one. We give a more detailed sketch of the argument for the perfect monitoring case in the next section. The most challenging part of the proof is to show that the arguments above continue to hold when monitoring is imperfect. We give the details of the argument for the imperfect monitoring case in the Appendix, and a less detailed outline of the steps of the proof before the Appendix.

Perfect Monitoring Case: A Sketch

Consider the product choice game, whose payoff matrix is depicted in figure 1. Player 1 is the row player and player 2 is the column player in the stage game. The set of actions for player 1 is $A_1 = \{H, L\}$: High quality production and Low quality production. The set of actions for player 2 is $A_2 = \{B, N\}$: Buying the product and Not buying the product.

Players who are playing at period t observe the entire history of the *action profiles* played before period t , hence this is a game of perfect monitoring. At the beginning of each period, player 1 is replaced with a new player 1 with probability $1 - \lambda \in (0, 1)$, and player 1's discount factor is $\delta \in (0, 1)$. There are 2 types for player 1; a simple commitment type who plays $a_1 = H$ at every period of the game in which he is in the game and a normal type who maximizes his expected discounted payoffs in the repeated game where the stage game payoffs are given in the payoff matrix above. The probability that player 1 is a commitment type at the beginning of the game is $\theta > 0$, and the probability that player 1 is a commitment type at any period t conditional on

¹⁰ $O(x)$ denotes at the order of x .

¹¹To see this, let $Y(\delta) = \delta^{\ln\left(\frac{1}{1-\delta}\right)}$. Taking logs on both sides gives $\ln Y(\delta) = \ln \delta \ln\left(\frac{1}{1-\delta}\right)$. Using L'Hopital's rule twice delivers $\lim_{\delta \rightarrow 1} \ln Y(\delta) = 0$, hence the desired result holds.

	<i>B</i>	<i>N</i>
<i>H</i>	1, 1	0, 0
<i>L</i>	2, -1	0, 0

Figure 1: The payoff matrix for the product choice game.

him being replaced at the beginning of period t is θ . Note that the commitment payoff of player 1 is 1, and in the remainder of this section we will find a lower bound on his equilibrium payoffs that is close to 1 and that holds uniformly over all equilibria and after all histories.

Bellman Equation: We start by introducing a Bellman equation, that serves a lower bound on the set of all Nash equilibrium payoffs for player 1. Let $\underline{\mu} = (1 - \lambda)\theta$, $\bar{\mu} = \lambda + (1 - \lambda)\theta$, $S \equiv [\underline{\mu}, \bar{\mu}]$ and let $V : S \rightarrow R$ be a function satisfying:¹²

$$\begin{aligned}
V(\mu) &= \inf_{1 \geq \alpha_1 \geq \mu} \{(1 - \delta)u_1(H, a_2(\alpha_1)) + \delta V(\mu'(\alpha_1, \mu))\} \text{ where} & (4) \\
\mu'(\alpha_1, \mu) &= \lambda \frac{\mu}{\alpha_1} + (1 - \lambda)\theta \text{ for } \alpha_1 \geq \mu. \\
s.t. \ a_2(\alpha_1) &= B \text{ for } \alpha_1 > 1/2 \text{ and } a_2(\alpha_1) = N \text{ for } \alpha_1 \leq 1/2
\end{aligned}$$

The interpretation of the above program is simple. V is the payoff player 1 would get if *i*) he always plays H (i.e. mimics the commitment type) and *ii*) player 2 has a belief about player 1's play that minimizes player 1's total payoff (i.e., his current payoff and continuation payoff) from mimicking the commitment type.

We will argue that V is a lower bound on player 1's equilibrium payoffs. Consider any belief α_1^t that player 2 has about her opponent's play at period t . In equilibrium this belief should coincide with player 1's equilibrium behavior (which incorporates σ_1 and player 1's reputation for being a commitment type, μ). In particular, the equilibrium requires that actions in the support of α_1^t be either optimal for player 1's normal type, or be coming from the commitment type's strategy. However, the Bellman program is relaxing the equilibrium constraint that player 2's belief indeed coincides with player 1's equilibrium play (hence the set from which player 2's belief is chosen is larger compared to the set of beliefs consistent with any equilibrium). However there is still the requirement that player 2's action choice is a best response to this "arbitrary" belief. Since V is the infimum of a function over all possible beliefs, relaxing a constraint might only make the objective function better: in this case make the value function lower. Since mimicking the commitment type

¹²Since $\lambda + (1 - \lambda)\theta \geq \mu'(\alpha_1) \geq (1 - \lambda)\theta$, our restriction to the state space $S \equiv [(1 - \lambda)\theta, \lambda + (1 - \lambda)\theta]$ is without loss of content.

is an available strategy for player 1, the set of equilibrium payoffs for player 1 is bounded from below by V .

This is a deterministic Bellman equation.¹³ Note that there is a policy function $\alpha_1(\mu)$, and $a_2(\mu)$ such that $V(\mu) = (1 - \delta)u_1(H, a_2(\mu)) + \delta V(\mu'(\alpha_1(\mu), \mu))$, $\alpha_1(\mu) \geq \mu$ and that $a_2(\alpha_1) = B$ for $\alpha_1 > 1/2$ and $a_2(\alpha_1) = N$ for $\alpha_1 \leq 1/2$. Moreover, for any $\mu \in S$,

$$V(\mu) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t u_1(H, a_2(\mu_t))$$

$$s.t. \mu_0 = \mu, \text{ and } \mu_t = \lambda \frac{\mu_{t-1}}{\alpha_1(\mu_{t-1})} + (1 - \lambda)\theta \text{ for } t \geq 1. \quad (5)$$

Every $\mu \in S$ induces an infinite sequence of reputation levels $\{\mu_t\}_{t=0}^{\infty}$ with $\mu_0 = \mu$ as described in Equation 5 above. We say that a sequence of reputation levels $\{\mu_t\}_{t=0}^{\infty}$ is admissible if $\mu_0 \in S$ and $\mu_t = \lambda \frac{\mu_{t-1}}{\alpha_1(\mu_{t-1})} + (1 - \lambda)\theta$ for $t \geq 1$. In the following, when we use the term period t , we mean the $(t + 1)$ st element of an admissible sequence.

We can actually solve for a closed form expression for the functional form of the function $V(\mu)$. However, to explain the way the proof for the imperfect monitoring case operates, we will first proceed by putting bounds on $V(\mu)$ at various subintervals of the state space S . We form our inequalities and bounds in a very generous way. Although we could make our bounds tighter in the case of perfect monitoring, the methodology we use to prove our main result becomes tight in the case of imperfect monitoring.

Belief Dynamics: We start by analyzing how reputations evolve on admissible sequences. If $a_2(\mu) = N$, then $\alpha_1(\mu) \leq 1/2$ and hence $\mu'(\alpha_1, \mu) = \lambda \frac{\mu}{\alpha_1} + (1 - \lambda)\theta \geq 2\lambda\mu + (1 - \lambda)\theta > 2\lambda\mu$. Intuitively, on the outcome paths where player 1 plays H , if at some period t player 2 is playing N , i.e., $a_2(\mu_t) = N$, then she is expecting player 1 to play H with a probability at most 0.5, i.e., $\alpha_1(\mu_t) \leq 1/2$, and therefore, $\mu_{t+1} > 2\lambda\mu_t$.

Similarly, if $a_2(\mu) = B$, then $\alpha_1(\mu) \leq 1$ and hence $\mu'(\alpha_1, \mu) = \lambda \frac{\mu}{\alpha_1} + (1 - \lambda)\theta \geq \lambda\mu + (1 - \lambda)\theta > \lambda\mu$. Intuitively, if player 2 is playing B at some period t , i.e., $a_2(\mu_t) = B$, then $\alpha_1(\mu_t) \leq 1$ and hence $\mu_{t+1} > \lambda\mu_t$.

Fix a number of periods $K > 2$. Then there exists a length T such that during a block of T periods, if player 2 plays N more than K periods, then at the end of T periods player 1's reputation level doubles up compared to its starting level. The following claim formalizes this reasoning. In the following, for any number x , $\lfloor x \rfloor$ denotes the largest integer not larger than x .

Claim 1 *For any $K > 2$, $\exists \underline{\lambda} < 1$ such that for every $\lambda \in [\underline{\lambda}, 1)$, $\exists$ a positive integer $T(\lambda) \equiv \left\lfloor \frac{(K-2)\ln(2)}{-\ln(\lambda)} \right\rfloor$ such that for any admissible sequence $\{\mu_t\}_{t=0}^{\infty}$: If $\mu_0 < 1/2$, and if the K^{th} time at*

¹³For the more general case of imperfect monitoring, the Bellman equation we work with is a stochastic Bellman equation.

which $a_2(\mu_t) = N$ happens at some period $t \leq T(\lambda)$, then $\mu_t \geq 2\mu_0$. If $\mu_0 \geq 1/2$, then there are a total of less than K times at which $a_2(\mu_t) = N$ at or before period $T(\lambda)$.

Proof. The reputation level gets multiplied by at least 2λ if $a_2(\mu_t) = N$ (i.e. $\mu_{t+1} \geq 2\lambda\mu_t$), and by at least λ if $a_2(\mu_t) = B$ (i.e. $\mu_{t+1} \geq \lambda\mu_t$). Since by the beginning of period t , there has been $K - 1$ times when player 2's action is N , $\mu_t \geq 2^{K-1}(\lambda)^t\mu_0$, so if $T(\lambda) = \left\lfloor \frac{(K-2)\ln(2)}{-\ln(\lambda)} \right\rfloor$ then the first assertion in the claim is satisfied. In order to ensure that $T(\lambda)$ is positive, we need $\frac{(K-2)\ln(2)}{-\ln(\lambda)} \geq 1$. If $\lambda > \underline{\lambda} = 2^{2-K}$ then indeed $\frac{(K-2)\ln(2)}{-\ln(\lambda)} \geq 1$. The second assertion in the claim follows from the fact that if there was a period t when $a_2(\mu) = N$ for the K^{th} time, then $\mu_t \geq 2\mu_0 = 1$. However $\mu_t \leq \lambda + (1 - \lambda)\theta < 1$, which is a contradiction. ■

Partitioning the State Space: We partition the state space by first defining a sequence of reputation thresholds: $\underline{\mu}_0 = 1/2$, $\underline{\mu}_1 = \frac{\mu_0}{2}, \dots, \underline{\mu}_k = \frac{\mu_0}{2^k}$. Let n be the smallest integer for which $\underline{\mu}_n \leq \underline{\mu}$. The collection of subintervals $[\underline{\mu}_0, 1], [\underline{\mu}_1, \underline{\mu}_0), \dots, [\underline{\mu}, \underline{\mu}_{n-1})$ cover the state space S . Next we will put lower bounds on $V(\mu)$ for every interval, by considering admissible sequences $\{\mu_t\}_{t=0}^\infty$, with various initial reputation levels. Let $\inf_{\mu \geq \underline{\mu}} V(\mu) = v$.

Bound on the Reward Phase: Fix $K > 2$ and let $T = T(\lambda)$ as in claim 1. Our first observation is that, for $\mu_0 \geq 1/2$,

$$V(\mu_0) \geq F_0 \equiv (1 - \delta) \sum_{s=K-1}^T \delta^s 1 + \delta^{T+1} v. \quad (6)$$

The inequality uses two observations. When the initial reputation level is sufficiently high, then for the initial $T + 1$ periods (from period 0 until period T), player 2's actions can be N at most $K - 1$ times. The permutation of payoffs across the periods that minimizes the average discounted payoffs is when during the first $K - 1$ periods, i.e., during $s \in \{0, 1, \dots, K - 2\}$ player 1 receives a payoff 0, and afterwards, until and including T if he receives a payoff 1. The continuation payoff at time $T + 1$ is bounded below by the least possible value of the value function across the whole range of the state space, delivering the inequality.

If $v \geq 1$, then we have a very tight lower bound on v . In the following development, we will put a less tight lower bound on v , hence we will assume that $v < 1$.

Bounds on the Reputation Building Phases If $\mu_0 \in [\underline{\mu}_1, \underline{\mu}_0)$, then either player 2 plays N less than K times during the first T periods, or plays N for the K^{th} time at some period $t \leq T$. In the first event, $V(\mu_0) \geq (1 - \delta) \sum_{s=K-1}^T \delta^s 1 + \delta^{T+1} v = F_0$ because there are at most $K - 1$ times when player 2 plays N , and the lowest possible permutation of payoffs is if these periods occur at the beginning of the sequence, i.e. at periods $t \in \{0, 1, \dots, K - 2\}$. Moreover, at period $T + 1$, the

continuation payoff is some value from the set of all possible values of the value function, we put the lowest value in this set. In the second event however,

$$V(\mu_0) \geq (1 - \delta) \sum_{s=K-1}^{t-1} \delta^s 1 + \delta^t \inf_{\mu \geq \underline{\mu}_0} V(\mu) \geq (1 - \delta) \sum_{s=K-1}^{t-1} \delta^s 1 + \delta^t F_0.^{14}$$

This is because until t , there are at most $K - 1$ times when player 1's payoff is zero, and his payoff is one for all the other times until t . Moreover, by the beginning of t , μ_t is at least $2\mu_0 \geq \underline{\mu}_0$. Since the continuation payoff belongs to the set of values where the state variable is at least $\underline{\mu}_0$, putting the lower bound of this set using inequality 6 establishes the lower bound above. Therefore,

$$V(\mu_0) \geq \min \left\{ \min_{K-1 \leq t \leq T} \left((1 - \delta) \sum_{s=K-1}^{t-1} \delta^s 1 + \delta^t F_0 \right), F_0 \right\}.$$

Since $F_0 < 1$, we deduce that

$$\min_{K-1 \leq t \leq T} \left((1 - \delta) \sum_{s=K-1}^{t-1} \delta^s 1 + \delta^t F_0 \right) \geq \delta^{K-1} F_0 \equiv F_1.$$

Therefore $V(\mu_0) \geq \delta^{K-1} F_0$ for $\mu_0 \in (\underline{\mu}_1, \underline{\mu}_0]$. A similar argument shows that if $\mu_0 \in [\underline{\mu}_{k+1}, \underline{\mu}_k)$ then

$$V(\mu_0) \geq \min \left\{ \min_{K-1 \leq t \leq T} \left((1 - \delta) \sum_{s=K-1}^{t-1} \delta^s 1 + \delta^t \inf_{\mu \geq \underline{\mu}_k} V(\mu) \right), F_0 \right\}.$$

At this point, we use induction (we'll skip the induction step in this section) to show that $V(\mu_0) \geq F_{k+1} \equiv (\delta^{K-1})^{k+1} F_0$.

