Indeterminacy of Reputation Effects in Repeated Games with Contracts*

Alexander Wolitzky

MIT

Cambridge, MA 02139

wolitzky@mit.edu

February 2011

Abstract

We study whether allowing players to sign binding contracts governing future play leads to reputation effects in repeated games with long-run players. We proceed by extending the analysis of Abreu and Pearce (2007) by allowing for the possibility that different behavioral types may not be immediately distinguishable from each other. Given any prior over behavioral types, we construct a modified prior with the same total weight on behavioral types and a larger support under which almost all efficient, feasible, and individually rational payoffs are attainable in perfect Bayesian equilibrium. Thus, whether reputation effects emerge in repeated games with contracts depends on details of the prior distribution over behavioral types other than its support.

JEL Codes: C70, C73, C78

Keywords: Bargaining, Contracts, Equilibrium Selection, Long-Run Players, Repeated Games, Reputation.

*I am grateful to Glenn Ellison for detailed comments, advice, and support; to Gabriel Carroll, Mehmet Ekmekci, David Pearce, and Muhamet Yildiz for helpful discussions; and to the editor and two anonymous referees for helpful comments. This research was supported by the NSF.

1 Introduction

Does game theory make strong predictions about the outcomes of long-run relationships? It has been known since the seminal papers of Kreps et al (1982), Kreps and Wilson (1982), and Milgrom and Roberts (1982) that reputation effects have important consequences for equilibrium selection in many dynamic games. Fudenberg and Levine (1989) famously showed that a patient long-run player facing a series of short-run opponents receives at least her Stackelberg payoff in any Nash equilibrium, if her "Stackelberg type" has positive prior probability, and similar results hold in two-player repeated games in the limit where one player becomes infinitely more patient than the other.[1] However, reputation effects are elusive in two-player games with comparably patient players:[2] indeed, it is not obvious what outcome one would expect reputation effects to select in such games. For this reason, the reputation result of Abreu and Pearce (2007, henceforth AP) is striking: AP show that, in two-player repeated games with common discounting in which players may offer each other binding commitments to future divisions of the surplus, all perfect Bayesian equilibrium (PBE) payoffs converge to the Nash bargaining with threats payoffs as the probability of behavioral types converges to zero, so long as the "Nash bargaining with threats type" has positive prior probability and different commitment types are distinguishable from each other from the start of the game. Thus, AP's results suggest that allowing players to sign binding contracts in repeated games—which seems very plausible in many applications, such as employer-employee and union-firm relationships—leads to extremely strong equilibrium selection results in the presence of an arbitrarily small amount of incomplete information. The current paper investigates whether this intuition is correct, or whether the ability to make such strong predictions about long-run relationships requires additional assumptions

[1] Schmidt (1993) and Cripps et al (1996) provide weaker payoff bounds than do Fudenberg and Levine (1989). Stronger results hold with trembles (Aoyagi (1996)), imperfect monitoring (Celentani et al (1996)), or complicated commitment types (Evans and Thomas (1997)).

[2] See Chan (2000) for a folk theorem and Chan (2000) and Cripps et al (2005) for uniqueness results in special games. Aumann and Sorin (1989) derive a uniqueness result for common interest games under additional assumptions. Recently, Atakan and Ekmekci (2009a, 2009b, 2009c) provide additional uniqueness results for a broad class of extensive-form games with perfect information and for a broad class of one-sided reputation-building games with imperfect monitoring of the non-reputation-builder. On reputation effects (or lack thereof) in bargaining, see Myerson (1991), Abreu and Gul (2000), Kambe (1999), Compte and Jehiel (2002), Lee and Liu (2010), and Wolitzky (2011).

about the nature of the incomplete information in the model.

Formally, we extend AP's model by allowing that different commitment types may not be immediately distinguishable, and show that whether or not reputation effects emerge depends on the relative probabilities of different behavioral types, rather than on only the support of the prior distribution over behavioral types.[3] In particular, given any prior over behavioral types, we construct a modified prior with a larger support under which almost all efficient, feasible, and individual rational payoffs are perfect Bayesian equilibrium payoffs (Theorem 1). Furthermore, the weight on any behavioral type under the original prior is at most K times its weight under the modified prior, where K is a constant that does not depend on the original prior and is non-decreasing in the discount rate; thus, there is a uniform bound on the extent to which any original prior must be modified to yield a new prior for which a folk theorem holds. Therefore, if the only assumption that a researcher is willing to make about the prior distribution of behavioral types is that some types have positive prior probability, she cannot rule out any efficient, feasible, individually rational payoffs. This stands in stark contrast with the case of one long-run player facing a series of short-run players (Fudenberg and Levine (1989, 1992)), where assumptions of this form lead to strong conclusions about equilibrium payoffs.

The essential intuition for our result is that, when different behavioral types are initially indistinguishable, imitating a "tough" behavioral type may not be profitable for a normal player (i, say), because doing so may lead her opponent (j) to believe that she is a "soft" behavioral type, at least for a long time. This is the key difference between our model and AP's, in which if player i imitates a tough behavioral type, player j believes that player i is either tough or normal, since in AP's model different behavioral types are immediately distinguishable. In particular, in our model there may be soft types of player i that play like tough types with some probability, but also concede to player j with high enough probability that player j will keep playing against an apparently tough type in the hope that it will turn out to be a soft type. As long as soft types continue to concede on the equilibrium path, player j will eventually become convinced that she is facing a tough type and concede. But

[3] We do, however, assume that normal types have the ability to distinguish themselves from behavioral types. We discuss the role of this assumption in Section 4.

if the prior probability of soft types is high enough relative to the prior probability of tough types, this will take long enough that player i will not be tempted to imitate a tough type.

We remark that our soft types reward one's opponent for failing to concede in much the same way as Evans and Thomas' (1997, 2001) commitment types punish one's opponent for failing to play a prescribed action. The reason why allowing complicated commitment types leads to multiplicity in our model and uniqueness in Evans and Thomas' is the difference in patience: with equal patience, the fact that player j thinks that player i may be a complicated commitment type may limit player i's ability to manipulate player j's beliefs quickly enough for her to benefit from doing so, while if player i is infinitely more patient than player j she can only benefit from player j's attributing to her a wide range of possible commitment types. This line of argument shows why a player cannot guarantee herself a high payoff in our model even if she is the only reputation-builder (i.e., if her opponent is known to be normal), despite her potentially useful ability to offer binding contracts. It also provides an intuition for why existing reputation results with equal patience rely on strong restrictions on the prior distribution over commitment types, even in the limited class of games for which such results apply,[4] while reputation results for games in which one player is infinitely more patient than the other do not require such restrictions.[5]

Finally, there is an interesting connection—suggested to me by an anonymous referee—between our results and the failure of reputation effects in some repeated games with a patient reputation-builder and a relatively impatient long-run opponent. Reputation effects may fail to obtain in that setting because the normal reputation-builder may punish her opponent for best-responding to her Stackelberg action (see chapter 16 of Mailath and Samuelson (2006) for an informative discussion of this point). However, the reputation-builder can circumvent this problem when she is allowed to offer binding contracts, as in AP, which makes AP's

[4] Chan's (2000) uniqueness result depends on there being only one commitment type; Cripps et al (2005) obtain uniqueness only in the limit as the weight on commitment types other than the Stackelberg type converges to zero; Atakan and Ekmekci (2009a) assume that non-Stackelberg types distinguish themselves from the Stackleberg type at a uniform rate; Atakan and Ekmekci (2009b) assume that there is only one commitment type; Atakan and Ekmekci (2009c) assume that all commitments types are finite-automata, which in their model is a similar restriction to that in Atakan and Ekmekci (2009a); and Aumann and Sorin (1989) assume that every commitment type that follows a pure strategy with finite memory has positive probability, but that no other commitment types have positive probability.

[5] For example, none of the papers cited in Footnote 1 relies on upper bounds on the prior probability of any type.

uniqueness result possible.[6] Introducing additional behavioral types, as in our model, can restore the opponent's incentive to fail to best-respond to the reputation builder's action, leading to the failure of reputation effects.

The remainder of our paper proceeds as follows: Section 2 introduces our model, which is very similar to AP's model, with the modification that distinct behavioral types are not immediately distinguishable from each other. Section 3 presents the main idea of the paper in the context of a simple example: the prisoner's dilemma with a single behavioral type on each side. It serves to build intuition and to contrast our results with AP's. Section 4 presents the main result, Theorem 1. Section 5 offers brief concluding remarks.

