Abstract: I study reputation effects when individuals have persistent private information that matters for their opponents’ payoffs. I examine a repeated game between a patient informed player and a sequence of myopic uninformed players. The informed player privately observes a persistent state, and is either a strategic type who can flexibly choose his actions or is one of the several commitment types that mechanically plays the same action in every period. Unlike the canonical models on reputation effects, the uninformed players’ payoffs depend on the state. This interdependence of values introduces new challenges to reputation building, namely, the informed player could face a trade-off between establishing a reputation for commitment and signaling favorable information about the state. My results address the predictions on the informed player’s payoff and behavior that apply across all Nash equilibria. When the stage game payoffs satisfy a monotone-supermodularity condition, I show that the informed long-run player can overcome the lack-of-commitment problem and secure a high payoff in every state and in every equilibrium. Under a condition on the distribution over states, he will play the same action in every period and maintain his reputation for commitment in every equilibrium. If the payoff structure is unrestricted and the probability of commitment types is small, then the informed player’s return to reputation building can be low and can provide a strict incentive to abandon his reputation.

Keywords: reputation, interdependent values, commitment type, payoff bound, unique equilibrium behavior

JEL Codes: C73, D82, D83

1 Introduction

Economists have long recognized that good reputations can lend credibility to people’s threats and promises. This intuition has been formalized in a series of works starting with Kreps and Wilson (1982), Milgrom and Roberts (1982), Fudenberg and Levine (1989) and others, who show that having the option to build a reputation dramatically affects a patient individual’s gains in long-term relationships. Their reputation results are robust as they apply across all equilibria, which enables researchers to make robust predictions in many decentralized markets where there is no mediator helping participants to coordinate on a particular equilibrium.

*Department of Economics, MIT. Email: harrydp@mit.edu. I am indebted to Daron Acemoglu, Drew Fudenberg, Juuso Toikka and Alex Wolitzky for guidance and support. I thank Jie Bai, Abhijit Banerjee, Vivek Bhattacharya, Seth Blumberg, Alessandro Bonatti, Gonzalo Cisternas, Dan Clark, Miaomiao Dong, Esther Dufo, Jetir Duraj, Mehmet Ekmekci, Glenn Ellison, Chishio Furukawa, Bob Gibbons, Ben Golub, Yingni Guo, Kevin He, Christian Hellwig, Daria Khromenkov, Annie Liang, Ernest Liu, Shuo Liu, Yusuke Narita, Alan Olivi, Parag Pathak, Ben Roth, Klaus Schmidt, Jean Tirole, Linh Tô, Xiaolu Wang, Birger Wernerfelt, Muhamet Yildiz and my seminar participants at MIT for helpful comments. All remaining errors are mine.
However, previous works on robust reputation effects all restrict attention to private value environments. This excludes situations where reputation builders have persistent private information that directly affects their opponents’ payoffs. For example in the markets for food and custom software, merchants can benefit from a reputation for providing good customer service, but they also want to signal their products have high quality. The latter directly affects consumers’ willingness to pay and is usually the merchants’ private information (Banerjee and Duflo 2000, Bai 2016). In the pharmaceutical, cable TV and passenger airline industries, incumbent firms could benefit from committing to fight potential entrants, but are also better informed about the market demand curve, such as the price elasticities, the effectiveness and spillover of advertising (Ellison and Ellison 2011, Seamans 2013, Sweeting, Roberts and Gedge 2016), etc. As a result, incumbent firms’ choices of prices, quantities and the intensity of advertising not only show their resolve to fight entrants but also signal their private information about demand. Understanding how the interactions between reputation building and signalling affect economic agents’ reputational incentives is important both for firms in designing business strategies and for policy makers in evaluating the merits of quality-control programs and anti-trust regulations.