The Number of Reputation Building Phases: Since $\mu = (1 - \lambda)\theta$, to back out n , we use the inequality that $\frac{(1/2)}{2^{n-1}} > (1 - \lambda)\theta$, and rearranging gives $n < \frac{-\ln[(1-\lambda)\theta]}{\ln 2}$ and so $n = \frac{-\ln[(1-\lambda)\theta]}{\ln 2} - 1$ if $\frac{-\ln[(1-\lambda)\theta]}{\ln 2}$ is an integer and $n = \left\lfloor \frac{-\ln[(1-\lambda)\theta]}{\ln 2} \right\rfloor$ otherwise.

Inequality on v : Since F_k is decreasing in k , $V(\mu) \geq (\delta^{K-1})^n F_0$ for every $\mu \in S$, and therefore $\inf_{\mu \in S} V(\mu) = v \geq (\delta^{K-1})^n F_0$. Rewriting, we get

$$v \geq (\delta^{K-1})^n \left((1 - \delta) \sum_{s=K-1}^T \delta^s 1 + \delta^{T+1} v \right). \quad (7)$$

An interpretation of Equation 7 is as follows. The state variable (reputation level) doubles up

¹⁴Note that $t \geq K - 1$. We use the convention that if $t = K - 1$, then $\sum_{s=K-1}^{t-1} \delta^s 1 = 0$.

if there are $K - 1$ periods within T periods where player 2 plays N and if player 1 has been playing the action H . The worst possible scenario for player 1 is if there are consecutive $K - 1$ periods where player 1 suffers from player 2's action N , but in this case he doubles up his reputation level. Since there is a positive lower bound on the state space, in a finite number of n steps, the reputation level reaches above a critical level, which is $1/2$. The number of steps n depends on the lower bound of the state space, and hence the replacement probability. The critical level of the state variable, $1/2$ has the feature that at this reputation level, for a long number of periods, player 1 gets his commitment payoff of one and at the end of these periods, possibly he gets some different payoff, that is non-negative. When his reputation level drops below $1/2$, obviously a lower bound on his continuation payoff is v .

We would like to stress again that this bound is very generous, since we know that once the reputation level drops below $1/2$, if it is still close to $1/2$, then there may be only one non-best response to H before player 1 may increase his reputation significantly (i.e. multiply by 2λ) and move it above $1/2$. However this non-tight lower bound reflects the methodology used to prove the imperfect monitoring case.

The ratio $(K - 2)/(T + 1)$ measures the adverse effect on player 1 of the replacement possibility when his reputation level is high. K is fixed and doesn't depend on the parameters of the model (but is at least 3), however T depends on the probability of replacement. The higher the probability of replacement is, the smaller T is, therefore the magnitude of the "guaranteed" rewards at high reputation levels becomes lower with more frequent replacements. In particular T increases at the order of $(1 - \lambda)^{-1}$.

The number of steps n required to reach the rewarding reputation level of $1/2$ is essentially the number of subintervals that should be passed before reaching the reward phase. The number of such intervals, n is at the order of $-\ln(1 - \lambda)$. As the replacement probability gets lower n gets larger. The multiplicative term $(\delta^{K-1})^n$ measures the time cost of building a high reputation. Whether reputation building is easy or not depends on the relative speeds by which the discount factor reaches one ($\delta \rightarrow 1$) and the replacement probability reaches zero ($-\ln(1 - \lambda) \rightarrow \infty$ or $n \rightarrow \infty$). Confirming the usual intuition, for any fixed probability of replacement, as player 1 gets arbitrarily patient ($\delta \rightarrow 1$), reputation building becomes arbitrarily costless. Recollecting the terms in equation 7 we get

$$v \left(1 - (\delta^{K-1})^n \delta^{T+1} \right) \geq (\delta^{K-1})^n (1 - \delta) \left(\sum_{s=K-1}^T \delta^s 1 \right). \quad (8)$$

In order for a sequence of lower bounds on v_r to converge to 1 along sequences of δ_r and λ_r , the following should be true: $(\delta_r^K)^{n_r} \rightarrow 1$, and $\frac{K}{T_r} \rightarrow 0$ along the sequence. Since $T(\lambda_r) = \left\lfloor \frac{(K-2)\ln(2)}{-\ln(\lambda_r)} \right\rfloor$, $\frac{K}{T_r} \rightarrow 0$ only if $\lambda_r \rightarrow 1$. On the other side, n is at the order of $\ln(1 - \lambda)$, so for a fixed K , $(\delta_r^{K-1})^{n_r} \rightarrow 1$ only if $(\delta_r)^{\ln(1-\lambda_r)} \rightarrow 1$, or $\ln \delta_r \ln(1 - \lambda_r) \rightarrow 0$. This is satisfied, for instance, when

$1 - \delta_r$ and $1 - \lambda_r$ disappear at the same rate, or if $1 - \delta_r$ disappears and λ_r either stays constant or goes to one more slowly than $1 - \delta_r$ disappears. Note that the critical property here is that $\ln \delta_r \ln(1 - \lambda_r) \rightarrow 0$, hence if $\liminf \frac{(1 - \lambda_r)^\beta}{1 - \delta_r} > 0$ for some $\beta > 0$, then $\ln \delta_r \ln(1 - \lambda_r) \rightarrow 0$.

On the other side, even as δ_r goes to 1, if λ_r is not converging to 1, then $\frac{K}{T_r}$ remains bounded away from zero. In this case, the reputation bound is weakened by the fact that when the reputation level is high, then replacement possibilities may indeed lower the reputation levels. On the other side, since $\frac{K}{T_r}$ approaches to zero as the replacement probability approaches to zero, we can make our result a bit stronger by saying that for every $\varepsilon > 0$, there exists a $\lambda^* < 1$ such that if $\lambda_r > \lambda^*$ for every $r > 1$, $\delta_r \rightarrow 1$ and if $\ln \delta_r \ln(1 - \lambda_r) \rightarrow 0$, then $\liminf v_r > 1 - \varepsilon$.

Imperfect Monitoring A Sketch:

In this section, we outline how the arguments work if the stage game actions are imperfectly observed.

Step 1: We express the stochastic Bellman Equation 10 using the likelihood ratio $L = \frac{1 - \mu}{\mu}$ as the state variable instead of μ . This equation is the imperfect monitoring counterpart of the Bellman equation 4. We do this in Appendix B.1. Lemma 1 shows that policy functions $\alpha_2(L)$ and $\alpha_1(L)$ which identify a stage game strategy for player 2 and a belief for her about player 1's play for each value of the state variable exist.

Step 2: We define the stochastic process $\{L_t\}_{t=0}^\infty$ induced by the infinite sum representation of $V(L)$ and by the policy functions $\alpha_2(L)$ and $\alpha_1(L)$ and the commitment type strategy. Next we define the "adjusted" process $\{S_t\}_{t=0}^T$ by $S_t = L_t + m - (1 - \lambda)zt$ for some constants m and z , and specify T to be the largest integer such that $m - (1 - \lambda)zT \geq 0$ (definition 4). We look at blocks of $T/2$ periods, and evaluate the payoff streams within each block.

Step 3: Lemma 5 shows that the adjusted process S_t is a supermartingale with respect to the probability measure induced by the policy functions and the commitment type strategy. The heart of this step is to show that if at any period t , player 1 receives a lower payoff than $u_1(w)$, i.e. if $v(w, a_2(L_t)) < u_1(w)$, then S_{t+1}/S_t is *expected* to "jump" by at least a certain amount $\psi > 0$ with at least probability ψ .

We prove this by first showing in Lemma 6 that if L_t is not expected to move, then $v(w, a_2(L_t)) \geq u_1(w)$. Lemmata 7 and 8 show that if L_t is expected to move, then S_{t+1}/S_t is *expected* to "jump" as in the supposition of this step. Hence we conclude that if $v(w, a_2(L_t)) < u_1(w)$ then L_t is expected to move, and if this is the case, then S_{t+1}/S_t is *expected* to "jump".

Step 4: We define a stopping time τ_k that picks certain periods, which includes periods in which S_{t+1}/S_t is expected to jump. In particular, if τ_k does not pick a period t , then by step 3 $v(w, a_2(L_t)) \geq u_1(w)$ (by lemma 11).

Step 5: We define a faster process $\widetilde{S}_k = S_{\tau_k}$. By the definition of the stopping process τ_k , $\widetilde{S}_k$ is an active super martingale. Hence we use Theorem 2 (this is FL's theorem that is restated in

Appendix A) to argue that, if within $T/2$ periods, the process S has stopped for the K^{th} time at a period t , then $S_t \leq S_0/2$ with a probability at least $1 - \varepsilon$. Combining this with step 4, and showing that $S_t \leq S_0/2$ implies $L_t \leq L_0/2$, we argue that if the supposition is true, then:

$$V(L) \geq (1 - \varepsilon) \left[(1 - \delta) \sum_{s=K-1}^{t-1} \delta^s u_1(w) + \delta^t \inf_{L' \leq L/2, \text{ and } L' \in S} V(L') \right] + \varepsilon \delta^t \inf_{L' \in S} V(L') \quad (9)$$

Step 6: We put bounds on the value function at several intervals, using inequality 9 in step 5.¹⁵ The first bound is for an interval that contains very high reputation levels (i.e. very low values of the state variable, likelihood ratio). This is similar to the lower bound for the reward phase of the perfect monitoring example, in Inequality 6. We put further bounds for lower reputation levels by successively using Inequality 9.

Step 7: In the last step, we collect terms, and get an inequality for $\inf_{L \in S} V(L)$, inequality 16. This is similar to Inequality 8 of the perfect monitoring case. In Appendix C we show that this lower bound converges to $u_1(w)$ when the discount factor goes to one and the replacement probability goes to zero not faster than the discount factor.

Discussion

In this section, we would like to discuss some of the assumptions of our model. We assume that a short run player at period t observes the public outcomes of the stage game play of every period that precedes period t . We form our bounds using the probability measure on histories induced by the commitment type strategies. Therefore our bound does not depend on whether the beliefs of the short run players (or the reputation of the long run player) are public (or are known by the long run player) or private. This is particularly important since Cripps et. al. (2007) show that reputation effects disappear even if the long run player doesn't know his own reputation (i.e., the short run players observe a different history than the long run player, and the long run player does not know this history, but makes inferences using his own history). Hence all our results and proofs would go through if the outcomes of the stage game were not publicly observed, but instead the short run players observed a signal that is informative about the long run player's actions.

In this paper we focus attention to stage games that are finite and simultaneous move. We could generalize our results easily (though with a lot of additional notation) to extensive form games, with the commitment bounds replaced by the self confirming equilibria and limit (as $\varepsilon \rightarrow 0$) of ε best responses as in FL. After all, the current paper applies the learning lemmata in FL repeatedly to achieve a sequence of bounds on the lowest possible equilibrium of the long run player uniformly across all Nash equilibria. In the same vein, we can dispense with identification assumption by defining the commitment payoff through the player 2's anticipation of the outcomes (i.e., the

¹⁵There is also the possibility that the process S has not stopped for K times within the first $T/2$ periods. This possibility gives us a lower bound which is less tighter than the bound in inequality 9 as we show in the Appendix.

probability distribution function he assigns to the outcomes), instead of the current form where they are defined as payoff functions of the action profiles.

We assume that replacement events are stationary, and transition probabilities from a type to another one is independent of the current type. The role of the replacement events is to prevent the uncertainty about the long run player to disappear. Any stochastic process that changes the type of the long run player with full support, and uniformly with a small probability (or a probability that vanishes uniformly with the impatience rate) would be sufficient for our results.

Appendix A: Active Supermartingales

Consider an abstract setting with a probability measure P on a set of histories H^∞ and a filtration $\{\tilde{H}_k\}_{k=0}^\infty$, where each $\tilde{H}_k$ is generated by a finite partition, with generic element denoted $\tilde{h}^k$.

Definition 1 *A positive supermartingale $\{S_k\}$ is active with activity $\psi > 0$ (under P) if*

$$P\{h^\infty : \left| \frac{S_k}{S_{k-1}} - 1 \right| \geq \psi | \tilde{h}^{k-1}\} \geq \psi$$

for almost all histories with $S_{k-1} > 0$.

Fudenberg and Levine (1992, Theorem A.1) showed the following remarkable result:

Theorem 2 *For every $\varepsilon > 0$, $\psi \in (0, 1)$, $R > 1$ there is a time $K < \infty$ such that*

$$P(h^\infty : \sup_{k \geq K} S_k \leq \frac{S_0}{R}) \geq 1 - \varepsilon$$

for every positive active supermartingale $\{S_k\}$ with activity ψ and $S_0 > 0$.

The power of the theorem comes from the fact that the integer K , which depends on the parameters ε , ψ , and R is otherwise independent of the underlying stochastic process P . Intuitively, S_k being a supermartingale is weakly decreasing in expectations. The assumption that it is active says that it must be jumping significantly up or down relative to S_{k-1} with a significant probability each period. The theorem says that if $\{S_k\}$ is an active supermartingale, then there is a fixed time K by which, with high probability, S_k drops below $\frac{S_0}{R}$ and remains below this number for all future periods.

Appendix B: Proof of Theorem 1

Theorem 3 *Let $v(\lambda, \delta)$ be the infimum of the set of all Nash equilibrium payoffs of the long run player in $G^\infty(\lambda, \delta)$ after all histories $h_1^t \in \bigcup_{t=0}^\infty H_1^t$. Let $(\lambda_r, \delta_r)_{r=0,1,\dots}$ be a sequence of replacement probabilities and discount factors such that i) $\lim_{r \rightarrow \infty} \lambda_r = 1$ ii) $\liminf_{r \rightarrow \infty} \frac{1-\lambda_r}{1-\delta_r} > 0$. Then $\liminf_{r \rightarrow \infty} v(\lambda_r, \delta_r) \geq u_1(w)$ for all $w \in W \setminus \{w_0\}$.*

First we normalize the stage game payoffs (without loss of generality) such that $u_1(a_1, y) \geq 0$ for any $a_1 \in A_1$ and $y \in Y$. Fix $w \in W \setminus \{w_0\}$.

Appendix B.1: Bellman Equation as a Lower Bound

In order to find a lower bound on the set of Nash equilibrium payoffs of player 1, we look at the values of the program described below. This program calculates the payoffs to player 1 from fully mimicking type w . Since mimicking type w is an available strategy for player 1, his equilibrium payoffs are weakly higher than this payoff. The payoff to player 1 from mimicking type w depends on the actions chosen by player 2 throughout the play of the game. There is a restriction, however, on player 2's actions that they each be a best response to the equilibrium strategy of their opponent. The only restriction the below program puts on player 2's actions is that they be a best response to *some* strategy of player 1, and this action is chosen to be the worst possible choice for player 1 supposing he is mimicking type w .

