Reputation with Analogical Reasoning*

Philippe Jehiel
Department of Economics
Paris School of Economics
48 Bd Jourdan
75014 Paris, France
and
University College London
Gower Street
London WC1E 6BT, United Kingdom
Jehiel@pse.ens.fr

Larry Samuelson
Department of Economics
Yale University
30 Hillhouse Avenue
New Haven, CT 06520-8281, USA
Larry.Samuelson@yale.edu

October 20, 2011

Abstract: We consider a repeated interaction between a long-run player and a sequence of short-run players, in which the long-run player may either be rational or may be a mechanical type who plays the same (possibly mixed) action in every stage game. We depart from the classic model, exemplified by Fudenberg and Levine [4, 5], in assuming that the short-run players make inferences by analogical reasoning, meaning that they correctly identify the average strategy of each type of long-run player, but do not recognize how this play varies across histories. Concentrating on 2×2 games, we provide a complete characterization of equilibrium payoffs, establishing a payoff bound for the rational long-run player that can be strictly larger than the familiar "Stackelberg" bound. We also provide a complete characterization of equilibrium behavior, showing that play begins with either a reputation-building or (depending on parameters) a reputation-spending stage, followed by a reputation-manipulation stage.

*We thank Olivier Compte, Drew Fudenberg, Rani Spiegler, and Asher Wolinsky for helpful discussions. Jehiel thanks the European Research Council for financial support. Samuelson thanks the National Science Foundation (SES-0850263) for financial support.

Reputation with Analogical Reasoning

October 20, 2011

Contents

Reputation with Analogical Reasoning

1 Introduction

1.1 Reputations

The literature on reputation, pioneered by Kreps, Milgrom, Roberts and Wilson [7, 8, 12], has shown that uncertainty about a player's type can have dramatic implications for equilibrium play in repeated games.[1] This is most effectively illustrated in the context studied by Fudenberg and Levine [4, 5]. A long-run player faces a sequence of short-run players. The long-run player is almost certainly a rational player interested in maximizing her discounted sum of payoffs, but may also be a mechanical type who plays the same (possibly mixed) stage-game action in every period. Fudenberg and Levine show that the rational long-run player, if sufficiently patient, can guarantee a payoff that corresponds to the payoff she would obtain by always playing like the mechanical type of her choice (with the short-run players playing a best response to this choice). The rational long-run player's payoff in the repeated game thus approximates the payoff she could achieve in a single-shot game in which the long-run player first chooses behavior matching that of a mechanical type, and then the short-run player chooses a best response.

The analysis of Fudenberg and Levine follows from a careful examination of the short-run players' updating of their beliefs as to which type of long-run player they are facing. These beliefs in turn follow from Bayes' rule, assuming that the short-run players have a *perfect* understanding of the equilibrium strategies of the various types of long-run players.

The initial reputation models of Kreps, Milgrom, Roberts and Wilson [7, 8, 12] were designed to explain certain types of seemingly intuitive behavior that does not appear in the equilibrium of a complete-information game, with entry deterrence in the finitely-repeated chain store game being the classic example. Subsequent developments, typified by Fudenberg and Levine [4, 5], have pushed reputation models toward characterizing equilibrium payoffs. Nonetheless, much of the more applied interest in reputations centers around behavioral questions. How do people build reputations? How do they manage or maintain their reputations, and when do they spend them?

This paper examines an alternative reputation model, centered around a simpler model of how short-run players formulate and update their beliefs. We view this behavior as a plausible alternative to the potentially demanding requirement that short-run players have a perfect understanding of equilibrium play. We characterize equilibrium payoffs, but are also able to obtain precise characterizations of equilibrium behavior.

[1] See Mailath and Samuelson [11] for a survey. Mailath and Samuelson refer to a player who necessarily chooses the same exogenously-specified action in every stage game as a *simple commitment* type. We refer to such players as *mechanical* types.

1.2 Analogical Reasoning

We assume that short-run players reason as if *all* types of the long-run players behave in a stationary fashion. This assumption is correct for mechanical types, but not necessarily so for the rational long-run player. The stationary behavior attributed by short-run players to each type of long-run player is assumed to match that type's true expected frequency of play, aggregated over the various periods.

We view this formulation as capturing a setting in which it is difficult for short-run players, who appear in the game just once, to obtain a detailed description of the actions of the long-run players after every possible history. Instead, we assume that short-run players can observe the aggregate frequency of play of the various types of long-run player in previous reputation games, but not how these frequencies depend on the exact history in the game. It is then plausible that short-run players will reason as if the behaviors of the various types of long-run players are stationary, and match the aggregate empirical frequencies of play. In particular, this is the simplest model making use of all of the information available to short-run players. Long-run players, on the other hand, play more often and are thus able to access data (about the play of short-run players) in a finer way.

The equilibrium approach pursued in this paper captures this intuitively-formulated scenario by assuming further that a steady state has been reached, so that the equilibrium play of the long-run players indeed matches the historical frequencies that gave rise to these beliefs. The resulting steady state corresponds to an analogy-based expectation equilibrium (Jehiel [6]), which has been defined for games with multiple types in Ettinger and Jehiel [3]. Observe that despite the coarseness of short-run players' understanding of the long-run player's strategy, short-run players still perform inferences using Bayes' rule as to which type of long-run player they are facing. However, this updating is based on a misspecified model of the long-run player, assuming behaviors are stationary, in contrast to the correct model used in the classic sequential equilibrium concept.[2]

1.3 Equilibrium Behavior

We concentrate throughout on the case in which both the short-run player and the long-run player must choose between two actions in every period, discussing extensions in Section 4. We provide a complete characterization of equilibrium behavior.

We find two types of equilibria—those in which the long-run frequency of play of the rational long-run player becomes concentrated around one frequency as the discount rate vanishes, which we refer to as unary equilibria, and those

[2]Liu and Skrzypacz [9] examine a model in which the short run players observe only the play of the long run player from the recent past, but are otherwise fully rational. They investigate the extent to which the long run player is still able to effectively commit to play matching that of a mechanical type.

which do not have this property. For a given constellation of mechanical types and specification of the stage game, there may exist multiple equilibria, with payoff differences that persist even as the agents become arbitrarily patient. However, if the set of possible mechanical types is sufficiently rich, then we (generically) have unique equilibria, and all such equilibria are unary.

Unary equilibria have an important implication for our learning interpretation of the equilibrium considered here. While our basic theory (implicitly) requires that short-run players have access to aggregate behavior from previous reputation games *by type*, there is no need for short-run players to have access to the types of previous long-run players in the case of unary equilibria. By observing the distributions of aggregate frequencies in past reputation games, short-run players can identify that there are different modes, and by identifying each mode with a type of long-run player, the short-run players would reason exactly as we assume they do in our current analysis.

We accordingly focus this introductory discussion on unary equilibria. Equilibrium behavior can then be divided into two phases. The game begins with an initial phase, in which (depending on parameters) the rational long-run player either builds her reputation (playing so as to inflate the belief of the short run player that she is the mechanical type with whom she wishes to be confused) or spends her reputation (exploiting the belief that with sufficiently high probability she is the mechanical type with whom she wants to be confused). If the set of commitment types satisfies a straightforward diversity condition, then this initial phase is relatively short, and converges to being an insignificant proportion of play as players get more patient.

The initial phase is followed by a manipulation phase. Here, the long-run player's behavior balances the interest of making the highest instantaneous payoff and the interest of maintaining the belief that the long-run player is mechanical. In the most interesting cases (in which the "reputation outcome" does not coincide with a Nash equilibrium of the stage game), player 1 manipulates player 2's belief so as to keep player 2 as close as possible to indifference between player 2's actions. In the process, player 1 introduces correlation into the actions of player 1 and 2, reducing the cost of manipulating 2's beliefs and boosting player 1's payoff.

For a fixed discount factor, and conditional on facing the rational long-run player, the short-run players' belief that the long-run player is mechanical need not converge to zero, in contract to the insight obtained by Cripps, Mailath and Samuelson [1, 2] in the classic rationality setup. Thus, even in the ultra long-run, short-run players remain uncertain as to which type of long-run player they are facing. This is so because the short-run players are working with a misspecified model, and hence even arbitrarily rich amounts of data need not lead them to correct beliefs.

1.4 Equilibrium Payoffs

A first result on equilibrium payoffs is straightforward. The rational long-run player can always guarantee a payoff that is no less than the bound derived in

Fudenberg and Levine [4, 5].[3] However, she can often ensure a strictly larger payoff. There are two reasons for this difference. First, the long-run player in our model can induce the short-run players to attach positive probability to multiple types, even in the limit as arbitrarily large amounts of data accumulate. This essentially allows the long-run player to "commit" to the behavior of a phantom mechanical type who does not actually appear in the list of possible mechanical types. Second, the correlation between player-1 and player-2 actions that appears in the reputation-manipulation phase allows an additional boost in the long-run player's payoff.

We can illustrate the second, somewhat more subtle, of these forces with the following game:

	Cheat	*Honest*
Audit	$4, -2$	$3, -1$
Not	$0, 4$	$4, 0$

(1)

Think of this as a game between a taxpayer (player 2) and the government (player 1). There is a potential surplus of 4, consisting of the taxpayer's liability, to be split between the two. If the government does not audit, the surplus is captured by the government if the taxpayer is honest, and by the taxpayer if the taxpayer cheats. Auditing simply reduces the payoffs of both agents by 1 if the taxpayer is honest. Auditing a cheating taxpayer imposes an additional penalty on the taxpayer, while allowing the government to appropriate the surplus and recover the auditing costs. This game has a unique mixed equilibrium, in which the government audits with probability 4/5 and the taxpayer cheats with probability 1/5, for payoffs (16/5, 4/5). In the classic model, commitment is of no value.[4] Suppose further that in addition to the normal or rational player 1, there is a mechanical type of player 1 who plays a stationary mixture giving *Audit* with probability strictly above $\frac{4}{5}$, as well as a mechanical type who plays *Audit* with probability strictly below $\frac{4}{5}$. As usual, we assume that the overall probability of the long-run player being mechanical is small. In our cognitive environment, the long-run player would choose *Audit* roughly 4/5 of the time and *Not* roughly 1/5 of the time, thereby ensuring that player 2 is always very close to indifferent between *Honest* and *Cheat*. However, player 1 would manage to *Audit* only when player 2 is cheating, and to not audit when player 2 is honest. A (very) patient player 1 thus always gets a per period payoff close to 4, which exceeds the stage-game Nash equilibrium payoff, as well as any conceivable equilibrium payoff in the classic approach, no matter what mechanical types are present. Player 1's reputation-manipulation stage allows player 1 not only to keep player 2 on the boundary between being honest and cheating, but

[3] This follows from the work of Watson [13], and can be established with the same sort of argument found in Fudenberg and Levine [4, 5].

[4] To verify this, note that inducing honest behavior from player 2 requires that player 1 audit with probability at least 4/5, which then ensures that player 1 gets no more than the equilibrium payoff. It is not a general property that commitment has no value when stage games have only mixed equilibria, but this property simplifies the current comparison without being necessary for the result.

to avoid miscoordination in doing so.

Interestingly, this is also the payoff that would be achieved by player 1 if the player 2s followed fictitious play dynamics, i.e. played best responses to the empirical frequencies of actions of the long-run player observed so far (in the current repeated game). This link to the payoff achievable in the fictitious play setup extends to all 2x2 games. More precisely, as players get patient, the best payoff the long-run player can achieve in an equilibrium of our cognitive setup corresponds to the payoff the long-run player would obtain by best responding to fictitious play dynamics on the part of short-run players.

This link is somewhat unexpected, to the extent that our approach is one with multiple types of long-run players in which short-run players make inferences as to which type they are facing, whereas the fictitious play setup has no such inference taking place. In addition, our approach is an equilibrium approach (in that the stationary behavior assumed for the long-run player matches the true aggregate empirical frequency), unlike the fictitious play setup. However, the equilibrium strategy of the long-run player in our environment involves manipulating the beliefs of the short-run players in much the same way that a long-run player would manipulate the experience of fictitious-play opponents.

2 The Model

2.1 The Reputation Game

We consider a repeated game of incomplete information, as in Fudenberg and Levine [4, 5]. A long-run player 1 faces a sequence of short-run player 2s. The interaction lasts over possibly infinitely many periods. Conditional upon reaching period t, there is a probability $1 - \delta$ that the game stops at t and probability δ that it continues.[5] As usual, we will be especially interested in the case in which δ is close to 1.

At the beginning of the game, the long-run player is chosen by Nature to be one of several types: either a rational type with probability $\mu_\emptyset^0$ or one of K possible mechanical types with probabilities $\mu_\emptyset^1, \ldots, \mu_\emptyset^K$, respectively. This choice is observed by player 1 but not player 2.

We let $a(t) \in A = A_1 \times A_2$ denote the pure stage-game actions chosen by players 1 and 2 in period $t \in \{0, 1, 2, \ldots\}$. A mechanical type plays the same completely-mixed stage-game action in every period. We denote by $\alpha^k \in \Delta(A_1)$ the mixed action played by type k of player 1 in each period t.

In period t, the rational player 1 and the new period-t short-run player observe the history of actions $h(t) \in A^t$, and the players then simultaneously

[5]It is a familiar observation that this specification of a repeated game with a random termination date is formally equivalent to a game that never terminates, but in which players discount payoffs with discount factor δ. In the current development, we commit throughout to the random-termination interpretation. This will have substantive implications when we place consistency conditions on players' information. We could also address the case in which the game lasts forever but players discount, with somewhat different details.

choose actions. The players receive stage-game payoffs $u(a(t))$, where $u : A \to \mathbb{R}^2$.

Player 1's expected payoff from the sequence of action profiles $\{a(t)\}_{t=0}^{\infty}$ is given by

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t u_1(a(t)).$$

Hence, the long-run player imposes no discounting other than that implicitly induced by the random termination of the game. We could also allow player 1 to discount payoffs, though this would complicate the notation. The short-run player values only the payoff she obtains in the stage where she is called to play.

2.2 The Solution Concept

The short-run players are initially uncertain about the long-run player's type, and will draw inferences about this type as play unfolds. Their inferences follow from Bayesian learning, but this learning is conducted in the context of a misspecified model. In particular, short-run players adopt a simplified model of the long-run players' behavior, assuming that this behavior is stationary.

Formally, we capture this by examining a *sequential analogy-based expectations equilibrium* (Jehiel [6] and Ettinger and Jehiel [3]). The long-run player distinguishes all of her information sets. The short-run players put all of their histories into a single analogy class, effectively assuming that any history they observe consists of a sequence of independent draws, from a common distribution given by the (in their view) stationary strategy of player 1. Ettinger and Jehiel [3] provide a general development of the solution concept. The remainder of this section makes the solution concept precise for the game considered in this paper.