Motivated by these applications, this paper addresses the robust predictions in reputation games where a player has persistent private information about his opponents’ payoffs. In my model, a patient long-run player (player 1, he, seller, incumbent) interacts with a sequence of short-run players (player 2, she/they, buyers, entrants). Unlike the canonical reputation models, I study interdependent value environments in which player 1 privately observes a perfectly persistent state (product quality, market demand) that directly affects player 2’s payoff. Player 1 is either one of the strategic types who maximizes his discounted payoff and will be referred to by the state he observes, or is committed to play a state-contingent stationary strategy. Player 2 updates her belief by observing all the past actions. I show that (1) the robust reputation effects on player 1’s payoffs extend to a class of interdependent value games despite the existence of a trade-off between commitment and signalling, (2) reputation can also lead to robust and accurate predictions on player 1’s equilibrium behavior.

To illustrate the challenges, consider an example of an incumbent firm (player 1) facing a sequence of potential entrants. Every entrant chooses between staying out (O) and entering the market (E). Her preference between O and E depends not only on the incumbent’s business strategy, which is either fight (F) or accommodate (A), but also on the market demand curve (the state \( \theta \), can be price elasticity, market size, etc.), which is fixed over time and is either high (H) or low (L). This is modeled as the following entry deterrence game:

<table>
<thead>
<tr>
<th>( \theta = \text{High} )</th>
<th>Out</th>
<th>Enter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fight</td>
<td>2, 0</td>
<td>0, −1</td>
</tr>
<tr>
<td>Accommodate</td>
<td>3, 0</td>
<td>1, 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \theta = \text{Low} )</th>
<th>Out</th>
<th>Enter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fight</td>
<td>2 − ( \eta ), 0</td>
<td>−( \eta ), 1</td>
</tr>
<tr>
<td>Accommodate</td>
<td>3, 0</td>
<td>1, 2</td>
</tr>
</tbody>
</table>

where \( \eta \in \mathbb{R} \) is a parameter. When \( \theta = H \) is common knowledge (call it the private value benchmark), the
The incumbent faces a lack-of-commitment problem in the stage game: His payoff from the unique Nash equilibrium \((A, E)\) is 1. This is strictly lower than his payoff by committing to fight, which provides his opponent an incentive to stay out and he will receive his \textit{commitment payoff} equal to 2. Fudenberg and Levine (1989) show that reputation can solve this lack-of-commitment problem by establishing the following commitment payoff \textit{theorem}: if the incumbent is non-strategic and fights in every period with positive probability, then a patient strategic incumbent can secure his commitment payoff in every Nash equilibrium of the repeated game. Intuitively, if the strategic incumbent imitates the non-strategic one, then he will eventually convince the entrants that \(F\) will be played with high enough probability and the latter will best respond by staying out.

This logic no longer applies when \(\theta\) is the incumbent’s private information. This is because an entrant’s best reply to \(F\) depends on \(\theta\) (it is \(O\) when \(\theta = H\) and \(E\) when \(\theta = L\)), which is signalled through the incumbent’s past actions. In situations where fighting is interpreted as a signal of state \(L\), an entrant will have an incentive to play \(E\) despite being convinced that \(F\) will be played. As a result, the incumbent’s return from always fighting will be low. Furthermore, obtaining robust and accurate predictions on the incumbent’s equilibrium behavior faces additional challenges as he is \textit{repeatedly signalling} the state. This could lead to multiple possible behaviors. Even the commitment payoff theorem cannot imply that he will maintain his reputation for fighting in every equilibrium, as a strategy that can secure himself a high payoff is not necessarily his optimal strategy.

In Section 3, I examine when the commitment payoff theorem applies to every payoff function of the long-run player (i.e. it is \textit{fully robust}) without any restrictions on the game’s payoff structure. Theorem provides a sufficient and (almost) necessary condition for full robustness, which requires that the prior likelihood ratio between each \textit{bad strategic type} and the commitment type be below a cutoff. According to this result, securing the commitment payoff from a mixed action occurs under more demanding conditions than that from a nearby pure action. This implies that small trembles of pure commitment types can lead to a large decrease in the strategic long-run player’s guaranteed payoff. Another interesting observation is that playing some actions in the support of the mixed commitment action can \textit{increase} the aforementioned likelihood ratios. Therefore, mixed commitment payoffs cannot be guaranteed by replicating the commitment strategy, making the existing techniques in Fudenberg and Levine (1989, 1992), Gossner (2011) inapplicable. To overcome these difficulties, my proof of the sufficiency part makes use of martingale techniques and the central limit theorem to construct a \textit{non-stationary} strategy such that player 1 can achieve three goals simultaneously: (1) avoiding negative...
inferences about the state, (2) matching the frequency of his actions to the mixed commitment action, (3) player 2’s prediction about his actions is close to the mixed commitment action in all but a bounded number of periods.