Note that because of the replacement possibility, player 1's reputation for being a type w can never be below $\underline{\mu} = (1 - \lambda)\eta^*(w)$. Let $\bar{L} \equiv \frac{1}{(1-\lambda)\eta^*(w)} - 1$, $S \equiv [0, \bar{L}]$ and consider the following Bellman equation $V : S \rightarrow R$:

$$\begin{aligned}
 V(L) &= \inf_{\alpha_1 \in W(L)} (1 - \delta)v_1(w, \alpha_2) + \delta E_{w, \alpha_2} \{V(L'(y, \alpha_1, a_2))\} \\
 &\quad \text{such that} \\
 1 &: \alpha_2 \in B(\alpha_1) \\
 2 &: W(L) = \{\alpha_1'' \in \Delta(A_1) \mid \exists \alpha_1' \in \Delta(A_1) \text{ s.t. } \alpha_1'' = \frac{1}{1+L}w + \frac{L}{1+L}\alpha_1'\} \\
 3 &: \frac{1}{1+L'(y, \alpha_1, a_2)} = \frac{\frac{1}{1+L}\rho(y|w, a_2)}{\rho(y|\alpha_1, a_2)}\lambda + (1 - \lambda)\eta^*(w) \text{ for } \alpha_2(a_2) > 0 \text{ and } \rho(y|\alpha_1, a_2) > 0.
 \end{aligned} \tag{10}$$

In this program, L is the state variable of the value function. This is the likelihood ratio for player 1 being a w type. If player 2 believes that player 1 is a w type with probability μ , then the likelihood ratio associated with this belief is $L = \frac{1-\mu}{\mu}$. The first constraint in the program says that player 2's strategy α_2 is a best response to some strategy α_1 of player 1. The second constraint says that this strategy, α_1 , should be feasible; in particular it should be consistent with the fact that player 1 is a commitment type w with a probability $\mu = \frac{1}{1+L}$. The last condition is the law of motion for the state variable. The likelihood ratio next period is obtained via Bayes' rule considering player 1's strategy to be α_1 . The objective function minimizes player 1's expected utility if he were to fully mimic type w given the conditions 1, 2 and 3. The term E_{w, α_2} corresponds to expectation over the set of possible realizations of public outcomes ($Y \times A_2$) under the strategy profile (w, α_2) . More precisely,

$$E_{w,\alpha_2}\{V(L'(y, \alpha_1, a_2))\} = \sum_{y \in Y, a_2 \in \alpha_2} V(L'(y, \alpha_1, a_2)) \alpha_2(a_2) \rho(y|w, a_2)$$

In lemma 1 we show that the value function there exists policy functions $\alpha_1 : S \rightarrow \Delta(A_1)$ and $\alpha_2 : S \rightarrow \Delta(A_2)$ such that $V(L) = (1 - \delta)v_1(w, \alpha_2(L)) + \delta E_{w,\alpha_2(L)}\{V(L'(y, \alpha_1(L), a_2))\}$.

Lemma 1 *There exists a policy function $\alpha_1(L)$ and $\alpha_2(L)$ such that:*

$$\begin{aligned} V(L) &= (1 - \delta)v_1(w, \alpha_2(L)) + \delta E_{w,\alpha_2(L)}\{V(L'(y, \alpha_1(L), a_2))\} \\ s.t. &: \alpha_2(L) \in B(\alpha_1(L)) \\ \alpha_1(L) &\in W(L) \end{aligned}$$

Proof. See Appendix D. ■

First observe that the policy functions $\alpha_1(L)$ and $\alpha_2(L)$ induce a Markovian stochastic process with a transition function Q on the state variable $L \in S$. Let the set of outcomes be $O \equiv Y \times A_2$, the set of t period outcomes be $O^t \equiv (Y \times A_2)^t$, the set of infinite sequence of outcomes be $O^\infty \equiv (Y \times A_2)^\infty$, and let $\{\tilde{O}_t\}_{t=0}^\infty$ be the filtration on O^∞ . Any $L_0 \geq 0$ induces an admissible stochastic process $\{L_t\}$ that is measurable with respect to the sigma algebra generated by $\{O^t\}_{t=0}^\infty$ with the corresponding probability measure Q^∞ and that for $\alpha_2(a_2(L_t)) > 0$ and $\rho(y|\alpha_1, a_2) > 0$,

$$L_{t+1}(o^{t+1} = (o^t, y, a_2)|o^t) = \frac{\frac{1}{1+L_t} \rho(y|w, a_2)}{\rho(y|\alpha_1(L_t), a_2)} \lambda + (1 - \lambda) \eta^*(w).$$

We can alternatively express $V(L)$ as an expected discounted sum of a sequence whose elements are $v_1(w, \alpha_2(L))$. In particular,

$$\begin{aligned} V(L) &= (1 - \delta) E_{Q^\infty} \left[\sum_{t=0}^{\infty} \delta^t v_1(w, \alpha_2(L_t)) \right] \\ L_0 &= L \text{ and } \{L_t\} \text{ is induced by } L_0 \text{ and } Q^\infty. \end{aligned}$$

In the remaining of the paper, P corresponds to probabilities of events and E denotes expectations under the probability measure Q^∞ .

Definition 2 *Let $\{\mu_t\}_{t \geq 0}$ be the stochastic process where*

$$\mu_t = \frac{1}{1 + L_t} \text{ for every } t \geq 0.$$

In the following lemma we show that the lowest possible value of the value function $V(L)$ on the state space S constitutes a lower bound for player 1's Nash equilibrium payoffs after any history of the repeated game with replacement probability $1 - \lambda$ and discount factor δ (i.e. $v(\lambda, \delta)$).

Lemma 2 $\inf_{L \leq \bar{L}} V(L) \leq v(\lambda, \delta)$

Proof. Note that $V(L)$ is the payoff to player 1 of the strategy profile where he always plays w and player 2 best responds to some belief about player 1's strategy which doesn't necessarily coincide with player 1's equilibrium strategy. This belief is chosen to minimize player 1's payoff, if he were to always play w . So any equilibrium strategy profile has one more restriction on how player 2 forms beliefs (they should coincide with the equilibrium strategies), and player 1's available strategies are chosen from a larger set. Since player 2 is choosing the beliefs to minimize player 1's payoff, making the set where player 2 chooses her beliefs from smaller and enlarging player 1's strategy set may only increase player 1's payoff. ■

Appendix B.2: Adjusted process $\{S_t\}$ is a supermartingale, and S_t "jumps" if $v(w, \alpha_2(L_t)) < u_1(w)$.

Definition of the adjusted process $\{S_t\}$ and showing $\{S_t\}$ is a supermartingale:

Definition 3 Let $f : [0, 1] \rightarrow [(1 - \lambda)\eta^*(w), \lambda + (1 - \lambda)\eta^*(w)]$ be the function:

$$f(a) = \lambda a + (1 - \lambda)\eta^*(w)$$

If player 2 assigns probability a to player 1 being a type w at the *end* of some period t , then in the *beginning* of period $t + 1$ she assigns probability $f(a) = \lambda a + (1 - \lambda)\eta^*(w)$ to player 1 being a type w . Alternatively, if player 2 assigns probability μ_{t+1} to his opponent being a type w in the beginning of period $t + 1$, then she assigns probability $f^{-1}(\mu_{t+1})$ to her opponent being a type w at the end of period t . In the following lemma we restate a well known fact that the expected value of the likelihood ratio (for player 1 being a type w) at the end of any period t is equal to the likelihood ratio at the beginning of t .

Lemma 3 $E \left[\frac{1 - f^{-1}(\mu_{t+1})}{f^{-1}(\mu_{t+1})} \middle| \mu_t \right] = \frac{1 - \mu_t}{\mu_t}$

Proof. This is a well known fact. We give a proof for completeness. Fix $o^t \in O^t$. Using Equation 3 in the Bellman program 10, and replacing $\frac{1}{1+L}$ with μ we obtain for every $o^{t+1} = (o^t, o) \in O^{t+1}$ with $P(o^{t+1}|o^t) > 0$ that (let $o = (y, a_2)$)

$$\mu_{t+1}(o^{t+1}) = \lambda \frac{P(o^{t+1}|o^t)\mu_t(o^t)}{\rho(y|\alpha_1(L_t), a_2)\alpha_2(a_2)} + (1 - \lambda)\eta^*(w).$$

Using the definition of f above, we get

$$f^{-1}(\mu_{t+1}) = \frac{P(o^{t+1}|o^t)\mu_t}{\rho(y|\alpha_1(L_t), a_2)\alpha_2(a_2)}, \text{ and}$$

$$\frac{1 - f^{-1}(\mu_{t+1})}{f^{-1}(\mu_{t+1})} = \frac{\rho(y|\alpha_1(L_t), a_2)\alpha_2(a_2) - P(o^{t+1}|o^t)\mu_t}{P(o^{t+1}|o^t)\mu_t} \text{ and}$$

$$\begin{aligned} E \left[\frac{1 - f^{-1}(\mu_{t+1})}{f^{-1}(\mu_{t+1})} \middle| \mu_t \right] &= \frac{1 - \mu_t}{\mu_t} E \left[\frac{\rho(y|\alpha_1(L_t), a_2)\alpha_2(a_2) - P(o^{t+1}|o^t)\mu_t}{P(o^{t+1}|o^t)(1 - \mu_t)} \middle| \mu_t \right] \\ &= \frac{1 - \mu_t}{\mu_t} \sum_{o \in O} \frac{\rho(y|\alpha_1(L_t), a_2)\alpha_2(a_2) - P(o^t, o|o^t)\mu_t}{P(o^t, o|o^t)(1 - \mu_t)} P(o^t, o|o^t) \\ &= \frac{1 - \mu_t}{\mu_t} \left(\frac{\sum_{o \in O \equiv (Y \times A_2)} \rho(y|\alpha_1(L_t), a_2)\alpha_2(a_2)}{(1 - \mu_t)} - \frac{\sum_{o \in O} P(o^t, o|o^t)\mu_t}{(1 - \mu_t)} \right) \\ &= \frac{1 - \mu_t}{\mu_t} \left(\frac{1}{(1 - \mu_t)} - \frac{\mu_t}{(1 - \mu_t)} \right) = \frac{1 - \mu_t}{\mu_t}. \end{aligned}$$

This last step follows from the fact that $\sum_{o \in O \equiv (Y \times A_2)} \rho(y|\alpha_1(L_t), a_2)\alpha_2(a_2) = 1$ and $\sum_{o \in O} P(o^t, o|o^t) = 1$. ■

In the following lemma, we attain an inequality that bounds how much in expectation the likelihood ratio at the beinning of a period $t + 1$ can exceed the likelihood ratio at the beginning of period t . Let

$$z = \frac{(1 - \eta^*(w))}{\eta^*(w)^2}$$

Lemma 4 $E[L_{t+1}|L_t] \leq L_t + \frac{(1-\lambda)}{\lambda}z$

Proof. For each number satisfying $(1 - \lambda)\eta^*(w) \leq x \leq \eta^*(w)$, we have $f^{-1}(x) = \frac{x - (1-\lambda)\eta^*(w)}{\lambda} \leq x$, hence

$$\frac{1 - x}{x} \leq \frac{1 - \left(\frac{x - (1-\lambda)\eta^*(w)}{\lambda} \right)}{\left(\frac{x - (1-\lambda)\eta^*(w)}{\lambda} \right)} = \frac{1 - f^{-1}(x)}{f^{-1}(x)} \leq \frac{1 - f^{-1}(x)}{f^{-1}(x)} + \frac{(1 - \lambda)}{\lambda}z.$$

For $1 > x > \eta^*(w)$, we have

$$\frac{1 - x}{x} - \frac{1 - \left(\frac{x - (1-\lambda)\eta^*(w)}{\lambda} \right)}{\left(\frac{x - (1-\lambda)\eta^*(w)}{\lambda} \right)} = \frac{(1 - \lambda)(x - \eta^*(w))}{x(x - (1 - \lambda)\eta^*(w))}.$$

Replacing x by 1 in the nominator, and by $\eta^*(w)$ in the denominator yields

$$\frac{1 - x}{x} \leq \frac{1 - \left(\frac{x - (1-\lambda)\eta^*(w)}{\lambda} \right)}{\left(\frac{x - (1-\lambda)\eta^*(w)}{\lambda} \right)} + \frac{(1 - \lambda)(1 - \eta^*(w))}{\eta^*(w)^2\lambda} = \frac{1 - f^{-1}(x)}{f^{-1}(x)} + \frac{(1 - \lambda)}{\lambda}z$$

Hence,

$$L_{t+1} = \frac{1 - \mu_{t+1}}{\mu_{t+1}} \leq \frac{1 - f^{-1}(\mu_{t+1})}{f^{-1}(\mu_{t+1})} + \frac{(1 - \lambda)}{\lambda} z = \frac{\rho(y|\alpha_1(L_t), a_2)\alpha_2(a_2) - P(o^{t+1}|o^t)\mu_t}{P(o^{t+1}|o^t)\mu_t} + \frac{(1 - \lambda)}{\lambda} z \quad (11)$$

$$\begin{aligned} E[L_{t+1}|L_t] &= E\left[\frac{1 - \mu_{t+1}}{\mu_{t+1}}|\mu_t\right] \leq E\left[\frac{1 - f^{-1}(\mu_{t+1})}{f^{-1}(\mu_{t+1})} + \frac{(1 - \lambda)}{\lambda} z|\mu_t\right] \\ &= \frac{1 - \mu_t}{\mu_t} + \frac{(1 - \lambda)}{\lambda} z = L_t + \frac{(1 - \lambda)}{\lambda} z, \end{aligned}$$

proving the claim. ■

If there was no replacement possibility, then the likelihood ratio would be a martingale under the measure induced by the commitment type strategy. In our model with replacements likelihood ratio is not a martingale, however we still have a bound on how much the likelihood ratio on average may increase. This bound becomes very tight when λ is close to one (or when replacements are very unlikely), and not tight when λ is away from one (when replacements are very likely).

In the following, we make use of change of variables technique and introduce a time dependent variable S_t that differs from the likelihood ratio L_t by a time dependent term. For any real number x , $\lfloor x \rfloor$ denotes the largest integer less than or equal to x . Fix any $m > 0$. We specify the exact value of m in definition 5.

