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

March 14, 2012

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 classical model 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 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 characterization of equilibrium behavior, showing that play begins with either a reputation-building or a reputation-spending stage (depending on parameters), followed by a reputation-manipulation stage.

Reputation with Analogical Reasoning

February 2, 2012

Contents

Reputation with Analogical Reasoning

1 Introduction

1.1 Reputations

The literature on reputation, pioneered by Kreps, Milgrom, Roberts and Wilson [11, 12, 17], 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 [7, 8]. 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 arbitrarily close 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. In effect, a small probability of being a mechanical type is as good as being known to be that type.

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.

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, which have been the focus of the reputation literature, but also characterize equilibrium behavior.

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. 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

---

[1]See Mailath and Samuelson [16] 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.

reason as if the behaviors of the various types of long-run players are stationary, and match the empirical frequencies of play. We assume further that a steady state has been reached, so that the induced *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 [10]), which has been defined for games with multiple types in Ettinger and Jehiel [6].[2] Notice 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 classical sequential equilibrium concept.

1.3   Preview

Focusing on 2×2 games, we develop conditions under which equilibrium behavior can 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 will take a particular action) or spends her reputation (exploiting the belief that with sufficiently high probability she will choose that action). 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 reputation manipulation phase. Here, the long-run player's behavior balances her interest in making the highest instantaneous payoff and her interest in maintaining the belief that she is likely to choose a certain action. 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 doing so, the long-run player's behavior will typically match the behavior of none of the mechanical types. Despite this, the short-run players' belief that the long-run player is mechanical need not converge to zero, in contrast to the insight obtained by Cripps, Mailath and Samuelson [3, 4] in the classical setup.

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 [7, 8].[3] The intuition here is that the long-run player can always simply mimic one of the mechanical types, leading the short-run player to eventually play a best response to such behavior. 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

---

[2]Kurz [13] introduces a similar model of beliefs.

[3]This follows from the work of Watson [18], and can be established with the same sort of argument found in Fudenberg and Levine [7, 8].

the behavior of a phantom mechanical type who does not actually appear in the list of possible mechanical types. Second, in the course of making player 2 indifferent between his actions, player 1 can introduce correlation into the actions of player 1 and 2. This correlation reduces the cost of manipulating 2's beliefs and allows an additional boost in the long-run player's payoff. Hence, unlike standard reputation models, neither the long-run player's behavior nor payoff need match that of the "Stackelberg type."

1.4   An Example: The Auditing Game

We can illustrate these results by introducing the first of two examples that we carry throughout the paper. Consider the following game:

$$
\begin{array}{c c}
 & \begin{array}{c c} Cheat\ (L) & Honest\ (R) \end{array} \\
\begin{array}{c} Audit\ (T) \\ Not\ (B) \end{array} & \begin{array}{|c|c|} \hline 4, -2 & 3, -1 \\ \hline 0, 4 & 4, 0 \\ \hline \end{array}
\end{array} \qquad . \qquad (1)
$$

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 $(16/5, -4/5)$.

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 the classical model, the presence of these mechanical types, or more generally the ability to commit, is of no value to player 1 in this game: no matter what the specification of mechanical types, the equilibrium payoff of a very patient player 1 is arbitrarily close to $16/5$.[4]

In our environment, the long-run player can achieve a higher payoff by exploiting the inferences of player 2. Overall, she will choose *Audit* roughly 4/5 of the time and *Not* roughly 1/5 of the time, as in the standard case (when the probabilities of mechanical types are small). However, player 1 will manage to *Audit* only when player 2 is cheating, and to not audit when player 2 is honest. A (very) patient player 1 thus secures a per-period payoff close to 4, which exceeds the stage-game Nash equilibrium payoff, as well as any conceivable equilibrium payoff in the classical approach. Player 1's reputation-manipulation stage allows

---

[4]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.

player 1 not only to keep player 2 on the boundary between being honest and cheating, but to avoid miscoordination in doing so.[5]

2 The Model

2.1 The Reputation Game

We consider a repeated game of incomplete information, as in Fudenberg and Levine [7, 8]. A long-run player 1 (she) faces a sequence of short-run player 2s (he). 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.[6]

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 realized stage-game actions of players 1 and 2 in period $t \in \{0, 1, 2, \ldots\}$. Mechanical type k of player 1 plays the same completely-mixed stage-game action $\alpha^k \in \Delta(A_1)$ in every period. 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 then simultaneously 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 $(1 - \delta) \sum_{t=0}^\infty \delta^t u_1(a(t))$.

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 [10] and Ettinger and Jehiel [6] 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.

Let σ_1 and σ_2 denote the strategies of the rational player 1 and of the short-run players 2, respectively, and let σ denote the strategy profile (σ_1, σ_2). We denote by $\sigma_{1h}(a_1)$ the probability that the rational player 1 selects action a_1

---

[5]This ability to avoid miscoordination arises out of the interaction of mechanical types, which prompt player 2 to shift between actions as his beliefs as to the type of player 1 shift, and analogical reasoning, which allows player 2's assumption of stationary behavior on the part of player 1 to mask the resulting correlation in actions.