Theorem 1 has two interpretations. First, starting with the private value reputation game in Fudenberg and Levine (1989), it evaluates the robustness of their main insight under a richer set of perturbations. Namely, player 2 can entertain the possibility that her opponent is another strategic type who has private information about her payoff. My result implies that their fully robust reputation result extends when these interdependent value perturbations are unlikely compared to the commitment types, and vice versa. Second, one can also start with a repeated incomplete information game with interdependent values and perturb it with commitment types. According to this view, every commitment type is arbitrarily unlikely compared to every strategic type. Theorem 1 then implies that in some equilibria, player 1’s return from reputation building is low, and in fact, he will have a strict incentive to abandon his reputation. Therefore, reputation cannot guarantee that player 1 can overcome the lack-of-commitment problem even when he is arbitrarily patient.

This second interpretation motivates the study of games with more specific payoff structures. In Section 4, I focus on stage games with monotone-supermodular payoffs (MSM for short). This requires that the states and every player’s actions be ranked such that (1) player 1’s payoff is strictly increasing in player 2’s action but is strictly decreasing in his own action (or monotonicity), and (2) the action profile and the state are complements in player 1’s stage game payoff function, and player 2 has a stronger incentive to play a higher action when the state is higher or when player 1’s action is higher (or supermodularity). In the entry deterrence example, if we rank the states and actions according to $H \succ L$, $F \succ A$ and $O \succ E$, then MSM translates into $\eta > 0$, which is the case when $\theta$ is the price elasticity of demand, the market size, the effectiveness of advertising, etc. MSM is also satisfied in buyer-seller games where providing good service is less costly for the seller when his product quality is high, which fits into the custom software industry and the restaurant industry.

My results establish robust predictions on player 1’s equilibrium payoff and behavior when there exists a commitment type that plays the highest action in every period. I consider two cases. When the high states are relatively more likely compared to the low states (the optimistic prior case), Theorem 2 shows that a patient player 1 can guarantee his commitment payoff from playing the highest action in every state and in every equilibrium. In the example, when state $H$ is more likely than state $L$, player 1 receives at least 2 in state $H$ and $\max\{2 - \eta, 1\}$ in state $L$. This payoff bound applies even when every commitment type is arbitrarily unlikely relative to every strategic type. It is also tight in the sense that no strategic type can guarantee himself a strictly higher equilibrium payoff by establishing a reputation for playing another pure commitment action.4

In the complementary scenario (the pessimistic prior case), Theorem 3 shows that when player 1 is patient

4This conclusion extends to any other mixed commitment action if in the stage game, the long-run player strictly prefers the highest action profile to the lowest action profile in every state.
and the probability of commitment is small (1) his equilibrium payoff equals to the highest equilibrium payoff in the benchmark game without commitment types (Theorem 3 and Proposition 4.2); (2) his on-path behavior is the same across all Nash equilibria.[5] According to this unique behavior, there exists a cutoff state (in the example, state \( L \)) such that the strategic player 1 plays the highest action in every period if the state is above this cutoff, plays the lowest action in every period if the state is below this cutoff, and mixes between playing the highest action in every period and playing the lowest action in every period at the cutoff state. That is to say, player 1 will behave consistently and maintain his reputation for commitment in all equilibria.