Definition 4 $S_t = \begin{cases} L_t + m - (\frac{1-\lambda}{\lambda})t & \text{if } t \leq T \equiv \left\lfloor \frac{m}{(\frac{1-\lambda}{\lambda})} \right\rfloor \\ S_T(\frac{1}{2})^{t-T} & \text{if } t > T \end{cases},$

The process S_t resembles the process L_t for $t \leq T$. In fact as we show below using lemma 4 S_t is a positive supermartingale.

Lemma 5 $E[S_{t+1}|S_t] \leq S_t$

Proof. For $t \leq T - 1$, $S_{t+1} - S_t = L_{t+1} - L_t - (\frac{1-\lambda}{\lambda})z$. Hence, $E[S_{t+1} - S_t|S_t] = E[L_{t+1} - L_t - (\frac{1-\lambda}{\lambda})z|L_t]$. By using lemma 4 we obtain $E[L_{t+1} - L_t - (\frac{1-\lambda}{\lambda})z|L_t] \leq 0$. Hence $E[S_{t+1} - S_t|S_t] \leq 0$. To prove the claim for $t \geq T$, we observe that $S_T \geq 0$ and then the result follows immediately. ■

To show that if $v_1(w, \alpha_2(L_t)) < u_1(w)$, then S_t is expected to jump: This subsection explores the relation between the expected amount by which the process S_t jumps at periods where player 1 receives low payoffs. Let $p_+(L_{t-1}) \in \Delta(O)$ be the probability distribution over outcomes generated by the strategy profile $(w, \alpha_2(L_{t-1}))$. Remember that $\alpha_1(L_{t-1}) \in W(L_{t-1})$, hence there is a $\alpha_1^*(L_{t-1}) \in \Delta(A_1)$ such that $\alpha_1(L_{t-1}) = w\mu_{t-1} + \alpha_1^*(L_{t-1})(1 - \mu_{t-1})$. In words, we can decompose the strategy $\alpha_1(L_{t-1})$ into two parts. The first part is the strategy of type w multiplied by the probability that player 1 is a type w , and the second part is a strategy $\alpha_1^*(L_{t-1})$

multiplied by the probability that player 1 is not a type w . Let $p_-(L_{t-1}) \in \Delta(O)$ be the probability distribution function over outcomes generated under the strategy profile $(\alpha_1^*(L_{t-1}), \alpha_2(L_{t-1}))$. In words, $p_-(L_{t-1})$ is the probability distribution function over outcomes conditional on player 1's type being different from w . For notational convenience we sometimes use $p_+(o)$ to denote the probability of outcome o when the state variable is L_{t-1} when it doesn't cause a confusion. Note that the strategy profile $(\alpha_1(L_{t-1}), \alpha_2(L_{t-1}))$ generates the probability distribution function $p(L_{t-1}) \equiv p_+(L_{t-1})\mu_{t-1} + p_-(L_{t-1})(1 - \mu_{t-1}) \in \Delta(O)$ over outcomes.

Let $d(L_{t-1})$ be the maximum difference in the probabilities of the outcomes between $p_+(L_{t-1})$ and $p_-(L_{t-1})$. $d(L_{t-1})$ serves as a measure of how the probability distribution functions generated by the commitment type strategy w and the strategy α_1 are apart from each other. In particular, let

$$\begin{aligned} d(L_{t-1}) &= \max_{o \in O} |p_+(o) - (\mu_{t-1}p_+(o) + (1 - \mu_{t-1})p_-(o))|, \text{ or} \\ d(L_{t-1}) &= \max_{o \in O} (1 - \mu_{t-1})|p_+(o) - p_-(o)|. \end{aligned} \tag{12}$$

Note that when the likelihood ratio L is very close to zero, μ is close to one, and $d(L)$ becomes very close to zero. This is very intuitive since if player 1 is believed to be type w with a very high probability, then the probability distribution function over outcomes generated by player 1's strategy will be very similar to the probability distribution over outcomes generated by type w . The following lemma shows that player 1 receives at least his commitment payoff $u_1(w)$ at period $t - 1$ if the player 2 believes that the probability distribution over outcomes at period t is very close to the probability distribution over outcomes generated by type w .

Lemma 6 *There exists $\xi > 0$ such that whenever $d(L_{t-1}) < \xi$, $v_1(w, \alpha_2(L_{t-1})) \geq u_1(w)$.*

Proof. This follows from assumption 1. To be more precise, the set of best responses is upper hemicontinuous hence we can choose $\varepsilon_1 > 0$ sufficiently small so that

$$\min_{\alpha_2 \in B(\alpha_1) \text{ for some } \alpha_1 \in \Delta(A_1) \text{ satisfying } \|\alpha_1(L) - w\| < \varepsilon_1} v_1(w, \alpha_2) \geq u_1(w)$$

Assumption 1 and the finiteness of the stage game imply that for every $\varepsilon > 0$, $\exists \gamma > 0$ such that $d(L) < \gamma$ implies that $\|\alpha_1(L) - w\| < \varepsilon$. Now choose $\varepsilon = \varepsilon_1$ and pick the corresponding $\gamma > 0$. If we set $\xi = \gamma$, then the claim is proven. ■

The following remark points out that if the reputation level is sufficiently high (or if the likelihood ratio is sufficiently close to zero), then the probability distribution function over outcomes will be very close to the probability distribution function induced by the strategy of type w .

Remark 1 *If $L_t \leq (1 - \xi)^{-1} - 1$ (or equivalently if $\mu_t \geq 1 - \xi$), then $d(L_t) < \xi$ by equation 12 above and by lemma 6 above $v_1(w, \alpha_2(L_t)) \geq u_1(w)$.*

Definition 5 Let m in definition 4 be equal to $\frac{(1-\xi)^{-1}-1}{2} > 0$.

In the following development we show that if player 1 gets a payoff lower than his commitment payoff $u_1(w)$, then the adjusted process S jumps by a non-trivial amount with a non-trivial probability. In particular, if $v_1(w, \alpha_2(L_{t-1})) < u_1(w)$ then $P\left(\frac{S_t}{S_{t-1}} - 1 \leq -(2/3)\frac{\xi}{|O|} | S_{t-1}\right) \geq (2/3)\frac{\xi}{|O|}$, or equivalently that $P\left(\frac{S_t}{S_{t-1}} - 1 \leq -(2/3)\frac{\xi}{|O|} | S_{t-1}\right) < (2/3)\frac{\xi}{|O|}$ implies $v_1(w, \alpha_2(L_{t-1})) \geq u_1(w)$.

Lemma 7 If $d(L_{t-1}) > \xi$ then $\exists o \in O$ such that $\frac{p_-(o)}{p_+(o)} \leq 1 - \frac{\xi}{|O|}$ and $p_+(o) \geq \frac{\xi}{|O|}$; and hence

$$P\left(\frac{p_-(o)}{p_+(o)} - 1 \leq -\frac{\xi}{|O|} | L_{t-1}\right) \geq \frac{\xi}{|O|}$$

Proof. If $d(L_{t-1}) > \xi$ then by equation 12, $\max_o |(1-\mu_{t-1})(p_+(o)-p_-(o))| > \xi$. Since $1-\mu_{t-1} \leq 1$, $\max_o |(p_+(o)-p_-(o))| > \xi$. Let $o_1 \in O$ be an outcome such that $|(p_+(o_1)-p_-(o_1))| > \xi$. Either $p_+(o_1) - p_-(o_1) > \xi$ or $p_-(o_1) - p_+(o_1) > \xi$.

case 1) If $p_+(o_1) - p_-(o_1) > \xi$, then $p_+(o_1) > \xi > \frac{\xi}{|O|}$, and $\frac{p_-(o_1)}{p_+(o_1)} \leq 1 - \xi < 1 - \frac{\xi}{|O|}$.

case 2) If $p_-(o_1) - p_+(o_1) > \xi$, then $\sum_{o \in O \setminus \{o_1\}} (p_+(o) - p_-(o)) > \xi$, and hence

$$(|O| - 1) \max_{o \in O \setminus \{o_1\}} (p_+(o) - p_-(o)) > \xi.$$

Let $o_2 \equiv \arg \max_{o \in O \setminus \{o_1\}} \{p_+(o) - p_-(o)\}$, then $p_+(o_2) - p_-(o_2) \geq \frac{\xi}{|O|-1} > \frac{\xi}{|O|}$ and so $p_+(o_2) > \frac{\xi}{|O|}$, and $\frac{p_-(o_2)}{p_+(o_2)} \leq 1 - \frac{\xi}{|O|}$.

Note that $P(o^{t-1}, o | L_{t-1}) = p_+(L_{t-1}(o^{t-1}))(o)$ so we have $P\left(\frac{p_-(o)}{p_+(o)} \leq 1 - \frac{\xi}{|O|} | L_{t-1}\right) \geq \frac{\xi}{|O|}$. ■

Lemma 7 shows that if the distribution over outcomes induced by strategies of type w and strategies of all remaining types differ from each other significantly, then it is significantly likely that the likelihood ratio will jump down significantly. We next prove a similar result for the adjusted process S_t . If the distribution over outcomes induced by the strategy of type w and that induced by the strategies of all remaining types differ significantly (i.e. $d(L_{t-1}) > \xi$), then it is significantly likely (i.e. with probability at least $(2/3)\frac{\xi}{|O|}$) that S_t will jump down significantly (i.e. will be multiplied by $1 - (2/3)\frac{\xi}{|O|}$).

Lemma 8 If $d(L_{t-1}) > \xi$ then $P\left(\frac{S_t}{S_{t-1}} - 1 \leq -(2/3)\frac{\xi}{|O|}\right) \geq \frac{\xi}{|O|} > (2/3)\frac{\xi}{|O|}$.

Proof. Consider an outcome $o^t = (o^{t-1}, o)$ for $t \leq T$,

$$\begin{aligned} \frac{S_t(o^t)}{S_{t-1}(o^{t-1})} &= \frac{L_t + m - \left(\frac{1-\lambda}{\lambda}z\right)t}{L_{t-1} + m - \left(\frac{1-\lambda}{\lambda}z\right)(t-1)} = \\ &\left(\frac{\frac{1-\mu_t}{\mu_t} + m - \left(\frac{1-\lambda}{\lambda}z\right)t}{\frac{1-\mu_{t-1}}{\mu_{t-1}} + m - \left(\frac{1-\lambda}{\lambda}z\right)(t-1)} \right) \leq \left(\frac{\frac{1-\mu_{t-1}}{\mu_{t-1}} \frac{p_-(L_{t-1})(o)}{p_+(L_{t-1})(o)} + m - \left(\frac{1-\lambda}{\lambda}z\right)(t-1)}{\frac{1-\mu_{t-1}}{\mu_{t-1}} + m - \left(\frac{1-\lambda}{\lambda}z\right)(t-1)} \right) \end{aligned}$$

The first two equalities follow from the definitions, and the last inequality is attained by the following argument. Note that

$$\frac{1 - \mu_t}{\mu_t} \leq \frac{\rho(y|\alpha_1(L_{t-1}), a_2)\alpha_2(a_2) - P(o^t|o^{t-1})\mu_{t-1}}{P(o^t|o^{t-1})\mu_{t-1}} + \frac{(1 - \lambda)}{\lambda}z$$

from inequality 11. Since player 1 and player 2's mixing probabilities are independent, and since $\alpha_1(L_{t-1}) = \mu_{t-1}w + (1 - \mu_{t-1})\alpha_1^*(L_{t-1})$, we have

$$\frac{\rho(y|\alpha_1(L_{t-1}), a_2)\alpha_2(a_2) - P(o^t|o^{t-1})\mu_{t-1}}{P(o^t|o^{t-1})\mu_{t-1}} = \frac{1 - \mu_{t-1}}{\mu_{t-1}} \frac{p_-(L_{t-1})(o)}{p_+(L_{t-1})(o)}.$$

Therefore we get the inequality by replacing $\frac{1 - \mu_t}{\mu_t} + m - (\frac{1 - \lambda}{\lambda}z)t$ by $\frac{1 - \mu_{t-1}}{\mu_{t-1}} \frac{p_-(L_{t-1})(o)}{p_+(L_{t-1})(o)} + m - (\frac{1 - \lambda}{\lambda}z)(t - 1)$. Note that $d(L_{t-1}) > \xi$ implies that $\mu_{t-1} < 1 - \xi$ by equation 12, and hence $L_{t-1} > (1 - \xi)^{-1} - 1 = 2m$

Let o be the outcome that satisfies $\frac{p_-(o)}{p_+(o)} \leq 1 - \frac{\xi}{|O|}$ and $p_+(o) \geq \frac{\xi}{|O|}$. Such an outcome exists by lemma 7. We have

$$\frac{p_-(o)}{p_+(o)} - 1 = \frac{\frac{1 - \mu_{t-1}}{\mu_{t-1}} \frac{p_-(L_{t-1})(o)}{p_+(L_{t-1})(o)}}{\frac{1 - \mu_{t-1}}{\mu_{t-1}}} - 1 \leq -\frac{\xi}{|O|} \quad (13)$$

Then,

$$\begin{aligned} \left(\frac{\frac{1 - \mu_{t-1}}{\mu_{t-1}} \frac{p_-(L_{t-1})(o)}{p_+(L_{t-1})(o)} + m - (\frac{1 - \lambda}{\lambda}z)(t - 1)}{\frac{1 - \mu_{t-1}}{\mu_{t-1}} + m - (\frac{1 - \lambda}{\lambda}z)(t - 1)} \right) &\leq \left(\frac{\frac{1 - \mu_{t-1}}{\mu_{t-1}} (1 - \frac{\xi}{|O|}) + m}{\frac{1 - \mu_{t-1}}{\mu_{t-1}} + m} \right) \leq \\ &\left(\frac{\frac{1 - \mu_{t-1}}{\mu_{t-1}} (1 - \frac{\xi}{|O|}) + \frac{1 - \mu_{t-1}}{\mu_{t-1}} \frac{1}{2}}{\frac{1 - \mu_{t-1}}{\mu_{t-1}} + \frac{1 - \mu_{t-1}}{\mu_{t-1}} \frac{1}{2}} \right) = \frac{(1.5 - \frac{\xi}{|O|})}{1.5} = 1 - (2/3) \frac{\xi}{|O|} \end{aligned}$$

The first inequality is attained by using inequality 13 to replace $\frac{p_-(L_{t-1})(o)}{p_+(L_{t-1})(o)}$ by $1 - \frac{\xi}{|O|}$, using the fact that $1 - \frac{\xi}{|O|} < 1$ and by adding $(\frac{1 - \lambda}{\lambda}z)(t - 1)$ to both the numerator and the denominator. The second inequality follows from setting $m = \frac{(1 - \xi)^{-1} - 1}{2}$, and using the fact that $\mu_{t-1} < 1 - \xi$ implies that $\frac{1 - \mu_{t-1}}{\mu_{t-1}} > (1 - \xi)^{-1} - 1 = 2m$, and then replacing m with the larger number $\frac{1 - \mu_{t-1}}{\mu_{t-1}} \frac{1}{2}$. Since $P(o^{t-1}, o|o^{t-1}) = p_+(L_{t-1})(o)$ we have that $P\left(\frac{S_t}{S_{t-1}} - 1 \leq -(2/3)\frac{\xi}{|O|}\right) \geq p_+(L_{t-1})(o) \geq \frac{\xi}{|O|}$ when $t \leq T$ and $d(L_{t-1}) > \xi$.