Given that the definition of the mechanical types dictates their strategies, we need only specify the strategies of the rational player 1 and of the short-run players 2. Let σ_1 and σ_2 denote these strategies, respectively, and let σ denote the strategy profile (σ_1, σ_2). These strategies specify, for every possible history of play $h \in \bigcup_{t=0}^{\infty} A^t$, a behavioral strategy $\sigma_{ih} \in \Delta A_i$ for $i = 1, 2$. We denote by $\sigma_{1h}(a_1)$ the probability that the rational player 1 selects action a_1 after history h and by $\sigma_{2h}(a_2)$ the probability that player 2 selects action a_2 after history h. We also denote by $P^\sigma(h)$ the probability that history h is reached when player 1 plays according to σ_1 and players 2 play according to σ_2 (taking into account the probability of breakdown after each period).

Given σ, we define

$$A^0 = \frac{\sum_h P^\sigma(h)\sigma_{1h}}{\sum_h P^\sigma(h)}. \tag{2}$$

Hence, A^0 is the aggregate strategy of the rational long-run player when this player's strategy is σ_1 and the strategy of players 2 is σ_2.[6]

[6] The strategy of players 2 affects A^0 insofar that it affects $P^\sigma(h)$.

In order to choose actions, short-run players must form beliefs about the action played by the long-run player. These beliefs incorporate two factors, namely beliefs about the action of the rational player 1, and updated beliefs as to which type of player 1 player 2 thinks she is facing. For the first component, player 2 assumes the rational player 1 chooses in each period according to a mixed action α^0. This again reflects the stationarity misconception built into player 2's beliefs by the sequential analogy-based expectations equilibrium. Turning to the second, let μ_h^k denote the belief that player 2 assigns to player 1 being type $k = 0, ...K$ after history h (where $k = 0$ refers to the rational type and $k > 0$ refers to the mechanical type k). For a history h_t we require:

$$\frac{\mu_h^0}{\mu_h^k} = \frac{\mu_\emptyset^0}{\mu_\emptyset^k} \prod_{\tau=0}^{t-1} \frac{\alpha^0(a_1(\tau))}{\alpha^k(a_1(\tau))}. \tag{3}$$

Player 2 thus updates her belief using Bayes rule, given the history h and the conjecture that player 1 with type 0 plays according to α^0, while player 1 with type k plays according to α^k.

To specify equilibrium actions, we then require that after every history h, player 2 plays a best-response to

$$\sum_{k=0}^{K} \mu_h^k \alpha^k, \tag{4}$$

which represents the expectation about player 1's play given the analogy-based reasoning of player 2. The rational long-run player 1 chooses a strategy σ_1 which is a best response to σ_2. A strategy profile σ is a sequential analogy-based expectation equilibrium if it satisfies these best-response requirements and also satisfies the consistency requirement that

$$A^0 = \alpha^0.$$

The unconventional aspect of a sequential analogy-based expectations equilibrium is the expectations formed by the short-run players. These players model the rational player-1's behavior as stationary, whereas it typically will not be stationary. Given this misspecified model, they behave as rationally as possible, forming posterior beliefs via Bayes' rule and choosing best responses to their beliefs.

Player 2's beliefs about player 1's actions must match the empirical frequencies of those actions. Hence, for the mechanical types, we assume that α^k is the mixed action attributed to mechanical type k by player 2 and is also mechanical type k's action. Similarly, we assume that player 2 attributes an action α^0 to the rational type that is indeed the empirical frequency A^0 of actions taken by that type.

We interpret the consistency requirement on player 2's beliefs as the steady-state result of a learning process. We assume the repeated game is itself played repeatedly, though by different players in each case. At the end of each game, a record is made of the frequency with which player 1 (and perhaps player 2 as well, but this is unnecessary) has played her various actions. This in turn is incorporated into a running record recording the frequencies of player 1's actions. A short-run player forms expectations of equilibrium play by consulting this record. As evidence accumulates, the empirical frequencies recorded in this record will match $\alpha^0, \alpha^1, \ldots, \alpha^K$, leading to the steady state captured by the sequential analogy-based expectations equilibrium.

The public record records the frequencies of the various actions played by player 1, but need not observe their order, with such information rendered irrelevant by player 2's stationarity assumption. We view this as consistent with the type of information typically available. It is relatively easy to find product reviews, ratings services, or consumer reports that give a good idea of the average performance of a product, firm or service, but much more difficult to identify the precise stream of outcomes.[7] Somewhat more demandingly, we assume that the record includes not only the empirical frequencies with which previous player 1's have played their various actions, but also that at the end of each (repeated) game the type of player 1 is identified and recorded. In some cases, one can readily find reports of performance by type. For example, one can find travel guides reporting that a certain airport has legitimate taxis, that provide good value-per-dollar with high probability, as well as pirate taxis, that routinely provide poor value-per-dollar. In other cases this assumption will be less realistic. We identify below an important class of games in which information about types is unnecessary, with empirical frequencies alone sufficing (see subsection 3.4.5).

3 Reputation Analysis

3.1 The Canonical Game

The stage game takes the following form:

	L	R
T	a, w	c, y
B	b, x	d, z

where the long-run player 1 must in each period choose between T and B and the short-run players 2 must choose between L and R.

Clearly, if player 2 has a strictly dominant strategy in the stage game, then every player 2 will play it (and player 1 will best respond to it) in every period. To avoid this trivial case, we assume throughout the analysis that player 2 has no dominant strategy, assuming (without loss of generality) that

$$y > w, \qquad x > z. \tag{5}$$

[7]Notice that with no change in the analysis, we could assume that player 2 can observe only the empirical frequencies of the actions taken by the *current* player 1 (and player 1's age), but not their order. Since player 2 reasons as if player 1's strategy is stationary, these empirical frequencies provide all the information player 2 needs.

We carry this assumption throughout the subsequent analysis without further mention.

We can then define p^* so that player 2's best response is R if player 1 plays T with probability at least p^*, and player 2's best response is L if 1 plays B with probability at least $1 - p^*$. That is,

$$p^* = \frac{x - z}{x - z + y - w}.$$

On the boundary, player 2 is indifferent when 1 plays T with probability p^*.

We can simplify the notation by taking α^k to be the probability with which player 1 of type k plays T. The analysis of the sequential analogy-based expectation equilibria in this $2x2$ case boils down to the determination of α^0, the average probability with which the rational player 1 chooses T. Once a candidate α^0 is fixed, the strategy σ_2 of players 2 after every history h is determined by (3) and (4), and the strategy σ_1 of the rational player 1 must be a best-response to σ_2. For such (σ_1, σ_2) to be a sequential analogy-based expectation equilibrium, it should be that the induced frequency A^0 with which the rational player 1 chooses T (as defined by (2)), equals α^0.

3.2 Best Responses

3.2.1 Player 2's Best Response

We begin with a characterization of the short-run players' best responses. For history h, let n_{hT} be the number of times action T has been played and let n_{hB} be the number of times action B has been played. The short-run player's posterior beliefs after history h depend only on (n_{hT}, n_{hB}), and not on the order in which the various actions have appeared. In addition, this posterior belief determines player 2's belief about player 1's current action. We then have:[8]

Lemma 1 *Fix a specification of player-1 types $(\alpha^0, \alpha^1, \ldots, \alpha^K)$ and prior probabilities $(\mu_\emptyset^0, \mu_\emptyset^1, \ldots, \mu_\emptyset^K)$. Then there exists an increasing function $N_B : \{0, 1, 2, \ldots\} \to \Re$, such that for every history h:*

- *Player 2 plays L if $n_{hB} > N_B(n_{hT})$;*
- *Player 2 plays R if $n_{hB} < N_B(n_{hT})$.*
- *Player 2 is indifferent between L and R when $n_{hB} = N_B(n_{hT})$.*

Proof. Let p_h be the probability player 2 attaches to the event that player 1 will play T, conditional on the history h. Then

$$p_h = \sum_{k=0}^{K} \mu_h^k \alpha^k,$$

[8]When $\alpha^k > p^*$ for all k, then player 2 always chooses R and we can set $N_B = +\infty$. When $\alpha^k < p^*$ for all k, then player 2 always chooses L and we can set $N_B = -\infty$.

where μ_h^k is the posterior attached to player-1 type k at history h. Player 2 will play R if she thinks player 2 will play T with probability $p_h > p^*$, and will play L if $p_h < p^*$. The secret to getting player 2 to play R is thus to have her expect T.

Intuitively, the rational player 1, together with the mechanical types of player 1, are characterized by an array of stationary (from player 2's point of view) probabilities of playing T, some exceeding p^* and some falling short of p^*. Whenever player 2 observes T, her posterior beliefs about the type of player 2 shift (in the sense of first-order stochastic dominance) toward types that are relatively likely to play T. Player 2 thus plays R if and only if she has observed enough T outcomes. Section 5.1 fills in the details. ■

Notice that this result holds regardless of what assumptions we make about the distribution of mechanical types or the strategy of the rational type, though different specifications of these strategies will give rise to different functions N_B. If there exists at least one mechanical type who plays T with probability less than p^*, and at least one mechanical type who plays T with probability greater than p^*, then for any specification of α^0 and any specification of the other mechanical types (if any), there will exist histories after which player 2 plays L as well as histories after which player 2 plays R.

3.2.2 Player 1's Best Response: $a > d > b, c$

Turning to player 1, the most interesting cases are those in which the largest of the four payoffs $\{a, b, c, d\}$ is neither b nor c, so that the largest feasible payoff to player 1 is not available as the outcome of a stage-game Nash equilibrium. It is then simply a matter of notation to assume that the largest payoff is a. We must still break the analysis into several cases, beginning here with that in which $a > d > b, c$. For example:

	L	R
T	$4, 0$	$1, 1$
B	$2, 1$	$3, 0$

. (6)

Player 1 then faces a trade-off in creating incentives for player 2 and exploiting 2's resulting play. For example, player 1 might consider inducing player 2 to choose L. Upon doing so, however, player 1 would like to exploit the result by choosing T. Unfortunately for player 1, player 2's best response to T is R, so that player 1 cannot exploit 2's play of L without also interfering with 2's incentive to choose L.

Lemma 2 *Let $a > d > b, c$ hold. Let N_B be an increasing function characterizing player 2's best-response behavior, with 2 playing L when $n_{hB} > N_B(n_{hT})$ and playing R when $n_{hB} < N_B(n_{hT})$. Then*

- *Player 1 plays T if $n_{hB} > N_B(n_{hT})$;*
- *Player 1 plays B if $n_{hB} < N_B(n_{hT})$.*

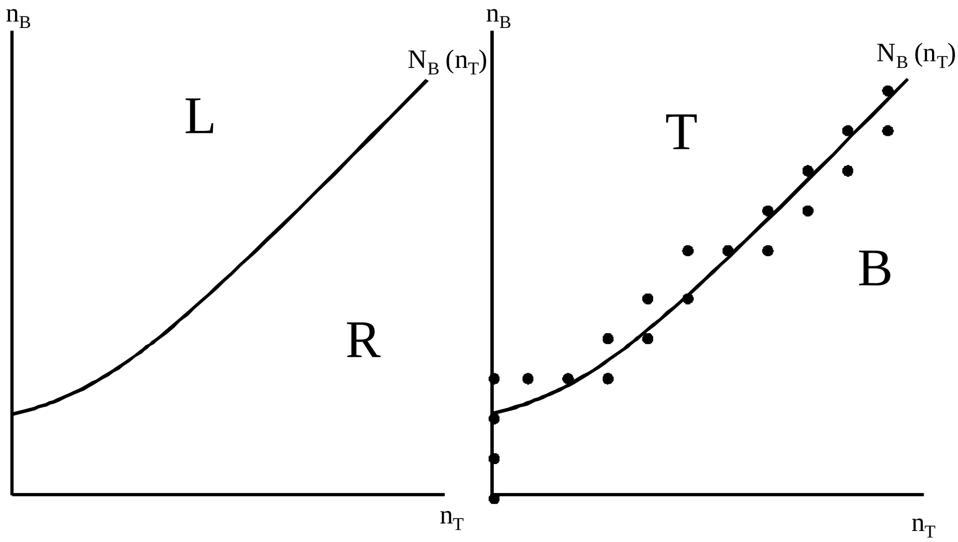

Figure 1: Strategies for the short-run player (left panel) and long-run player (right panel) given $a > d > b, c$. The function N_B, taken from Lemma 1 and characterizing player 2's best responses, does not require $a > d > b, c$. Lemma 2 shows that under $a > d > b, c$, this function also characterizes player 1's best responses. The depiction of N_B is schematic; it is increasing but we cannot in general restrict its intercept or curvature. The outcome is shown in the right panel, consisting of a succession of dots identifying successive (n_T, n_B) values, starting at the origin and proceeding upward (whenever B is chosen) and to the right (whenever T is chosen).

Figure 1 illustrates player 1's and player 2's best responses, as well as the equilibrium path of play. Intuitively, the strategy adopted by player 1 has the effect of keeping player 2 as close as possible to being indifferent between L and R as often as possible.

Let N_B be the function describing player's 2's best response, characterized in 1. Let $\tilde{n}_B(t)$ and $\tilde{n}_T(t)$ be the values of n_B and n_T induced by player 1's best response to the player-2 strategy induced by N_B, after t periods. Then the following limit will exist:

$$q := \lim_{t \to \infty} \frac{\tilde{n}_T(t)}{\tilde{n}_B(t)} \in (0, 1), \tag{7}$$

Furthermore, the limit (as $\delta \to 1$) of the value of α^0 induced by player 1's strategy will equal q. The value of q is the long-run average proportion of T in such a strategy.