[6]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. We could also address the case in which the game lasts forever but players discount, with somewhat different details.

after history $h \in \bigcup_{t=0}^{\infty} A^t$, with $\sigma_{2h}(a_2)$ being analogous for player 2. 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 then 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.[7]

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 he 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 his 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

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

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,

---

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

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 includes the frequencies of the various actions played by player 1, but need not identify 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.[8] 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

We focus on games of 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 that player 2 has no dominant strategy, stipulating (without loss of generality) that

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

---

[8]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 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,

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

We simplify the notation by taking α^k to be the probability with which player 1 of type k plays T. We assume there is at least one mechanical type with $0 < \alpha^k < p^*$, and one with $p^* < \alpha^k < 1$. We could work without this assumption, but then some of the results would become cluttered with additional special cases.

The analysis of the sequential analogy-based expectation equilibria in this 2×2 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 player 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 the "state" (n_{hT}, n_{hB}), and not on the order in which the various actions have appeared.

Whenever player 2 observes T, her beliefs about player 1's type shift (in the sense of first-order stochastic dominance) toward types that are more likely to play T, with the reverse holding for an observation of B. Hence, for any given number n_{hT} of T actions, the probability attached by player 2 to player 1 playing B is larger, the larger the number n_{BT} of B actions. Player 2 will then play L if and only if she has observed enough B actions. More precisely, we have:

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 $n_{hB} = N_B(n_{hT})$.*

This result holds regardless of what assumptions we make about the distribution of mechanical types or the strategy of the rational type, as long as we have at least one type playing T with probability greater than p^* and one playing T with probability less than p^*, though different specifications of these strategies will give rise to different functions N_B.

3.2.2 Player 1's Best Response: The Auditing Game

Player 1's best-response behavior is slightly more nuanced, and depends upon the specification of the game. We first consider the auditing game given by (1).

Lemma 2 *Consider the auditing game, and 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 illustrates player 1's and player 2's best responses, as well as an equilibrium path of play. Player 1 plays T whenever 2 is playing L, and 1 plays B when 2 plays R. 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.

Proof. The argument is a straightforward application of 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 she plays T (and player 2 plays L) until she 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 player 1 plays B instead of T at h and then follows σ_1, she must play T after her 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 paths of play (with subsequent play being identical, and hence being irrelevant for this comparison):

Equilibrium path	Payoff	Deviation path	Payoff
TL	4	BL	0
TL	4	TL	4
$\vdots$	$\vdots$	$\vdots$	$\vdots$
TL	4	TL	4
BR	4	TL	4

We thus see that the only difference in play between the two scenarios is in the first and last period within this sequence of periods (the deviation gives
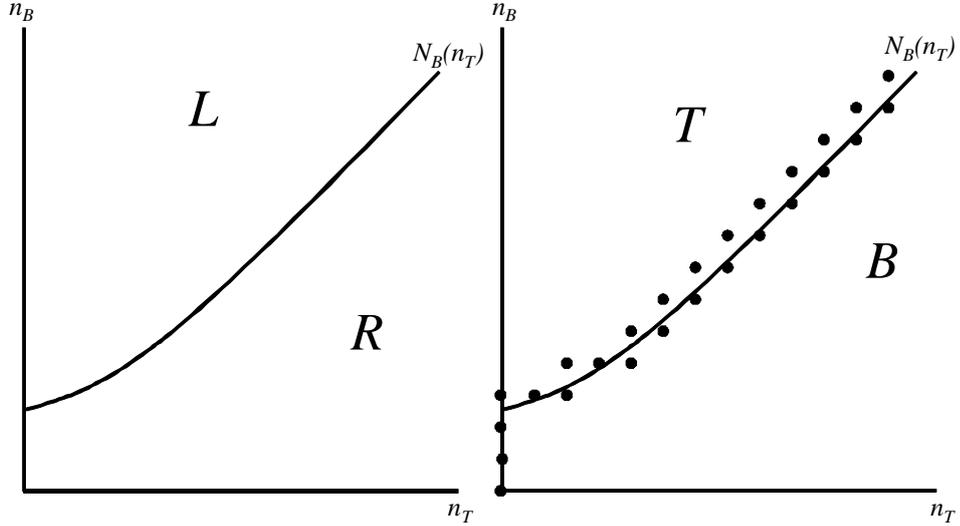
8

Figure 1: Strategies for the short-run player (left panel) and long-run player (right panel) in the auditing game (cf. (1)). The function N_B is identified in Lemma 1. Lemma 2 shows that in the auditing game, this function also characterizes player 1's best responses. The function N_B is increasing, but we cannot in general restrict its intercept or curvature. An 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).

$u_1(B, L) = 0$ and $u_1(T, L) = 4$ in the first and last period, whereas in the absence of the deviation, player 1 receives $u_1(T, L) = 4$ and $u_1(B, R) = 4$ in the first and last period). It is then immediate that the deviation lowers player 1's payoff.

An analogous argument argument applies for histories h such that $n_{hB} < N_B(n_{hT})$. ■

This result does not depend on the fortuitous payoff tie $a = d$ in the auditing game, and holds for any game in which $a, d > b, c$.

3.2.3 Player 1's Best Response: The Product Choice Game

Consider the product-choice game of Mailath and Samuelson [15], transcribed here as:

$$
\begin{array}{c c}
 & \begin{array}{c c} L & R \end{array} \\
\begin{array}{c} T \\ B \end{array} &
\begin{array}{|c|c|}
\hline
3,0 & 1,1 \\
\hline
2,3 & 0,2 \\
\hline
\end{array}
\end{array}
\qquad . \tag{6}
$$

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. Low quality is a dominant strategy for the firm, presumably because it is cheaper. The firm would like the consumer to buy the custom product, which is a best response for the consumer if the firm chooses high (but not low) quality.