2 Model

We begin with the hybrid discrete-time/continuous-time model developed by AP. There are two players. At each integer time $n = 0, 1, 2, \ldots$, players choose actions in a finite stage game $G = (S_i, U_i)_{i=1}^2$ and also make demands ("contracts," "offers") $u_i \in \Pi_i$, where Π_i is the convex hull of the set of player i's feasible payoffs in G, and u_i is interpreted as the lowest payoff that player i is willing to accept in the continuation game. Actions determine flow payoffs until the next integer time, assuming neither player accepts the other's contract offer. That is, if players use actions (s_1, s_2), player i's period payoff is $U_i(s_1, s_2) \int_0^1 e^{-rt} dt$, where r is the common discount rate. We also assume, as in AP, that players can select mixed actions (m_i, m_j) at integer times, in which case mixing occurs continuously throughout the period, so it is as if mixed actions are observable; let M_i be the set of player i's mixed actions. At any time (not just integer times), either player j $(= -i)$ can accept the other player's standing offer u_i ("concede"), in which case the players receive $\left(u_i, \phi_j(u_i)\right)$, where $\phi_j(u_i)$ is the highest feasible payoff for j consistent with i getting u_i, and the game ends (each player only has one standing offer at a time—these may change on the integers). As in AP, there is a first and last date at which player j can accept each offer of player i's (i.e., "just after

[6] Indeed, the ability to offer binding contracts makes reputation-building easier in many settings, which provides another motivation for our indeterminacy result. Games with a patient reputation-builder facing a relatively impatient opponent is one example. Another is common-interest games with two equally patient players (Cripps and Thomas, 1997), where it is again easy to see that allowing binding contracts leads to reputation effects.

n" and "just before $n+1$"), and the players move sequentially in an arbitrary, pre-specified order at each integer time n; see AP for more details of this formulation of time. We also assume that the function ϕ_j is strictly decreasing, which rules out common-interest games, and use ϕ_j^{-1} and ϕ_i interchangeably. The game ends immediately if the standing offers ever satisfy $(u_1, u_2) \in \Pi$, in which case both players get their demands. Thus, the game can be thought of as a "repeated game with contracts" or as "bargaining with payoffs as you go." At time t, the (disagreement) history h^t of mixed actions (m_i, m_j) and demands (u_i, u_j) is publicly observed.

At the beginning of the game, there is a chance that each player is one of a number of behavioral types, which are simply repeated game strategies (i.e., arbitrary automata that may condition their player on the entire history h^t). Player i is of behavioral type γ_i (i.e., is committed to strategy γ_i) with prior probability $\pi_i(\gamma_i)$, and π_i is assumed have countable support; since we do not assume that $\pi_i(\gamma_i)$ is positive for any γ_i, this formulation allows for both one-sided and two-sided reputation-formation.[7] We assume that each γ_i plays a pure strategy over (Π_i, M_i) but may mix over accepting or rejecting j's offer. This restriction is made to simplify notation, and is without significant loss of generality, since a mixed strategy over (Π_i, M_i) can be approximated by a lottery over countably many pure strategies over (Π_i, M_i); in addition, an element of M_i is already a lottery over S_i, so the only restriction here is that behavioral types do not mix over uncountably many elements of Π_i.[8] Players' types are drawn independently. Let z_i be the probability that i is one of the behavioral types, i.e., $z_i = \sum_{\gamma_i \in \operatorname{supp} \pi_i} \pi_i(\gamma_i)$. Let (G, π) describe the stage game together with the common prior over behavioral types.

AP assume that, before play over (Π, M) begins, there is an initial "announcement" stage, where each player simultaneously announces a behavioral type γ_i. AP assume that

[7] In AP, $\pi_i(\gamma_i)$ is the probability of player i's being of type γ_i conditional on being a behavioral type. We let $\pi_i(\gamma_i)$ be the unconditional probability of player i's being of type γ_i.

[8] One difference between our model and AP is that AP do not allow behavioral types to play mixed strategies or concede at non-integer times. Our assumption that behavioral types can mix and concede at non-integer times is not crucial, as each type we consider that mixes and concedes at non-integer times can be replaced by a set of types, each one of which concedes with probability 1 at a different integer time without substantially affecting our results. Furthermore, AP's results do not rely on their assumption that behavioral types do not mix or concede at non-integer times. Thus, this difference in assumptions—which substantially simplifies our exposition—does not drive the difference in results between the current paper and AP.

behavioral types (but not normal types) announce their types truthfully; this is why behavioral types are instantly distinguishable from each other in their model. We dispense with the announcement stage almost entirely: we assume only that there is an initial "revelation" stage, in which each normal player has the option to "reveal rationality", i.e., to costlessly and certifiably reveal to the other player that she is normal. Formally, we assume that before players choose their initial (Π, M), they publicly announce an element of the set $\{0, 1\}$, and that all behavioral types announce 0; we refer to announcing 1 as "revealing rationality". We also assume that behavioral types do not condition their play on whether their opponents reveal rationality.[9,10]

Unlike the announcement stage of AP, the revelation stage is included in our model essentially for convenience, and indeterminacy of reputation effects persists without the revelation stage; see the discussion preceding the proof of Theorem 1 for a discussion of the role of the revelation stage in our model. In addition, our analysis goes through if behavioral types also have the ability to certifiably reveal their types, because a normal type cannot mimic a behavioral type that certifiably reveals itself. Hence, the revelation stage can also be given a positive justification if players can exhibit hard information that reveals their types. For example, an incumbent firm may be able to publicly exhibit its production costs by letting potential entrants tour its factories and look at its financial records, and an employee may be able to publicly exhibit her outside option by producing job offers from rival employers. Thus, even with the revelation stage, our analysis does not rely on normal and behavioral players having different abilities to reveal their types (as it is as if every player can either reveal her true type or reveal "nothing"), in contrast to the analysis of AP (as in their model behavioral players are forced to reveal their true types but normal players are not).

[9]The assumption that normal players may "reveal rationality" to each other is also present, roughly speaking, in AP. Technically, AP require normal players to announce a behavioral type, rather than allowing them to announce that they are normal, but they show in their Footnote 17 that this assumption is immaterial in their model. More substantively, AP also assume that behavioral types do not condition their play on announcements.

[10]One can check that Theorem 1 continues to hold if behavioral types can condition their play on whether their opponents reveal rationality, provided that the prior probability that each player is behavioral is sufficiently small (proof available upon request).

3 Example

In this section, we illustrate the main idea of our paper in the context of a simple example. Let G be the prisoner's dilemma:

	C	D
C	$1,1$	$-1,3$
D	$3,-1$	$0,0$

We first consider this stage game with a single behavioral type (the Nash bargaining with threats type, analyzed by AP) on each side, and note that a uniqueness result applies as in AP. We then add an additional "soft" behavioral type on each side and show that, for suitably chosen priors, almost all efficient, feasible, and individually rational payoffs can be attained in PBE. The argument for this fact contains many of the ideas of the proof of our main result (Theorem 1) in a much simpler context.

First, consider the case where each player is normal with probability $1-z$, and with probability z is the behavioral type that in every period plays D and demands 1, and never accepts an offer of less than 1. Call this behavioral type γ; note that γ is the Nash bargaining with threats type. AP show that, when z is small, a normal player's expected payoff in any PBE is close to her Nash bargaining with threats payoff. In this simple example, an even stronger result applies: for any $z>0$, both normal players receive payoff 1 in any PBE. This follows from an argument similar to the proof of Lemma 1 of AP, which we sketch here. Suppose there exists a PBE in which normal player i's payoff is less than 1. Then she must receive payoff less than 1 from imitating γ, i.e., from not revealing rationality and then playing D and demanding 1 in every period. Player i receives 1 from this strategy when player j is of type γ, so she must receive less than 1 from this strategy when player j is normal. However, if player i imitates γ, there exists some finite time T such that normal player j accepts her demand with probability 1 by T.[11] Let T_0 be the infimum over all T such that this is the case, and suppose towards a contradiction that $T_0>0$. Then there exists $\varepsilon>0$ such that player j's offer at $T_0-\varepsilon$, $\phi_i(u_j)$, is less than 1 (as otherwise the

[11] This follows by a fairly standard reputation-building argument. For the details, see the proof of Lemma 1 of AP.

two demands would be compatible and the game would have ended), and $e^{-r\varepsilon}(1) > \phi_i(u_j)$. Therefore, upon reaching time $T_0 - \varepsilon$ player i will not concede until after time T_0 (as this yields payoff at least $e^{-r\varepsilon}(1)$, whereas conceding yields payoff $\phi_i(u_j)$). This implies that normal player j must concede at time $T_0 - \varepsilon$, contradicting the hypothesis that $T_0 > 0$. Hence, player i's offer of 1 must be accepted with probability 1 by normal player j at time 0, which implies that player i can guarantee herself a payoff of 1 in any PBE by imitating γ. And, of course, the same argument applies to player j.