Proof. We first show that the player-1 behavior described in Lemma 2 is a best response, using the one-shot deviation principle. Assume first that we are at a history h such that $n_{hB} > N_B(n_{hT})$. If player 1 follows σ_1 he plays T (and player 2 plays L) until he first reaches a history h' with $n_{h'B} = n_{hB}$ and $n_{h'B} < N_B(n_{h'T})$, with player 1 playing B (while player 2 plays R) at h' and with the next history being characterized by $(n_{h'T}, n_{hB} + 1)$. If at h, player 1 plays B instead of T and then follows σ_1, he must play T after his first B (and player 2 plays L) until reaching a history h'' featuring $n_{h''} = (n_{h'T}, n_{hB} + 1)$. Summarizing, we have the following:

Equilibrium path	Payoff	Deviation path	Payoff
TL	a	BL	b
TL	a	TL	a
TL	a	TL	a
$\vdots$	$\vdots$	$\vdots$	$\vdots$
TL	a	TL	a
BR	d	TL	a

We thus see that the only difference between the two scenarios is in the first and last period within this sequence of periods (the deviation gives $u_1(B, L) = b$ and $u_1(T, L) = a$ in the first and last period whereas in absence of the deviation, player 1 receives $u_1(T, L) = a$ and $u_1(B, R) = d$ in the first and last period). We now compare the payoffs in two steps. We can think of transforming the equilibrium outcome to the deviation outcome by first switching the final-period payoff from the d earned under the equilibrium payoff to b, and then operating on this intermediate payoff sequence by moving its initial a to the end and bringing its terminal b to the beginning. Because $d > b$, the initial alteration in the final-period payoff that transforms the equilibrium path to the intermediate path decreases player 1's payoff. But then, since $a > b$, the second adjustment that delays the receipt of a in order to accelerate the receipt of b also decreases player 1's payoff. Hence, the deviation must yield a lower payoff.

The same argument applies for histories h such that $n_{hB} < N_B(n_{hT})$. The details of this argument and uniqueness are given in Section 5.2.

■

3.2.3 Player 1's Best Response: $a > c > d > b$

Now suppose $a > c > d > b$, so that a is again the largest payoff for player 1, but d is not the next largest. For example,

	L	R
T	$4, 0$	$3, 1$
B	$1, 1$	$2, 0$

. (8)

The difference between this and the previous case is that here player 1 pays a cost when playing B against player 2's R. We can view playing B against R

as an investment in pushing player 2 to the point that 2 will play L. When B is a best response to R, as in the previous case, this investment is costless. In the current case B is not a best response to R, and if too many plays of BR are required to elicit a play of L from player 2, then player 1 may simply give up on manipulating player 2's behavior and instead settle for the perpetual play of TR.

Let q continue to be defined as in (7), i.e., let q be the limiting proportion with which player 1 chooses T if player 1 ignores the possibility that the perpetual play of TR may be optimal and instead plays as in Lemma 2. Then:

Lemma 3 *Let $a > c > d > b$. Let N_B be an increasing function characterizing player 2's best-response behavior, with 2 playing L when $n_{hB} > N_B(n_{hT})$ and playing R when $n_{hB} < N_B(n_{hT})$. Then there exists a function $\underline{N}_B(n_T, \delta) \leq N_B(n_T)$ such that for every h:*

- *Player 1 plays T if $n_{hB} > N_B(n_{hT})$;*
- *Player 1 plays B if $N_{hB}(n_T) > n_B > \underline{N}_B(n_{hT}, \delta)$;*
- *Player 1 plays T if $n_{hB} < \underline{N}_B(n_{hT}, \delta)$.*
- *If $qa + (1-q)d > c$ (cf. (7)), then*

$$\lim_{\delta \to 1} \sup\{n_T : \underline{N}_B(n_T, \delta) < 0\} = \infty. \tag{9}$$

Section 5.3 contains the proof. Figure 2 illustrates these strategies. The first two items in the lemma describe player-1 behavior matching that of Lemma 2, culminating in a reputation-manipulation stage in which player 1 manages to keep player 2 close to indifferent between T and B. However, reaching this reputation-manipulation stage may now require a costly initial sequence of BR plays. The third item captures the possibility that player 1 might find this investment too costly. The more patient is player 1, the more investing 1 is willing to do. The final statement of the lemma indicates that if $qa+(1-q)d > c$ holds, then a sufficiently patient player 1 never settles for the perpetual play of TR, with every outcome (at least eventually) calling for player 1 to invest in enough BR plays to bring about a history in which $n_{nB} > N_B(n_{hT})$, and allowing 1 to earn the payoff a from TL.

3.2.4 Player 1's Best Response: $a > b > c, d$ or $a > c > b > d$

Suppose one of the following holds:

$$a > b > c, d$$
$$a > c > b > d.$$

For example,

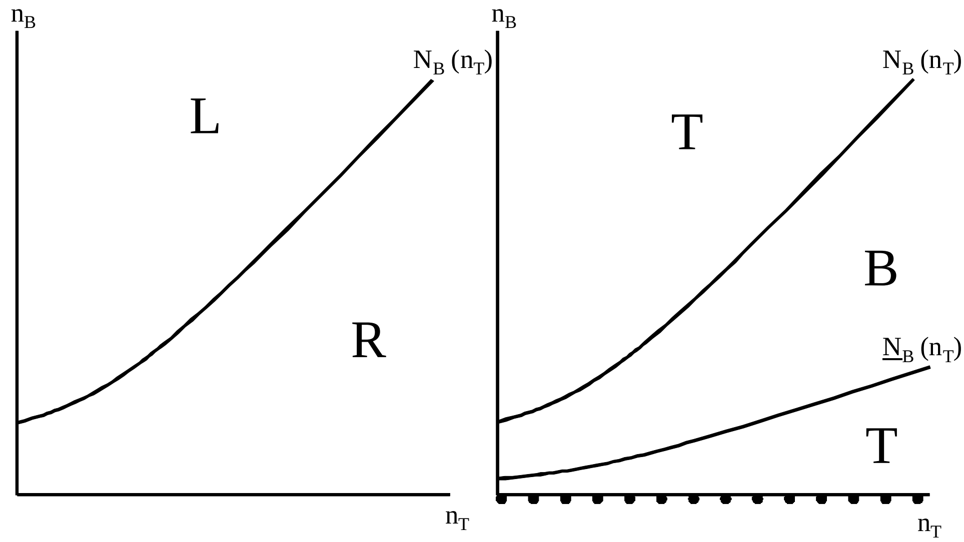

Figure 2: Strategies for the short-run player (left panel) and long-run player (right panel) given $a > c > d > b$. The function N_B is taken from Lemma 1. Condition $a > c > d > b$ is needed only for deriving player 1's best responses. Again, the qualitative properties of the functions are general; the detailed specification is not. An outcome path is shown, beginning at the origin and in this case proceeding relentlessly to the right via the consistent play of T. In this case, player 1 eschews any attempt to build a reputation, settling instead for a payoff of c in every period. From Lemma 3, this can happen only if $qa + (1-q)d < c$ or δ is not too large.

$$\begin{array}{c|c|c|} & L & R \\ \hline T & 4,0 & 2,1 \\ \hline B & 3,1 & 1,0 \\ \hline \end{array} \quad . \tag{10}$$

This is similar to the game presented in (8). In both games, player 1 endeavors to induce TL as often as possible, and hence keeps player 2 nearly indifferent between playing L and R. The difference is that in (8), player 1 prefers 2 to play R on those occasions when 1 must play B, while in (10), 1 will contrive to have player 2 invariably play L.

Again let q be defined as in (7), i.e., let q be the limiting proportion with which player 1 chooses T if player 1 plays as in Lemma 2. Then:

Lemma 4 *Let $a > b > c, d$ or $a > c > b > d$ hold. Let N_B be an increasing function characterizing player 2's best-response behavior, with 2 playing L when*

$n_{hB} > N_B(n_{hT})$ and playing R when $n_{hB} < N_B(n_{hT})$. Then there exists a function $\underline{N}_B(n_T) \leq N_B(n_T)$ such that, for sufficiently large δ and history h:

- *Player 1 plays T if $n_{hB} > N_B(n_{hT}+1)$;*
- *Player 1 plays B if $\underline{N}_B(n_{hT},\delta) < n_{hB} < N_B(n_{hT}+1)$;*
- *Player 1 plays T if $n_{hB} < \underline{N}_B(n_{hT},\delta)$.*
- *If $qa+(1-q)b > c$ (cf. (7), then*

$$\lim_{\delta\to 1} \sup\{n_T : \underline{N}_B(n_T,\delta) < 0\} = \infty.$$

Section 5.4 contains the proof. The new development here is that once player 1 has induced player 2 to choose L, 1 ensures that 2 thereafter always plays L. The state never subsequently crosses the border $N_B(n_T)$. Instead, whenever the state comes to the brink of this border, 1 drives the state away with a play of B before 2 has a chance to play R.

Figure 3 illustrates these strategies.

3.3 Equilibrium: Examples

We now combine these characterizations of best-response behavior to illustrate equilibria.

3.3.1 Example I: The Product Choice Game

Consider the product-choice game of Mailath and Samuelson [10], transcribed here as:[9]

	L	R
T	$3,0$	$1,1$
B	$2,3$	$0,2$

.

Player 2 is indifferent between L and R when $p = p^* = \frac{1}{2}$. Player 2's best response is described by Lemma 1, and player 1's by Lemma 4.

Let us assume there is a single mechanical type, characterized by the action $\alpha^1 = \frac{1}{10}$. Hence, the mechanical type plays B with high probability. Action B is the pure "Stackelberg" type for player 1 in this game, i.e., the pure action to which player 1 would most like to be committed, conditional on player 2 playing a best response.

A first observation is that, in equilibrium, we must have $\alpha^0 > p^*$. If not, player 2's expectation is that player 1 will always choose T with probability less than p^*, and hence player 2 would always choose L. The rational player 1 would

[9]To interpret the labels, we think of player 1 as a firm who can choose either high quality (B) or low quality (T), and player 2 as a consumer who can choose to buy either a custom product (L) or generic product (R) from the firm.

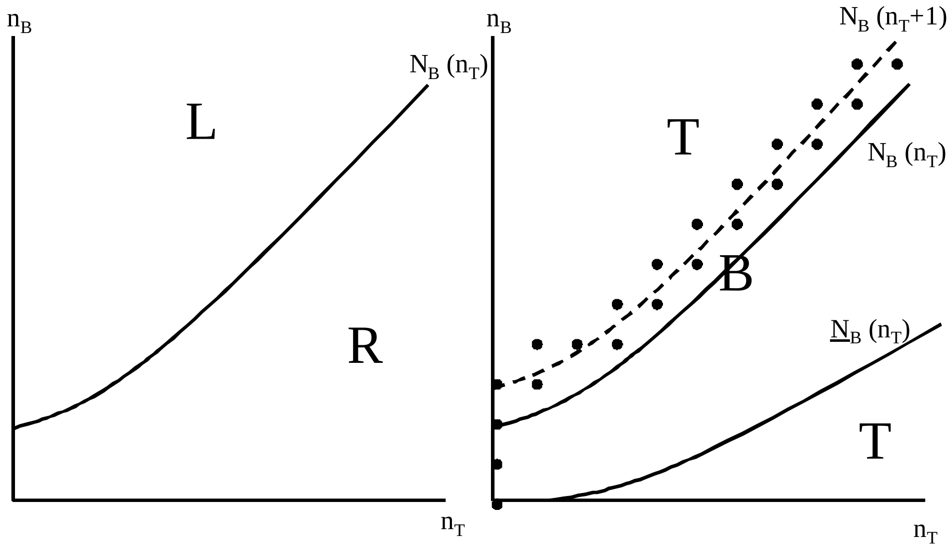

Figure 3: Strategies for the short-run player (left panel) and long-run player (right panel) given $a > b > d > c$. The function N_B is taken from Lemma 1. Condition $a > b > d > c$ is needed only for deriving player 1's best responses. Again, the qualitative properties of the functions are general; the detailed specification is not. An outcome path is shown, beginning at the origin and proceeding upward (whenever B is chosen) and to the right (whenever T is chosen). In this case, the function $\underline{N}_B$ from Lemma 4 is never reached in equilibrium and plays no role in shaping equilibrium behavior. Notice that once player 2 plays L, player 2 plays L in every subsequent period, with player 1 choosing T as often as is consistent with such player-2 behavior.

respond with the perpetual play of T, ensuring that the consistency condition (2) cannot hold and vitiating the existence of an equilibrium.

Given that $\alpha^0 > p^*$, let

$$\mu^{1*}\alpha^1 + (1-\mu^{1*})\alpha^0 = p^*. \tag{11}$$

If $\mu_h^1 > \mu^{1*}$, player 2 will choose L, while $\mu_h^1 < \mu^{1*}$ will cause player 2 to choose R. From Bayes' rule, the function N_B, defining player 2's strategy, must solve

$$\mu^{1*} = \frac{\mu_\emptyset^1(\alpha^1)^{n_{hT}}(1-\alpha^1)^{N_B(n_T)}}{\mu_\emptyset^1(\alpha^1)^{n_{hT}}(1-\alpha_1)^{N_B(n_{hT})} + \mu_\emptyset^0(\alpha^0)^{n_{hT}}(1-\alpha_0)^{N_B(n_{hT})}} = \frac{\mu_\emptyset^1}{\mu_\emptyset^1 + \mu_\emptyset^0\left(\frac{\alpha^0}{\alpha^1}\right)^{n_{hT}}\left(\frac{1-\alpha^0}{1-\alpha^1}\right)^{N_B(n_{hT})}}.$$

Hence, the function N_B is linear, and is of the form

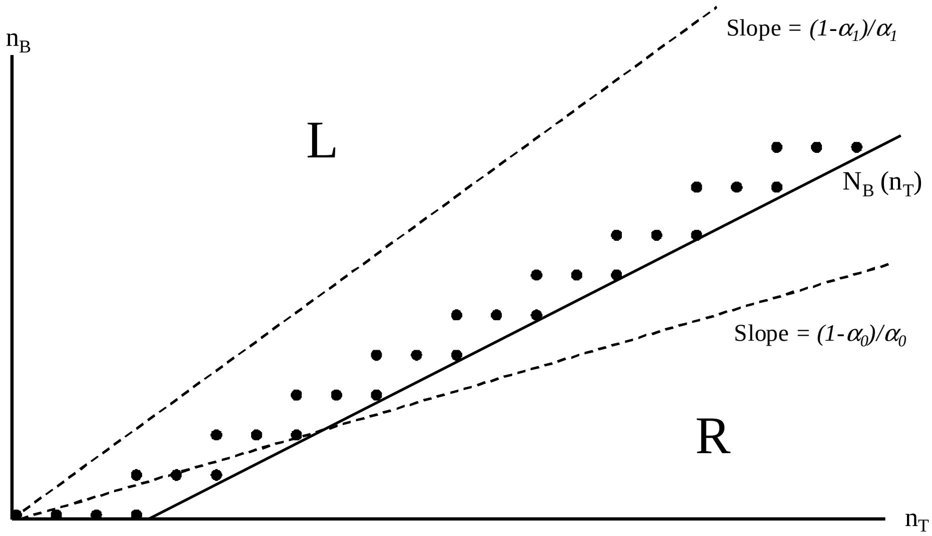

Figure 4: Player 2's strategy for the case of one mechanical type with α^1 (probability of T) falling short of p^*, along with a sample outcome path. Play begins with player 1 choosing T and player 2 choosing L, with the state climbing the horizontal axis (via a sequence of T actions on the part of player 1) until first threatening to cross the line $N_B(n_T)$. Play then ascends along the line $N_B(n_T)$. The boundary $\underline{N}_B$ is irrelevant to equilibrium behavior in this case, and is omitted.

$$N_B(n_{hT}) = -\frac{\ln \frac{\alpha^0}{\alpha^1}}{\ln \frac{1-\alpha^0}{1-\alpha^1}} n_{hT} + C$$

for some constant C. One can calculate that the slope of this line is larger than $(1-\alpha^0)/\alpha^0$ but smaller than $(1-\alpha^1)/\alpha^1$:

$$\frac{1-\alpha^0}{\alpha^0} < -\frac{\ln \frac{\alpha^0}{\alpha^1}}{\ln \frac{1-\alpha^0}{1-\alpha^1}} < \frac{1-\alpha^1}{\alpha^1}. \tag{12}$$

This is intuitive. The function N_B identifies ratios of B to T observations that keep the posterior on the mechanical type fixed at μ^{1*}. This ratio must lie between the ratio of B to T observations under the rational type, or $(1-\alpha^0)/\alpha^0$, and the ratio under the mechanical type, or $(1-\alpha^1)/\alpha^1$. Figure 4 illustrates these strategies.