The intuition behind this behavioral uniqueness result is the following disciplinary effect: (1) player 1 can obtain a high continuation payoff by playing the highest action, (2) but it is impossible for him to receive a high continuation payoff after he has failed to do so, as player 2’s belief about the state will become even more pessimistic than her prior. The first part is driven by the commitment type and the second is because the low states are more likely. This contrasts with Fudenberg and Levine (1989) and the optimistic prior case where deviating from the commitment action may lead to an optimistic posterior, after which a patient player 1 can still receive a high continuation payoff. As a result, player 1 can have multiple on-path behaviors, and in many sequential equilibria, he may have a strict incentive to behave inconsistently and abandon his reputation.

Conceptually, the above comparison suggests that interdependent values can contribute to the sustainability of reputation. This channel is novel compared to those proposed in the existing literature, such as impermanent commitment types (Mailath and Samuelson 2001, Ekmekci, et al. 2012), competition between informed players (Hörner 2002), incomplete information about the informed player’s past behavior (Ekmekci 2011) and others.⁶

A challenge to prove Theorems 2 and 3 comes from the observation that a repeated supermodular game is not supermodular. This is because player 1’s action today can have persistent effects on future equilibrium play. I apply a result in a companion paper (Liu and Pei 2017) which states that if a 1-shot signalling game has MSM payoffs, then the sender’s equilibrium action must be non-decreasing in the state. In a repeated signalling game with MSM stage game payoffs, this result implies that in equilibria where playing the highest action in every period is optimal for player 1 in a low state (call them regular equilibria), then he must be playing the highest action with probability 1 at every on-path history in every higher state. Therefore, in every regular equilibrium, player 2’s posterior about the state will never decrease if player 1 has always played the highest action. Nevertheless, there can also exist irregular equilibria where playing the highest action in every period is not optimal in any low state, and it is possible that at some on-path histories, it will lead to a deterioration

---

5Theorem 3 states the result when there is one commitment action. Theorem 3' (Appendix D.2) allows for multiple commitment actions and shows that when the total probability of commitment is small enough, player 1’s payoff and on-path behavior are almost the same across all equilibria. If all commitment types are pure, then payoff and on-path behavior are the same across all equilibria.

6In contrast to these papers and Cripps, Mailath and Samuelson (2004), I adopt a more robust standard for reputation sustainability by requiring that it be sustained in every equilibrium.
of player 2’s belief about the state. To deal with this complication, my proof shows that in every irregular equilibrium, if player 1 has never deviated from the highest action, then player 2’s belief about the state can never fall below a cutoff.

To summarize, we know that in the optimistic prior case, player 2’s posterior cannot become too pessimistic given that player 1 has always played the highest action, no matter whether the equilibrium is regular or irregular. Therefore, if player 1 plays the highest action in every period, he can convince player 2 that the highest action will be played in the future and at the same time, player 2’s posterior belief about the state will remain optimistic, which leads to the commitment payoff theorem. However, due to the existence of irregular equilibria, player 1 has multiple equilibrium behaviors. In the pessimistic prior case, the necessary condition for irregular equilibria is violated in the first period. Therefore, irregular equilibria do not exist and every regular equilibrium will lead to the same equilibrium payoff and on-path behavior.

My work contributes to the existing literature from several different angles. From a modeling perspective, it unifies two existing approaches to the study of reputation, differing mainly in the interpretation of the informed player’s private information. Pioneered by Fudenberg and Levine (1989), the literature on reputation refinement focuses on private value environments and studies how a reputation for commitment affects a patient informed player’s payoff in all equilibria.7 A separate strand of works on dynamic signalling games, including Bar-Isaac (2003), Lee and Liu (2013), Pei (2015) and Toxvaerd (2017), examines the effects of persistent private information about payoff-relevant variables (such as talent, quality, market demand) on the informed player’s behavior. However, these papers have focused on some particular equilibria rather than on the common properties of all equilibria. In contrast, I introduce a framework that incorporates commitment over actions and persistent private information about the uninformed players’ payoffs. In games with MSM payoffs, I derive robust predictions on the informed player’s payoff and behavior that apply across all Nash equilibria.