For $t > T$, $\frac{S_t}{S_{t-1}} \leq 1/2$, so for $\xi < \frac{3|O|}{4}$ we have $P\left(\frac{S_t}{S_{t-1}} - 1 \leq -(2/3)\frac{\xi}{|O|}\right) \geq \frac{\xi}{|O|}$. Hence, $P\left(\frac{S_t}{S_{t-1}} - 1 \leq -(2/3)\frac{\xi}{|O|}\right) \geq \frac{\xi}{|O|} > (2/3)\frac{\xi}{|O|}$ when $d(L_{t-1}) > \xi$ as desired. ■

In the following lemma we show that if the adjusted process S_t is not expected to jump down (i.e. multiplied by $1 - (2/3)\frac{\xi}{|O|}$) with a non-trivial probability $((2/3)\frac{\xi}{|O|})$ at time t , then player 2's play $(\alpha_2(L_{t-1}))$ gives player 1 at least his commitment payoff $u_1(w)$.

Lemma 9 *If $P\left(\frac{S_t}{S_{t-1}} - 1 \leq -(2/3)\frac{\xi}{|O|}\right) \leq (2/3)\frac{\xi}{|O|}$ then $v_1(w, \alpha_2(L_{t-1})) \geq u_1(w)$.*

Proof. Suppose to the contrary that $v_1(w, \alpha_2(L_{t-1})) < u_1(w)$, then by lemma 6 we have $d(L_{t-1}) > \xi$. Then using lemma 8 we have $P\left(\frac{S_t}{S_{t-1}} - 1 \leq -(2/3)\frac{\xi}{|O|}\right) > (2/3)\frac{\xi}{|O|}$ which is a contradiction. ■

Appendix B.3: Generating active supermartingales and the faster process $\tilde{S}_k$.

We define an increasing sequence of stopping times $\{\tau_k\}_{k=0}^\infty$ relative to $\{S_t\}_{t=0}^\infty$ and ξ as follows.

Definition 6 Let $\psi = (2/3)\frac{\xi}{2|O|} > 0$, and set $\tau_0 = 0$. Let $\tau_k(o^\infty) = \infty$ whenever $\tau_{k-1}(o^\infty) = \infty$. If $\tau_{k-1}(o^\infty) < \infty$, let $\tau_k(o^\infty)$ be the smallest integer $t > \tau_{k-1}(o^\infty)$ such that either:

Criterion 1 $P_Q(o^{t-1}) > 0$ and $P_Q(o^\infty : S_t/S_{t-1} - 1 \leq -2\psi|o^{t-1}) > 2\psi$ or

Criterion 2 $S_t/S_{\tau_{k-1}} - 1 \geq \psi$.

If there is no such t that satisfies either Criterion 1 or 2, then set $\tau_k(o^\infty) = \infty$.

Let the process $\{\tilde{S}_k\}_{k=0}^\infty$ be the stopped process with the stopping times $\{\tau_k\}_{k=0}^\infty$ and with the associated filtration $\tilde{\sigma}^t$, i.e., $\tilde{S}_k = S_{\tau_k}$ if $\tau_k < \infty$ and $\tilde{S}_k = 0$ otherwise.

Lemma 10 The stochastic process $\{\tilde{S}_k\}_{k=0}^\infty$ is a positive active supermartingale with activity $\psi > 0$ (under P).

Proof. $\{S_t\}$ is a positive supermartingale by lemma 5, $\{\tau_k\}_{k=0}^\infty$ is a stopping time, so the stopped process $\{\tilde{S}_k\}_{k=0}^\infty$ is a positive supermartingale. We will show in the following development that $\{\tilde{S}_k\}_{k=0}^\infty$ has activity ψ , i.e., that $P\left(\left|\frac{\tilde{S}_k}{\tilde{S}_{k-1}} - 1\right| \geq \psi | \tilde{S}_{k-1}\right) \geq \psi$ when $\tilde{S}_{k-1} > 0$.

If $\tau_{k-1} = \infty$, then $\tilde{S}_{k-1} = 0$, hence the claim is true vacuously. Suppose $\tau_{k-1} < \infty$. If $\tau_k = \infty$ then the claim is again true, so suppose $\tau_k < \infty$.

Let e_1 be the event that $\tau_k - 1 = \tau_{k-1}$. Then either Criterion 1 or 2 is used to pick τ_k . Conditional on Criterion 2 being used to pick τ_k we have

$$P\left(\left|\frac{\tilde{S}_k}{\tilde{S}_{k-1}} - 1\right| \geq \psi | \tilde{S}_{k-1}, e_1, \text{Criterion 2}\right) \geq \psi$$

trivially. Conditional on Criterion 1 being used to pick τ_k , we have

$$P\left(\frac{\tilde{S}_k}{\tilde{S}_{k-1}} - 1 \leq -2\psi | \tilde{S}_{k-1}, e_1, \text{Criterion 1}\right) \geq 2\psi,$$

and hence $P\left(\left|\frac{\tilde{S}_k}{\tilde{S}_{k-1}} - 1\right| \geq \psi | \tilde{S}_{k-1}, e_1, \text{Criterion 1}\right) \geq \psi$.

Let e_2 be the event that $\tau_k - 1 \neq \tau_{k-1}$ and $S_{\tau_k}/S_{\tau_{k-1}} - 1 \geq \psi$. Then again trivially

$$P\left(\left|\frac{\tilde{S}_k}{\tilde{S}_{k-1}} - 1\right| \geq \psi | \tilde{S}_{k-1}, e_2\right) \geq \psi.$$

Let e_3 be the event that $\tau_k - 1 \neq \tau_{k-1}$, $S_{\tau_k}/S_{\tau_{k-1}} - 1 \leq \psi$. The fact that $\tau_k - 1 \neq \tau_{k-1}$ implies $S_{\tau_k-1}/S_{\tau_{k-1}} - 1 < \psi$ because otherwise Criterion 2 would have been satisfied for $\tau_k - 1$, and the process would have stopped at $\tau_k - 1$. We know that Criterion 1 is used to pick S_{τ_k} , so

$$P\left(\frac{S_{\tau_k}}{S_{\tau_{k-1}}} - 1 \leq -2\psi \mid S_{\tau_{k-1}}, e_3\right) > 2\psi.$$

Note that $\frac{\tilde{S}_k}{\tilde{S}_{k-1}} = \frac{S_{\tau_k}}{S_{\tau_{k-1}}} = \frac{S_{\tau_k}}{S_{\tau_{k-1}}} \frac{S_{\tau_{k-1}}}{S_{\tau_{k-1}}}$. So if $\frac{S_{\tau_k}}{S_{\tau_{k-1}}} \leq 1 - 2\psi$ and $\frac{S_{\tau_{k-1}}}{S_{\tau_{k-1}}} < 1 + \psi$, then $\frac{\tilde{S}_k}{\tilde{S}_{k-1}} = \frac{S_{\tau_k}}{S_{\tau_{k-1}}} \frac{S_{\tau_{k-1}}}{S_{\tau_{k-1}}} < (1 - 2\psi)(1 + \psi) = 1 - \psi - 2\psi^2 < 1 - \psi$. Therefore

$$P\left(\left|\frac{\tilde{S}_k}{\tilde{S}_{k-1}} - 1\right| \geq \psi \mid \tilde{S}_{k-1}, e_3\right) > 2\psi > \psi.$$

Since the claim is true conditional on each of the above three events (e_1 , e_2 and e_3), and the three events exhaust all possibilities, the claim is true unconditionally. ■

Lemma 11 *If time $t \leq T$ is omitted (not picked) by $\tilde{S}$ from S , then $v_1(w, \alpha_2(L_t)) \geq u_1(w)$.*

Proof. Note that if $t \leq T$ is omitted (not picked) by $\tilde{S}$ from S then Criterion 1 of the 2 criteria used to pick the stopping time is not true for t . Therefore $P\left(\frac{S_{t+1}}{S_t} - 1 \leq -2\psi \mid o^t\right) \leq 2\psi$. Since $2\psi = 2/3 \frac{\xi}{|O|}$, by lemma 9 we have $v_1(w, \alpha_2(L_t)) \geq u_1(w)$ as desired. ■

Appendix B.4: Bounds on the value function by successive use of Theorem 2

In the following development, we will put bounds on the value function for different regions of the state variable L . For a reminder, let's express $V(L)$ as the infinite discounted sum of the expected payoffs of player 1, when he is mimicking type w , and player 2's action at each period t is determined by the policy function $\alpha_2(L_t)$, and the stochastic process $\{L_t\}_{t=0}^\infty$ follows Q^∞ defined on the space of histories of outcomes O^∞ , with an associated filtration $\{\tilde{O}_t\}_{t=0}^\infty$. In particular,

$$V(L) = E\left[(1 - \delta) \sum_{s=0}^{\infty} \delta^s v_1(w, \alpha_2(L_s))\right].$$

We start by putting a bound when L is very small, i.e., when the reputation level of player 1 is very close to 1. One thing to remember before we proceed is that, we restricted our state space to $[0, \bar{L}]$, and $V(L)$ is bounded, so $v = \inf_{L \leq \bar{L}} V(L)$ exists and $-\infty < v < \infty$. Note that if $\inf_{L' \leq \bar{L}} V(L') = v \geq u_1(w)$, then we have our desired result. So from hereon we assume that $v \leq u_1(w)$. Moreover for the following development, we'll make a working hypothesis that $v \leq (1 - \delta) \sum_{s=K-1}^{T/2-1} \delta^s u_1(w) + \delta^{T/2} v$.

Condition 1 (*Working Hypothesis*) $v \leq (1 - \delta) \sum_{s=K-1}^{T/2-1} \delta^s u_1(w) + \delta^{T/2} v$.

Let $\underline{L}_0 = (1 - \xi)^{-1} - 1$, $\underline{L}_1 = 2\underline{L}_0, \dots$, $\underline{L}_k = 2^k \underline{L}_0$, and let n be the integer such that $\bar{L} \in (\underline{L}_{n-1}, \underline{L}_n]$.¹⁶ In the following development, we'll partition the state space into intervals of the form $[0, \underline{L}_0]$, $(\underline{L}_0, \underline{L}_1]$, $\dots$, $(\underline{L}_{n-1}, \bar{L}]$, and put lower bounds on $V(L)$ for each of these intervals. Also for the following, we fix $\varepsilon > 0$ and choose $R = 2$ in theorem 2, and pick the K whose existence is claimed in the theorem. Let the process $\{S_t\}$ be defined as in Definition 4, stopping times $\{\tau_k\}$ and the corresponding stopped times $\{\tilde{S}_t\}$ be defined as in Definition 6.

Claim 2 $V(L) \geq E \left[(1 - \delta) \left(\sum_{s=K-1}^{\min\{\tau_K-1, T/2-1\}=N} \delta^s u_1(w) \right) + \delta^{N+1} V(L_{N+1}) \right]$

Proof. For any $t \leq T/2$, if t is not picked by the stopping time, then $v_1(w, \alpha_2(L_t)) \geq u_1(w)$. Until time N , the process has stopped at most $K - 1$ times since $N < \tau_K$. Therefore, the summation is lower when we replace player 1's payoff by 0 at exactly $K - 1$ periods, and if these $K - 1$ periods are at the beginning of the sequence, at periods $\{0, 1, \dots, K - 2\}$ instead of some later periods, delivering the desired inequality. ■

Definition 7 We define the following four events:

- $E_1 : \tau_K \leq T/2$ and $\tilde{S}_K > \tilde{S}_0/2$.
- $E_2 : \tau_K > T/2$.
- $E_3 : \tau_K \leq T/2$ and $\tilde{S}_K \leq \tilde{S}_0/2$.
- $E_4 \subset E_3 : \tau_K \leq T/2$ and $\tilde{S}_k \leq \tilde{S}_0/2$ for every $k \geq K$.

Also for an event e , e^c denotes the complement of e and let $V(L|e)$ denote

$$E \left[(1 - \delta) \sum_{s=0}^{\infty} \delta^s v_1(w, \alpha_2(L_s)) | e \right].$$

Claim 3

$$\begin{aligned} i) \quad V(L|E_1) &\geq \min_{K-1 \leq k \leq T/2} (1 - \delta) \left(\sum_{s=K-1}^{k-1} \delta^s u_1(w) \right) + \delta^k v \\ ii) \quad V(L|E_2) &\geq (1 - \delta) \left(\sum_{s=K-1}^{T/2-1} \delta^s u_1(w) \right) + \delta^{T/2} v \\ iii) \quad V(L|E_3) &\geq \min_{K-1 \leq k \leq T/2} (1 - \delta) \left(\sum_{s=K-1}^{k-1} \delta^s u_1(w) \right) + \delta^k \inf_{L' \leq \frac{L}{2}} V(L') \end{aligned}$$

¹⁶We can actually calculate the integer n , by using the inequalities that $2^{n-1} \underline{L}_0 < \bar{L}$ and $2^n \underline{L}_0 \geq \bar{L}$.