Figure 4 assumes that $C < 0$. We now argue that this must be the case, for sufficiently large δ. Suppose $C > 0$. Then play would begin with a sequence of B plays from player 1. This seems an intuitive reputation-building stage.

Upon first reaching a period t with $t > N_B(0)$ (and hence $n_{hB} > N_B(0)$), player 1 would subsequently play T whenever $n_{hB} > N_B(n_{hT+1})$ and play B otherwise. This reputation-exploiting phase gives a proportion of B plays equal to the slope of N_B, and hence (from 12) a proportion of B greater than would the stationary strategy α^0. However, the consistency condition is that player 1's actions average to α^0. One cannot combine an initial phase in which only action B is played with a steady-state phase playing action B more than α^0 to get an average frequency of α^0. Hence, $C < 0$.

Play thus necessarily begins with a reputation-spending phase, in which player 1 continually plays T, until entering a reputation-manipulation phase. In the reputation-manipulation phase, player 1 chooses T as often as possible, consistent with inducing player 2 to always choose L. Whenever player 2's posterior brings player 2 to the brink of playing R, player 1 chooses B in order to keep player 2's beliefs inside the region in which L is a best response. Lemma 4 indicates that there are histories under which player 1 will eschew reputation building entirely and simply play T at every opportunity. It is clear that these histories cannot arise in equilibrium.

How can player 1 afford a reputation-spending stage if μ^1 (upon which we have placed no restrictions) is very small? During the reputation-manipulation stage, player 2 attaches positive probability to type α^0 as well as α^1, with the posteriors on these two types hovering around the level that makes 2 indifferent between L and R. If μ^1 is very small, then the equilibrium value of α^0 will be very close to p^*, so that the posterior attached to type α^1 will have to drop yet further in order to induce indifference.

For any fixed value of δ, player 2 remains perpetually uncertain as to which type of agent he faces. This stands in contrast to the findings of Cripps, Mailath and Samuelson [1, 2] that the types of long-run players in reputation games eventually must become known. To reconcile these seemingly contrasting results, notice that the short-run players in this case have a misspecified model of the long-run player's behavior. It need not be surprising that even overwhelming amounts of data do not suffice for players with a misspecified model to learn the true state of nature. At the same time, the short run players have a correct understanding of the aggregate play of the long run players. One might have thought that eventually the long run frequency of actions of the normal long run player would coincide with this aggregate frequency, thereby leading to an identification of the rational type of player 1. This intuition is correct for the limiting case in which δ gets close to 1, but away from this limit, the long run frequency need not coincide with the aggregate behavior, and hence the convergence of beliefs need not obtain.

3.3.2 Example II: A Zero-Sum Game

Consider the zero-sum game:

	L	R
T	$7, -7$	$0, 0$
B	$3, -3$	$5, -5$

. (13)

Player 2's best-response behavior is described by Lemma 1, and player 1's behavior is now described by Lemma 2.

In the standard model, reputations have no value in this game. The minmax strategies for the two players in this game are $(\frac{2}{9}T, \frac{7}{9}B)$ for player 1 and $(\frac{5}{9}L, \frac{4}{9}R)$ for player 2. This gives player 1 a payoff of $\frac{35}{9}$. No commitment type can give player 1 a larger payoff. For example, a commitment to T gives a payoff of 0, while a commitment to B gives a payoff of 3. Nonetheless, player 1 is able to effectively exploit the presence of the mechanical types in the current context.

We consider the case in which there is only one mechanical type playing T with probability α_1 strictly smaller than p^* ($= \frac{2}{9}$ in (13)).

The equilibrium is determined by a function N_B with the property that TL is played at histories h in which $n_{hB} > N_B(n_{hT})$ and BR is played at histories in which $n_{hB} < N_B(n_{hT})$. We can repeat the reasoning of the previous example to conclude that

$$N_B(n_{hT}) = -\frac{\ln \frac{\alpha^0}{\alpha^1}}{\ln \frac{1-\alpha^0}{1-\alpha^1}} n_{hT} + C$$

for some constant C.

In equilibrium, we must again have $\mu^{1*} \leq \mu^1_\emptyset$ (cf. (11)). The game begins with a sequence of TL plays. Player 1 here starts with an excess reputation, and exploits player 2 while pushing this reputation downward. Once μ^1_h hits μ^{1*}, play moves back and forth between TL and BR (as the posterior of the mechanical type dips below or moves above μ^{1*}). The former appears in weighted proportion less than α_0 (so that the consistency condition, that player 1's play averages to α_0, is satisfied). We can then again divide these strategies into a reputation spending stage and a second "reputation-manipulation" stage, the latter beginning the first time the posterior belief μ^1_h crosses μ^{1*}.

The payoff obtained by the long-run player in this game is above the value. Equilibrium play includes some periods in which $n_{hB} > N_B(n_{hT})$ and the outcome is TL for a player-1 payoff of 7, and some periods in which $n_{hB} < N_B(n_{hT})$ and the outcome is BR for a player-1 payoff of 5. Both payoffs exceed player 1's value, and hence so must player 1's equilibrium payoff.[10]

3.3.3 Example III: Multiple Mechanical Types

Consider once again the zero-sum game of Section 3.3.2, but now with two or more mechanical types, with at least one playing T with probability exceeding p^* and one playing T with probability falling short of p^*. The important feature of this configuration is that there are mechanical types on both sides of p^*. Given that this is the case, introducing additional mechanical types does nothing but complicate the calculations.

[10] We have not specified what actions the players take when $n_{hB} = N_B(n_{hT})$. In particular, player 2 might mix after such a history. However, whatever (possibly mixed) action player 2 takes at such a history, player 1 has an action available that ensures player 1 at least her value, and that leads to continuation payoffs above 1's value, ensuring that 1's equilibrium payoff exceeds her value.

Lemmas 1 and 2 again describe the players' best response functions, and so the equilibrium here will share many of the features of the equilibrium calculated in Section 3.3.2. The function N_B now need not be linear, though it will be increasing.

Once again, there will be an initial phase, consisting of either a reputation-building phase consisting of a string of B actions, or a reputation-spending phase consisting of a string of T actions. This initial phase will last until either the first time a B action causes the state (n_{hT}, n_{hB}) to cross from below to above the function N_B, or until the first time a T action causes the state (n_{hT}, n_{hB}) to cross from above to below the function N_B. Thereafter, play will enter a reputation-manipulation phase in which the state hovers as close as possible to the graph of the function N_B.

In Section 3.3.2, the initial phase necessarily involved a string of T actions, which have the effect of diminishing the probability attached to the mechanical type. When there are mechanical types on both sides of p^*, the initial phase may consist of a string of either B or T plays, depending on the prior distribution over mechanical types.[11]

What determines whether player 1 initially builds or spends her reputation? Suppose (counterfactually) that only the mechanical player-1 types $\{\alpha^1, \ldots, \alpha^K\}$ were present, with the probabilities $\{\mu_0^1, \ldots, \mu_0^K\}$ scaled up to sum to one. If player 2 would play R against this mixture, then equilibrium play in the original game must constitute a sequence of BR plays. If player 2 would play L against this mixture, then equilibrium play in the initial phase of the original game must constitute a sequence of TL plays. Player 1's initial play must then push player 2 away from the action player 2 would choose against the mechanical types.

To see that this is the case, suppose player 2 would play R if facing only the mechanical types. Our best-response characterizations give us two possible configurations for equilibrium play. First, it may be that $\alpha^0 > p^*$. Given that player 2 would initially play R against the mechanical types, the addition of such a rational type reinforces 2's initial proclivity to play R. From Lemmas 1–2, play then consists of an initial sequence of BR plays, until reaching the reputation-manipulation phase. Here, player 2's beliefs become concentrated on the two types of player 1 that straddle p^*, namely $\min\{\alpha^0, \overline{\alpha}\}$ and $\underline{\alpha}$, where $\underline{\alpha}$ denote the largest α^k such that $\alpha^k < p^*$ and $\overline{\alpha}$ denotes the smallest α^k such that $\alpha^k > p^*$. Player 1's average play $\hat{\alpha}^0$ during this reputation-manipulation phase must then be below α^0. However, there is no way to combine an initial sequence of B plays with a subsequent phase that is relatively heavy on B ($\hat{\alpha}^0 < \alpha^0$) and get average play of α^0. The equilibrium must then feature $\alpha^0 < p^*$, and hence the reputation-manipulation phase must be comprised of behavior that is relatively heavy on T ($\hat{\alpha}^0 > \alpha^0$). Player 1's initial phase of play must then

[11] Section 3.3.2 examined a case in which there was a single mechanical type with $\alpha^1 < p^*$, finding that equilibrium required a string of initial "reputation-spending" T actions. If we examine an equilibrium with a single mechanical type with $\alpha^1 > p^*$, equilibrium would again require a string of initial "reputation-spending" B plays. An initial reputation-spending phase is thus a robust feature of specifications in which all of the mechanical types lie on the same side of p^*, which no longer holds when there are mechanical types on both sides of p^*.

feature B, and hence (from Lemmas 1–2), play must begin with a sequence of BR plays.

For a fixed distribution of mechanical types, there is an upper bound M on the length of the introductory phase, that holds over all equilibria for all discount factors. To see that this is the case, note that the function N_B depends only on the set of types $(\alpha^0, \alpha^1, \ldots, \alpha^K)$ and the prior $(\mu^0_\emptyset, \mu^1_\emptyset, ..., \mu^K_\emptyset)$, and not the discount factor. We must accordingly ask the following questions. Suppose the initial specification is such that $N_B(0) > 0$, i.e., player 2 initially plays R. What is the maximum (over α^0) of the minimum number t of successive observations of B it would take for player to play L, i.e., what is the smallest t for which $t > N_B(0)$? This minimum number t is maximized either when $\alpha^0 = 0$ or $\alpha^0 = 1$, where it takes a finite value that we can bound by M. Similarly, we can suppose the initial specification is such that $N_B(0) < 0$, i.e., player 2 initially plays L. What is the maximum (over α^0) of the minimum number t of successive observations of T it would take for player to play R, i.e., what is the smallest t for which $0 < N_B(t)$? This value is again maximized at either $\alpha^0 = 0$ or $\alpha^0 = 1$, where it takes a finite value that we can bound by M.

In the limit as $\delta \to 1$, player 2's limiting (as t gets large) posterior on the rational type must converge to 1. Eventually the empirical frequency of past play of the rational type must be very close to α^0, by the consistency condition, and thus α^0 must converge to p^* as $\delta \to 1$ (as otherwise, player 2s would not be kept to be playing either L or R in the long run).

More precisely, as $\delta \to 1$, the bounded initial phase is followed by a reputation manipulation phase of arbitrarily long (expected) length. The consistency condition given by (2) can then be satisfied only if $\hat{\alpha}^0$, the average play of the rational type during the reputation-manipulation phase, approaches α^0, the overall average play of the rational type. During the reputation-manipulation phase, the average play $\hat{\alpha}^0$ of the rational type must balance player 2's posterior over types so that player 2 remains nearly indifferent over player 2's actions. In the limit, player 2's posterior will be concentrated on only two types of player 1, being the two closest types on either side of $\hat{\alpha}^0$. One of these types will be α^0, and one will be a mechanical type. But if $\hat{\alpha}^0$ is converging to α^0, the consistent play of $\hat{\alpha}^0$ can maintain player 2's near indifference only if $\hat{\alpha}^0$ and α^0 are both very close to p^*, and if the posterior probability placed on the rational type converges to 1.

The result that player 2 learns player 1's type is a double limiting result, referring to the limit of player 2's beliefs as t gets large, in a sequence of equilibria for games in which δ gets large. For any fixed δ, player 2 remains perpetually uncertain as to player 1's type.

3.4 Equilibrium: Analysis

We now characterize equilibria for the general case with many mechanical types. We retain (5), ensuring that player 2 does not have a dominant strategy, but do not restrict player 1's payoffs. We fix a specification of the mechanical types' actions $(\alpha^1, \ldots, \alpha^K)$ and prior probabilities $(\mu^1_\emptyset, \ldots, \mu^K_\emptyset)$. We assume there is at least one mechanical type that plays T with probability greater than p^*, and one that plays T with probability less than p^*. There may be many mechanical types.

Let $\underline{\alpha}$ denote the strategy of the mechanical type who attaches the largest probability less than p^* to T. Let $\overline{\alpha}$ be the strategy of the mechanical type who attaches the smallest probability larger than p^* to T. Then let q^* satisfy

$$\underline{\alpha}^{q^*}(1-\underline{\alpha})^{1-q^*} = \overline{\alpha}^{q^*}(1-\overline{\alpha})^{1-q^*}.$$

Intuitively, a collection of actions featuring q^* proportion of T is equally likely to have come from mechanical type $\underline{\alpha}$ as from mechanical type $\overline{\alpha}$. When observing a sufficiently long string of such data, player 2 will rule out the other mechanical types, but will retain both $\underline{\alpha}$ and $\overline{\alpha}$ as possibilities.

3.4.1 Existence of Equilibrium

Proposition 1 *There exists a sequential analogy-based expectation equilibrium.*

Proof. Intuitively, we think of (1) fixing α^0, an average strategy for the long-run player, (2) deducing player 2's best responses, (3) deducing player 1's best responses, and (4) calculating the values of A^0 implied by these best responses. This gives us a map from values of α^0 to values of A^0. A fixed point of this map gives us an equilibrium. The details of this argument are fairly standard, and are given in Section 5.5. ■

3.4.2 Pure-Outcome Equilibria

We say that the equilibrium outcome is pure if either $\alpha^0 = 0$ or $\alpha^0 = 1$. This does not mean that the equilibrium features pure strategies, since player 1 may mix at out-of-equilibrium histories. However, player 2 models player 1 as playing a pure strategy, and will receive no contradictory evidence along the equilibrium path. In other cases, we say the equilibrium outcome is mixed.