Next, suppose that each player is still normal with probability $1 - z$, but is now of type γ with probability $z\left(\frac{1}{K}\right)$ for some $K \geq 1$, and with probability $z\left(\frac{K-1}{K}\right)$ is of type $\tilde{\gamma}$, defined as follows: $\tilde{\gamma}$ plays D and demands 1 in every period, but also accepts any non-negative offer at time t at hazard rate $r/\chi(t)$, where $\chi(t)$ is the probability that player i is of type $\tilde{\gamma}$ at time t conditional on her being behavioral (i.e., of type γ or $\tilde{\gamma}$) and having played D, demanded 1, and not conceded until time t. That is,

$$\chi(t) \equiv \frac{e^{-rt} - \frac{1}{K}}{e^{-rt}}$$

for all t such that this is nonnegative. This is illustrated in Figure 1, where $\chi(t)$ is the ratio (at time t) of the distance from the curve e^{-rt} (which is the probability that player i does not concede before time t conditional on her being behavioral) to the dotted line at $\frac{1}{K}$ to the distance from the curve to the x-axis. $\tilde{\gamma}$ always rejects negative offers, and also rejects any offer at any time t such that $e^{-rt} \leq \frac{1}{K}$. We claim that, for any $u_i^* \in \left(\frac{1}{K}, 2 - \frac{1}{K}\right)$, there is now a PBE in which player i receives payoff u_i^* when both players are normal.

To see this, consider the following strategy profile: Normal player j reveals rationality. If player i reveals rationality, player j plays D and demands $\phi_j(u_i^*)$ in every period, and accepts player i's demand if and only if $\phi_j(u_i) \geq \phi_j(u_i^*)$. If player i does not reveal rationality, player j plays D and demands 2 in every period up to time $\frac{1}{r}\log K$, and never accepts an offer of less than 2 until time $\frac{1}{r}\log K$. At time $\frac{1}{r}\log K$, player j continues playing D and demanding 2 in every period, but switches to accepting player i's demand if and only if $\phi_j(u_i) \geq 1$. To complete the description of on-path play, we specify that normal player i's strategy is identical to player j's, except that in the subgame after both players reveal rationality player

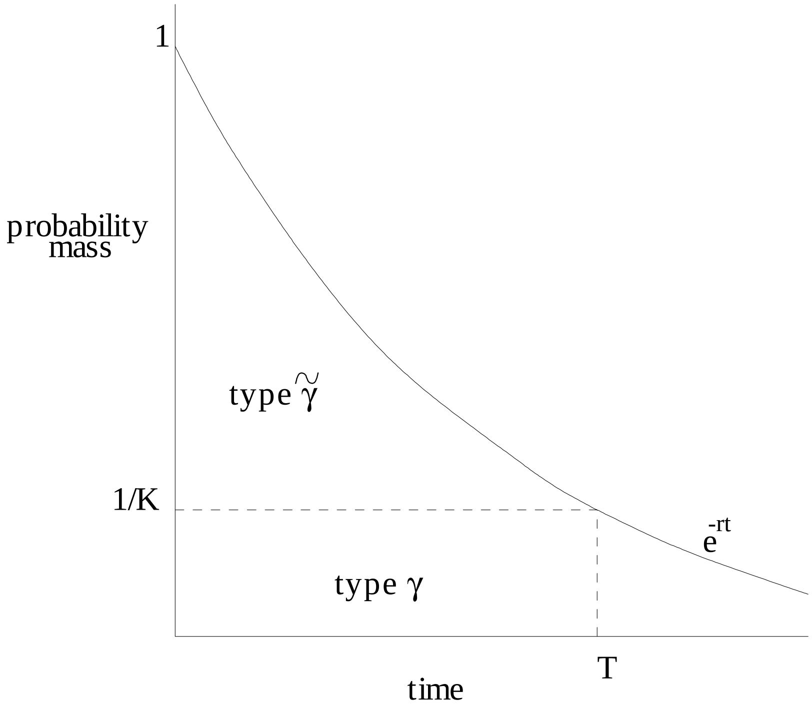

Figure 1: Evolution of Beliefs Conditional on Facing a Behavioral Opponent

i demands u_i^* and accepts player j's demand if and only if $\phi_i(u_j) \geq u_i^*$. Without going into the details, we specify that off-path behavior is as in the proof of Theorem 1, and assert that this behavior is sequentially rational. Thus, to check that this strategy profile is a PBE, we must only check that there are no profitable one-shot deviations at on-path histories.

Clearly, both players' play is optimal after both players reveal rationality. It remains only to check that their play is optimal (at on-path histories) after one's opponent fails to reveal rationality, and that it is optimal to reveal rationality initially. We first verify that player j's play is optimal after player i does not reveal rationality (the argument for player i is symmetric). If player i does not reveal rationality, then she must be one of the behavioral types, which implies that it is optimal for player j to play D in every period. Also, since player j always has the option of accepting her opponent's offer of 1, and only type $\tilde{\gamma}$ ever accepts a demand of more than 1, player j can do no worse than always demanding 2, the highest demand that offers player i a non-negative payoff. Furthermore, note that if player j's assessment that player i is of type $\tilde{\gamma}$ is $\chi(t) > 0$, then player j expects player i to accept her demand of 2 at rate $\frac{r}{\chi(t)}\chi(t) = r$. Therefore, for any τ such that $\chi(\tau) > 0$, player j's expected payoff from rejecting player i's offer until time τ and then accepting is

$$\int_0^\tau e^{-(r+r)t} \left(r\left(2\right) + 0\right) dt + e^{-(r+r)\tau}\left(1\right) = 1,$$

which is the same as player j's payoff from accepting player i's offer immediately. Therefore, player j's decision to reject player i's offer whenever $\chi\left(t\right) > 0$ is (weakly) optimal. Next, observe that $\chi\left(t\right)$ reaches 0 at time T satisfying

$$e^{-rT} = \frac{1}{K},$$

or

$$T = \frac{1}{r}\log K.$$

At this time, player j becomes certain that player i is of type γ, and therefore must accept player i's offer of 1. Thus, player j's continuation strategy is optimal after player i does not reveal rationality.

Finally, we must verify that revealing rationality is optimal for both players. Consider player i first. If player j is behavioral, then player i's payoff is not affected by whether or not she reveals rationality, so it is optimal for her to reveal rationality if and only if it is optimal for her to do so conditional on the event that player j is normal. If player j is normal and player i reveals rationality, player i receives payoff u_i^*. If player j is normal and player i does not reveal rationality, player i can never receive a positive flow payoff and cannot accept a positive offer $\phi_i\left(u_j\right)$ or have a positive demand of her own accepted until time T. Furthermore, the highest offer she ever receives is 1, and the highest demand of hers that player j ever accepts is also 1. Therefore, in the event that player j is normal, player i's payoff in the subgame after she does not reveal rationality is no more than

$$e^{-rT}\left(1\right) = \frac{1}{K}.$$

Since $u_i^* > \frac{1}{K}$, it follows that player i's decision to reveal rationality is optimal. Finally, the same argument applies to player j, and the fact that $u_i^* < 2 - \frac{1}{K}$ implies that $\phi_j\left(u_i^*\right) > \frac{1}{K}$, so player j's decision to reveal rationality is optimal as well. This completes the argument that each player can receive any payoff in $\left(\frac{1}{K}, 2-\frac{1}{K}\right)$ when her opponent is normal in some PBE.

We make four brief remarks to conclude our analysis of this example: First, the fact that each player can receive any payoff in $\left(\frac{1}{K}, 2-\frac{1}{K}\right)$ when her opponent is normal in a PBE implies that she can receive any ex ante expected payoff in this range when the probability that her opponent is behavioral (z) is sufficiently small. Second, taking K large yields a single prior distribution over behavioral types under which almost any efficient, feasible, and individually rational payoff vector is attainable in PBE; that is, there is no need to tailor the prior distribution to the target payoff vector. Third, the above argument does not require type $\tilde{\gamma}$'s concession rate to be exactly $r/\chi(t)$; all that is needed is that type $\tilde{\gamma}$ concedes at least this quickly.[12] Fourth, the smallest K required for a given payoff vector to be attainable in PBE with the above prior is independent of the discount rate, r. All but the last of these observations also apply to our general model, as will become clear in the following section. Furthermore, the constant K used in the construction of the modified prior in the general analysis is non-decreasing in r, so K remains bounded as the players become more patient.