Lemma 3 *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 $\underline{N}_B(n_T) \le N_B(n_T)$ such that, for any history h and sufficiently large δ,*

- *(3.1) Player 1 plays T if $n_{hB} > N_B(n_{hT} + 1)$;*

- *(3.2) Player 1 plays B if $\underline{N}_B(n_{hT}, \delta) < n_{hB} < N_B(n_{hT} + 1)$;*

- *(3.3) Player 1 plays T if $n_{hB} < \underline{N}_B(n_{hT}, \delta)$;*

- *(3.4) $\lim_{\delta \to 1} \underline{N}_B(n_T) < 0$ for all n_T.*

Section 5.1 contains the proof and Figure 2 illustrates these strategies. There are two new developments here. First, 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. (This is the content of (3.1)–(3.2).) Second, player 1 would like player 2 to play L, and can build a reputation that ensures this by playing B. However, it is now costly to play B when player 2 plays R. If the number of B plays required to build a reputation is too large, player 1 may surrender all thoughts of a reputation and settle for the continual play of TR (cf. (3.3)). However, a patient player 1 inevitably builds a reputation (cf. (3.4)).

3.3 Equilibrium: Examples

We now combine these best-response results to study equilibrium behavior.

3.3.1 Example I: The Product Choice Game

We begin with the product-choice game (cf. (6)). Player 2 is indifferent between L and R when $p = p^* = 1/2$. Action B is the pure "Stackelberg" action for
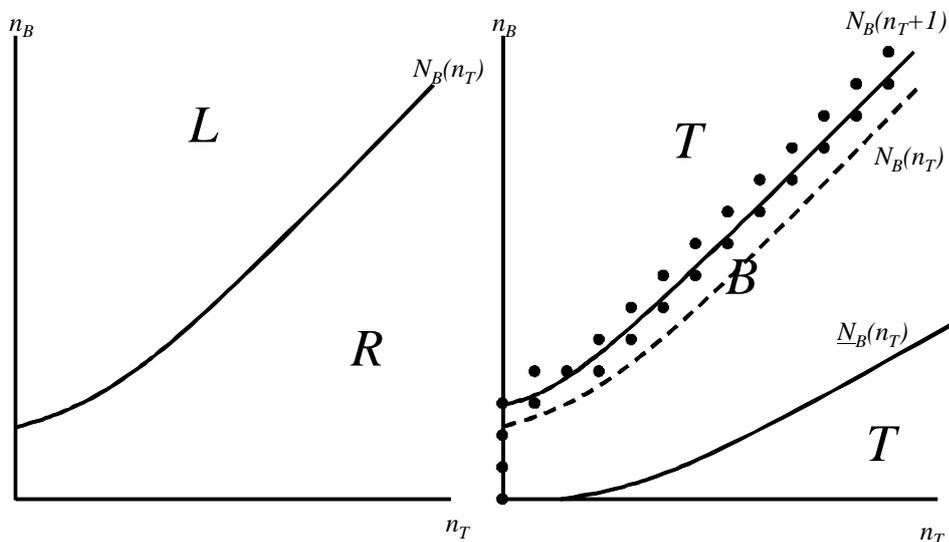
Figure 2: Strategies for the short-run player (left panel) and long-run player (right panel) in the product choice game. The function N_B is taken from Lemma 1. The functions N_B and $\underline{N}_B$ are both increasing, with the former lying above the latter, but with shapes and intercepts that depend on the specific game. 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, as will occur whenever player 1 is sufficiently patient, the function $\underline{N}_B$ from Lemma 3 is never reached in equilibrium and plays no role in shaping equilibrium behavior. 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.

player 1 in this game, i.e., the pure action to which player 1 would most like to be committed to, conditional on player 2 playing a best response.

Lemmas 1 and 3 tell us much about the equilibrium outcome in this case, but leave three possibilities, that correspond to three possibilities for the intercepts of the functions $\underline{N}_B$ and N_B in Figure 2.

First, it may be that $\underline{N}_B(0) > 0$. In this case, the equilibrium outcome is that rational player 2 chooses T and player 2 chooses R in every period. Player 1 thus abandons any hope of building a reputation, settling instead for the perpetual play of the stage-game Nash equilibrium TR. This is potentially optimal because building a reputation is costly, requiring player 1 to play B and hence settle for a current payoff of 0 rather than 1. If player 1 is sufficiently impatient, this current cost will outweigh any future benefits of reputation building, and player 1 will indeed forego reputation building. By the same token, this will not be an equilibrium if δ is sufficiently large (but still less then one), and we accordingly

11

turn to the other two possibilities.

Second, it may be that $\underline{N}_B < 0 < N_B$, as in Figure 2. In this case, play begins with a reputation-building stage, in which player 1 chooses B and the outcome is BR. This continues until player 2 finds L a best response (intuitively, until the state has climbed above the function N_B). Thereafter, we have a reputation-manipulation stage in which player 1 sometimes chooses T and sometimes B, selecting the latter just often enough to keep player 2 always playing L.

Alternatively, if $N_B(0) < 0$, then play begins with a reputation-spending phase in which player 1 chooses T, with outcome TL, in the process shifting player 2's beliefs towards types that play T. This continues until player 2 is just on the verge of no longer finding L a best response (intuitively, until the state just threatens to cross the function N_B). Thereafter, we again have a reputation-manipulation stage in which player 1 sometimes chooses T and sometimes B, again selecting the latter just often enough to keep player 2 always playing L.