When does a pure-outcome equilibrium exist? We can assume $c > b$ without losing any generality, with the case $c < b$ simply being a relabeling.

Proposition 2 *Let $c > b$. Then there exists a pure-outcome equilibrium for sufficiently large δ (i.e., there exists a $\underline{\delta} \in (0,1)$ such that a pure-outcome equilibrium exists for any $\delta > \underline{\delta}$) if and only if*

$$c > d \quad (14)$$

$$c > q^* \max\{a, c\} + (1-q^*) \max\{b, d\}. \quad (15)$$

If (14)–(15) hold, then for any $\varepsilon > 0$, there is a $\underline{\delta}(\varepsilon) < 1$ such that for all $\delta > \underline{\delta}(\varepsilon)$, every pure-outcome equilibrium payoff for player 1 exceeds $c - \varepsilon$, as does every equilibrium payoff.

Proof. [*Necessity*] Suppose $c < d$ and we have a candidate pure-outcome equilibrium with either $\alpha^0 = 0$ or $\alpha^0 = 1$. As $\delta \to 0$, the payoff from such an equilibrium approaches b in the first case and c in the second. Consider a strategy in which player 1 chooses T if the cumulative frequency which which he has played T falls short of $\overline{\alpha}$, and otherwise plays B. Then after a finite number of periods, Player 2 will attach sufficiently large probability to player 1 being type $\overline{\alpha}$ as to thereafter always play R. Hence, except for a bounded number of initial periods, which become insignificant as $\delta \to 1$, player 1 earns a payoff of $\overline{\alpha}c + (1 - \overline{\alpha})d$ which exceeds both of b and c, and hence exceeds the payoff of any pure equilibrium. This ensures that there for sufficiently large δ, there are no pure-outcome equilibria.

Alternatively, suppose $c > d$ but (15) fails, in which case (given $c > b, d$) we must have $a > c$ and $c < q^*a+(1-q^*)\max\{b, d\}$. Fix a candidate pure-outcome equilibrium yielding payoff c. Now suppose player 1 undertakes a strategy of initially playing B, until player 2's posterior belief is pushed to indifference between L and R. Thereafter, player 1 plays actions that keep the realized histories near the function $N_B(n_T)$. If $b < d$, player 1 allows the history to cross back and forth over the line, giving a mixture between payoffs a and d. If $b > d$, player 1 ensures that the history lies always just above this boundary, giving a mixture between payoffs a and b. The probability attached to a in either of these mixtures is q^*, which suffices for the result.

The sufficiency result is similar, and is relegated along with the payoff characterization to Section 5.6. ∎

3.4.3 Mixed-Outcome Equilibria

When will there exist a mixed-outcome equilibrium? Sections 3.3.1 and 3.3.2 have illustrated two mixed equilibria. The limiting payoff in the first of these equilibria is given by $p^*a+(1-p^*)b$, and in the second is given by $p^*a+(1-p^*)d$. In each case, it was important that this payoff exceeded c, since otherwise a sufficiently patient long-run player 1 could ensure a payoff arbitrarily close to c simply by always playing T. This suggests the conjecture that (retaining our convention that $c > b$) there exists a mixed-outcome equilibrium as long as

$$c < p^* \max\{a, c\} + (1 - p^*) \max\{b, d\},$$

and that the payoff in this equilibrium is given by $p^* \max\{a, c\}+(1-p^*) \max\{b, d\}$. This is indeed a sufficient condition for existence, but a glance at Proposition 2 suggests that it is not the only sufficient condition. There is no pure-outcome equilibrium if $c < q^* \max\{a, c\} + (1-q^*) \max\{b, d\}$, and the latter is indeed also sufficient for the existence of a mixed-outcome equilibrium. Section 5.7 proves:

Proposition 3 *Let $c > b$. Then for sufficiently large δ, a mixed-outcome equilibrium exists if and only if at least one of the following holds:*

$$c < p^* \max\{a, c\} + (1 - p^*) \max\{b, d\} \quad (16)$$

$$c < q^* \max\{a, c\} + (1 - q^*) \max\{b, d\}. \quad (17)$$

The first step in the proof is straightforward. If $c < p^* \max\{a, c\} + (1 - p^*) \max\{b, d\}$, we show that there exists a mixed-outcome equilibrium analogous to that of Sections 3.3.1 and 3.3.2. A fixed point argument establishes the existence of such an equilibrium. This leaves one case to be addressed, namely that in which

$$p^* \max\{a, c\} + (1 - p^*) \max\{b, d\} < c < q^* \max\{a, c\} + (1 - q^*) \max\{b, d\}. \quad (18)$$

Notice, from Proposition 2, that there is no pure equilibrium for this case, while the first inequality ensures there is no mixed equilibrium analogous to those of Sections 3.3.1–3.3.2.

To see how we proceed, it is helpful to acquire some notation. We say that a sequence of equilibria, as δ converges to 1, with player-1 average strategy $\{\alpha^0_\ell\}_{\ell=0}^\infty$ is unary if $P\{h : |\alpha^0_\ell(h) - \alpha^0_\ell| > \varepsilon\}$ converges to zero, for all $\varepsilon > 0$. Hence, in the limit the average strategy of player 1 is independent of history. Otherwise, the equilibrium is binary (a term justified by the following lemma). A pure equilibrium is obviously unary. The mixed equilibria of Sections 3.3.1–3.3.2 are unary.

We have already concluded that when (18) holds, the (only) equilibrium is a binary, mixed equilibrium. Notice that (18) can hold only if $b, d < c < a$ and $q^* > p^*$, the former placing constraints on the payoffs in the game and the latter on the distribution of mechanical types.

We can then construct an equilibrium as follows. In the first period, player 1 is indifferent between T and B, and mixes, placing probability ζ on T. If the first action is T, then player 1 plays T thereafter. If the first action is B, then player 1 plays B until making player 2 indifferent between L and R, after which point player 1 maintains this indifference. This gives a long-run average T play of $\underline{\alpha}^0 > p^*$. We have aggregate play for player 1 of

$$\alpha^0 = \zeta + (1 - \zeta)\underline{\alpha}^0.$$

It is then a straightforward calculation, following from the facts that $p^* \max\{a, c\} + (1 - p^*) \max\{b, d\} < c < q^* \max\{a, c\} + (1 - q^*) \max\{b, d\}$ and $q^* > p^*$ that we can choose $\underline{\alpha}^0$ and ζ so that

- Player 1 is indifferent over the actions T and B in 1's initial mixture, This requires adjusting $\underline{\alpha}^0$ so that the payoff to player 1 from building and maintaining player 2's indifference is c,

- Probability $\underline{\alpha}^0$ makes player 2 indifferent between L and R. This requires adjusting ζ and hence α^0 so that $\underline{\alpha}^0$ causes player 2's posterior to concentrate probability on types $\underline{\alpha}$ and α^0,

completing the specification of the equilibrium.

Are there equilibria in which player 1 mixes over more than two continuation paths? The answer is no:

Lemma 5 *Let $c > b$. There is a value $\underline{\delta}$ such that for all $\delta \in (\underline{\delta}, 1)$, in any equilibrium that is not unary, there are two long-run averages of play for player 1. Player 1's payoff in any sequence of such equilibria converges to c as $\delta \to 1$.*

Proof. Fix a candidate equilibrium and the associated value α^0. Then let $\underline{\underline{\alpha}}$ and $\overline{\overline{\alpha}}$ be (respectively) the largest frequency smaller than p^* with which a type of player 1 plays T, and smallest frequency larger than p^* with which a type of player 1 plays T. These differ from $\underline{\alpha}$ and $\overline{\alpha}$ in that we now include the rational type of player 1 in the set of possibilities. Let $\hat{q}$ satisfy

$$\underline{\underline{\alpha}}^{\hat{q}}(1-\underline{\underline{\alpha}})^{1-\hat{q}} = \overline{\overline{\alpha}}^{\hat{q}}(1-\overline{\overline{\alpha}})^{1-\hat{q}}.$$

Then player 1 can attain a payoff arbitrarily close (for large δ) to c by always playing T, and otherwise the largest payoff player 1 can obtain is

$$\hat{q}\max\{a,c\} + (1-\hat{q})\max\{b,d\}.$$

These are accordingly the only two payoffs that can be attached positive probability in a mixed equilibrium that is not unitary, and will both appear only if equal. But then only two long-run averages of play for player 1 can appear. ■

3.4.4 Payoffs

We can now collect our results to characterize equilibrium payoffs and behavior. It is convenient to start with payoffs. To conserve on notation, let

$$P^* := p^*\max\{a,c\} + (1-p^*)\max\{b,d\} \quad (19)$$

$$Q^* := q^*\max\{a,c\} + (1-q^*)\max\{b,d\} \quad (20)$$

Section 5.8 proves:

Proposition 4 *let $c > b$. For sufficiently large δ:*

[4.1] If $P^*, Q^* < c$, then the only equilibrium is pure, featuring payoff c.

[4.2] If $c < P^*, Q^*$, then there exist unary mixed equilibria. The rational player 1's behavior in a unary sequence of mixed-outcome equilibrium satisfies $\lim_{\delta\to 1}\alpha^0(\delta) = p^*$, and the limiting equilibrium payoff of the rational player 1 is given by P^*. If $c > d$, there may also exist a binary mixed equilibria, with payoff c for the rational player 1.

[4.3] If $Q^* < c < P^*$, then there exists a pure equilibrium if $c > d$. There also exists a unary mixed equilibria, and the rational player 1's behavior in a sequence of unary mixed-outcome equilibrium satisfies $\lim_{\delta\to 1}\alpha^0(\delta) = p^*$, and the limiting equilibrium payoff of the rational player 1 is given by P^*.

[4.4] If $P^* < c < Q^*$, then the only equilibrium is a binary mixed equilibrium, and the rational player 1's payoff in any such equilibrium approaches c as $\delta \to 1$.

We can summarize our results as follows.

Parameters	Equilibria	Payoffs
$P^*, Q^* < c$	Pure	c
$c < P^*, Q^*$	Unary mixed	P^*
$c < P^*, Q^*$	Binary mixed (possibly, and only if $c > d$)	c
$Q^* < c < P^*$	Pure (if and only if $c > d$)	c
$Q^* < c < P^*$	Unary mixed	P^*
$P^* < c < Q^*$	Binary mixed	c

.

Example. If $Q^* < c < P^*$ and $c > d$, there exists both a pure and mixed equilibrium. For example, consider the game:

	L	R
T	$6, 0$	$3, 1$
B	$2, 1$	$0, 0$

.

Notice that $p^* = 1/2$. Let there be two mechanical types, characterized by the probability they attach to playing T, with this probabilities being .01 and .51. We thus have $Q^* < c < P^*$. There is then a pure equilibrium, with payoff $c = 3$, and a mixed equilibrium, whose payoff converges to $p^*a + (1-p^*)b = p^*6 + (1-p^*)2 = P^* = 4$ as $\delta \to 1$.

The question is then why the possibility of obtaining payoff P^* does not preclude the existence of a pure equilibrium, which yields a payoff of 3 for player 1. Let us consider a candidate pure equilibrium, in which $\alpha^0 = 1$, so that the rational type of player 1 always chooses T, for a payoff of 3. Why cannot player 1 earn a higher payoff, namely P^*, given this candidate equilibrium? Player 1 can endure an initial phase in which the posterior that 2 attaches to 1 playing T can be pushed to the point at which 2 is indifferent between L and R, with this indifference thereafter maintained. In calculating 1's payoff, we can ignore the initial phase (by focussing on $\delta \to 1$), and the payoff will be very close to

$$q^*6 + (1-q^*)2,$$

where q^* is the frequency required to maintain the posterior near p^*, namely

$$q^* \ln \frac{.01}{.51} = (1-q^*) \ln \frac{.49}{.99}.$$

This gives us a value of q^* equal to approximately .15, and Q^* equal to approximately 2.6. As a result, player 1 will get a higher payoff from simply playing T all of the time and receiving 3, rather than the mix $q^*6 + (1-q^*)2$. Hence $p^*a + (1-p^*)b$ is not available as a payoff to player 1 given the equilibrium hypothesis of $\alpha^0 = 1$, and we thus have multiple equilibria. ■

Remark 1 The value of q^* must lie between the probabilities $\underline{\alpha}$ and $\overline{\alpha}$, the former the largest probability less than p^* attached to T by a mechanical type, and the latter the smallest probability larger than p^* attached to T by a mechanical type. If the set of mechanical types becomes rich, such as would be the case with a sequence of increasingly dense grids of mechanical types, the value of q^* must then approach p^*. This will eventually (generically) ensure that P^* and Q^* are on the same side as c, precluding the type of coexistence of pure and mixed equilibria exhibited in the preceding example. ■

Remark 2 If $c < P^*, Q^*$, there always exists a unary mixed equilibrium (with payoff approaching P^* as $\delta \to 1$), and there may or may not also exist a binary mixed equilibrium (with payoff c). First, it is clear that such existence requires $c > d$, since otherwise it can not be optimal to put positive probability on always playing T (as does the binary mixed equilibrium), and this in turn implies that we must have $b, d < c < a$. Second, the lower probability $\underline{\alpha}^0$ in this equilibrium must be enough smaller than p^* as to push the expected payoff from this outcome down to c. This in turn requires that $\underline{\alpha}$, largest probability less than p^* attached to T by a mechanical type, must be sufficiently small. Hence, if $\underline{\alpha}$ is sufficiently close to p^*, perhaps because the set of mechanical types is sufficiently rich, then binary mixed equilibria will not exist for this case.[12] ■

We can combine the insights of Remarks 1 and 2. Let us say that the set of mechanical types is ε-rich if there is no interval subset of $[0, 1]$ of length exceeding ε that does not contain a mechanical type.

Corollary 1 *Let $c > b$. Consider the (generic) set of games for which $c \neq P^*$.*

[1.1] There is an $\varepsilon > 0$ such that if the set of mechanical types is at least ε-rich, then any equilibrium is either pure or unary mixed.