Proof. *i)* The event E_1 says that the period $t = \tau_K$ at which the process stopped for the K^{th} time is not larger than $T/2$, and $\tilde{S}_K = S_t > \tilde{S}_0/2 = \frac{S_0}{2}$.¹⁷ In such a case, using claim 2 we get

$$\begin{aligned} V(L|E_1) &\geq E \left[(1-\delta) \left(\sum_{s=K-1}^{t-1} \delta^s u_1(w) \right) + \delta^t V(L_{t+1}) | \tau_K = t \right] \geq (1-\delta) \left(\sum_{s=K-1}^{t-1} \delta^s u_1(w) \right) + \delta^t v \\ &\geq \min_{K-1 \leq k \leq T/2} (1-\delta) \left(\sum_{s=K-1}^{k-1} \delta^s u_1(w) \right) + \delta^k v. \end{aligned}$$

ii) The event E_2 says that the period t at which the process stopped K^{th} time is larger than $T/2$. Then, using Claim 2 we get:

$$V(L|E_2) \geq (1-\delta) \left(\sum_{s=K-1}^{T/2-1} \delta^s u_1(w) \right) + \delta^{T/2} v.$$

iii) The event E_3 says that the period t at which the process stopped K^{th} time is not larger than $T/2$, and $\tilde{S}_{\tau_K} \leq \tilde{S}_0/2$, so $S_t \leq \frac{S_0}{2}$. For $t \leq T/2$, $S_t \leq \frac{S_0}{2}$ implies $L_t + m - t(\frac{1-\lambda}{\lambda}z) \leq \frac{L+m}{2}$. Rearranging, we have $L_t - \frac{L}{2} \leq -\frac{m}{2} + t(\frac{1-\lambda}{\lambda}z)$. Note that since $t \leq T/2$, we have $-\frac{m}{2} + t(\frac{1-\lambda}{\lambda}z) \leq 0$, hence $L_t \leq \frac{L}{2}$. Then, using the bound in Claim 2 we get:

$$V(L|E_3) \geq (1-\delta) \left(\sum_{s=K-1}^{t-1} \delta^s u_1(w) \right) + \delta^t V(L_t) \geq \min_{K-1 \leq k \leq T/2} (1-\delta) \left(\sum_{s=K-1}^{k-1} \delta^s u_1(w) \right) + \delta^k \inf_{L' \leq \frac{L}{2}} V(L').$$

■

Claim 4 *Interval $[0, \underline{L}_0]$: For $L \leq (1-\xi)^{-1} - 1$, we have the following lower bound:*

$$V(L) \geq (1-\varepsilon) \left[(1-\delta) \sum_{s=K-1}^{T/2-1} \delta^s u_1(w) + \delta^{T/2} v \right] + \varepsilon \delta^{K-1} v.$$

Proof. By Theorem 2 and the fact that $\tilde{S}_k$ is a positive supermartingale with activity ψ (by Lemma 10) we deduce that $P(\tilde{S}_k \leq \tilde{S}_0/2 \text{ for every } k \geq K) > 1-\varepsilon$. The event $\tilde{S}_k \leq \tilde{S}_0/2$ for every $k \geq K$ is a subset of $E_2 \cup E_4$, therefore $P(E_2 \cup E_4) > 1-\varepsilon$. Now note that, conditional on the event E_4 , during the first $T/2$ periods, the process stops at least K times, and let t be the K^{th} time the process stops. At any time t' such that $T/2 \geq t' \geq t$, if t' is not picked by the stopping process, then $v_1(w, \alpha_2(L_{t'})) \geq u_1(w)$. If t' is picked by the process, let's say for the k'^{th} time where $k' \geq K$, then by the definition of the event E_4 , $\tilde{S}_{k'} \leq \tilde{S}_0/2$. The implication of this is $S_{t'} \leq S_0/2$, and $L_{t'} + m - t'(\frac{1-\lambda}{\lambda}z) \leq \frac{L+m}{2}$. Using $t' \leq T/2$ yields $L_{t'} \leq L/2$. Now notice that since $L \leq (1-\xi)^{-1} - 1$, $L_{t'} \leq \frac{(1-\xi)^{-1}-1}{2}$ and therefore $d(L_{t'}) < \xi$ by equation 12 and $v_1(w, \alpha_2(L_{t'})) \geq u_1(w)$ by lemma 6.

¹⁷Remember that $\tau_0 = 0$, hence $\tilde{S}_0 = S_0$.

Therefore, conditional on E_4 , during the first $T/2$ periods, player 1 gets a payoff less than $u_1(w)$ for at most $K - 1$ periods. Conditional on E_2 , there are less than K periods where the process is stopped before $T/2$. Conditional on the complement of $(E_2 \cup E_4)$, during the first $T/2$ periods (note that periods start from zero), the first time when the process is stopped for some k^{th} time where $k' \geq K$ and $\tilde{S}_{k'} > \tilde{S}_0/2$ occurs at some period $t \leq T/2$. Until period t , player 1's payoffs are less than $u_1(w)$ at ost $K - 1$ times since a period $t' < t$ is either not picked by the process, or if picked, then $S_{t'} > S_0/2$. We have therefore:

$$\begin{aligned} V(L|(E_2 \cup E_4)) &\geq (1 - \delta) \sum_{s=K-1}^{T/2-1} \delta^s u_1(w) + \delta^{T/2} v, \\ V(L|(E_2 \cup E_4)^c) &\geq \min_{K-1 \leq k \leq T/2} \left((1 - \delta) \sum_{s=K-1}^{k-1} \delta^s u_1(w) + \delta^k v \right) = \delta^{K-1} v. \end{aligned}$$

The last equality follows from the assumption that $v < u_1(w)$. Note that under Condition 1, $\delta^{K-1} v \leq (1 - \delta) \sum_{s=K-1}^{T/2-1} \delta^s u_1(w) + \delta^{T/2} v$. The total probabily of E_2 and E_4 is at least $1 - \varepsilon$, therefore we get:

$$\begin{aligned} V(L) &\geq P(E_2 \cup E_4) \left[(1 - \delta) \sum_{s=K-1}^{T/2-1} \delta^s u_1(w) + \delta^{T/2} v \right] + (1 - P(E_2 \cup E_4)) (\delta^{K-1} v) \\ &\geq (1 - \varepsilon) \left[(1 - \delta) \sum_{s=K-1}^{T/2-1} \delta^s u_1(w) + \delta^{T/2} v \right] + \varepsilon \delta^{K-1} v. \end{aligned}$$

■

Definition 8 Let F_m be a sequence defined recursively by $F_m = (1 - \varepsilon) \delta^{K-1} \min\{F_0, \dots, F_{m-1}\} + \varepsilon \delta^{K-1} v$ for $m \geq 1$ and

$$F_0 = (1 - \varepsilon) \left[(1 - \delta) \sum_{s=K-1}^{T/2-1} \delta^s u_1(w) + \delta^{T/2} v \right] + \varepsilon \delta^{K-1} v.$$

Note that $F_0 \leq u_1(w)$ because $v \leq u_1(w)$.¹⁸ Moreover $F_m \leq u_1(w)$ for every $m \geq 0$ from the definition of the sequence.

Claim 5 Interval $(\underline{L}_i, \underline{L}_{i+1}]$: For $L \in (\underline{L}_i, \underline{L}_{i+1}]$, if Condition 1 holds true, then we have the following lower bound:

$$V(L) \geq F_{i+1} = (1 - \varepsilon) \delta^{K-1} \min\{F_0, F_1, \dots, F_i\} + \varepsilon \delta^{K-1} v$$

¹⁸To see this more clearly, note that the sum of the coeffiecients that multiply period payoffs and the continuation payoff v is less than one, and period payoffs are not more than $u_1(w)$ and v is less than $u_1(w)$.

Proof. We'll prove the claim by induction. E_1, E_2, E_3 exhaust all possibilities, so $P(E_1 \cup E_2 \cup E_3) = 1$, moreover by Theorem 2 we have $P(E_1) < \varepsilon$, therefore $P(E_2 \cup E_3) > 1 - \varepsilon$.

Using claim 3, we have

$$\begin{aligned} i) \quad V(L|E_1) &\geq \min_{K-1 \leq k \leq T/2} (1 - \delta) \left(\sum_{s=K-1}^{k-1} \delta^s u_1(w) \right) + \delta^k v \\ ii) \quad V(L|E_2) &\geq (1 - \delta) \left(\sum_{s=K-1}^{T/2-1} \delta^s u_1(w) \right) + \delta^{T/2} v \\ iii) \quad V(L|E_3) &\geq \min_{K-1 \leq k \leq T/2} (1 - \delta) \left(\sum_{s=K-1}^{k-1} \delta^s u_1(w) \right) + \delta^k \inf_{L' \leq \frac{L}{2}} V(L') \end{aligned}$$

Note that $V(L|E_i) \geq \delta^{K-1} v$ for $i \in \{1, 2, 3\}$ since $v \leq u_1(w)$.

For $i = 0$: $V(L|E_2) \geq F_0$ using the definition and the Condition 1. Hence we have $V(L|E_2) \geq \delta^{K-1} F_0$. We also have $V(L|E_3) \geq \min_{K-1 \leq k \leq T/2} (1 - \delta) \left(\sum_{s=K-1}^{k-1} \delta^s u_1(w) \right) + \delta^k \inf_{L' \leq \frac{L}{2}} V(L') \geq \min_{K-1 \leq k \leq T/2} (1 - \delta) \left(\sum_{s=K-1}^{k-1} \delta^s u_1(w) \right) + \delta^k F_0$ using claim 4. Since $F_0 < u_1(w)$, we have $V(L|E_3) \geq \delta^{K-1} F_0$. Hence we have $V(L) \geq P(E_2 \cup E_3) \max\{\delta^{K-1} v, \delta^{K-1} F_0\} + P(E_1)(\delta^{K-1} v)$. Combining with the fact that $P(E_2 \cup E_3) > 1 - \varepsilon$ yields $V(L) \geq (1 - \varepsilon)\delta^{K-1} F_0 + \varepsilon\delta^{K-1} v$.

Assume the claim is true for all $i \leq I$. Then for $i = I + 1$, we first have $V(L|E_3) \geq \min_{K-1 \leq k \leq T/2} (1 - \delta) \left(\sum_{s=K-1}^{k-1} \delta^s u_1(w) \right) + \delta^k \inf_{L' \leq \frac{L}{2}} V(L') \geq \min_{K-1 \leq k \leq T/2} (1 - \delta) \left(\sum_{s=K-1}^{k-1} \delta^s u_1(w) \right) + \delta^k \min\{F_0, F_1, \dots, F_I\}$ by the induction hypothesis, and then $V(L|E_3) \geq \delta^{K-1} \min\{F_0, F_1, \dots, F_I\}$ by the fact that $F_0 < u_1(w)$. Second, we have $V(L|E_2) \geq (1 - \delta) \left(\sum_{s=K-1}^{T/2-1} \delta^s u_1(w) \right) + \delta^{T/2} v$. Note that under Condition 1, the definition of the sequence $F_m = (1 - \varepsilon)\delta^{K-1} \min\{F_0, \dots, F_{m-1}\} + \varepsilon\delta^{K-1} v$ implies that $F_m \leq (1 - \delta) \left(\sum_{s=K-1}^{T/2-1} \delta^s u_1(w) \right) + \delta^{T/2} v$ for every $m \geq 0$. Combining the above two lines, we get $V(L|E_2) \geq \delta^{K-1} \min\{F_0, F_1, \dots, F_I\}$. Now using the fact that $V(L) \geq P(E_2 \cup E_3) \max\{\delta^{K-1} \min\{F_0, F_1, \dots, F_I\}, \delta^{K-1} v\} + P(E_1)\delta^{K-1} v$, and that $P(E_2 \cup E_3) \geq 1 - \varepsilon$, we get:

$$\begin{aligned} V(L) &\geq (1 - \varepsilon)\delta^{K-1} \min\{F_0, F_1, \dots, F_I\} + \varepsilon\delta^{K-1} v \\ &= F_{I+1} \end{aligned}$$

■

Proposition 1 *Under Condition 1 we have:*

$$v \geq ((1 - \varepsilon)\delta^{K-1})^n F_0 + \varepsilon\delta^{K-1} v \sum_{s=0}^{n-1} ((1 - \varepsilon)\delta^{K-1})^s$$

Proof. Suppose the Proposition is not true. Then collecting terms and rearranging, we get

$$\begin{aligned}
& v \left[1 - \varepsilon \delta^{K-1} \sum_{s=0}^{n-1} ((1-\varepsilon)\delta^{K-1})^s - ((1-\varepsilon)\delta^{K-1})^n \left((1-\varepsilon)\delta^{T/2} + \varepsilon\delta^{K-1} \right) \right] < \\
& (1-\varepsilon)((1-\varepsilon)\delta^{K-1})^n(1-\delta) \sum_{s=K-1}^{T/2-1} \delta^s u_1(w) \\
\iff v & < \frac{(1-\varepsilon)((1-\varepsilon)\delta^{K-1})^n(1-\delta) \sum_{s=K-1}^{T/2-1} \delta^s u_1(w)}{1 - \varepsilon \delta^{K-1} \sum_{s=0}^{n-1} ((1-\varepsilon)\delta^{K-1})^s - ((1-\varepsilon)\delta^{K-1})^n \left((1-\varepsilon)\delta^{T/2} + \varepsilon\delta^{K-1} \right)}. \quad (14)
\end{aligned}$$

Our first observation is that $g(i) \equiv \frac{(1-\varepsilon)((1-\varepsilon)\delta^{K-1})^i(1-\delta) \sum_{s=K-1}^{T/2-1} \delta^s u_1(w)}{1 - \varepsilon \delta^{K-1} \sum_{s=0}^{i-1} ((1-\varepsilon)\delta^{K-1})^s - ((1-\varepsilon)\delta^{K-1})^i \left((1-\varepsilon)\delta^{T/2} + \varepsilon\delta^{K-1} \right)}$ is a decreasing function in i . Therefore, $v < g(i)$ for every $i \leq n$.

The first step of the proof is to show that $F_m \leq F_{m-1}$, $m = 0, 1, \dots, n$. We will use induction. To see this, let's first consider $m = 0$ and $m = 1$. We can see that $v < F_0$, $F_1 < F_0$ and $v < F_1$ from direct calculation, using inequality 14. Moreover,

$$\begin{aligned}
F_2 &= (1-\varepsilon)\delta^{K-1} \min\{F_1, F_0\} + \varepsilon\delta^{K-1}v \\
&= (1-\varepsilon)\delta^{K-1}F_1 + \varepsilon\delta^{K-1}v \\
&= ((1-\varepsilon)\delta^{K-1})^2 F_0 + \varepsilon\delta^{K-1}v \sum_{s=0}^1 ((1-\varepsilon)\delta^{K-1})^s
\end{aligned} \quad (15)$$

We can thus see that $F_2 \leq F_1$ (from the second line), as well as $F_2 > v$ using inequality 14. Now, suppose by induction hypothesis that $F_l \leq F_{l-1}$, $l = 3, \dots, m$. This implies

$$F_l = (1-\varepsilon)\delta^{K-1}F_{l-1} + \varepsilon\delta^{K-1}v, \quad l = 1, \dots, m$$

and recursive substitution leads to

$$F_l = ((1-\varepsilon)\delta^{K-1})^l F_0 + \varepsilon\delta^{K-1}v \sum_{s=0}^{l-1} ((1-\varepsilon)\delta^{K-1})^s > v \text{ by inequality 14 for every } l = 1, \dots, m$$

In particular, $F_m > v$. Therefore, $F_{m+1} = (1-\varepsilon)\delta^{K-1}F_m + \varepsilon\delta^{K-1}v \leq F_m$, proving the induction step.