In the study of repeated Bayesian games with interdependent values,8 my reputation results can be interpreted as an equilibrium refinement, just as Fudenberg and Levine (1989) did for the repeated complete information games studied in Fudenberg, Kreps and Maskin (1990). By allowing the informed long-run player to be non-strategic and mechanically playing a state-contingent stationary strategy, Theorems 1 and 2 show that reputation effects can sharpen the predictions on a patient player’s equilibrium payoff. Theorem 3 advances this research

---


8. This is currently a challenging area and not much is known except for 0-sum games (Aumann and Maschler 1995, Pésky and Toikka 2017), undiscounted games (Hart 1985), belief-free equilibrium payoff sets in games with two equally patient players (Hörner and Lovo 2009, Hörner et al. 2011). In ongoing work (Pei 2016), I characterize the limiting equilibrium payoff set in a repeated Bayesian game between a patient long-run player and a sequence of short-run players when the stage game has MSM payoffs.
agenda one step further by showing that reputation effects can also lead to accurate predictions on a patient player’s equilibrium behavior, which is a distinctive feature of interdependent value models.

In terms of the applications, my result offers a robust explanation to Bain (1949)’s classical observation that “...established sellers persistently ... forego high prices ... for fear of thereby attracting new entry to the industry and thus reducing the demands for their outputs and their own profit”. This will only happen in some non-renegotiation proof equilibria under private values, but will happen in every equilibrium when the incumbent has private information about demand and the potential entrants are optimistic about their prospects of entry. Similarly, in the study of firm-consumer relationships, my result provides a robust foundation for Klein and Leffler (1981)’s reputational capital theory, which assumes that consumers will coordinate and punish the firm after observing low effort. This will happen in every equilibrium when consumers are skeptical enough about the product quality, which the firm privately knows. I will elaborate more on these in subsection 4.5.

2 The Model

Time is discrete, indexed by $t = 0, 1, 2,...$. An infinitely-lived long-run player (player 1, he) with discount factor $\delta \in (0, 1)$ interacts with a sequence of short-run players (player 2, she), one in each period. In period $t$, players simultaneously choose their actions $(a_{1,t}, a_{2,t}) \in A_1 \times A_2$. Both $A_1$ and $A_2$ are finite sets with $|A_i| \geq 2$ for $i \in \{1, 2\}$. Players have access to a public randomization device, with $\xi_t \in \Xi$ as the realization in period $t$.

States, Strategic Types & Commitment Types: Let $\theta \in \Theta$ be the state of the world, which is perfectly persistent and is player 1’s private information. I assume that $\Theta$ is a finite set. Player 1 is either strategic, in which case he can flexibly choose his action in every period, or he is committed to play the same action in every period, which can be pure or mixed and can be state contingent.

I abuse notation by using $\theta$ to denote the strategic type who knows that the state is $\theta$ (or type $\theta$). As for commitment, every commitment type is defined based on the (mixed) action he plays. Formally, let $\Omega \equiv \Theta \cup \Omega^m \subset \Delta(A_1)$ be the set of actions player 1 could possibly commit to, which is assumed to be finite. I use $\alpha_1 \in \Omega^m$ to represent the commitment type that plays $\alpha_1$ in every period (or commitment type $\alpha_1$). Let $\phi_{\alpha_1} \in \Delta(\Theta)$ be the distribution of $\theta$ conditional on player 1 being commitment type $\alpha_1$, with $\phi \equiv \{\phi_{\alpha_1}\}_{\alpha_1 \in \Omega^m}$.

Let $\Omega \equiv \Theta \cup \Omega^m$ be the set of types, with $\omega \in \Omega$ a typical element. Let $\mu \in \Delta(\Omega)$ be player 2’s prior belief, which I assume has full support. The pair $(\mu, \phi)$ induces a joint distribution over $\theta$ and player 1’s characteristics (committed or strategic), which I call a distributional environment.