4 Indeterminacy of Reputation Effects

This section contains the formal statement and proof of our main result. The analysis is complicated by the possibility that players may imitate behavioral types with arbitrary repeated game strategies, rather than only the stationary type considered in the above example. However, the idea that a player will not imitate a given behavioral type if there is a high prior probability on a particular "soft" type whose play initially resembles that type carries over from the example.

Let $\underline{u}_i$ be i's (mixed action) minmax payoff, let $\bar{u}_i \equiv \phi_i\left(\underline{u}_j\right)$, and let $\underline{\hat{u}}_i$ and $\hat{u}_i$ be i's lowest and highest feasible payoffs, respectively. Let $\underline{m}_i$ be a mixed action of i's that minmaxes j. We say that u is a "PBE payoff of (G, π) when both players are normal" if there is a

[12] However, if type $\tilde{\gamma}$ concedes faster than this, then K must be larger to support the same range of target payoffs in PBE.

PBE of (G, π) that yields expected payoff u conditional on both players' being normal. Of course, if z_1 and z_2 are small, then u is a PBE payoff of (G, π) when both players are normal if and only if u is close to an ex ante expected PBE payoff of the normal players in (G, π) (since G is finite), but we need not assume that z_1 and z_2 are small. We write $u > (\geq)\, u'$ if $u_i > (\geq)\, u'_i$ for $i \in \{1, 2\}$.

Our main result is the following:

Theorem 1 *For any finite game G, vector $\tilde{u} > \underline{u}$, and number $\bar{r} > 0$, there exists a number $K \geq 1$ such that, for every prior π, there exists a modified prior π' with the following three properties:*

1. *$z'_i = z_i$ for $i \in \{1, 2\}$.*
2. *$\pi'_i(\gamma_i) \geq \frac{1}{K}\pi_i(\gamma_i)$ for all $\gamma_i \in \operatorname{supp} \pi_i$ and $i \in \{1, 2\}$.*
3. *For any discount rate $r \in (0, \bar{r})$, the set of PBE payoffs of (G, π') when both players are normal contains any efficient $u^* \in \Pi$ such that $u^* \geq \tilde{u}$.*

Theorem 1 says that there exists a single modified prior for which the PBE set contains almost any efficient, feasible, and individually rational payoff, for any discount rate below an arbitrary fixed number (in particular, the constant $\bar{r}$ is chosen freely, and need not be "small"). Also, the extent to which the original prior must be modified to yield such a new prior, measured by K, does not depend on the original prior, π, but only on G, $\tilde{u}$, and $\bar{r}$. Finally, Theorem 1 does *not* show that the uniqueness result of AP is sensitive to the addition of a "small" mass of behavioral types that initially pool with another behavioral type. Rather, it shows that their result is sensitive to the addition of a "large" mass of such types, where "largeness" is determined only by G and $\tilde{u}$.

A noteworthy consequence of Theorem 1 is the existence of a bound on the extent to which the original prior must be modified (K) that is uniform over discount rates below $\bar{r}$; in particular, this bound does not explode as the discount rate goes to zero. This contrasts with the results of Fudenberg and Levine (1989), which imply that the prior probability of the Stackelberg type must converge to zero if payoff multiplicity is to persist as r goes to zero.[13] The key difference is the presence of equal discounting in our model. In particular, the rate at which player i must be conceding to player j for player j to be willing to reject her demand scales with r. This implies that the time T required for player i to convince player j that she is a "tough" type scales with $1/r$. Hence, the resulting cost of delay to player i, e^{-rT}, is independent of r. This argument is exactly as in the example of Section 3.[14] Indeed, K only depends on the discount rate at all in Theorem 1 due to a technical issue resulting from the hybrid discrete-time/continuous-time nature of the model.[15]

We now outline the proof of Theorem 1. The first step is constructing the modified prior π' for given π, G, $\tilde{u}$, and $\bar{r}$. The goal is constructing a "soft" (henceforth, "offsetting") type $\tilde{\gamma}_i$ for every $\gamma_i \in \operatorname{supp} \pi_i$ such that $\tilde{\gamma}_i$ has the following properties: On-path, $\tilde{\gamma}_i$ follows the same strategy over (Π_i, M_i) as γ_i does for a long time; $\tilde{\gamma}_i$ concedes to player j quickly enough that player j does not accept any offer less than $\phi_j(\tilde{u}_i)$ when she is not confident whether she is facing type γ_i or type $\tilde{\gamma}_i$, but slowly enough that it takes a long time for player j to learn whether she is facing type γ_i or type $\tilde{\gamma}_i$; and $\tilde{\gamma}_i$ induces player j to play either $\left(\bar{u}_j, \underline{m}_j\right)$ or some other "tough" action for a long time. These "tough" actions of player j that are induced by type $\tilde{\gamma}_i$ are called *admissible* in the proof of Theorem 1. The point of this construction is that, if player j assigns sufficient weight to his facing type $\tilde{\gamma}_i$, then player j will play an admissible action and reject all offers of less than $\phi_j(\tilde{u}_i)$ until some distant time T_i.[16] Therefore, if normal player i receives at least $\tilde{u}_i$ in some strategy profile and T_i is sufficiently large, then she does not want to pretend to be of type γ_i, since

[13] It is easy to see that Fudenberg and Levine's results continue to apply if contracts are allowed, as a patient long-run player could simply imitate the Stackelberg type who rejects all contract offers and demands her highest feasible payoff every period. This relates to our observation in Footnote 6 that allowing contracts often makes reputation results easier to obtain.

[14] In contrast, if in the current model player i were made much more patient than player j, she could guarantee herself nearly $\bar{u}_i$ in any PBE by making a demand close to $\bar{u}_i$ and playing $\underline{m}_i$ every period, as long as the behavioral type that follows this strategy is present with positive probability.

[15] The issue is that, if player j acts second at an integer time t, it is impossible to punish player j for deviating from his prescribed action until time $t+1$. This friction is larger when r is larger, which necessitates a larger K in the construction in the proof of Theorem 1. This friction would be entirely absent if the continuous-time model were replaced by the limit of discrete-time models as actions become frequent, which is why the fact that K depends on $\bar{r}$ may be viewed as an artifact of the hybrid discrete time/continuous time nature of the model.

[16] An exception to this is that player j may play an inadmissible action in response to an offer by player i of at least $\phi_j(\tilde{u}_i)$. However, it can be shown that player i also receives a low payoff in this case, essentially because either player j accepts her generous offer or she accepts a correspondingly aggressive demand of player j's.

she is guaranteed to receive a low flow payoff until time T_i. Finally, the modified prior π' is defined so that player j's beliefs after player i fails to reveal rationality are that, whichever strategy $\gamma_i \in \operatorname{supp} \pi_i$ player i follows, player i is initially very likely to be an offsetting type.

The second step is verifying that it is indeed optimal for player j to play an admissible action and reject any offer less than $\phi_j(\tilde{u}_i)$ until time T_i. This might at first appear to be difficult, because it is very difficult to determine player j's entire optimal strategy under prior π'. However, to show that it is optimal for player j to play an admissible action and reject any offers less than $\phi_j(\tilde{u}_i)$ at some history h^t, it suffices to exhibit a single continuation strategy of player j's that involves playing an admissible action and rejecting any offer less than $\phi_j(\tilde{u}_i)$ and yields a higher payoff than any continuation strategy that involves playing any inadmissible action or accepting an offer less than $\phi_j(\tilde{u}_i)$. And, if type $\tilde{\gamma}_i$ concedes at a high enough rate until time T_i, then it can be verified that playing the best admissible action in the current period, then playing $(\bar{u}_j, \underline{m}_j)$ and rejecting player i's offer until time T, and finally accepting player i's offer just after time T yields a higher payoff for player j than does any continuation strategy involving playing any inadmissible action or accepting an offer less than $\phi_j(\tilde{u}_i)$ at history h^t.

Finally, we construct a PBE in which normal players attain the target payoffs (u_1^*, u_2^*). In this PBE, when both players reveal rationality, they then demand their target payoffs, ending the game immediately with the target payoffs. If player i deviates by failing to reveal rationality, then, as we have seen, she faces only admissible actions until time T_i and her offer is rejected unless she demands less than $\tilde{u}_i$ (with the exception described in Footnote 16), so, regardless of her continuation strategy, she receives a payoff below her target payoff. Hence, both players reveal rationality. The specification of off-path play supporting this behavior is somewhat complicated, and builds on a construction in AP.