Which of these will be the case? For sufficiently patient players, whether one starts with a reputation-building or reputation-spending phase depends on the distribution of mechanical types. If player 2's best response conditional on player 1 being mechanical is R, then the rational player 1 must start with a reputation building phase, and we have the first of the preceding possibilities. Alternatively, If player 2's best response conditional on player 1 being mechanical is T, then the rational player 1 must start with a reputation spending phase, and we have the second of the preceding possibilities.

To see why this is the case, recall that α^0 is the frequency with which the rational player 1 chooses T, and let $\hat{\alpha}^0$ be the frequency with which she chooses T during the reputation-manipulation stage. We first note a consistency condition. We cannot have either $p^* < \alpha^0 < \hat{\alpha}^0$ or $\hat{\alpha}^0 < \alpha^0 < p^*$. In the first case, for example, the continued play of $\hat{\alpha}^0$ during the reputation-manipulation stage would cause player 2's posterior probability to concentrate on types at least as large as α^0 ($> p^*$), ensuring that player 2 is no longer nearly indifferent between L and R, a contradiction.[9]

Now let us suppose that R is player 2's best response against the mechanical types, and argue that play must begin with reputation building. Suppose first that $\alpha^0 > p^*$. This reinforces the optimality of R for player 2, ensuring that player 2 will initially find R optimal and that the game will open with a sequence of BR plays. To balance this sequence, we must have $\hat{\alpha}^0 > \alpha^0$, since player 2's initial sequence of B plays must combine with the behavior summarized by $\hat{\alpha}^0$ to average to α^0. But then we have $p^* < \alpha^0 < \hat{\alpha}^0$, which we have just seen is impossible. Hence, it must instead be that $\alpha^0 < p^*$ and therefore $\alpha^0 < \hat{\alpha}^0$, which ensures that player 1 must indeed begin the game with a reputation-building string of B (so that again the initial string of B, averaged with $\hat{\alpha}^0$, gives α^0), as claimed. Reversing this argument ensures that if player 2 finds L

---

[9]This argument uses the fact that the initial reputation-building or reputation-spending phase remains bounded as δ gets large, so that this initial phase does not overwhelm inferences based on $\hat{\alpha}^0$, which we establish.

a best response conditional on the mechanical types, then the game begins with a series of reputation-spending TL plays.

Now let us examine the equilibrium outcome when play begins with a reputation-building stage. Player 1 will initially continually play B, with player 2 playing R, in the process building up player 2's posterior belief that player 1 will play B, until pushing 2 to play L. Then the reputation-manipulation stage begins. The reputation-manipulation stage will feature a frequency of T given by $\hat{\alpha}^0 > \alpha^0$, to balance the reputation-building string of B plays. In addition, we will have $\alpha^0 < p^*$, as explained in the preceding paragraph. In the long run, player 2's belief will then become concentrated on types α^0 and $\overline{\alpha} = \min(\alpha^k, \alpha^k > p^*)$. Player 1 balances 2's beliefs on these two types so as to keep 2 just willing to play L.

When δ approaches 1, the initial reputation building stage must become short relative to the reputation-manipulation stage (because there is an upper bound on the number of B required for player 2 to be close to indifferent between actions L and R, that applies uniformly over all specifications of α^0). As a result, $\hat{\alpha}^0$ and α^0 must get close together, since otherwise α^0 could not be the average of play characterized almost entirely by $\hat{\alpha}^0$. In addition, α^0 must approach p^*, since otherwise the prolonged play of $\hat{\alpha}^0$ would not render player 2 continually nearly different.

How does this compare to the outcome when the short-run players are perfectly rational, and hence the analysis of Fudenberg and Levine [7, 8] applies? First, with perfectly rational player 2s, the result is that a patient player 1 can attain the payoff very close to that she would receive if she were known to be type $\underline{\alpha} = \max(\alpha^k, \alpha^k < p^*)$. More generally, player 1 can receive a payoff close to that she would receive if known to be her favorite mechanical type. In contrast, our player 1 receives a payoff close to that she would receive if known to be a type very close to (but less than) p^*. In effect, our player 1 can commit to a phantom mechanical type. As a result, player 1's payoff is relatively insensitive to the precise specification of mechanical types.

Second, 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 [3, 4] that the types of long-run players in reputation games eventually must become known.[10] 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, and so 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

---

[10]Cripps, Mailath and Samuelson [3, 4] prove this result for the case of imperfect monitoring, but it applies equally to the case of perfect monitoring with mechanical types playing mixed strategies, as considered here.

δ gets close to 1. Indeed, in the limit as $\delta \to 1$, the initial reputation-spending stage will become arbitrarily short compared to the reputation-manipulation stage, and as a result, $\hat{\alpha}^0$ and α^0 will both converge to p^*. Away from this limit, however, 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: The Auditing Game

Now consider the auditing game of 1. Lemmas 1 and 2 describe the behavior of the best responses. Recall that commitment is of no value here, and hence that conventional reputation arguments have no force.

There are now two possibilities. Play might begin with a reputation-building phase consisting of a string of BR outcomes (in this case N_B has a positive intercept, as in Figure 1), until player 2 is just on the verge of choosing L. Alternatively, play might begin with a reputation-spending phase consisting of a string of TL outcomes (in this case N_B has a negative intercept), until player 2 is just on the verge of choosing R. In either case, play then enters a reputation-manipulation phase in which the state hovers as close as possible to the graph of the function N_B. Play in this reputation-manipulation stage includes TL and BR outcomes.