Note that the above formulation of commitment accommodates the one in which player 1 commits to play a state-contingent stationary strategy. To see this, let $\gamma : \Theta \rightarrow \Delta(A_1)$ be a state-contingent commitment plan,
with $\Gamma$ the finite set of commitment plans. Player 2 has a prior over $\Theta$ as well as the chances that player 1 is strategic or is committed to follow each plan in $\Gamma$. To convert this to my formulation, let

$$\Omega^m \equiv \{\alpha_1 \in \Delta(A_1) \mid \text{there exist } \gamma \in \Gamma \text{ and } \theta \in \Theta \text{ such that } \gamma(\theta) = \alpha_1\},$$

which is the set of actions that are played under (at least) one commitment plan. The probability of every $\alpha_1 \in \Omega^m$ and its correlation with the state $\phi_{\alpha_1}$ can be computed via player 2’s prior. My formulation is more general, as it allows for arbitrary correlations between the state and the probability of being committed.

**Histories & Payoffs:** All past actions are perfectly monitored. Let $h^t = \{a_{1,s}, a_{2,s}, \xi_s\}_{s=0}^{t-1} \in \mathcal{H}^t$ be the public history in period $t$ with $\mathcal{H} \equiv \bigcup_{t=0}^{+\infty} \mathcal{H}^t$. Let $\sigma_\omega : \mathcal{H} \to \Delta(A_1)$ be type $\omega$’s strategy, with the restriction that $\sigma_{\alpha_1}(h^t) = \alpha_1$ for every $(\alpha_1, h^t) \in \Omega^m \times \mathcal{H}$. Let $\sigma_1 \equiv (\sigma_\omega)_{\omega \in \Omega}$ be player 1’s strategy. Let $\sigma_2 : \mathcal{H} \to \Delta(A_2)$ be player 2’s strategy. Let $\sigma \equiv (\sigma_1, \sigma_2)$ be a typical strategy profile, and let $\Sigma$ be the set of strategy profiles.

Player $i$’s stage game payoff in period $t$ is $u_i(\theta, a_{1,t}, a_{2,t})$, with $i \in \{1, 2\}$, which is naturally extended to the domain $\Delta(\Theta) \times \Delta(A_1) \times \Delta(A_2)$. Unlike Fudenberg and Levine (1989), my model has *interdependent values* as player 2’s payoff depends on $\theta$, which is player 1’s private information. Strategic type $\theta$ maximizes $\sum_{t=0}^{+\infty}(1-\delta)^t u_1(\theta, a_{1,t}, a_{2,t})$. The player 2 who arrives in period $t$ maximizes his expected stage game payoff.

Let $\text{BR}_2(\alpha_1, \pi) \subset A_2$ be the set of player 2’s pure best replies when $\alpha_1$ and $\theta$ are independently distributed with marginal distributions $\alpha_1 \in \Delta(A_1)$ and $\pi \in \Delta(\Theta)$, respectively. For every $(\alpha_1^*, \theta) \in \Omega^m \times \Theta$, let

$$v_\theta(\alpha_1^*) \equiv \min_{a_2^* \in \text{BR}_2(\alpha_1^*, \theta)} u_1(\theta, \alpha_1^*, a_2^*)[9]$$

be type $\theta$’s (complete information) commitment payoff from playing $\alpha_1^*$. If $\alpha_1^*$ is pure, then $v_\theta(\alpha_1^*)$ is a pure commitment payoff. Otherwise, $\alpha_1^*$ is mixed and $v_\theta(\alpha_1^*)$ is a mixed commitment payoff.

**Solution Concept & Questions:** The solution concept is Bayes Nash equilibrium (or *equilibrium* for short). The existence of equilibrium follows from Fudenberg and Levine (1983), as $\Theta, A_1$ and $A_2$ are all finite sets and the game is continuous at infinity. Let $\text{NE}(\delta, \mu, \phi) \subset \Sigma$ be the set of equilibria under parameter configuration $(\delta, \mu, \phi)$. Let $V_\theta^\sigma(\delta)$ be type $\theta$’s discounted average payoff under strategy profile $\sigma$ and discount factor $\delta$. Let $V_\theta(\delta, \mu, \phi) \equiv \inf_{\sigma \in \text{NE}(\delta, \mu, \phi)} V_\theta^\sigma(\delta)$ be type $\theta$’s worst equilibrium payoff.