Finally, since F_m is a decreasing sequence, by recursive substitution we get:

$$F_n = ((1-\varepsilon)\delta^{K-1})^n F_0 + \varepsilon\delta^{K-1}v \sum_{s=0}^{n-1} ((1-\varepsilon)\delta^{K-1})^s.$$

On the other side since $v = \inf_{L \in S} V(L)$, using claim 5 for every interval contained in S we get $v \geq \min\{F_0, F_1, \dots, F_n\}$ and we have

$$v \geq ((1 - \varepsilon)\delta^{K-1})^n F_0 + \varepsilon\delta^{K-1}v \sum_{s=0}^{n-1} ((1 - \varepsilon)\delta^{K-1})^s$$

which is a contradiction. ■

Claim 6 $n - 1 < \frac{-[\ln(1-\lambda) + \ln(\eta^*(w)) + \ln((1-\xi)^{-1}-1)]}{\ln 2}$

Proof. since $\bar{L} = \frac{1-(1-\lambda)\eta^*(w)}{(1-\lambda)\eta^*(w)}$ and $\underline{L}_0 = (1 - \xi)^{-1} - 1$,

$$\begin{aligned} 2^{n-1}\underline{L}_0 &\leq \bar{L} = \frac{1 - (1 - \lambda)\eta^*(w)}{(1 - \lambda)\eta^*(w)} < \frac{1}{(1 - \lambda)\eta^*(w)}, \text{ hence} \\ n - 1 &< \frac{-[\ln(1 - \lambda) + \ln(\eta^*(w)) + \ln((1 - \xi)^{-1} - 1)]}{\ln 2}. \end{aligned}$$

■

Note that the choice of ξ is independent of δ and λ , so we can do the same analysis for every discount factor δ and replacement probability λ . Moreover, remember that we fixed ε and picked the K from Theorem 2, so the only variable that ε affects is K . Let $\{v_r\}$ be the sequence of numbers, each corresponding to the infimum of player 1's Nash equilibrium payoffs in a sequence of repeated games parameterized by λ_r and δ_r . Let

$$m_r = (1 - \varepsilon)\delta_r^{K-1} \text{ and } C(\lambda_r, \delta_r) = (1 - \delta_r) \left(\sum_{s=K-1}^{T_r/2-1} \delta_r^s \right).$$

Then we have a sequence of inequalities,

$$v_r \geq \varepsilon\delta_r^{K-1} \left(\sum_{s=0}^{n-1} m_r^s \right) v_r + m_r^n \left[(1 - \varepsilon)C(\lambda_r, \delta_r)u_1(w) + (1 - \varepsilon)\delta_r^{T_r/2}v_r + \varepsilon\delta_r^{K-1}v_r \right] \quad (16)$$

for every $\varepsilon > 0$, where $T_r = \frac{m}{z} \frac{\lambda_r}{1-\lambda_r}$ and K depends on ε and neither n nor T_r depends on ε . Recollecting terms we have,

$$v_r \left(1 - \frac{m_r \varepsilon}{1 - \varepsilon} \frac{1 - m_r^n}{1 - m_r} - m_r^n \left((1 - \varepsilon)\delta_r^{T_r/2} + \varepsilon\delta_r^{K-1} \right) \right) \geq m_r^n (1 - \varepsilon)C(\lambda_r, \delta_r)u_1(w).$$

Note that, if the replacement rate is constant along the sequence r , $\lambda_r = \lambda < 1$ and $T_r = T$ and if $\delta_r \rightarrow 1$ then we have $\lim_{\delta_r \rightarrow 1} \frac{m_r^n (1-\varepsilon)C(\lambda_r, \delta_r)}{\left(1 - \frac{m_r \varepsilon}{1-\varepsilon} \frac{1-m_r^n}{1-m_r} - m_r^n \left((1-\varepsilon)\delta_r^{T_r/2} + \varepsilon\delta_r^{K-1}\right)\right)} = \frac{T/2-(K-1)}{(1-\varepsilon)T/2+\varepsilon(K-1)}$. Note that the limit points approach to one if $\lambda \rightarrow 1$ (and hence if $T \rightarrow \infty$). For the following, we will however show that, even if the replacement rate is approaching to zero as the discount factor is approaching

to one, v_r in the limit is bounded below by $u_1(w)$. In particular assume that $\lambda_r \rightarrow 1$ and that $\inf \lim \frac{(1-\lambda_r)}{(1-\delta_r)} > 0$. We show in Appendix C that in this case

$$\liminf_{r \rightarrow \infty} v_r \geq u_1(w).$$

Note if Condition 1 is not true, then we have a tighter bound (a larger lower bound) for v then inequality 16, hence $\liminf_{r \rightarrow \infty} v_r \geq u_1(w)$.

Appendix C

Suppose $\liminf_{r \rightarrow \infty} v_r < u_1(w)$. Then there exists a subsequence $\{r^k\}$ of $\{r\}$ such that the subsequence $\{v_{r^k}\}$ is convergent and that $\lim_{k \rightarrow \infty} v_{r^k} < u_1(w)$. Rename the sequence to be $\{r\}$. Now we can pick a subsequence of the new sequence r, r^k , such that either $\lim_{k \rightarrow \infty} \frac{1-\lambda_{r^k}}{1-\delta_{r^k}} = k^*$ with $\infty > k^* > 0$ or $\lim_{k \rightarrow \infty} \frac{1-\lambda_{r^k}}{1-\delta_{r^k}} = \infty$. We can choose this subsequence since $\liminf_{r \rightarrow \infty} \frac{1-\lambda_r}{1-\delta_r} > 0$. Rename this subsequence r again, and note that since this is a subsequence of the second sequence we have $\lim_{r \rightarrow \infty} v_r < u_1(w)$. We will now prove that, on the contrary, along the sequence r , $\lim_{r \rightarrow \infty} v_r \geq u_1(w)$. Fix $\varepsilon > 0$. In the following, for any sequence of numbers x_r , $\lim x_r$ denotes $\lim_{r \rightarrow \infty} x_r$. Note that $\lim m_r = 1 - \varepsilon$.

Rearranging inequality 16 we get:

$$v_r \left(1 - \frac{m_r \varepsilon}{1 - \varepsilon} \frac{1 - m_r^n}{1 - m_r} - m_r^n \left((1 - \varepsilon) \delta_r^{T/2} + \varepsilon \delta_r^{K-1} \right) \right) \geq m_r^n (1 - \varepsilon) C(\lambda_r, \delta_r) u_1(w),$$

or:

$$v_r \left(\frac{(1 - \delta_r^{K-1})}{(1 - m_r) m_r^n (1 - \varepsilon)} - \left(\delta_r^{T/2} + \frac{\varepsilon}{1 - \varepsilon} \delta_r^{K-1} \right) + \frac{\varepsilon m_r}{(1 - \varepsilon)(1 - m_r)(1 - \varepsilon)} \right) \geq C(\lambda_r, \delta_r) u_1(w).$$

Case 1: $\lim \frac{1-\lambda_r}{1-\delta_r} \in (0, \infty)$:

Claim 7 $\lim C(\lambda_r, \delta_r) = \lim 1 - \delta_r^{T/2} > 0$.

Proof. First note that $C(\lambda_r, \delta_r) - (1 - \delta_r^{T/2}) = (1 - \delta_r) \sum_{s=0}^{K-2} \delta_r^s$ and $\lim (1 - \delta_r) \sum_{s=0}^{K-2} \delta_r^s = 0$, so

$\lim C(\lambda_r, \delta_r) = \lim 1 - \delta_r^{T/2}$. Moreover, $\lim \delta_r^{\frac{1}{1-\lambda_r}} < 1$ by L'Hopital's rule and by $\lim \frac{1-\lambda_r}{1-\delta_r} > 0$.

¹⁹Therefore $\lim 1 - \delta_r^{T/2} > 0$. ■

Claim 8 $\lim \frac{(1 - \delta_r^{K-1})}{(1 - m_r) m_r^n (1 - \varepsilon)} = 0$ if ε is sufficiently small.

¹⁹To see this, let $y_r = \delta_r^{\frac{1}{1-\lambda_r}}$. Taking logs gives $\ln y_r = \frac{1}{1-\lambda_r} \ln \delta_r$. Note that under case 1, $\lim d\lambda_r/d\delta_r \in (0, \infty)$ and hence applying L'Hopital's rule yields $\lim \ln y_r = \lim \frac{\delta_r^{-1}}{-d\lambda_r/d\delta_r} < 0$ and hence $\lim y_r < 1$.

Proof. $\lim \frac{(1-\delta_r^{K-1})}{(1-m_r)m_r^n(1-\varepsilon)} = \lim \frac{1}{(1-m_r)(1-\varepsilon)} \lim \frac{(1-\delta_r^{K-1})}{m_r^n}$, and $\lim \frac{1}{(1-m_r)(1-\varepsilon)} = \frac{1}{(1-\varepsilon)\varepsilon} < \infty$. Let $\frac{(1-\delta_r^{K-1})}{m_r^n} = y_r$, then $\ln y_r = \ln(1 - \delta_r^{K-1}) - n \ln m_r$. Note that $\lim \ln y_r \leq \lim \{\ln(1 - \delta_r^{K-1}) + \ln(1 - \lambda_r) \ln(1 - \varepsilon) + (K - 1) \ln \delta_r \ln(1 - \lambda_r)\} = -\infty$ if ε is sufficiently close to zero.²⁰ Therefore, $\lim y_r = 0$, and hence $\lim \frac{(1-\delta_r^{K-1})}{(1-m_r)m_r^n(1-\varepsilon)} = 0$. ■

Claim 9 $\lim \frac{\varepsilon m_r}{(1-\varepsilon)(1-m_r)(1-\varepsilon)} - \frac{\varepsilon}{1-\varepsilon} \delta_r^{K-1} = 1$ for each $\varepsilon > 0$.

Proof. $\lim m_r = 1 - \varepsilon$ and $\lim \frac{\varepsilon}{1-\varepsilon} \delta_r^{K-1} = \frac{\varepsilon}{1-\varepsilon}$, therefore the result follows. ■

Claim 10 $\lim v_r \geq u_1(w)$

Proof. We first get for every $\varepsilon > 0$, $\lim v_r \geq \frac{\lim C(\lambda_r, \delta_r)}{\lim 1 - \delta_r^{T_r/2}} u_1(w)$ and noting that $\lim C(\lambda_r, \delta_r) = \lim 1 - \delta_r^{T_r/2+1}$, we conclude that $\lim v_r \geq u_1(w)$. ■

Case 2: $\lim \frac{1-\lambda_r}{1-\delta_r} = \infty$: This case is a bit more tricky.

Claim 11 $\lim \frac{(1-\delta_r^{K-1})}{(1-m_r)m_r^n(1-\varepsilon)} \frac{1}{C(\lambda_r, \delta_r)} = 0$ if ε is sufficiently small.

Proof. $\lim \frac{(1-\delta_r^{K-1})}{(1-m_r)m_r^n(1-\varepsilon)} \frac{1}{C(\lambda_r, \delta_r)} = \lim \frac{1}{(1-m_r)(1-\varepsilon)} \lim \frac{(1-\delta_r^{K-1})}{m_r^n C(\lambda_r, \delta_r)}$ and $\lim \frac{1}{(1-m_r)(1-\varepsilon)} = \frac{1}{(1-\varepsilon)\varepsilon} < \infty$. Since $C(\lambda_r, \delta_r) = \delta_r^{K-1} - \delta_r^{T_r/2}$, $\limsup \frac{T_r}{2} (1 - \lambda_r) < \infty$ and $\liminf \delta_r^{T_r/2} = 1$, $\limsup \frac{(1-\delta_r^{K-1})}{C(\lambda_r, \delta_r)} \frac{1}{1-\lambda_r} < \infty$. Moreover if $\lambda_r \rightarrow 1$, then $\lim \frac{(1-\lambda_r)}{(1-\varepsilon) - \ln(1-\lambda_r)\delta_r^{-(K-1)\ln(1-\lambda_r)}} = 0$ if ε is sufficiently small.²¹ Therefore, $\limsup \frac{(1-\delta_r^{K-1})}{m_r^n C(\lambda_r, \delta_r)} = \limsup \frac{(1-\delta_r^{K-1})}{C(\lambda_r, \delta_r)} \frac{1}{1-\lambda_r} \lim \frac{(1-\lambda_r)}{(1-\varepsilon) - \ln(1-\lambda_r)\delta_r^{-(K-1)\ln(1-\lambda_r)}} = 0$, and therefore $\lim \frac{(1-\delta_r^{K-1})}{m_r^n C(\lambda_r, \delta_r)} = 0$. ■

Claim 12 $\lim v_r \geq u_1(w)$.

Proof. Let's first show that $\lim \frac{\frac{\varepsilon m_r}{(1-\varepsilon)(1-m_r)(1-\varepsilon)} - \left(\delta_r^{T_r/2} + \frac{\varepsilon}{1-\varepsilon} \delta_r^{K-1}\right)}{C(\lambda_r, \delta_r)} = 1$. Simplification delivers that

$$\frac{\frac{\varepsilon m_r}{(1-\varepsilon)(1-m_r)(1-\varepsilon)} - \left(\delta_r^{T_r/2} + \frac{\varepsilon}{1-\varepsilon} \delta_r^{K-1}\right)}{C(\lambda_r, \delta_r)} = 1 - \frac{1 - \delta_r^{K-1}}{\left(1 - \delta_r^{T_r/2-(K-1)}\right)},$$

and $\lim \frac{1-\delta_r^{K-1}}{\left(1 - \delta_r^{T_r/2-(K-1)}\right)} = 0$ delivering the first step. Since

$$\lim \frac{(1 - \delta_r^{K-1})}{(1 - m_r)m_r^n(1 - \varepsilon)} \frac{1}{C(\lambda_r, \delta_r)} = 0$$

²⁰This step uses the fact that $\liminf \frac{\ln(1-\delta_r^{K-1})}{\ln(1-\lambda_r)} > 0$.