What determines whether player 1 initially builds or spends her reputation? We can apply reasoning mimicking that of the product-choice game. If player 2 finds R a best response conditional on facing a mechanical type, then equilibrium play in the original game must begin with a sequence of BR plays. If player 2 would find L a best response conditional on the mechanical types, 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.

We find here another contrast with the case of a perfectly-rational player 2. Player 1 cannot earn more that the stage-game equilibrium payoff of $16/5$ against rational player 2s. In our case, a patient player 1 earns a payoff in the auditing game arbitrarily close to 4. The key here is that in equilibrium, player 1 plays a mixture of T and B, while player 2 plays a mixture of L and R. However, player 1 is effectively able to correlate these mixtures, ensuring that T appears against L and B against R, thereby boosting her payoff above that which can be achieved with perfectly rational player 2s. A similar result holds in any zero-sum game.

As in the product choice game, we find that in the limit as δ goes to 1, the initial phase must be relatively short relative to the reputation-manipulation stage. As a result, $\hat{\alpha}^0$ and α^0 approach one another and approach p^*. Hence, TL outcomes appear with probability p^* and BR outcomes with probability $1 - p^*$, ensuring that player 1's payoff approaches $p^* u_1(T, L) + (1 - p^*) u_1(B, R) = 4$ as δ gets close to 1.

3.4 Equilibrium: Analysis

We now characterize equilibria for general payoffs. 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 $(\mu_\emptyset^1, \ldots, \mu_\emptyset^K)$. 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 sets (since there may be multiple best responses) of values of A^0. A fixed point of this map gives us an equilibrium, and the existence of such a fixed point follows from a relatively straightforward application of Kakutani's fixed point theorem. ∎

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*

$$c > d \tag{7}$$
$$c > q^* \max\{a, c\} + (1 - q^*) \max\{b, d\}. \tag{8}$$

If (7)–(8) 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 $\delta \to 1$, 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, i.e., in which player 1 mimics type $\overline{\alpha}$. 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, a contradiction to the optimality of α^0. Hence, for sufficiently large δ, there are no pure-outcome equilibria. The logic here is analogous to that lying behind the reputation bounds of Fudenberg and Levine [7].

Alternatively, suppose $c > d$ but (8) 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 strategy $\alpha^0 = 1$, 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 (as in Section 3.2.2). If $b > d$, player 1 ensures that the history lies always just above this boundary, giving a mixture between payoffs a and b (as in Section 3.2.3). The limiting probability attached to a in either of these mixtures must be q^*, which (because (8) fails) exceeds the payoff from α^9, a contradiction.

The sufficiency result is similar, and is relegated along with the payoff characterization to Section 5.2. ∎

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 also sufficient for the existence of a mixed-outcome equilibrium.

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:*

$$c \; < \; p^* \max\{a,c\} + (1-p^*)\max\{b,d\} \tag{9}$$

$$c \; < \; q^* \max\{a,c\} + (1-q^*)\max\{b,d\}. \tag{10}$$

Section 5.3 provides a proof, while we sketch the reasoning here. The first step 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

$$p^* \max\{a,c\} + (1-p^*)\max\{b,d\} < c < q^* \max\{a,c\} + (1-q^*)\max\{b,d\}. \tag{11}$$

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 and 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 $\{\alpha_\ell^0\}_{\ell=0}^\infty$ is unary if $P\{h : |\alpha_\ell^0(h) - \alpha_\ell^0| > \varepsilon\}$ 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 and 3.3.2 are unary.

We have already concluded that when (11) holds, the (only) equilibrium is a binary, mixed equilibrium. Notice that (11) 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 denoted by $\underline{\alpha}^0$. 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,

17

completing the specification of the equilibrium.

Are there equilibria in which player 1 mixes over more than two continuation paths? The following, proven in Section 5.4, shows that the answer is no:

Lemma 4 *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 $\delta \to 1$.*

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

$$P^* := p^* \max\{a, c\} + (1 - p^*) \max\{b, d\} \tag{12}$$

$$Q^* := q^* \max\{a, c\} + (1 - q^*) \max\{b, d\} \tag{13}$$

We can collect the results of the preceding two propositions to give (with Section 5.5 filling some proof details):

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 these 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 these 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 $\delta \to 1$.

Why doesn't the possibility of obtaining payoff P^* preclude the existence of a pure equilibrium, which yields a payoff of $3 < P^*$ for player 1? Player 1 could deviate from the pure-equilibrium strategy of always choosing T by initially playing B, pushing the posterior that 2 attaches to 1 playing T to the point at which 2 is indifferent between L and R. By subsequently maintaining this indifference, player 1 can achieve a payoff very close (for large δ) to

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

where we have defined q^* to be the frequency required to maintain the posterior near p^*, namely

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

We can solve for 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 exists a binary mixed equilibrium only if the lower probability $\underline{\alpha}^0$ in this equilibrium is enough smaller than p^* as to push the expected payoff from this outcome down to c. This in turn requires that $\underline{\alpha}$ 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.

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

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

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 much of 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] 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.

[2.3] The action profile played in the initial phase of a unary mixed equilibrium is BR if player 2's best response to the mechanical types (only) is R, 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.

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

[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.