[1.2] There is an $\varepsilon > 0$ such that if the set of mechanical types is at least ε-rich and δ is sufficiently large, then player 1's equilibrium payoff, is at least

$$\max\{b, c, p^* \max\{a, c\} + (1 - p^*) \max\{b, d\}\} - \varepsilon.$$

It is natural to compare our reputation results to those of Fudenberg and Levine [4]. Their result is that

$$\lim_{\delta \to 1} U_1^*(\delta) \geq \max_{\alpha^k \in \{\alpha^1, \ldots, \alpha^K\}} \min_{a_2 \in BR(\alpha^k)} u_1(\alpha^k, a_2),$$

[12] To see the issues here, consider the game:

	L	R
T	$5, 0$	$2, 1$
B	$0, 1$	$1, 0$

Then $p^* = 1/2$. A binary mixed equilibrium requires $\underline{\alpha}^0$ to equal approximately $1/4$, so that player 1 is indifferent (and hence willing to mix) between always playing T, for a payoff of approximately 2, and mixture $\underline{\alpha}^0$, for a payoff of $5\underline{\alpha}^0 + (1 - \underline{\alpha}^0)$. We can construct such an equilibrium if (for example) $\underline{\alpha} = 1/10$ and $\overline{\alpha} = 9/10$. However, since we must have $\underline{\alpha} < \underline{\alpha}^0 < \overline{\alpha}$, a binary mixed equilibrium does not exist if $\underline{\alpha} \geq 1/4$.

where $U_1^*(\delta)$ is player 1's equilibrium payoff and $BR(\alpha^k)$ is the set of best responses for player 2 to the player-1 action α^k. Intuitively, player 1 can choose her favorite mechanical type, and then receive the payoff she would earn if she were known to be that type, given that player 2 plays the best response to that type that is least favorable to player 1.

A proof virtually identical to that used to establish Fudenberg and Levine's lower bound ensures that player 1 in our setting is assured a payoff at least as high. The following result is a corollary of Watson [13]:

Proposition 5 *In any sequential analogy-based expectation equilibrium,*

$$\lim_{\delta \to 1} U_1^*(\delta) \geq \max_{\alpha^k \in \{\alpha^1, \ldots, \alpha^K\}} \min_{a_2 \in BR(\alpha^k)} u_1(\alpha^k, a_2).$$

Proof. Suppose there exists a sequential analogy-based expectations equilibrium with an equilibrium payoff for player 1 that falls short of $u_1(\alpha^k, a_2)$ for some mechanical type α^k and $a_2 \in \arg\min_{a_2 \in BR(\alpha^k)} u_1(\alpha^k, a_2)$. Then one feasible strategy available to player 1 is to choose T with probability α^k in each period. The argument then follows from the fact that within a finite number of periods (independent of δ), player 2 will place sufficiently high probability on mechanical type α^k as to play a best response, which suffices for the result. ■

Our lower bound is often tighter than that of Fudenberg and Levine. In particular, the payoff of a unary mixed equilibrium satisfies

$$\lim_{\delta \to 1} U_1^*(\delta) \geq P^* = p^* \max\{a, c\} + (1 - p^*) \max\{b, d\}.$$

This limit payoff typically exceeds the lower bound of Fudenberg and Levine, and it does so for potentially two reasons. First, P^* is independent of the specifications of the mechanical types α^k. So unless there are mechanical types characterized by actions arbitrarily close to p^*, the bound here will be higher. Second, even if there are mechanical types characterized by actions close to p^*, the bound found by Fudenberg and Levine will not exceed $\max\{p^*a + (1 - p^*)b, p^*c + (1 - p^*)d\}$ (corresponding to the Stackelberg payoff when the long run player can commit to a behavior either slightly above or below p^*), and this bound is (strictly) smaller than P^* for a range of games (including zero-sum games).

To illustrate this, we offer two examples. First, consider the product-choice game:

	L	R
T	$3, 0$	$1, 1$
B	$2, 3$	$0, 2$

.

Recall that $p^* = \frac{1}{2}$. If there is a mechanical type on each side of p^*, then there exists a mixed-outcome equilibrium in which player 1 earns a payoff very close to $\frac{5}{2}$. If these mechanical types are roughly equally spaced around p^*, then every equilibrium has this property. However, if the types are not too close to p^*, the Fudenberg-and-Levine bound will be smaller.

Second, consider our familiar zero-sum game. In such games, there is no value to commitment and thus the limiting per period payoff obtained by the rational long-run player when there are very few mechanical types must be close to his value when short-run players are fully rational. By contrast, when short-run players use the mode of reasoning analyzed above, the rational long-run player will typically get a payoff strictly above the value.[13] For the sake of illustration, consider the zero-sum game of (6):

	L	R
T	$7, -7$	$0, 0$
B	$3, -3$	$5, -5$

.

We have already calculated that the rational long-run player can guarantee a per period payoff of $P^* = 7p^* + 5(1-p^*)$, whereas the value of this game is only $7p^* + 3(1-p^*)$ with $p^* = \frac{2}{9}$. This latter phenomenon is not unique to zero-sum games. The auditing game given by (1) has the same property.

3.4.5 Equilibrium Behavior

This subsection turns the attention from equilibrium payoffs to equilibrium behavior, stressing three aspects of such behavior. We have characterized this behavior in the course of proving Propositions 2–4, and we need only summarize this characterization here:

Corollary 2

[2.1] Player 1's play, in any equilibrium, can be divided into two phases, including an initial "reputation-building" or "reputation-spending" phase and a subsequent "reputation-manipulation" phase.

[2.2] Throughout the reputation-manipulation phase, player 2 remains nearly indifferent over L and R. Player 1 manages to correlate her actions with those of player 2, allowing a higher payoff than is possible under uncorrelated mixtures.

[2.3] The reputation-manipulation phase is nonexistent in a pure equilibrium. In a unary mixed equilibrium, the length of the initial phase remains bounded as $\delta \to 1$, while the expected length of the reputation-manipulation phase grows arbitrarily long.

[2.4] The action profile played in the initial phase of a unary equilibrium is BR is player 2's best response to the mechanical types (only) is B, and otherwise is TL. Player 2's initial action in such an equilibrium is a best response to the mechanical types, and player 1's initial sequence of actions pushes player 2 away from this behavior and toward indifference.

[2.5] For any fixed δ, in any unary mixed equilibrium, player 2 remains uncertain throughout the game as to the type of player 1.

[13] Such an observation about zero-sum games was anticipated by Ettinger and Jehiel [3] in the simpler context of a two-period interaction yet requiring a significant chunk of non-rational long-run players. Our present analysis allows us to provide a full analysis under the same conditions as the ones in Fudenberg and Levine's [4] classic contribution on reputation.

In general, we think of player 2 as having access to historical information concerning the types of past player 1s, and their average frequency of play. In many cases, we think this is quite reasonable. We can well imagine consumer reporting agencies indicating that there are low-quality providers who often provide bad service, as well as high quality providers, who rarely provide bad service. In the case of unary equilibria, however, the informational demands are weaker. Here, player 2 need only have access to the average frequency of play of past types. If there exist mechanical types α^1 and α^2 as well as a rational type characterized by α^0, player 2 will observe in the historical record a collection of cases in which player 1's play matched α^1, a collection in which 1's play matched α^2, and a collection in which play matched α^0. Player 2 can then interpret this evidence as indicating there are three types of player 1, in prior probabilities equal to their relative frequency in the data. Player 1 has no way of knowing which is the rational and which the mechanical types, but also has no need of knowing this.[14]

3.5 Fictitious Play

The distinguishing feature of our model is that player 2 models player 1's behavior as stationary, even if (as in the case of a rational player 1) this need not be the case. Another setting in which players potentially mistakenly model the play of their opponents as stationary is that of fictitious play. A comparison is instructive. Consider a model in which there are no mechanical types of player 1, but player 2 plays a best response to a fictitious-play model of player 1.

Having reached period t with history h, player 2 computes the empirical frequency with which player 1 has played T, or

$$\frac{n_{hT}}{t}.$$

Player 2 then plays L if this empirical frequency falls short of p^*, and plays R if this empirical frequency exceeds p^*. Intuitively, player 2 views player 1 as playing a stationary strategy corresponding with the empirical frequency of 1's play, to which 2 plays a best response.

We can describe this behavior in more familiar terms.

Lemma 6 *Let player 2 play the game in period t with history h. Then 2's action is given by*

$$\begin{array}{lll} L & \text{if} & n_{hT} < p^*t \\ R & \text{if} & n_{hT} > p^*t. \end{array}$$

[14]In the case of binary equilibria, the record must include types, since the rational player 1 will sometimes give rise to one long-run average behavior and sometimes to another, and player 2 must amalgamate both into a single type of player 1. We note that if each mode in a non-unary equilibrium is interpreted as a different (stationary) type, we may run into existence issues.

Player 2's behavior is thus once again described by a function

$$N_B(n_{hT}) = \frac{1-p^*}{p^*} n_{hT}.$$

The function N_B is thus a ray through the origin.

Player 1's best response behavior is again characterized by Section 3.2. Now, however, there is no equilibrium condition to be satisfied. Player 2 is an automaton, and characterizing player 1's behavior is equivalent to characterizing equilibrium behavior. The fact that N_B is a ray through the origin indicates that there is now no initial reputation-building or reputation-spending phase. Instead, player 1 moves immediately to reputation manipulation. We then immediately have:

Proposition 6 *Suppose player 1 faces a fictitious-play opponent and that in case of indifference on the part of player 2, player 1 is free to pick player 2's behavior. Then:*

(6.1) Player 1's equilibrium payoff, in the limit as $\delta \to 1$ is given by

$$\max\{b, c, p^* \max\{a, c\} + (1-p^*) \max\{b, d\}\}. \tag{21}$$

(6.2) The frequency with which player 1 plays T is given by 0 (if b is the maximizer in (21)), 1 (if c is the maximizer in (21)), or p^* (if $p^* \max\{a, c\} + (1-p^*) \max\{b, d\}$ is the maximizer in (21)).

For generic games, instances will not arise in which player 2 is indifferent, allowing us to dispense with the assumption that player 1 can then choose player 2's behavior.

From Corollary 1, as the the set of mechanical types in our model becomes rich, player 1's payoff approaches the payoff player 1 could achieve against a fictitious-play opponent.

4 Discussion

We have examined reputation models in which short run players reason as if all types of long run players behaved in a stationary way. This belief is correct for most types of player 1, but will typically not be true of the (in most models) most likely type, namely the rational player 1. Player 2's beliefs about the rational type are not arbitrary, instead being required to match the long run empirical frequency of play of the type. We view such beliefs as natural for cases in which player 2 can most readily collect information about average frequencies of play. Player 2's model of player 1 then makes use of all of the information at 2's disposal, and is contradicted by nothing that player 2 could observe.[15]

[15] Player 2 observed the frequencies with which player 1 has played her various actions in previous games, with our equilibrium conditions ensuring that these match the strategies that player 2 attributes to the various types of player 1. If player 2 observes the sequence of actions

The most interesting cases are those in which player 1's payoff is not maximized by a stage-game Nash equilibrium, in which case attention turns to what we have called unary mixed equilibria. In these equilibria, play consists of an initial stage, whose relative length becomes insignificant as player 1 becomes patient, in which player 1 either builds or spends down her reputation, depending on the payoffs of the game and the prior distribution over mechanical types. This is followed by a reputation-manipulation stage in which play 1 essentially controls player 2's belief, keeping player 2 as close as possible to being indifferent between player 2's actions. Doing so requires player 1 to switch back and forth between her actions, but she can correlate her actions with those of player 2. As a result, there are two forces that allow player 1 to push her payoff above the conventional bound that can be obtained by committing to the behavior of player 1's favorite mechanical type. Player 1 can manipulate 2's beliefs so as to effectively commit to mechanical types that don't appear in the prior distribution, and player 1 can exploit the correlation induced during the manipulation phase.

The obvious direction for extending these results is to consider larger stage games. Our analysis of 2×2 games has relied heavily on the best-response structure of the stage game. There are not too many variations on this structure in 2×2 games, allowing a reasonably succinct comprehensive treatment. In particular, the key component of this structure, exploited throughout the analysis, is that player 2 becomes more anxious to play L the more 1 plays T, and more anxious to play R the more 1 plays B. This is what lies behind the manipulative strategy of player 1, with the manipulation then taking the form of ensuring that player 1 receives payoff $\max\{a, c\}$ when playing T and payoff $\max\{b, d\}$ when playing B.

The analysis in larger games will again be tied to the best response structure, but there are now many more possibilities for this structure, making general results quite tedious to state. Indeed, what it means to manipulate player 2 will depend on the best response structure. However, we can establish one obvious lower bound on payoffs, in that player 1's payoff will be at least as high (and possibly strictly higher) than the payoff bound established by Fudenberg and Levine [4, 5].

We can go further to sketch the types of results that are available in larger games. To build some intuition, first consider the following 2×2 game:

	L	R
T	$5, 0$	$2, 1$
B	$2, 1$	$4, 0$

.

The player-1 mixture that makes player 2 indifferent between L and R plays T and B each with probability $\frac{1}{2}$, and our result is that player 1 can ensure a payoff

played by the rational player 1 in the current game, he may (after sufficiently many periods) have cause to question the independence that he attributes to player 1's actions. Player 2 could not do so if 2 observes only the aggregate frequencies with which player 1 has played her actions, a formulation that suffices for our results and that strike us as realistic in many cases.

close to $\frac{1}{2}\cdot 5+\frac{1}{2}\cdot 4$, effectively keeping player 2 indifferent while coordinating play so that only the outcomes TL and BR appear. Let $T \rightarrow_{BR_2} R$ be interpreted as strategy T for player 2 causes R to be a best response for player 2. Then the structure that we exploit in constructing player 1's manipulative strategy is that

$$T \rightarrow_{BR_2} R \rightarrow_{BR_1} B \rightarrow_{BR_2} L \rightarrow_{BR_1} T.$$

Suppose in a large game that we could find a sequence of actions $\{T, M, B\}$ for player 1 and $\{L, C, R\}$ for player 2, with

$$T \rightarrow_{BR_2} R \rightarrow_{BR_1} M \rightarrow_{BR_2} C \rightarrow_{BR_1} B \rightarrow_{BR_2} L \rightarrow_{BR_1} T.$$

Let p^* be the mixture for player 1 that makes player 2 indifferent between L, C, and R, and suppose that L, C and R are best responses to this mixture.[16] Then player 1 can achieve a limiting payoff of

$$p^*(T)u_1(T, L) + p^*(M)u_1(M, R) + p^*(B)u_1(B, M).$$

This is the outcome of a manipulation phase, in which player 1 maintains player 2's indifference over the three actions $\{L, C, R\}$, while correlating play so as to play 1's best response against each action of player 2.