²¹To see this, we take the log of the expression whose limit we are taking, and we get $\ln(1 - \lambda_r) + \ln(1 - \lambda_r) \ln(1 - \varepsilon) + (K - 1) \ln(1 - \lambda_r) \ln \delta_r$. Or equivalently, $\ln(1 - \lambda_r)[1 + \ln(1 - \varepsilon) + \ln \delta_r]$. The term within the square brackets is positive for small ε and δ_r close to one, and $\ln(1 - \lambda_r)$ goes to $-\infty$. Since this is the log of the expression, the expression goes to zero.

for small ε , and

$$\lim_{\varepsilon \rightarrow 0} \frac{\frac{\varepsilon m_r}{(1-\varepsilon)(1-m_r)(1-\varepsilon)} - \left(\delta_r^{T_r/2} + \frac{\varepsilon}{1-\varepsilon} \delta_r^{K-1} \right)}{C(\lambda_r, \delta_r)} = 1, \text{ so}$$

$$\lim v_r \geq u_1(w).$$

■

Appendix D

Proof of lemma 1

There exists a policy function $\alpha_1(L)$ and $\alpha_2(L)$ such that:

$$\begin{aligned} V(L) &= (1 - \delta)v_1(w, \alpha_2(L)) + \delta E_{w, \alpha_2(L)} \{V(L'(y, \alpha_1(L), a_2))\} \\ \text{s.t.} &: \alpha_2(L) \in B(\alpha_1(L)) \\ \alpha_1(L) &\in W(L) \end{aligned}$$

Proof: Let's define the following

$$w(L) = \left\{ \alpha \in \Delta(A_1) : \exists \alpha' \in \Delta(A_1) \text{ s.t. } \alpha = \frac{1}{1+L}w + \frac{L}{1+L}\alpha' \right\}$$

and

$$\Gamma(L) = \{(\alpha, \alpha_2) \in \Delta(A_1) \times \Delta(A_2) : \alpha \in w(L) \text{ and } \alpha_2 \in B(\alpha)\}$$

and

$$\frac{1}{1 + L'(y, a_2, L, \alpha)} = \lambda \frac{\frac{1}{1+L}\rho(y|w, a_2)}{\rho(y|\alpha, a_2)} + (1 - \lambda)\eta^*(w)$$

We can reformulate the optimization problem as

$$V(L) = \inf_{(\alpha, \alpha_2) \in \Gamma(L)} \{(1 - \delta) v_1(w, \alpha_2) + \delta E_{w, \alpha_2} [V(L'(y, a_2, L, \alpha))]\}$$

Lemma 12 $\Gamma(L)$ is nonempty, compact and uhc.

Proof. First, note that $w(L)$ is compact. Nonemptiness is obvious. Next, since $\Gamma(L) \subseteq \Delta(A_1) \times \Delta(A_2)$, we just need to show closedness. Take a convergent sequence from $\Gamma(L)$, $(\alpha^n, \alpha_2^n) \rightarrow (\alpha, \alpha_2)$. By definition, for each n , $\alpha^n \in w(L)$ and $\alpha_2^n \in B(\alpha^n)$. Since $w(L)$ is compact, $\alpha \in w(L)$. So, we are left to show that $\alpha_2 \in B(\alpha)$ but this follows from the u.h.c. of $B(\cdot)$, which is a consequence of the Maximum Theorem. Finally, uhc follows from uhc of $B(\cdot)$.

However, to apply Theorem 9.6 in S&L we need lhc as well. So, we will generalize that result. First of all, Let $LSC_b(X) = \{f : X \rightarrow \mathbb{R} : f \text{ is lsc and bounded}\}$, where X is a metric space.

Lemma 13 *Let $f : X \rightarrow \mathbb{R}$ be bounded. Then, f is lsc $\iff$ for any sequence $\{x_n\} \subseteq X$ with $x_n \rightarrow x$ and $\forall \epsilon > 0, \exists N$ such that $f(x) - f(x_n) < \epsilon, \forall n \geq N$.*

Proof. Take a sequence $\{x_n\} \subseteq X$ such that $x_n \rightarrow x \in X$. By lemma 2.39 in Aliprantis & Border (A&B from now on), $f \in LSC(X)$ if and only if

$$\liminf_n f(x_n) \geq f(x)$$

$\Rightarrow$) Suppose not, that is, $\exists \epsilon > 0$ such that $f(x) - f(x_n) \geq \epsilon$ infinitely often. Let $\{x_{n_k}\}$ be such a subsequence. Then, $-K \leq f(x_{n_k}) \leq f(x) - \epsilon, \forall k$. By Bolzano-Weierstrass, we can extract a further subsequence, let's call it $\{f(x_{n_k})\}$ as well, so that $f(x_{n_k}) \rightarrow z^*$ but

$$z^* \leq f(x) - \epsilon < f(x) \leq \liminf_n f(x_n)$$

thus reaching a contradiction.

($\Leftarrow$) Suppose by way of contradiction that $\liminf_n f(x_n) < f(x)$ and let $\epsilon > 0$ be such that $\liminf_n f(x_n) < f(x) - \epsilon$. By assumption, $\exists N$ such that $f(x_n) > f(x) - \epsilon, \forall n \geq N$ but this implies $\liminf_n f(x_n) \geq f(x) - \epsilon$, which is a contradiction.

Lemma 14 *The uniform limit of a sequence of lsc real-valued functions is lsc.*

Proof. Let $K > 0$ be such that $\sup_{x \in X} |f(x)| < K$. Next, take a sequence $\{x_n\} \subseteq X$ such that $x_n \rightarrow x \in X$. Suppose $\{f_n\} \subseteq LSC(X)$ is such that $f_n \xrightarrow{unif} f$. Fix $\epsilon > 0$, then $\exists N$ such that

$$\sup_{z \in X} |f_n(z) - f(z)| < \epsilon, \quad \forall n \geq N$$

Furthermore, since f_N is lhc, by lemma 13, there exists M such that $f_N(z_m) - f_N(z_m) < \epsilon, \forall m \geq M$, whenever $z_m \rightarrow z$. So,

$$\begin{aligned} f(x) - f(x_m) &= [f(x) - f_N(x)] + [f_N(x) - f_N(x_m)] + [f_N(x_m) - f(x_m)] \\ &\leq |f(x) - f_N(x)| + [f_N(x) - f_N(x_m)] + |f_N(x_m) - f(x_m)| \\ &\leq 3\epsilon \end{aligned}$$

for any $m \geq M$. Since this is true $\forall \epsilon$, applying lemma 13 again, we can conclude that $f \in LSC(X)$.

Lemma 15 $LSC_b = \{f \in LSC(X) : f \text{ is bounded}\}$ is a complete metric space under the supnorm.

Proof. See page 74 in A&B together with lemmas 13 and 14.

Lemma 16 *If $f, g \in LSC_b(X)$, then $f + g \in LSC_b(X)$*

Proof. Take $x_n \rightarrow x$, then

$$\begin{aligned} \liminf_n [f(x_n) + g(x_n)] &\geq \liminf_n f(x_n) + \liminf_n g(x_n) \\ &\geq f(x) + g(x) \end{aligned}$$

and the result follows from lemma 2.39 in A&B.

Lemma 17 *Let $\phi : X \rightrightarrows Y$ be a nonempty, uhc and compact-valued correspondence and $f : \text{Graph}(\phi) \rightarrow \mathbb{R}$ be lsc. Define $V : X \rightarrow \mathbb{R}$ by*

$$V(x) = \min_{y \in \phi(x)} f(x, y)$$

Then, $V(\cdot)$ is well-defined and lsc.

Proof. This is just a modification of the proof of lemma 16.30 in A&B.

Lemma 18 *If $f \in LSC_b([0, \bar{L}])$, then*

$$(Mf)(\alpha, \alpha_2, L) = \int_{Y \times A_2} f(L'(y, a_2, L, \alpha)) Q(\alpha_2, dy, da_2)$$

is bounded and lsc, where

$$Q(\alpha_2, y, a_2) = \rho(y|w, a_2)\alpha_2(a_2)$$

Proof. Boundedness is obvious. Next, let's take a sequence $(\alpha^n, \alpha_2^n, L^n) \rightarrow (\alpha, \alpha_2, L)$. Then,

$$\begin{aligned} &(Mf)(\alpha, \alpha_2, L) - (Mf)(\alpha^n, \alpha_2^n, L^n) \leq \\ &|(Mf)(\alpha, \alpha_2, L) - (Mf)(\alpha, \alpha_2^n, L)| + [(Mf)(\alpha, \alpha_2^n, L) - (Mf)(\alpha^n, \alpha_2^n, L^n)] \\ &= |(Mf)(\alpha, \alpha_2, L) - (Mf)(\alpha, \alpha_2^n, L)| \\ &+ \int [f(L'(y, a_2, L, \alpha)) - f(L'(y, a_2, L^n, \alpha^n))] Q(\alpha_2^n, dy, da_2) \end{aligned} \tag{17}$$

Let's consider the first term on the rhs. Since $|Y \times A_2| < \infty$ and $\alpha_2^n \rightarrow \alpha_2$, there must exist N such that $|(Mf)(\alpha, \alpha_2) - (Mf)(\alpha, \alpha_2^n)| < \frac{\epsilon}{2}$, $\forall n \geq N$. Next, since $L'(\cdot)$ is continuous in both α and L and f is lsc, by lemma 13, $\forall \epsilon > 0$, $\exists N' > 0$ such that $f(L'(y, a_2, L, \alpha)) - f(L'(y, a_2, L, \alpha^n)) < \frac{\epsilon}{2}$, $\forall n \geq N'$. Therefore,

$$(Mf)(\alpha, \alpha_2, L) - (Mf)(\alpha^n, \alpha_2^n, L^n) < \epsilon, \quad \forall n \geq \max\{N, N'\}$$

and the same lemma implies that Mf is lsc.

Remark 2 *We need Mf to be lsc in L as well to apply lemma 17.*

We can now state the following result.

Theorem 4 *Define the operator T on $LSC_b([0, \bar{L}])$ by*

$$(Tf)(L) = \min_{(\alpha, \alpha_2) \in \Gamma(L)} \{(1 - \delta)v_1(w, \alpha_2) + \delta (Mf)(\alpha, \alpha_2, L)\}$$

Then, $T : LSC_b([0, \bar{L}]) \rightarrow LSC_b([0, \bar{L}])$ is well-defined and has a unique fixed point.

Proof. The fact that the operator is well-defined follows from the previous lemmas. For the rest of the proof, we can just apply Blackwell's theorem (see the proof of Theorem 9.6 in S&L).

Corollary 1 *There exist policy functions $(\alpha(L), \alpha_2(L)) \in \Gamma(L)$ such that*

$$V(L) = (1 - \delta) v_1(w, \alpha_2(L)) + \delta E_{w, \alpha_2(L)} [V(L'(y, a_2, L, \alpha(L)))]$$

References

- [1] Aliprantis, C. D., K. Border (1999): "Infinite Dimensional Analysis," Springer, Heidelberg, Germany.
- [2] Atakan, A., and M. Ekmekci (2009): "Reputation in Long Run Relationships," working paper.
- [3] Atakan, A., and M. Ekmekci (2009): "Reputation with Long Run Players and Imperfect Observability," working paper.
- [4] Bar-Isaac, H., and S. Tadelis (2008): "Seller Reputation," *Foundations and Trends in Microeconomics*, 4:4, 273-351.
- [5] Celentani, M., and W. Pesendorfer (1996): "Reputation in Dynamic Games," *Journal of Economic Theory*, 70, 109-132.
- [6] Cole, H. L., J. Dow, and W.B. English (1995): "Default, Settlement, and Signalling: lending Resumption in a Reputational Model of Sovereign Debt," *International Economic Review*, 36, 365-385.
- [7] Cripps, M. W., G.J. Mailath, and L. Samuelson (2004): "Imperfect Monitoring and Impermanent Reputations," *Econometrica*, 72, 407-432.

- [8] ————— (2007): "Disappearing Private Reputations in Long-Run Relationships," *Journal of Economic Theory*, 134(1), 287-316.
- [9] Mehmet Ekmekci (2010): "Sustainable Reputations with Rating Systems," *Journal of Economic Theory*, forthcoming.
- [10] Eduardo Faingold (2008): "Building a Reputation under Frequent Decisions," working paper.
- [11] Faingold, E., Y. Sannikov (2007): "Reputation Effects and Equilibrium Degeneracy in Continuous-Time Games," working paper.
- [12] Fudenberg, D., and D. K. Levine (1989): "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57, 759-778.
- [13] Fudenberg, D., and D. K. Levine (1992): "Maintaining a Reputation When Strategies are Imperfectly Observed," *Review of Economic Studies*, 59, 561-579.
- [14] Fudenberg, D., and D. K. Levine (1994): "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, 62, 103-135.
- [15] Holmstrom, B. (1999): "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66, 169-182.
- [16] Klein, B. and K. Leffler (1981): "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89: 615-41.
- [17] Kreps, D., P. R. Milgrom, D. J. Roberts, and R. Wilson (1982): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 27, 245-252.
- [18] Kreps, D., and R. Wilson (1982): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253-279.
- [19] Qingmin Liu (2009): "Information Acquisition and Reputation Dynamics", working paper.
- [20] Liu, Q., and A. Skrzypacz: "Limited Records and Reputation," working paper.
- [21] Lucas, G., and N. Stokey (1993): *Recursive Methods in Economic Dynamics*.
- [22] Mailath, G. J., and L. Samuelson (2001): "Who Wants a Good reputation?" *Review of Economic Studies*, 68, 415-441.
- [23] Mailath, G. and L. Samuelson (2007): "Repeated Games and Reputations: Long-Run Relationships," MIT Press, Cambridge, MA
- [24] Milgrom, P. R., and D.J. Roberts (1982): "Predation, Reputation and Entry Deterrence," *Journal of Economic Theory*, 27, 280-312.

- [25] Phelan, C. (2006): "Public Trust and Government Betrayal," *Journal of Economic Theory*, 130, 27-43.
- [26] Resnick, P., R. Zeckhauser, E. Friedman, and K. Kuwabara (2000): "Reputation Systems," *Communications of the ACM*, 43, 45-48.
- [27] Sorin, S. (1999): "Merging, Reputation, and Repeated Games with Incomplete Information," *Games and Economic Behavior*, 29, 274-308.
- [28] Bernardita Vial (2008): "Competitive Equilibrium and Reputation under Imperfect Public Monitoring," Working Paper.
- [29] Thomas Wiseman (2008): "Reputation and Impermanent Types," *Games and Economic Behavior*, 62, 190-210.
- [30] Thomas Wiseman (2009): "Reputation and Exogenous Private Learning," *Journal of Economic Theory*, 144, 1352-1357.