The bounded length of the initial reputation-building or reputation spending phase is an immediate implication of our maintained assumption that there is a mechanical type that plays T with probability less than p^* (the probability that makes player indifferent), as well as one that plays T with probability greater than p^*. Notice that whether this initial phase consists of reputation building or reputation spending depends on the distribution of types *conditional on be-*

ing mechanical, but not on the total prior probability attached to mechanical types.[11]

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.[12]

3.5 Comparisons

3.5.1 Rational Short-Run Players

It is natural to compare our reputation results to those of Fudenberg and Levine [7], obtained when short run players are rational. Their result is that

$$\lim_{\delta \to 1} U_1^*(\delta) \geq \max_{\alpha^k \in \{\alpha^1, \ldots, \alpha^K\}} \min_{a_2 \in BR(\alpha^k)} u_1(\alpha^k, a_2), \tag{14}$$

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.

A proof virtually identical to that used to establish Fudenberg and Levine's ensures that player 1 in our setting is assured a payoff at least as high. Alternatively, this result can be viewed as a corollary of Watson [18].

We can point to three differences, concerning payoffs, behavior, and beliefs.

---

[11]The argument given for this result in the context of the product-choice game in Section 3.3.1 holds in general. How can player 1 afford a reputation-spending stage if mechanical types are very unlikely? During the reputation-manipulation stage, player 2's posterior probability becomes concentrated on type α^0 as well as a mechanical type k, with the posteriors on these two types hovering around the level that makes 2 indifferent between L and R. If μ_0^k is very small, then the equilibrium value of α^0 will be very close to p^*, so that the posterior attached to the mechanical type will have to drop yet further in order to induce indifference.

[12]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.

First, 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, for potentially two reasons. First, the Fudenberg-and-Levine bound is tied to the payoff player 1 would receive if known to be her favorite mechanical type. In contrast, P^* is independent of the specifications of the mechanical types α^k. Unless there are mechanical types characterized by actions arbitrarily close to p^*, the bound here will be higher. In effect, the short-run players' analogical reasoning allows player 1 to commit to the play of advantageous but phantom types. 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^*). This is (strictly) smaller than P^* for a range of games (including zero-sum games) that exhibit the properties of our auditing game. The key to the long-run player's payoff in the latter is the ability to introduce correlation into the actions of player 1 and 2.

Second, standard reputation models say little about equilibrium behavior. There are examples of equilibrium play for finitely-repeated games, in which player 1's reputation invariably decreases (Kreps, Milgrom, Roberts and Wilson [11, 12, 17], Mailath and Samuelson [16, Chapter 17]). Player 1's payoff in our model is relatively insensitive to the specification of mechanical types, but player 1's *behavior* is not. Player 1 will undertake an initial reputation-building or reputation-spending stage, followed by a phase of reputation-manipulation. Whether player 1 initially builds or spends her reputation does not depend upon the total prior probability attached to mechanical types, but does hinge upon the distribution of this probability, with player 1 typically pushing player 2's beliefs away from his best response to the mechanical types.

Third, Cripps, Mailath and Samuelson [3, 4] establish conditions under which the short-run players in a standard reputation model must eventually learn the type of the long-run player. In contrast, our short-run players typically never learn this type. This failure to learn is the key to the long-run player's ability to "commit" to being a phantom type.

3.5.2 Limited Observability

Liu and Skrzypacz [14] examine a model in which a long-run player faces a succession of short-run players in a game in which each player has a continuum of actions, but which gives rise to incentives reminiscent of our product-choice game. To emphasize the similarities, we interpret player 1's action as a level of quality to produce, and player 2's action as a level of sophistication in the product he purchases. Player 1 may be rational, or may be a (single) mechanical type who always chooses some fixed quality c. The players in their game are

fully rational, but the short-run players can observe only actions taken by the long-run player, and can only observe such actions in the last K periods for some finite K.

Liu and Skrzypacz [14] show that their rational long-run player invariably chooses either quality c (mimicking the mechanical type) or quality 0 (the "most opportunistic" quality level). After any history in which the short run-player has received K observations of c and no observations of 0 (or has received as many c observations as periods in the game, for the first $K-1$ short-run players), the long-run player chooses quality 0, effectively burning her reputation. This pushes the players into a reputation-building stage, characterized by the property that the short-run players have observed at least one quality level 0 in the last K periods. During this phase the long-run player mixes between 0 and c, until achieving a string of K straight c observations. Her reputation has then been restored, only to be promptly burned. Liu and Skrzypacz [14] then establish that as long as the record length K exceeds a finite lower bound, then the limiting payoff as $\delta \to 1$ is given by the Fudenberg-and-Levine payoff. Moreover, they show that this bound holds after every history.

In terms of payoffs, the long-run player again potentially earns a higher payoff in the current setting than under the limited records of Liu and Skrzypacz [14], again both because her payoff is not tied to a particular mechanical types and because she may be able to induce correlation in actions. Our model gives an initial reputation-building or reputation-spending stage, followed by consistent reputation-manipulation, while the long-run player in Liu and Skrzypacz [14] continually alternates between building and then burning her reputation. In both models, the short-run players fail to become sure of the long-run players' type, in our case because of their misspecified model and in their case because of the limited records.

3.5.3 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.

Consider a model in which 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 to the empirical frequency of 1's play, to which 2 plays a best response.[13]

---

[13]Alternatively, there exists a specification of mechanical types for player 1 and a prior distribution over these types such that a Bayesian player 2, believing that player 1's type is drawn from this distribution, would duplicate 2's play in the fictitious play model.

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

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

which in this case is thus a ray through the origin.

Player 1's best response behavior is again characterized in 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 5 *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:*

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

$$\max\{b, c, p^* \max\{a, c\} + (1-p^*)\max\{b, d\}\}. \tag{15}$$

(5.2) The frequency with which player 1 plays T is given by 0 (if b is the maximizer in (15)), 1 (if c is the maximizer in (15)), or p^* (if $p^* \max\{a, c\} + (1-p^*)\max\{b, d\}$ is the maximizer in (15)).