The role of our assumption that normal players can "reveal rationality" to each other is as follows: Suppose that, at the beginning of the game, normal player i is supposed to demand u_i^* and normal player j is supposed to demand $\phi_j(u_i^*)$, ending the game. If some behavioral player j offers player i more than u_i^*, player i is tempted to demand more than u_i^* in the hope that player j is of this type. If player j turns out to make the normal demand $\phi_j(u_i^*)$, player i can simply accept this offer an instant after it is made and thus go unpunished for her experimentation. Allowing normal players to reveal rationality to each other before the beginning of play eliminates this problem, since, if player j reveals rationality, normal player i has no reason to experiment with higher initial demands. Since this is the only point in the proof where the ability of normal players to reveal rationality matters, any modification of the game that prevents normal player i from experimenting in this way allows us to eliminate the revelation stage.[17] The revelation stage can also be eliminated without such a modification if z_1, z_2, and r are sufficiently small. The idea is that the strategy profile constructed in the proof of Theorem 1 can be modified by prescribing that normal player 1 initially demands $\hat{u}_1$ and normal player 2 initially demands $\phi_2(u_1^*)$, player 1 immediately accepts player 2's demand if it equals $\phi_2(u_1^*)$, and the first player who deviates receives her minmax payoff if agreement is not reached by time 1. Under this strategy profile, neither player has an incentive to experiment in the above manner, because player 1 is already demanding her highest feasible payoff, and, when z_1 is small, player 2 receives less than her minmax payoff if she deviates and then accepts when player 1 demands $\hat{u}_1$. The requirement that r is also small is needed to ensure that a player is willing to wait and receive her highest feasible and individually rational at the next integer time when her opponent deviates. The details of this construction are available upon request.

Proof of Theorem 1. *Step 1: Construction of π'*

We begin by constructing, for any type γ_i, an offsetting type $\tilde{\gamma}_i$ that at every instant either follows the strategy of type γ_i or concedes. The motivation for the details of the specification of $\tilde{\gamma}_i$ will become clear in Step 2 of the proof.

First, observe that the theorem is trivial if there is no vector $u^* \in \Pi$ such that $u^* \geq \tilde{u}$, so assume that such a vector exists. Since $\tilde{u} > \underline{u}$ and ϕ_j is decreasing, this implies that $\bar{u}_j > \phi_j(\tilde{u}_i) > \underline{u}_j$. Let

$$u_i^0 \equiv \frac{\tilde{u}_i + \underline{u}_i}{2},$$

[17] For example, rather than allowing normal players to reveal rationality we could impose an ε penalty on both players for failing to come to agreement immediately, and assume that z_1 and z_2 are sufficiently small that this penalty outweighs any incentive to experiment, i.e., that $\varepsilon(1 - z_j) > (\hat{u}_i - \underline{u}_i) z_j$ for all i (proof available upon request).

let

$$\rho_i \equiv \max\left\{\frac{\phi_j\left(u_i^0\right)-\underline{\hat{u}}_j}{\bar{u}_j-\underline{\hat{u}}_j}, \frac{\hat{u}_j-\phi_j\left(u_i^0\right)}{\hat{u}_j-\left(\left(1-e^{-\bar{r}}\right)\phi_j\left(u_i^0\right)+e^{-\bar{r}}\phi_j\left(\tilde{u}_i\right)\right)}, \frac{\hat{u}_j-\phi_j\left(\tilde{u}_i\right)}{\hat{u}_j-\underline{u}_j}\right\},$$

and let

$$K \equiv \frac{\left(\frac{\hat{u}_i-\underline{u}_i}{u_i^0-\underline{u}_i}\right)^{\frac{\phi_j\left(u_i^0\right)-\underline{\hat{u}}_j}{\bar{u}_j-\phi_j\left(u_i^0\right)}}}{1-\rho_i}.$$

Observe that $\rho_i < 1$, so K is finite. This number K will suffice for the proof.

Fix a discount rate $r \in (0, \bar{r})$. Let

$$\lambda_i \equiv r\frac{\phi_j\left(u_i^0\right)-\underline{\hat{u}}_j}{\bar{u}_j-\phi_j\left(u_i^0\right)},$$

and let

$$\chi_i(t) \equiv 1-\frac{e^{\lambda_i t}}{K}.$$

To understand the definition of $\chi_i(t)$, suppose that initially player i plays strategy γ_i with probability $\frac{1}{K}$, and with probability $\frac{K-1}{K}$ plays a different strategy that up to time t plays the same (π_i, m_i) as does γ_i and also accepts player j's offer at a hazard rate that makes the unconditional hazard rate that player i accepts player j's offer equal λ_i. Then $\chi_i(t)$ is a lower bound on the probability that player i is not playing strategy γ_i at time t conditional on the event that she has not accepted player j's offer by time t (indeed, $\chi_i(t)$ is exactly this probability if and only if strategy γ_i never accepts player j's offer before time t).

Next, let T be the time at which $\chi_i(t)$ reaches ρ_i; that is,

$$T_i \equiv \frac{1}{\lambda_i}\log\left(K\left(1-\rho_i\right)\right).$$

Finally, let

$$\hat{\lambda}_i(t) \equiv \frac{\lambda_i(t)}{\chi_i(t)}$$

if $\chi_i(t) > 0$, and let $\hat{\lambda}_i(t) = 0$ otherwise. If $\hat{\lambda}_i(t)$ is player i's acceptance rate conditional on not playing γ_i, and $\chi_i(t)$ is the probability that player i is not playing γ_i, then $\lambda_i(t)$ is player i's unconditional acceptance rate.

We now define *admissible* actions, as previewed in the outline of the proof:

Definition 1 *At an integer time t at which player j acts first, the action $\left(\bar{u}_j, \underline{m}_j\right)$ is* admissible *and all other actions are* inadmissible*. At an integer time t at which player j acts second, the action $\left(\bar{u}_j, \underline{m}_j\right)$ is* admissible*, as is the action $\left(\max\left\{\bar{u}_j, U_j\left(m_i\left(t\right), m_j\right)\right\}, m_j\right)$ for any m_j that satisfies $m_j \in \arg\max_{m_j'} U_j\left(m_i\left(t\right), m_j'\right)$ and $U_j\left(m_i\left(t\right), m_j\right) > \phi_j\left(u_i^0\right)$, where $m_i\left(t\right)$ is player i's realized time-t action; and all other actions are* inadmissible*. A history h^t is* admissible *if player j has never played an inadmissible action, and is* inadmissible *otherwise.*

We are ready to define type $\tilde{\gamma}_i$:

Definition 2 *Given any type γ_i, the γ-offsetting type, $\tilde{\gamma}_i$, is the strategy defined as follows: If history h^t is admissible, $\tilde{\gamma}_i$ plays the same $\left(u_i, m_i\right)$ as does γ_i, and $\tilde{\gamma}_i$ accepts player j's demand at hazard rate $\hat{\lambda}_i\left(t\right)$; if in addition $t = T_i$, then $\tilde{\gamma}_i$ accepts player j's demand with probability 1. If h^t is inadmissible, $\tilde{\gamma}_i$ plays $\left(\bar{u}_i, \underline{m}_i\right)$ and rejects player j's demand.*

Finally, we define the modified prior π_i'. In this definition, $\tilde{\gamma}_i'$ is the γ_i'-offsetting type defined above.

Definition 3 *The modified prior π_i' is given by $\pi_i'\left(\gamma_i\right) = \frac{1}{K}\pi_i\left(\gamma_i\right) + \sum_{\gamma_i' \in \operatorname{supp}\pi_i : \tilde{\gamma}_i' = \gamma_i} \frac{K-1}{K}\pi_i\left(\gamma_i'\right)$ for all types γ_i.*

Observe that $z_i' = z_i$ and $\pi_i'\left(\gamma_i\right) \geq \frac{1}{K}\pi_i\left(\gamma_i\right)$ for all $\gamma_i \in \operatorname{supp}\pi_i$. Thus, to prove the theorem it suffices to show that the set of PBE payoffs of $\left(G, \pi'\right)$ when both players are normal contains any efficient $u^* \in \Pi$ such that $u^* \geq \tilde{u}$. This is done in Steps 2 and 3 of the proof.

Step 2: Behavior under π_i'

We now establish the key property of the modified prior π_i'. Under a strategy profile in the game $\left(G, \pi'\right)$ in which normal players reveal rationality, player j's optimal continuation strategy at any history h^t that is consistent with player i following some strategy $\gamma_i \in \operatorname{supp}\pi_i'$ is determined up to indifference by sequential rationality. We say that an action of player j's is (weakly) optimal at such a history if it is part of an optimal continuation strategy.