I am interested in two sets of questions. First, can we find good lower bounds for a patient long-run player’s guaranteed payoff, i.e. $\lim_{\delta \to 1} V_\theta(\delta, \mu, \phi)$? In particular, can we extend Fudenberg and Levine (1989)’s insights that reputation can overcome the lack-of-commitment problem (when the reputation builder is patient)

---

9 Abusing notation, I will use $\theta$ to denote the Dirac measure on $\theta$. The same rule applies to degenerate distributions on $A_1$ and $A_2$. 
to interdependent value environments. Formally, for a given $(\alpha^*_1, \theta) \in \Omega^m \times \Theta$, is it true that:

$$\liminf_{\delta \to 1} V^\phi(\delta, \mu, \phi) \geq v^\theta(\alpha^*_1)?$$

(2.2)

Furthermore, when is the above commitment payoff bound fully robust, that is, inequality (2.2) applies to every payoff function of the long-run player? Second, can we obtain robust predictions on player 1’s equilibrium behavior? In particular, will he play the commitment strategy and maintain his reputation in every equilibrium?

My first set of questions examines player 1’s guaranteed payoff when he can build a reputation. When $u_2$ does not depend on $\theta$, inequality (2.2) is implied by the results in Fudenberg and Levine (1989, 1992) and player 1 can guarantee the payoff on the RHS by playing $\alpha^*_1$ in every period. In interdependent value environments, however, player 1 may receive a low payoff by playing $\alpha^*_1$ in every period, as convincing player 2 that $\alpha^*_1$ will be played does not determine her best reply. I address the robustness against equilibrium selection and against misspecifications of the long-run player’s payoff function. Both are desirable properties of the results in Fudenberg and Levine (1989, 1992) as (1) reputation models are often applied to decentralized markets where there are no mediators helping participants to coordinate on a particular equilibrium, and (2) the modeler and the short-run players may entertain incorrect beliefs about the informed long-run player’s payoff function.

My second set of questions advances the reputation literature one step further by examining the robust predictions on the long-run player’s equilibrium behavior. Nevertheless, delivering robust behavioral predictions in this infinitely-repeated signalling game is challenging, as the conventional wisdom suggests that both infinitely-repeated games and signalling games have multiple equilibria with diverging behavioral predictions. Note that the commitment payoff bound does not imply that the long-run player will play his commitment strategy in every equilibrium, as a strategy that can secure him a high payoff is not necessarily his optimal strategy.

## 3 Fully Robust Commitment Payoff Bounds

I characterize the set of distributional environments under which the commitment payoff bound is fully robust. My conditions require that the likelihood ratios between some strategic types and the commitment type be below some cutoffs. My result evaluates the robustness of the commitment payoff bound in private value games against interdependent value perturbations. It also examines the validity of the commitment payoff bound in interdependent value environments without any restrictions on the long-run player’s payoff function.
3.1 Saturation Set & Strong Saturation Set

In this subsection, I make the generic assumption that for every \((\alpha_1^*, \theta) \in \Omega^m \times \Theta\), \(BR_2(\alpha_1^*, \theta)\) is a singleton.\(^{10}\) Let \(a_2^*\) be the unique element in \(BR_2(\alpha_1^*, \theta)\). For every \((\alpha_1^*, \theta) \in \Omega^m \times \Theta\), let

\[
\Theta^b_{(\alpha_1^*, \theta)} \equiv \{ \tilde{\theta} \in \Theta \mid a_2^* \not\in BR_2(\alpha_1^*, \tilde{\theta}) \},
\]

be the set of bad states (with respect to \((\alpha_1^*, \theta)\)). Let \(k(\alpha_1^*, \theta) \equiv |\Theta^b_{(\alpha_1^*, \theta)}|\) be its cardinality, with all private value models satisfying \(k(\alpha_1^*, \theta) = 0\). If \(\tilde{\theta} \in \Theta^b_{(\alpha_1^*, \theta)}\), then type \(\tilde{\theta}\) is a ‘bad’ strategic type. For every \(\hat{\mu} \in \Delta(\Omega)\) with \(\hat{\mu}(\alpha_1^+) > 0\), let \(\hat{\lambda}(\tilde{\theta}) \equiv \hat{\mu}(\tilde{\theta})/\hat{\mu}(\alpha_1^+)\) be the likelihood ratio between type \(\tilde{\theta}\) and commitment type \(\alpha_1^+\). Let