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 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.[14]

4   Discussion

4.1   Summary

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 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

_____

[14]An approximate form of this result holds if exact fictitious play is replaced by stochastic fictitious play, with player 1's payoff under fictitious play converging to the payoff in our model as the error in the stochastic fictitious play gets small. The convergence will be relatively rapid in games like those of Section 3.3.1, where player 1 ultimately induces only one action from player 2, and will be slower in games like those of Section 3.3.2, where player induces correlation in the two players' actions.

view such beliefs as natural for cases in which player 2 can most readily collect information about average frequencies of play.

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 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.

4.2 Extensions

A first obvious direction for extending these results is to consider larger stage games. Return momentarily to either the auditing game or the product-choice game. Let $T \to_{BR_2} R$ be interpreted as "strategy T for player 1 causes R to be a best response for player 2." Then

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

Our analysis relied heavily on this best-response structure, with player 2 becoming 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 behavior of player 1.

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

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

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.[15] 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. We can thus extend our

---

[15]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^*.

ideas to larger games, but the best-response structures in such games can be considerably more complicated, as will be the details of the analysis.

A second interesting extension would be to consider cases in which some player 2s are analogical reasoners, while others are rational. In this case, the manipulative strategy that works well with analogical reasoners may be costly when facing rational player 2s, as the deterministic nature of the manipulative strategy makes it predictable for rational player 2s. Obviously, if the share of rational player 2s is small, our analysis will carry over (as the benefit vis a vis analogical reasoners would outweigh the cost vis a vis rational player 2s). Understanding how the reputation-manipulation strategy will depend more generally on the distribution of player 2's cognitive sophistications should be the subject of further research.

4.3 Implications

To see the importance of the various reputation models we have examined, consider a classical application of reputation idea, the Backus and Driffill's [2] analysis of inflation. They consider the following stage game:

$$
\begin{array}{c|c|c|}
 & \textit{Low} & \textit{High} \\
\hline
\textit{Low} & 0,0 & -2,-1 \\
\hline
\textit{High} & 1,-1 & -1,0 \\
\hline
\end{array}
\tag{16}
$$

Player 1 is the government, and can choose either high or low inflation. Player 2 represents the citizens in the economy, and can choose either high or low inflationary expectations. The citizens would like their expectations to be correct. The government would prefer its citizens to expect low inflation, a goal complicated by the fact that the government can then gain by surprising its citizens with an inflationary burst.

This game is strategically equivalent to our product choice game. Backus and Driffill [2] pursue a standard analysis, examining a finitely-repeated version of this game, along the lines of Kreps, Milgrom, Roberts and Wilson [11, 12, 17], with there being some probability that the government is a mechanical type who always chooses low inflation. In an infinitely-repeated such game (or a sufficiently long finitely-repeated game), the standard result is that a sufficiently patient government can secure a payoff close to zero, as it would if it where known to be the mechanical type. Alternatively, the bounded-records model of Liu and Skrzypacz [14] leads to cyclical behavior, with the government continually (stochastically) refraining from inflation just long enough to allow consumers to think the government might be the mechanical type, at which point the government disabuses them of this notion with an inflationary burst.

In our case, the government would earn a payoff close to 1/2, by combining an initial reputation-building or reputation-spending phase (depending on the specification of mechanical types) with a reputation-manipulation stage in which low inflation is chosen just often enough for low-inflation expectations to be optimal for citizens. Analogical reasoning is beneficial for the government in

this case, at the expense of citizens, who endure as much inflation as they are willing to take without changing their behavior.

Alternatively, our auditing game is a special case of the class of inspection games (Avenhaus, Von Stengel and Zamir [1]). Variations on this game have cast player 1 as, among many other applications, a law enforcement official, an environmental regulator, a teacher, a customs official, or a quarterback, each of whom must deter player 2 from committing crimes, polluting, neglecting homework, smuggling, or neglecting to defend against the long pass. It is a well-known feature of such games that the penalty imposed in case player 2 is caught in a transgression has no effect on the equilibrium probability of such a transgression. This is a direct implication of the logic of mixed strategies, but is sufficiently counterintuitive as to be known as the "Dixit-Skeath conundrum" (Dixit and Skeath [5], Grant, Kajii and Polak [9]).

Rewrite the auditing game as

	$Cheat\ (L)$	$Honest\ (R)$
$Audit\ (T)$	$4, -z$	$3, -1$
$Not\ (B)$	$0, 4$	$4, 0$

.

How does equilibrium play in our model vary as does z, the payoff from having been caught cheating? As the penalty z increases, the probability p^* that player 1 must attach to T in order to render player 2 indifferent decreases. Since the reputation manipulation stage mixes TL and BR outcomes in proportions p^* and $1 - p^*$, so does the incidence of cheating. We thus recover an intuitive link between the severity of punishment and the incidence of cheating.

5 Appendix: Proofs

5.1 Proof of Lemma 3

Suppose first that player 1 faces a history at which $n_{hB} > N_B(n_{hT} + 1)$ 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	3	BL	2
TL	3	TL	3
$\vdots$	$\vdots$	$\vdots$	$\vdots$
TL	3	TL	3
BL	2	TL	3

.

This looks precisely like the comparison we made in proving Lemma 2, and ensures the optimality of the candidate equilibrium strategies.