This result would generalize to longer cycles. We thus extend our main result for 2×2 games to larger games. However, establishing this result for 2×2 games required considering a number of cases, and this number grows as does the size of the game, making a complete enumeration of results significantly more tedious. We believe that a more productive approach would be to concentrate on particular applications.

5 Appendix: Proofs

5.1 Proof of Lemma 1

We must establish the stochastic-dominance claim. Renumber the types of player 1 so that type $k \in \{0, \ldots, K\}$ plays T with probability $\hat{\alpha}_k$, with $\hat{\alpha}_{k+1} > \hat{\alpha}_k$. Hence, we order the types by the probability that they play T, with the rational type fit into the appropriate place in the list. Then the stochastic-dominance claim requires, for any $K' < K$ and interior beliefs $\mu_h^0, \ldots, \mu_h^K$,

$$\sum_{k=0}^{K'} \mu^k_{(h,T)} < \sum_{k=0}^{K'} \mu^k_h,$$

where (h, T) denotes the history consisting of h followed by an observation of T. Using Bayes' rule, this is

$$\sum_{k=0}^{K'} \mu^k_{(h,T)} = \sum_{k=0}^{K'} \frac{\hat{\alpha}^k \mu^k_h}{\sum_{j=0}^{K} \hat{\alpha}^j \mu^j_h} < \sum_{k=0}^{K'} \frac{\mu^k_h}{\sum_{j=0}^{K} \mu^j_h} = \sum_{k=0}^{K'} \mu^k_h.$$

[16] We are here ruling out the existence of yet a fourth strategy that is superior to L, C, and R, when 1 mixes according to p^*.

Dropping the notation for the history h, this is

$$\sum_{k=0}^{K'} \hat{\alpha}^k \mu^k \sum_{j=0}^{K} \mu^j < \sum_{k=0}^{K'} \mu^k \sum_{j=0}^{K} \hat{\alpha}^j \mu^j$$

or, deleting terms common to both sides

$$\sum_{k=0}^{K'} \hat{\alpha}^k \mu^k \sum_{j=K'+1}^{K} \mu^j < \sum_{k=0}^{K'} \mu^k \sum_{j=K'+1}^{K} \hat{\alpha}^j \mu^j,$$

which follows from the fact that the $\hat{\alpha}^k$ are increasing in k.

5.2 Proof of Lemma 2

The same argument applies for histories h such that $n_{hB} < N_B(n_{hT})$. The relevant comparison is

Equilibrium path	Payoff	Deviation path	Payoff
BR	d	TR	c
BR	d	BR	d
BR	d	BR	d
$\vdots$	$\vdots$	$\vdots$	$\vdots$
BR	d	BR	d
TL	a	BR	d

The result then follows from the facts that $a > c$ and $d > c$.

We now note that σ_1 is the unique best response. The one-shot deviation principle ensures there is no history h and alternative strategy σ_1' that gives player 1 a payoff *superior* to that of σ_1 after history h. Could there by an alternative σ_1' that gives player 1 a payoff identical to that of strategy σ_1 after some history h', but prescribes a different action? Notice that σ_1' and σ_1 must yield identical payoffs after every history, since (i) σ_1' can never yield a higher payoff, (i) σ_1' can never yield a lower payoff at a history that is reached (since then it would not be a best response), and (iii) σ_1' can without loss of generality be taken to yield at least as high a payoff at histories that are unreached. Then similarly without sacrificing generality, we can take σ_1 and σ_1' to prescribe identical behavior after every history other than h'. But now our previous argument, which generates strict inequalities, assures that the action prescribed by σ_1 at history h' is uniquely optimal.

5.3 Proof of Lemma 3

Suppose first that player 1 faces a history at which $n_{hB} > N_B(n_{hT})$ and hence T is prescribed. Then analogously to the proof of Lemma 2, we can compare the following two paths:

"Equilibrium path$''$	Payoff	Deviation path	Payoff
TL	a	BL	b
TL	a	TL	a
TL	a	TL	a
$\vdots$	$\vdots$	$\vdots$	$\vdots$
TL	a	TL	a
BR	d	TL	a

This looks precisely like the comparison we made in proving Lemma 2. The difference is that here we do not know that the final play of B on the alleged equilibrium path is indeed a prescription of the equilibrium (and hence the quotation marks). The previous play of TL may have produced a point between $N_B(n_{hT})$ and $\underline{N}_B(n_{hT})$ (in which case B is prescribed) but may also have produced a point below $\underline{N}_B(n_{hT})$ (in which case T is prescribed). Nonetheless, the play listed under "equilibrium path" is feasible and, since it gives a higher payoff than does the deviation (by the same argument as offered in proving Lemma 2), the deviation is suboptimal.

Now suppose $n_{hB} < N_B(n_{hT})$. First, we fix n_{hT}, and argue that if player 1 chooses B at $(n_{hT}, n_{hB} - 1)$, then player 1 must also choose B at (n_{hT}, n_{hB}). Suppose this is not the case. Then player 1's strategy specifies T at (n_{hT}, n_{nB}), and we can consider the following equilibrium path and proposed deviation, beginning at history $(n_{hT}, n_{hB} - 1)$,:

Equilibrium path	Payoff	Deviation path	Payoff
BR	d	TR	c
TR	c	BR	d

These two paths both terminate at $(n_{hT} + 1, n_{hB})$, and hence thereafter can be taken to generate identical continuation payoffs. Because $c > d$, the proposed deviation yields higher payoff. This establishes that if player 1 chooses B at $(n_{hT}, n_{hB} - 1)$, then player 1 must also choose B at (n_{hT}, n_{hB}), and hence establishes the existence of a function $\underline{N}_B(n_T)$ satisfying the second two bullet points of the lemma.

Suppose $qa + (1 - q)d > c$, but (9) fails. Then there exists a history h and corresponding (n_{hT}, n_{hB}), as well as a sequence $\{\delta_k\}_{k=0}^{\infty}$ with $\delta_k \to 1$ such that the equilibrium outcome following history h is the play of TR in each subsequent period, for a continuation payoff of c, for each k. In addition, there exists a smallest integer n such that history h followed by n plays of B gives a history h' with $(n_{h'T}, n_{h'B}) = (n_{hT}, n_{hB} + n)$ and $n_{h'B} > N_B(n_{h'T})$. Furthermore, as δ

gets large, the continuation revenue given history h' approaches $(qa+(1-q)d)$.[17] But since $qa+(1-q)d>c$, we have, for sufficiently large δ,

$$(1-\delta^n)d+\delta^n(qa+(1-q)d)>c,$$

giving a contradiction.

It is once gain straightforward that this is the unique best response for player 1.

5.4 Proof of Lemma 4

The arguments are similar to those of Lemma 3, and we consider here only cases where differences arise. Fix a history h and suppose $n_{hB} > N_B(n_{hT})$ but $n_{hB} < N_B(n_{hT}+1)$. Then another play of T would lead to a history $(n_{hT}+1, n_{hB})$ with $n_{hB} < N_B(n_{hT}+1)$, prompting player 2 to play R in the next period. In cases 1 and 2, this is the equilibrium prescription for player 1. In cases 3 or 4, a play of R is relatively more costly for player 1, and so player 1 preempts the appearance of R by playing B at history h. In particular, given history h with (n_{hT}, n_{hB}), we have the following equilibrium path (initiated by a preemptive B at history h) and possible deviation (initiated by playing T at h):

Equilibrium path	Payoff	Deviation path	Payoff
BL	b	TL	a
TL	a	BR	d

These two paths both terminate at $(n_{hT}+1, n_{hB}+1)$, and hence thereafter can be taken to generate identical continuation payoffs. Because $b > d$, the equilibrium path is optimal for sufficiently patient players.

To establish the last claim in the lemma, we must confirm that, given a history (n_{hT}, n_{hB}) with $n_{hT} \in (N_B(n_{hT}, N_B(n_{hT}+1)$, that T appears in the continuation play with average proportion that approaches q as δ approaches 1. This is the case for the strategies given in Lemma 3, and we hence need only show that the current strategies induce the same limit. Now we note that, beginning with history h, these strategies induce the following outcomes:

Lemma 3 strategies	Current strategies
TL	BL
BR	BL
$\vdots$	$\vdots$
BR	BL
BR	TL

, (22)

[17] Once the history h' is reached, continuation play features instances of TL and BR. As $\delta \to 1$, the continuation payoff then is an average of payoffs a and d, with the probability attached to a in this average being the limiting proportion of time a is played. This limit is q.

at which point each strategy arrives at a history h' with $n_{h'T} \in (N_B(n_{h'T}, N_B(n_{h'T}+1)$. The argument can then be repeated, beginning with strategy h'. The result then follows from noting that T and L appear in the same proportions under both scenarios in (22).

5.5 Proof of Proposition 1

Let Σ_i be the set of functions of the form $\mathbb{N}^2 \to [0,1]$ (where $\mathbb{N}$ is the set of natural numbers, including 0). We interpret an element of $\mathbb{N}^2$ as a pair (n_{hT}, n_{hB}) characterizing a history h, and a strategy for player i gives the probability of playing T (player 1) or L (player 2) at each such history. Notice that we are assuming that players make use only of the information in a history identifying how often T and how often B has been chosen. This is restrictive, but suffices to establish equilibrium. Strictly order $\mathbb{N}^2$ such that shorter histories precede longer histories in the order, and then view an element of Σ_i as an element of $[0,1]^\infty$, where the zth component of $[0,1]^\infty$ is the mixture chosen by player i at the zth history (which will in general come well before period z). Endowed with the product topology, this space is compact.

Fix a specification of the game, including a discount factor and specification of the mechanical types, to be held constant throughout. Let $f_2 : [0,1] \to \Sigma_2$ be a correspondence, with $f_2(\alpha^0)$ identifying the set of player-2 best responses to the rational player-1 frequency α^0. Let $f_1 : \Sigma_2 \to \Sigma_1$ be a correspondence, with $f_1(\sigma_2)$ identifying the set of player-1 best responses to the player-2 strategy σ_2. Let $A : \Sigma_1 \to [0,1]$ be a function, calculating the value of A^0 (given the discount factor) for the player-1 strategy σ_1. Notice that the function A depends on σ_1 alone. In particular, we are restricting attention to a class of strategies in which $\sigma_1(h)$ and $\sigma_2(h)$ both condition only on (n_{hT}, n_{hB}), the number of times player 1 has played T and the number of times player 1 has played B. As a result, the players' behavior does not depend on what player 2 has chosen in the past, and does not depend on the order in which player 1 has played T and B. Fixing σ_1 alone does not fix the distribution over histories—for this, we also need to know σ_2. However, for any period t, fixing σ_1 fixes the distribution over values (n_{hT}, n_{hB}) for histories of length t, and hence fixing σ_1 alone suffices to determine A.

It suffices for the existence of an equilibrium to show that $Af_1f_2 : [0,1] \to [0,1]$ has a fixed point. This is obviously a correspondence from a compact set into itself, so we must show that it is upper-hemicontinuous and convex-valued.

[*Upper-hemicontinuity*] Let $\{\alpha^0(\ell)\}_{\ell=1}^\infty$ be a converging sequence with $\alpha^0(\ell) \to \overline{\alpha}^0$. Let $\{A^0(\ell)\}_{k=1}^\infty$ be a converging sequence of values with $A^0(\ell) \in Af_1f_2(\alpha^0(\ell))$. We need to show:

$$\lim_{\ell\to\infty} A^0(\ell) \in Af_1f_2(\overline{\alpha}^0).$$

Since $A^0(\ell) \in Af_1f_2(\alpha^0(\ell))$, there exist sequences $\{\sigma_1(\ell)\}_{\ell=1}^\infty$ and $\{\sigma_2(\ell)\}_{\ell=1}^\infty$

with (taking subsequences if necessary)

$$\begin{aligned}
\sigma_2(\ell) &\in f_2(\alpha^0(\ell)) \\
\sigma_1(\ell) &\in f_1(\sigma_2(\ell)) \\
A^0(\ell) &= A(\sigma_1(\ell)) \\
\lim_{\ell\to\infty} \sigma_2(\ell) &= \overline{\sigma}_2 \\
\lim_{\ell\to\infty} \sigma_1(\ell) &= \overline{\sigma}_1.
\end{aligned}$$

Because player 2's payoff function is continuous, we have $\overline{\sigma}_2 \in f_2(\overline{\alpha}^0)$. It then suffices to show

$$\lim_{\ell\to\infty} A^0(\ell) \in Af_1(\overline{\sigma}_2).$$

Because player 1's payoff function is continuous, we have $\overline{\sigma}_1 \in f_1(\overline{\sigma}_2)$. It then suffices to show

$$\lim_{\ell\to\infty} A^0(\ell) = A(\overline{\sigma}_1).$$

This follows from the continuity of A.

[Convex-valuedness] Fix α^0 and let A^0 and $A^{0'}$ be distinct members of $Af_1f_2(\alpha^0)$. We need to show that $Af_1f_2(\alpha^0)$ contains every convex combination of A^0 and $A^{0'}$. There exist player-2 strategies σ_2 and σ_2' with $A^0 \in Af_2(\sigma_2)$ and $A^{0'} \in Af_1(\sigma_2')$, and $\sigma_2, \sigma_2' \in f_2(\alpha^0)$. Because f_2 is a best-response correspondence, it is convex-valued, and hence $\{\lambda\sigma_2 + (1-\lambda)\sigma_2'\}_{\lambda\in(0,1)} \in f_2(\alpha^0)$. It then suffices to show that $Af_1(\{\lambda\sigma_2 + (1-\lambda)\sigma_2'\}_{\lambda\in(0,1_)})$ is convex. Let $g : [0,1] \to [0,1]$ be a correspondence with $g(\lambda) = Af_1(\lambda\sigma_2 + (1-\lambda)\sigma_2')$. Then $g(\lambda)$ is upper-hemicontinuous, because f_1 is upper-hemicontinuous and A is continuous. It follows from the fact that f_1 is a best response correspondence that $f_1(\lambda\sigma_2 + (1-\lambda)\sigma_2')$ is convex-valued, and the continuity of A then ensures that $Af_1(\lambda\sigma_2 + (1-\lambda\sigma_2') = g(\lambda)$ is convex valued. But an upper-hemocontinuous, convex-valued correspondence from $\Re$ into $\Re$ preserves convex sets, giving the result.

5.6 Proof of Proposition 2

[*Sufficiency*] Suppose $c > d$ and $c > q^* \max\{a,c\} + (1-q^*)\max\{b,d\}$. Consider a candidate equilibrium in which $\alpha^0 = 1$. Then playing T in each period gives player 1 a payoff of c. The highest alternative payoff is $q^* \max\{a,c\} + (1-q^*)\max\{b,d\}$, which gives the result.