Lemma 1 *Fix a strategy profile in the game (G, π') in which normal players reveal rationality. Suppose that history h^t is admissible, player i's past play at h^t is consistent with her following some strategy $\gamma_i \in \operatorname{supp} \pi'_i$, and $t < T_i$. Then the following two statements hold at h^t:*

1. *Suppose that player i's demand is at least $\tilde{u}_i$. Then it is optimal for player j to reject player i's demand. If in addition t is an integer, then it is optimal for player j to play an admissible action.*

2. *If t is an integer at which player j acts second and it is optimal for player j to reject player i's demand and play an inadmissible action, then under any optimal continuation strategy agreement is reached by time $t+1$ and player j receives continuation payoff at least $\phi_j\left(u_i^0\right)$.*

Proof of Lemma 1. First, suppose that player i's time-t demand is at least $\tilde{u}_i$. We show that rejecting player i's demand at time t is optimal by exhibiting a strategy that involves rejected player i's demand at time t and yields a weakly higher payoff than accepting this demand. In particular, suppose that player j plays $\left(\bar{u}_j, \underline{m}_j\right)$ from the next integer time onward and rejects player i's demand until just after time T_i, and then accepts player i's demand. When player j follows this strategy, player i concedes at unconditional rate at least λ_i at all times earlier than T_i (since at such times there is probability at least $\chi_i(t)$ that she is conceding at rate $\hat{\lambda}_i(t)$, as player j is always playing an admissible action), and concedes at time T_i with unconditional probability at least

$$\chi_i\left(T_i\right) = \rho_i \geq \frac{\phi_j\left(u_i^0\right) - \underline{\hat{u}}_j}{\bar{u}_j - \underline{\hat{u}}_j}.$$

Therefore, player j's continuation payoff from this strategy is at least

$$\begin{aligned}
& \int_0^{T_i - t} e^{-(r+\lambda_i)t}\left(\lambda_i \bar{u}_j + r\underline{\hat{u}}_j\right) dt + e^{-(r+\lambda_i)(T_i - t)}\left(\frac{\phi_j\left(u_i^0\right) - \underline{\hat{u}}_j}{\bar{u}_j - \underline{\hat{u}}_j}\left(\bar{u}_j\right) + \frac{\bar{u}_j - \phi_j\left(u_i^0\right)}{\bar{u}_j - \underline{\hat{u}}_j}\left(\underline{\hat{u}}_j\right)\right) \\
= \; & \left(1 - e^{-(r+\lambda_i)(T_i - t)}\right)\frac{1}{r+\lambda_i}\left(\lambda_i \bar{u}_j + r\underline{\hat{u}}_j\right) + e^{-(r+\lambda_i)(T_i - t)}\phi_j\left(u_i^0\right) \\
= \; & \left(1 - e^{-(r+\lambda_i)(T_i - t)}\right)\phi_j\left(u_i^0\right) + e^{-(r+\lambda_i)(T_i - t)}\phi_j\left(u_i^0\right) \\
= \; & \phi_j\left(u_i^0\right).
\end{aligned}$$

On the other hand, player j's continuation payoff from accepting player i's demand at time t is at most $\phi_j(\tilde{u}_i)$. This is less than $\phi_j(u_i^0)$, so it is optimal for player j to reject player i's demand at time t.

Second, suppose that t is an integer at which player j acts first. Player j's continuation payoff from rejecting player i's demand and playing any action other than $(\bar{u}_j, \underline{m}_j)$ (the unique admissible action) is at most

$$\chi_i(t)\,\underline{u}_j + (1 - \chi_i(t))\,\hat{u}_j,$$

because if player i is of type $\tilde{\gamma}_i$ (which is the case with probability at least $\chi_i(t)$) she immediately minmaxes player j. Since $\chi_i(t)$ is decreasing and $t < T$, this is less than

$$\begin{aligned}
& \chi_i(T_i)\,\underline{u}_j + (1 - \chi_i(T_i))\,\hat{u}_j \\
= \;& \rho_i \underline{u}_j + (1 - \rho_i)\,\hat{u}_j \\
\leq \;& \frac{\hat{u}_j - \phi_j(\tilde{u}_i)}{\hat{u}_j - \underline{u}_j}\left(\underline{u}_j\right) + \frac{\phi_j(\tilde{u}_i) - \underline{u}_j}{\hat{u}_j - \underline{u}_j}\left(\hat{u}_j\right) \\
= \;& \phi_j(\tilde{u}_i) \\
< \;& \phi_j\left(\tilde{u}_i^0\right).
\end{aligned}$$

As we have seen, $\phi_j(\tilde{u}_i^0)$ is a lower bound on player i's continuation from rejecting player i's demand and following (at least) one strategy that involves playing $(\bar{u}_j, \underline{m}_j)$. Therefore, if player i's time-t demand is at least $\tilde{u}_i$, then it is optimal for player j to reject this demand and play $(\bar{u}_j, \underline{m}_j)$.

Third, suppose that t is an integer at which player j acts second. Let $U_j^* \equiv \max_{m_j'} U_j\left(m_i(t), m_j'\right)$ and let $m_j^* \in \arg\max_{m_j'} U_j\left(m_i(t), m_j'\right)$. Then playing $\left(\max\left\{\bar{u}_j, U_j^*\right\}, m_j^*\right)$ at time t and subsequently playing $(\bar{u}_j, \underline{m}_j)$ and rejecting player i's demand until just after time T_i (and then accepting) yields continuation payoff at least

$$\left(1 - e^{-r}\right)\max\left\{U_j^*, \phi_j\left(u_i^0\right)\right\} + e^{-r}\phi_j\left(u_i^0\right). \tag{1}$$

This follows because, if $U_j^* \geq \phi_j(u_i^0)$, then player j receives flow payoff U_j^* and demands U_j^*

from time t to time $t+1$; and we have already seen that player j's continuation payoff from playing $\left(\bar{u}_j, \underline{m}_j\right)$ and rejecting player i's demand until time T_i is at least $\phi_j\left(u_i^0\right)$. On the other hand, playing any inadmissible action at time t yields continuation payoff at most

$$\left(1-\chi_i(t)\right) \hat{u}_j + \chi_i(t)\left(\left(1-e^{-r}\right) \max\left\{U_j^*, \phi_j\left(u_i\right)\right\} + e^{-r} \phi_j\left(u_i\right)\right), \tag{2}$$

where u_i is player i's time-t demand, because type $\tilde{\gamma}_i$ responds to an inadmissible action by always rejecting player j's demand and minmaxing player j starting at time $t+1$. If $u_i \geq \tilde{u}_i$, then (2) is at most

$$\left(1-\chi_i(t)\right) \hat{u}_j + \chi_i(t)\left(\left(1-e^{-r}\right) \max\left\{U_j^*, \phi_j\left(\tilde{u}_i\right)\right\} + e^{-r} \phi_j\left(\tilde{u}_i\right)\right),$$

and therefore the difference between (1) and (2) is at least

$$\left(1-\chi_i(t)\right)\left(1-e^{-r}\right) \phi_j\left(u_i^0\right) + e^{-r}\left(\phi_j\left(u_i^0\right) - \chi_i(t) \phi_j\left(\tilde{u}_i\right)\right) - \left(1-\chi_i(t)\right) \hat{u}_j.$$

This expression is non-negative if and only if

$$\chi_i(t) \geq \frac{\hat{u}_j - \phi_j\left(u_i^0\right)}{\hat{u}_j - \left(\left(1-e^{-r}\right) \phi_j\left(u_i^0\right) + e^{-r} \phi_j\left(\tilde{u}_i\right)\right)}.$$

Since $r \leq \bar{r}$, a sufficient condition for this inequality is

$$\chi_i(t) \geq \frac{\hat{u}_j - \phi_j\left(u_i^0\right)}{\hat{u}_j - \left(\left(1-e^{-\bar{r}}\right) \phi_j\left(u_i^0\right) + e^{-\bar{r}} \phi_j\left(\tilde{u}_i\right)\right)}$$

Now $\chi_i(t) \geq \chi_i(T) = \rho_i \geq \frac{\hat{u}_j - \phi_j\left(u_i^0\right)}{\hat{u}_j - \left(\left(1-e^{-\bar{r}}\right) \phi_j\left(u_i^0\right) + e^{-\bar{r}} \phi_j\left(\tilde{u}_i\right)\right)}$, so this sufficient condition holds. Hence, it is optimal for player j to play an admissible action at time t whenever player i's demand is at least $\tilde{u}_i$.