Next fix a history h and suppose $n_{hB} > N_B(n_{hT})$ but $n_{hB} < N_B(n_{hT} + 1)$, and so B is prescribed. 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. Hence, 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	2	TL	3
TL	3	BR	0

These two paths both terminate at $(n_{hT} + 1, n_{hB} + 1)$, and hence thereafter can be taken to generate identical continuation payoffs. As a result, it is clear from this comparison that the equilibrium path is optimal for sufficiently patient players.

Now suppose $n_{hB} < N_B(n_{hT})$. We begin with a preliminary result. 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	0	TR	1
TR	1	BR	0

These two paths both terminate at $(n_{hT} + 1, n_{hB})$, and hence thereafter can be taken to generate identical continuation payoffs. A comparison of the payoffs then shows that 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}).

With this result in hand, suppose that we have a state (n_{hT}, n_{hB}) with $n_{hB} < N_B(n_{hT})$ and with the equilibrium prescription being BR. Then we show that player 1 must also find B optimal at $(n_{hT} - 1, n_{hB})$. This ensures that there exists an increasing function $\underline{N}_B$ with the asserted properties. Supposing this is not the case, then we have the following equilibrium path and possible deviation at $(n_{hT} - 1, n_{hB})$:

"Equilibrium path"	Payoff	Deviation path	Payoff
TR	1	BR	0
BR	0	BR	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$
BR	0	BR	0
BR	0	TL	3

with identical subsequent play. The assumption that BR is played at (n_{hT}, n_{hB}) fixes the play after the first period in the alleged "equilibrium path." Moreover,

the length of the intervening BR sequence is bounded. As a result, sufficiently patient player 1's will find the deviation superior.

The observation that $\lim_{\delta \to 1} \underline{N}_B(n_T) < 0$ for all n_T follows from noting that for a fixed specification of mechanical types and any history, the number of B observations required to push the state above N_B is bounded (over values of α^0 and δ).

5.2 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 involves keeping player 2 nearly indifferent between L and R, which requires playing T with probability (arbitrarily close to, as δ gets large) q^* and yields payoff $q^* \max\{a, c\} + (1 - q^*) \max\{b, d\}$, which gives the result.

[*Payoff Characterization*] This result follows from noting that if player 1 plays T in every period, then there will only be 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.

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

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

This payoff can exceed c only if $\alpha_\tau^0 > \max\{q^*, p^*\}$. Hence, if $\alpha_\tau^0 \leq \max\{q^*, p^*\}$, 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 $\tau \to \infty$)) 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 $\tau \to \infty$) number of periods), a contradiction.

29

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

Upon reaching history h_τ, 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 τ. Once again, an upper bound on this continuation payoff is given by

$$\alpha^0_\tau(h_\tau)a + (1 - \alpha^0_\tau(h_\tau))\max\{b, d\}.$$

This payoff falls short of c if $\alpha^0_\tau(h_\tau) \leq \max\{q^*, p^*\}$, 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 $\alpha^0_\tau(h_\tau) > \max\{q^*, p^*\} + \varepsilon$. Suppose this is the case. Then beginning at h_τ, 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^0_τ, given that we know $\alpha^0_\tau > q^*$.) This ensures that player 1's continuation payoff at history h_τ must fall short of c. This in turn is a contradiction since playing T in every period after every history in h_τ gives a payoff approaching (as $\delta_\tau \to 1$) 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.) Now consider a correspondence $h(\alpha^0)$ that for any value of α^0, identifies the set of player-2 best responses, then identifies the set of player-1 best responses to these player-2 best responses, and then calculates the induces values of A^0. Our preceding calculations ensure that for sufficiently large δ, we have $h(\underline{p}) > \underline{p}$. 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 $\bar{p}$ of the time must lead to a payoff of c, we must have $h(\bar{p}) < \bar{p}$, for sufficiently large δ. It then follows from a version of the intermediate value theory and the fact that h is an upper-hemicontinuous, convex-valued correspondence that it has a fixed point on $[\underline{p}, \bar{p}]$, 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

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

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.

5.4 Proof of Lemma 4

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 $\bar{\alpha}$ 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.

5.5 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 no 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 $\alpha^0, q^* < \underline{\alpha}^0$, 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, consider the unary mixed equilibrium. Let $\{\delta_r\}_{r=1}^\infty$ be 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] Rudolf Avenhaus, Bernhard Von Stengel, and Shmuel Zamir. Inspection games. In Robert J. Aumann and Sergiu Hart, editors, *Handbook of Game Theory, Volume 3*, pages 1947–1987. North Holland, New York, 2002.

[2] David Backus and John Driffill. Inflation and reputation. *American Economic Review*, 75(3):530–538, 1985.

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

[4] 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.

[5] Avinahs Dixit and Susan Skeath. *Games of Strategy*. W. W. Norton and Company, New York, 1999.

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

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

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

[9] Simon Grant, Atsushi Kajii, and Ben Polak. Third down and a yard to go: Recursive expected utility and the Dixit-Skeath conundrum. *Economics Letters*, 73(3):275–286, 2001.

[10] Philippe Jehiel. Analogy-based expectation equilibrium. *Journal of Economic Theory*, 123(2):81–104, 2005.

[11] 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.

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

[13] Mordecai Kurz. On the structure and diversity of rational beliefs. *Economic Theory*, 4(6):877–900, 1994.

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

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

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

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

[18] Joel Watson. A 'reputation' refinement without equilibrium. *Econometrica*, 61(1):199–206, 1993.