[*Payoff Characterization*] This result follows from noting that if player 1 plays T in every period, then there will only a finite number of periods in which player 2 can play L, after which 2 plays R and player 1's payoff is c in every subsequent period. The length of the initial string depends on the specification of mechanical types, but is independent of the discount factor. Hence, as δ gets large, this initial string becomes insignificant in player 1's payoff, which approaches c.

More precisely, suppose we have a value $\varepsilon > 0$ and a sequence of discount factors $\{\delta_k\}_{k=1}^{\infty}$ with $\lim_{k\to\infty} \delta_k = 1$ and a sequence of equilibria, each with its corresponding value α_k^0 describing player 1's behavior and with its corresponding function N_{Bk} describing player 2's behavior, in each which player 1 earns a payoff less than $c - \varepsilon$. Fix one such equilibrium, and consider the alternative strategy for player 1 of always playing T. If $N_{Bk}(0) > 0$, then the result is that TR appears in every period, giving a payoff of c and a contradiction to the claim that the purported equilibrium is an equilibrium. Suppose $N_{Bk}(0) = 0$. Then define $n^*(\alpha_k^0)$ to be the smallest integer with the property $N_{Bk}(n^*(\alpha_k^0)) > 0$. Notice that the function N_{Bk} depends on α_k^0 but not δ_k, so that we can indeed write n^* as a function only of α_k^0.

The payoff from playing T in every period is bounded below by

$$(1 - \delta_k^{n^*(\alpha_k^0)}) \min\{a, c\} + \delta_k^{n^*(\alpha_k^0)} c.$$

It thus suffices to show that there exists an integer M such that

$$\sup_{\alpha_0 \in [0,1]} n^*(\alpha^0) < M. \tag{23}$$

In particular, if this relationship holds, then no matter what the equilibrium value of α^0, player 1 can ensure (by always playing T) an outcome path featuring at most M periods of TL, followed by the perpetual play of TR. As $\delta_k \to 1$, the payoff from this path approaches c.

We thus need only to verify (23). For this, we need only recognize that $n^*(\alpha^0)$ is maximized when $\alpha^0 = 0$, and this maximum is finite. In particular, there is at least one mechanical type k with $\alpha^k > p^*$. As a result, there is a finite number of observations of T that will cause player 2 to attach sufficiently large posterior probability to mechanical types who play T with probability greater than p^* as to make R a best response.

5.7 Proof of Proposition 3

Proof. [*Necessity*] Suppose that $c > p^* \max\{a, c\} + (1 - p^*) \max\{b, d\}$ and $c > q^* \max\{a, c\} + (1 - q^*) \max\{b, d\}$. Then it must be that c exceeds both of b and d. We show that there exists no mixed-outcome equilibrium.

If we have $c > a$, then the c is the largest stage-game payoff available to player 1. In addition, by consistently playing T, player 1 can ensure that player 2 will play R in all but a bounded (independently of α^0) number of periods, delivering a payoff that converges to c as $\delta \to 0$. No mixed-outcome equilibrium can provide as high a payoff, giving the result.

Suppose instead that $c < a$. As before we must have $c > b, d$, and the function $q \max\{a, c\} + (1 - q) \max\{b, d\}$ is increasing in q.

Fix a sequence of discount factors $\{\delta_k\}_{k=1}^{\infty}$ with $\delta_k \to 1$ and a corresponding sequence of equilibria featuring values $\{\alpha_k^0\}_{k=0}^{\infty}$ with $\alpha_k^0 \to \overline{\alpha}^0 < 1$. Notice first that we must have $\lim_{k\to\infty} \alpha_k^0 > \max\{q^*, p^*\}$. This follows from noting that, given α_k^0, an upper bound on player 1's payoff is given by

$$\alpha_k^0 a + (1 - \alpha_k^0) \max\{b, d\}.$$

This payoff can exceed c only if $\alpha_k^0 > \max\{q^*, p^*\}$. Hence, if $\alpha_k^0 \leq \max\{q^*, p^*\}$, then player 1's current strategy gives an expected payoff falling short of c. But the perpetual play of T gives a payoff that converges (as $k \to \infty$)) to c (since, regardless of α^0, such a strategy must induce player 2 to play T in all but a finite (and bounded, as $k \to \infty$) number of periods), a contradiction.

This leave open the possibility that we my have $\max\{q^*, p^*\} < \lim_{k\to\infty} \alpha_k^0 < 1$. If this is to be the case, then for each k, there must then be a history h_k after which player 1 plays B for a finite number of periods, earning payoff d in each such period, until reaching history a h'_k with $n_{h'_k T} = n_{h_k T}$ and $n_{h'_k B} > N_B(n_{h'_k T})$. Indeed, the first time that player 1 plays B gives rise to such a history. Hence, we can take each h_k to be a history of the form $T \cdots T$. Let $\alpha_k^0(h_k)$ be the value of A^0, calculated in the continuation game beginning with history h_k. We must have (taking a subsequence if necessary to ensure the existence of the limit) $\lim_{k\to\infty} \alpha_k^0(h_k) \leq \overline{\alpha}^0$.

Upon reaching history h_k, the consistent play of T would generate a payoff of c. It suffices for a contradiction to show that the continuation payoff falls short of c for sufficiently large k. Once again, an upper bound on this continuation payoff is given by

$$\alpha_k^0(h_k)a + (1 - \alpha_k^0(h_k)) \max\{b, d\}.$$

This payoff falls short of c if $\alpha_k^0(h_k) \leq \max\{q^*, p^*\}$, since $q \max\{a, c\} + (1 - q) \max\{b, d\}$ increases in q and falls short of c for $q \leq \max\{q^*, p^*\}$. Hence, we avoid a contradiction only if there exists ε with $\alpha_k^0(h_k) > \max\{q^*, p^*\} + \varepsilon$. Suppose this is the case. Then beginning at h_k, player 2 will within a finite number of periods play R in every subsequent period. (This is true no matter what the value or α_k^0, given that we know $\alpha_k^0 > q^*$.) This ensures that player 1's continuation payoff at history h_k must fall short of c. This in turn is a contradiction since playing T in every period after every history in h_k gives a payoff approaching (as $\delta_k \to 1$) c.

[*Sufficiency*] Suppose $p^* \max\{a, c\} + (1 - p^*) \max\{b, d\} > c$. Then there is an interval of probabilities $[\underline{p}, \overline{p}]$ with the property that for any $p \in [\underline{p}, \overline{p}]$, we have $p \max\{a, c\} + (1 - p) \max\{b, d\} > c$. Now fix a value α^0 and consider a strategy in which player 1 first plays a sequence of B or T, as needed, to make player 2 nearly indifferent between L and R, and player 1 thereafter alternates between T and B, playing T q proportion of the time, so as to maintain 2's near indifference and to achieve payoff $q \max\{a, c\} + (1 - q) \max\{b, c\}$ for some q. What will the value of q be in this mixture? The answer depends on α^0, but we must have $q > \alpha^0$ when $\alpha^0 < p^*$, and must have $q < \alpha^0$ when $\alpha^0 > p^*$. (If, for example, $\alpha^0 < p^*$ and player 1 plays so that $q < \alpha^0$, then player 2 will eventually come to attach arbitrarily high probability to types less than p^*, prompting 2 to consistently play L. Similarly, if $\alpha^0 > p^*$ and player 1 plays so that $q > \alpha^0$, then player 2 will eventually come to attach arbitrarily high probability to types larger than p^*, prompting 2 to consistently play L.) Next, consider the correspondence Af_1f_2, defined in the proof of Proposition 1 (f_2 associated player-2 best responses to values of α^0, f_1 associates player 1 best responses with player-2 strategies, and A identifies the resulting empirical frequency of player-

1 actions.). Our preceding calculations ensure that for sufficiently large δ, we have $Af_1f_2(\underline{p}) > \underline{p}$. In particular, any strategy with $\alpha^0 \leq \underline{p}$ must induce player 2 to eventually always play L, which is suboptimal (because always playing T ensures a payoff arbitrarily close to c). Alternatively, for the case of $\alpha^0 = \overline{p}$, we have just established there is a strategy for player 1 with $q < \overline{p}$ (and hance a smaller average probability of T, when δ is large, that gives a payoff larger than c. Since any strategy that plays T more than $\overline{p}$ of the time must lead to a payoff of c, we must have $Af_1f_2(\overline{p}) < \overline{p}$, for sufficiently large δ. It then follows from a version of the intermediate value theory and the fact that Af_1f_2 is an upper-hemicontinuous, convex-valued correspondence that it has a fixed point on $[\underline{p}, \overline{p}]$, which corresponds to an equilibrium.

Suppose $p^* \max\{a, c\} + (1-p^*) \max\{b, d\} < c < q^* \max\{a, c\} + (1-q^*) \max\{b, d\}$. Then it must be that $d < c < a$. Notice that this in turn ensures that $q^* > p^*$. We can construct an equilibrium for this case as follows. In the first period, player 1 is indifferent between T and B, and mixes, placing probability ζ on T. If the first action is T, then player 1 plays T thereafter. If the first action is B, then player 1 plays B until making player 2 indifferent between L and R, after which point player 1 maintains this indifference. This gives a long-run average T play of $\underline{\alpha}^0 > p^*$. We have aggregate play for player 1 of

$$\alpha^0 = \zeta \overline{\alpha}^0 + (1 - \zeta)\underline{\alpha}^0.$$

It is then a straightforward calculation, following from the facts that $p^* \max\{a, c\} + (1 - p^*) \max\{b, d\} < c < q^* \max\{a, c\} + (1 - q^*) \max\{b, d\}$ and $q^* > p^*$ that we can choose $\underline{\alpha}^0$ and ζ so that

- Player 1 is indifferent over the actions T and B in 1's initial mixture,
- Probability $\underline{\alpha}^0$ makes player 2 indifferent between L and R, in the process causing player 2's posterior to concentrate probability on types $\underline{\alpha}$ and α^0,

completing the specification of the equilibrium.

5.8 Proof of Proposition 4

Proposition 4.1 is a restatement of parts of Propositions 2 and 3.

Consider Proposition 4.4. Propositions 2 and 3 ensure that there is no pure equilibrium, in this case, and that there exists a binary mixed equilibrium, with payoff c. The fact that $P^* < c$ ensures there is not unary mixed equilibrium, since the payoff of such an equilibrium must approach P^*, while c is always a feasible payoff.

Consider Proposition 4.3, so that $Q^* < c < P^*$. This configuration is consistent with $c > d$, and if and only if this is the case, there is a pure equilibrium with payoff c (by Proposition 2). Proposition 3 ensures the existence of a unary mixed equilibrium. We need then only argue that there is no binary mixed equilibrium. Notice first that a binary mixed equilibrium can exist only if $c > d$. In particular, a binary equilibrium gives payoff c. If $c < d$, then for any configuration of player 1 strategies (including a proposed strategy for the rational player

1), mimicking type $\overline{\overline{\alpha}}$ gives play 1 a payoff that is a mixture of c and d, which is larger than c, and hence vitiates the proposed binary mixed equilibrium. Hence, suppose $c < d$. Then it must be that $a > c$ (in order for $P^* > c$ to be possible), and $Q^* < c < P^*$ is possibly only if $q^* < p^*$. But then the construction of a binary equilibrium must feature $\alpha^0 < \underline{\alpha}^0 > q^*$, a contradiction.

Consider Proposition 4.2. The fact that $Q^* > c$ ensures there is no pure equilibrium, and Proposition 3 has constructed a unary mixed equilibrium. We can mimic the construction of Proposition 3 to obtain a binary mixed equilibrium if $c > d$. Finally, we characterize the unary mixed equilibrium. Let $\{\delta_r\}_{r=1}^{\infty}$ is a sequence of discount factors with $\delta_r \to 1$. Suppose $p^* \max\{a, c\} + (1 - p^*) \max\{b, d\} > c$. Let there be a sequence of equilibria with corresponding values $\{\alpha_r^0\}_{r=0}^{\infty}$, with limit $\overline{\alpha}^0 \neq p^*$. Let $\alpha_r^0(h)$ be the continuation value of α^0, after history h, under the rth equilibrium.

We have supposed the sequence is unary, meaning that as r gets large, $P\{h : |\alpha_r^0(h) - \overline{\alpha}_r^0| > \varepsilon\}$ converges to zero for all $\varepsilon > 0$. Then for sufficiently large r, player 2 learns player 1's type and, since α_r^0 is bounded away from p^*, player 2 eventually either always plays L or always plays R. Player 1 thus receives either a mixture of payoffs c and d or a mixture of a and b. Suppose the first is the case, (the second is similar). If $c > d$, then we have a contradiction, since player 1 would be better off always playing T for a payoff arbitrarily close to c, ensuring that the candidate equilibrium is in fact not an equilibrium. If $d > c$, then player 1 would be better off playing a mixture arbitrarily close to p^*. ■

References

[1] Martin W. Cripps, George J. Mailath, and Larry Samuelson. Imperfect monitoring and impermanent reputations. *Econometrica*, 72(2):407–432, 2004.

[2] Martin W. Cripps, George J. Mailath, and Larry Samuelson. Disappearing private reputations in long-run relationships. *Journal of Economic Theory*, 134(1):287–316, 2007.

[3] David Ettinger and Philippe Jehiel. A theory of deception. *American Economic Journal: Microeconomics*, 2(1):1–20, 2010.

[4] Drew Fudenberg and David K. Levine. Reputation and equilibrium selection in games with a patient player. *Econometrica*, 57(4):759–778, 1989.

[5] Drew Fudenberg and David K. Levine. Maintaining a reputation when strategies are imperfectly observed. *Review of Economic Studies*, 59(3):561–579, 1992.

[6] Philippe Jehiel. Analogy-based expectation equilibrium. *Journal of Economic Theory*, 123:81–104, 2005.

[7] David M. Kreps, Paul R. Milgrom, John Roberts, and Robert J. Wilson. Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, 27(2):245–252, 1982.

[8] David M. Kreps and Robert J. Wilson. Reputation and imperfect information. *Journal of Economic Theory*, 27(2):253–279, 1982.

[9] Qingmin Liu and Andrzej Skrzypacz. Limited records and reputation. Technical report, NYU annd Stanford Univerity, 2011.

[10] George J. Mailath and Larry Samuelson. Who wants a good reputation? *Review of Economic Studies*, 68(1):425–442, 2001.

[11] George J. Mailath and Larry Samuelson. *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, Oxford, 2006.

[12] Paul R. Milgrom and John Roberts. Predation, reputation and entry deterrence. *Journal of Economic Theory*, 27(2):280–312, 1982.

[13] Joel Watson. A 'reputation' refinement without equilibrium. *Econometrica*, 61:199–206, 1993.