Finally, if player j plays an inadmissible action at time t, then just before time $t+1$ his continuation payoff from rejecting player i's demand is at most

$$\chi_i(t) \underline{u}_j + \left(1-\chi_i(t)\right) \hat{u}_j,$$

because the probability that player i is of type $\tilde{\gamma}_i$ at this time is at least $\chi_i(t)$, and type $\tilde{\gamma}_i$ minmaxes player j starting at time $t+1$. Since $\chi_i(t) \leq \chi_i(T) \leq \frac{\hat{u}_j - \phi_j(\tilde{u}_i)}{\hat{u}_j - \underline{u}_j}$, this is no more than

$$\frac{\hat{u}_j - \phi_j(\tilde{u}_i)}{\hat{u}_j - \underline{u}_j}\underline{u}_j + \left(1 - \frac{\hat{u}_j - \phi_j(\tilde{u}_i)}{\hat{u}_j - \underline{u}_j}\right)\hat{u}_j = \phi_j(\tilde{u}_i).$$

Hence, if $u_i < \tilde{u}_i$, then agreement is reached by time $t+1$ under any optimal continuation strategy. Finally, as we have seen, it is optimal for player j to play an admissible action at time t if $u_i \geq \tilde{u}_i$, and this yields continuation payoff at least $\phi_j(u_i^0)$. Therefore, at time t it is optimal for player j to reject player i's demand and play an inadmissible action only if, under any optimal continuation strategy, agreement is reached by time $t+1$ and he receives continuation payoff at least $\phi_j(u_i^0)$. ■

Step 3: Equilibrium Construction

We now construct strategy profiles for the normal types in (G, π') that yield the desired range of PBE payoffs. The construction builds on that in Lemma 24 of AP. The next paragraph specifies on-path play, and the three paragraphs after it specify off-path play.

Fix $u_i^* \in [\tilde{u}_i, \phi_i(\tilde{u}_j)]$. Normal players reveal rationality. If both players reveal rationality, normal player i initially plays $(u_i^*, \underline{m}_i)$ and normal player j initially plays $\left(\phi_j(u_i^*), \underline{m}_j\right)$. Thus, if both players are normal and follow their equilibrium strategies, the game ends immediately with payoffs $\left(u_i^*, \phi_j(u_i^*)\right)$. If player i does not reveal rationality and at history h^t her play is consistent with some type $\gamma_i \in \operatorname{supp} \pi_i'$, then normal player j is certain that her opponent is behavioral, and her on-path continuation play is pinned down up to indifference by sequential rationality. At any history at which player j is certain that player i is behavioral and is indifferent between accepting and rejecting player i's offer, we specify that she rejects; and at any history at which player j is certain that player i is behavioral and is indifferent between playing $\left(\bar{u}_j, \underline{m}_j\right)$ and playing any other action, we specify that she plays $\left(\bar{u}_j, \underline{m}_j\right)$. Player j's play at other histories at which he is certain that player i is behavioral and is indifferent between any two actions is irrelevant and can therefore be specified arbitrarily.

Off-path play in the subgame after both players reveal rationality is as in Lemma 24 of AP. In particular, if player i deviates to an incompatible demand, then player i plays $(\underline{u}_i, \underline{m}_i)$ and player j plays $\left(\bar{u}_j, \underline{m}_j\right)$ at the next integer time, and in the interim player i

accepts player j's demand and player j rejects player i's demand. The same continuation play follows any single-player deviation. Next, suppose that both players deviate from their prescribed play (to incompatible demands) at integer time t. If player 1's flow payoff given the realized time-t actions is weakly less than player 2's offer to her, then she accepts player 2's demand, and player 2 rejects her demand. If this condition fails for player 1 but holds for player 2, then player 2 accepts and player 1 rejects. If both players' flow payoffs are strictly greater than their opponents' offers to them, then neither player accepts until the next integer time, at which point continuation play is as at the beginning of the subgame after both players reveal rationality, with u_i^* replaced by player i's flow payoff.

Next, consider the subgame after player j reveals rationality and player i does not. Suppose that at time t player i's play becomes inconsistent with all types $\gamma_i \in \operatorname{supp} \pi_i'$ (i.e., at time t player i either makes or rejects a demand that makes the history inconsistent with all $\gamma_i \in \operatorname{supp} \pi_i'$). Then player i plays $(\underline{u}_i, \underline{m}_i)$ and player j plays $\left(\bar{u}_j, \underline{m}_j\right)$ at the next integer time. In the interim, if player i's flow payoff given the realized time-t actions is weakly less than player j's offer to her, then she accepts player j's demand, and player j rejects her demand. If this condition fails for player i, then player j accepts player i's demand if and only if this yields a higher payoff than receiving his flow payoff until the next integer time and then receiving $\bar{u}_j$. If both of these conditions fail, then neither player accepts until just before the next integer time, at which point player i accepts player j's demand, if this demand is no more than $\bar{u}_j$. Continuation play following the next integer time is as at the beginning of the subgame after both players reveal rationality, with u_i^* replaced by $\underline{u}_i$. Now, as long as player i's play is consistent with some type $\gamma_i \in \operatorname{supp} \pi_i'$, player j's play is pinned down by sequential rationality and an arbitrary rule for breaking indifferences. Hence, the above specification of play after player i deviates from any type $\gamma_i \in \operatorname{supp} \pi_i'$ determines player i's entire optimal continuation strategy, again up to indifference.

Finally, consider the subgame after neither player reveals rationality. If at time t player j's play becomes inconsistent with all types $\gamma_j \in \operatorname{supp} \pi_j'$, then continuation strategies at the resulting history h^t are identical to continuation strategies at history $h^{t\prime}$, defined to equal h^t with the modification that player j initially revealed rationality (which are specified in the previous paragraph). That is, continuation play is "as if" player j had revealed rationality.

Similarly, if at time t both players' play simultaneously becomes inconsistent with all of their types in $\operatorname{supp} \pi'$, then continuation play is specified to be the continuation play at the corresponding history where both players have revealed rationality.

It is clear that each player's strategy is optimal at every history except the null history before the players do or do not reveal rationality. Therefore, to check that the above strategy profile is a PBE, it suffices to check that revealing rationality is optimal for both players. This in turn requires only checking that revealing rationality is optimal for player i conditional on player j's being normal, as revealing rationality has no effect on play if player j is behavioral, by the assumption that behavioral types do not condition their play on whether their opponents reveal rationality.

Suppose that player j is normal and normal player i does not reveal rationality. At any admissible history, if player i takes an action that is not consistent with any type $\gamma_i \in \operatorname{supp} \pi'_i$, then her continuation payoff equals $\underline{u}_i$ (by the above specification of off-path play). Thus, suppose that player i's play remains consistent with some type $\gamma_i \in \operatorname{supp} \pi'_i$ at all admissible histories. By Lemma 1, continuation play prior to time T_i falls into one of two categories: either player j always plays an admissible action and only accepts demands of no more than $\tilde{u}_i$; or, at some integer time t at which player j acts second, player j play an inadmissible action and agreement is reached by the next integer time. We now show that in either case player i's payoff is no more than u_i^*, regardless of her strategy.

In the former case, the fact that $u_i^* \geq \tilde{u}_i > \underline{u}_i$ implies that player i may receive a payoff strictly above u_i^* only if the game does not end before T_i. Since in this case player j plays either $\underline{m}_j$ or some action m_j such that $U_j\left(m_i\left(t\right), m_j\right) > \phi_j\left(u_i^0\right)$ (and thus $U_i\left(m_i\left(t\right), m_j\right) \leq u_i^0$) at every time $t < T_i$, player i's payoff is at most

$$\left(1 - e^{-rT_i}\right) u_i^0 + e^{-rT_i} \hat{u}_i.$$

Now

$$\begin{aligned}
e^{-rT_i} &= \left(K\left(1-\rho_i\right)\right)^{-r/\lambda_i} \\
&= \left(\left(\frac{\hat{u}_i-\underline{u}_i}{u_i^0-\underline{u}_i}\right)^{\lambda_i/\bar{r}}\right)^{-r/\lambda_i} \\
&= \frac{\tilde{u}_i-\underline{u}_i}{\hat{u}_i-\underline{u}_i}.
\end{aligned}$$

Hence,

$$\left(1-e^{-rT_i}\right)\underline{u}_i+e^{-rT_i}\hat{u}_i \leq \left(\frac{\hat{u}_i-\tilde{u}_i}{\hat{u}_i-\underline{u}_i}\right)\underline{u}_i+\left(\frac{\tilde{u}_i-\underline{u}_i}{\hat{u}_i-\underline{u}_i}\right)\hat{u}_i=\tilde{u}_i\leq u_i^*.$$

In the latter case, recalling that agreement is reached before time $t+1$ if player j plays an inadmissible action at time t, leaving player j with payoff at least $\phi_j\left(u_i^0\right)$, player i's continuation payoff cannot be higher than the maximum of player i's time-t demand, denoted u_i, and the continuation payoff of type $\tilde{\gamma}_i$, denoted $u_i^{\tilde{\gamma}_i}$ (this is because player i's time-t action m_i is the same as the time-t action of types γ_i and $\tilde{\gamma}_i$, because player j acts second at time t). We claim that both of these values are weakly less than $\tilde{u}_i$ (and thus weakly less than u_i^*). First, the fact that player j's continuation payoff is at least $\phi_j\left(u_i^0\right)$ implies that

$$\left(1-\chi_i(t)\right)\hat{u}_j+\chi_i(t)\left(\left(1-e^{-r}\right)\hat{u}_j+e^{-r}\phi_j\left(u_i\right)\right)\geq\phi_j\left(u_i^0\right),$$

and therefore

$$\phi_j\left(u_i\right)\geq\hat{u}_j-\frac{e^r}{\chi_i(t)}\left(\hat{u}_j-\phi_j\left(u_i^0\right)\right).$$

Hence,

$$\begin{aligned}
u_i &\leq \phi_i\left(\hat{u}_j-\frac{e^r}{\chi_i(t)}\left(\hat{u}_j-\phi_j\left(u_i^0\right)\right)\right) \\
&\leq \phi_i\left(\hat{u}_j-\frac{e^r}{\chi_i(T)}\left(\hat{u}_j-\phi_j\left(u_i^0\right)\right)\right) \\
&\leq \tilde{u}_i.
\end{aligned}$$

Second, because player i cannot receive continuation payoff greater than $\phi_j\left(u_i^{\tilde{\gamma}_i}\right)$ when player i is of type $\tilde{\gamma}_i$,

$$\left(1-\chi_i\left(t\right)\right)\hat{u}_j+\chi_i\left(t\right)\phi_j\left(u_i^{\tilde{\gamma}_i}\right)\geq\phi_j\left(u_i^0\right).$$

Hence,

$$\phi_j\left(u_i^{\tilde{\gamma}_i}\right)\geq\hat{u}_j-\frac{1}{\chi_i\left(t\right)}\left(\hat{u}_j-\phi_j\left(u_i^0\right)\right),$$

which implies that $u_i^{\tilde{\gamma}_i}\leq\tilde{u}_i$ by the same argument as above.

We conclude that player i receives payoff weakly below u_i^* if she does not reveal rationality, which implies that failing to reveal rationality is not a profitable deviation for player i. The same argument applies to player j, because the fact that $u_i^*\leq\phi_i\left(\tilde{u}_j\right)$ implies that $\phi_j\left(u_i^*\right)\geq\tilde{u}_j$. Therefore, the above strategy profile is a PBE. ■

5 Conclusion

This paper shows that allowing players to sign binding contracts governing future play does not lead to reputation effects in the absence of assumptions on the relative probabilities of different behavioral types. This suggests that equilibrium selection due to reputation effects is substantially weaker in games with two long-run players than in games with a single long-run player, even in the presence of contracts, and that existing results do not provide a completely convincing equilibrium selection argument for applications in which different behavioral types may not be immediately distinguishable.

However, we reiterate that AP's uniqueness result is robust to introducing a small mass of behavioral types that initially pool with other behavioral types; in particular, AP's result continues to hold when behavioral types are not immediately distinguishable if the prior probability of the Nash bargaining with threats is high enough relative to the prior probability of "softer" types whose early play resembles that of the Nash bargaining with threats type. This raises the intriguing question of where the boundary between AP's uniqueness result and our multiplicity results lies. That is, for what prior distributions of behavioral types do repeated games with contracts have unique equilibria, and for what priors does the folk theorem apply? What happens in the transitional region between these regimes? Relatedly, our arguments suggest that some behavioral types may be more profitably imitated for a wide range of prior distributions than others; for example, a player must be very confident that her opponent is really a soft type for her to keep playing when her opponent imitates a type that gets "tougher" over time, as this behavior penalizes her for failing to concede. Ongoing research suggests that this approach may be tractable: in Wolitzky (2011), I characterize the behavioral type that is most profitably imitated in bargaining by a player who holds "worst-case" beliefs about her opponent's prior belief about her own strategy, and show that this type does indeed get "tougher" over time. I view these ideas as interesting directions for future research.

References

[1] Abreu, D. and Pearce, D. (2007), "Bargaining, Reputation, and Equilibrium Selection in Repeated Games with Contracts," *Econometrica*, 75, 653-710.

[2] Abreu, D. and Gul, F. (2000), "Bargaining and Reputation," *Econometrica*, 2000, 85-117.

[3] Aoyagi, M. (1996), "Reputation and Dynamic Stackelberg Leadership in Infinitely Repeated Games," *Journal of Economic Theory*, 71, 378-393.

[4] Atakan, A., and Ekmekci, M. (2009a), "Reputation in Long Run Relationships," *mimeo.*

[5] Atakan, A., and Ekmekci, M. (2009b), "A Two Sided Reputation Result with Long Run Players," *mimeo.*

[6] Atakan, A., and Ekmekci, M. (2009c), "Reputation with Long Run Players and Imperfect Observation," *mimeo.*

[7] Aumann, R.J. and Sorin, S. (1989), "Cooperation and Bounded Recall," *Games and Economic Behavior*, 1, 5-39.

[8] Celentani, M., Fudenberg, D., Levine, D.K., and Pesendorfer, W. (1996), "Maintaining a Reputation Against a Long-Lived Opponent," *Econometrica*, 64, 691-704.

[9] Chan, J. (2000), "On the Non-Existence of Reputational Effects in Two-Person Infinitely-Repeated Games," Working Paper, Department of Economics, Johns Hopkins University.

[10] Compte, O., and Jehiel, P. (2002), "On the Role of Outside Options in Bargaining with Obstinate Parties," *Econometrica*, 70, 1477-1517.

[11] Cripps, M.W., Dekel, E., and Pesendorfer, W. (2005), "Reputation with Equal Discounting in Repeated Games with Strictly Conflicting Interests," *Journal of Economic Theory*, 121, 259-272.

[12] Cripps, M.W., Schmidt, K., and Thomas, J.P. (1996), "Reputation in Perturbed Repeated Games," *Journal of Economic Theory*, 69, 387-410.

[13] Cripps, M.W., and Thomas, J.P. (1997), "Reputation and Perfection in Repeated Common Interest Games," *Games and Economic Behavior*, 18, 141-158.

[14] Evans, R., and Thomas, J.P. (1997), "Reputation and Experimentation in Repeated Games with Two Long-Run Players," *Econometrica*, 65, 1153-1173.

[15] Evans, R., and Thomas, J.P. (2001), "Cooperation and Punishment," *Econometrica*, 69, 1061-1075.

[16] Fudenberg, D., and Levine, D.K. (1989), "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57, 759-778.

[17] Fudenberg, D., and Levine, D.K. (1992), "Maintaining a Reputation When Strategies are Imperfectly Observed," *Review of Economic Studies*, 59, 561-579.

[18] Kambe, S. (1999), "Bargaining with Imperfect Commitment," *Games and Economic Behavior*, 28, 217-237.

[19] Kreps, D.M., Milgrom, P.R., Roberts, J., and Wilson, R.J. (1982), "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 27, 245-252.

[20] Kreps, D.M. and Wilson, R.J. (1982), "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253-279.

[21] Lee, J., and Liu, Q. (2010), "Gambling Reputation: Repeated Bargaining with Outside Options," *mimeo.*

[22] Mailath, G.J., and Samuelson, L. (2006), *Repeated Games and Reputations: Long-Run Relationships*, Oxford: Oxford University Press.

[23] Milgrom, P.R. and Roberts, J. (1982), "Predation, Reputation, and Entry Deterrence," *Journal of Economic Theory*, 27, 280-312.

[24] Myerson, R. (1991), *Game Theory: Analysis of Conflict.* Cambridge, MA: Harvard University Press.

[25] Schmidt, K. (1993), "Reputation and Equilibrium Characterization in Repeated Games of Conflicting Interests," *Econometrica*, 61, 325-351.

[26] Wolitzky, A. (2011), "Reputational Bargaining Under Knowledge of Rationality," *mimeo.*