\[
\hat{\lambda}(\tilde{\theta}) = \left(\hat{\lambda}(\tilde{\theta})\right)_{\tilde{\theta} \in \Theta^b_{(\alpha_1^+, \theta)}} \in \mathbb{R}^{k(\alpha_1^+, \theta)}
\]

be the likelihood ratio vector. The best response set for \((\alpha_1^+, \theta) \in \Omega^m \times \Theta\) is defined as:\(^{11}\)

\[
\overline{\Lambda}(\alpha_1^+, \theta) \equiv \left\{ \hat{\lambda} \in \mathbb{R}^{k(\alpha_1^+, \theta)} \mid \{a_2^*\} = \arg \max_{a_2 \in A_2} \left\{ u_2(\phi_1^+, \alpha_1^+, a_2) + \sum_{\tilde{\theta} \in \Theta^b_{(\alpha_1^+, \theta)}} \hat{\lambda}(\tilde{\theta}) u_2(\tilde{\theta}, \alpha_1^+, a_2) \right\} \}
\]

Intuitively, a likelihood ratio vector belongs to the best response set if \(a_2^*\) is player 2’s strict best reply to \(\alpha_1^+\) when she only counts the bad strategic types and the commitment type \(\alpha_1^+\) in her calculations, while ignoring all the other strategic types and commitment types.

**Definition 1** (Saturation Set). The saturation set for \((\alpha_1^+, \theta) \in \Omega^m \times \Theta\) is:

\[
\Lambda(\alpha_1^+, \theta) \equiv \{ \hat{\lambda} \mid \lambda' \in \overline{\Lambda}(\alpha_1^+, \theta) \text{ for every } 0 \ll \lambda' \ll \hat{\lambda} \},
\]

in which ‘\(\ll\)’ denotes weak dominance in product order on \(\mathbb{R}^{k(\alpha_1^+, \theta)}\) and 0 is the null vector in \(\mathbb{R}^{k(\alpha_1^+, \theta)}\).

Intuitively, \(\hat{\lambda}\) belongs to the saturation set if and only if every likelihood ratio vector equal or below \(\hat{\lambda}\) belongs to the best response set \(\overline{\Lambda}(\alpha_1^+, \theta)\). By definition, \(\Lambda(\alpha_1^+, \theta) \neq \{ \emptyset \}\) if and only if 0 \(\in\) \(\Lambda(\alpha_1^+, \theta)\), or equivalently, \(BR_2(\alpha_1^+, \theta) = BR_2(\alpha_1^+, \phi_{\alpha_1^+}) = \{a_2^*\}\).

If \(\Lambda(\alpha_1^+, \theta) \neq \{ \emptyset \}\), then for every \(\tilde{\theta} \in \Theta^b_{(\alpha_1^+, \theta)}\), let \(\psi(\tilde{\theta})\) be the largest \(\psi \in \mathbb{R}_+\) such that:

\[
a_2^* \in \arg \max_{a_2 \in A_2} \left\{ u_2(\phi_1^+, \alpha_1^+, a_2) + \psi u_2(\tilde{\theta}, \alpha_1^+, a_2) \right\}.
\]

By definition, \(\psi(\tilde{\theta})\) is the intercept of \(\Lambda(\alpha_1^+, \theta)\) on the \(\lambda(\tilde{\theta})\)-coordinate, which is strictly positive and finite.

\(^{10}\)\(BR_2(\alpha_1^+, \theta)\) being a singleton is satisfied under generic \(u_2(\theta, a_1, a_2)\). This assumption will be relaxed in Online Appendix B, where I develop generalized fully robust commitment payoff bounds.

\(^{11}\)A relevant argument \(\phi_{\alpha_1^+}\) is suppressed in the expression \(\overline{\Lambda}(\alpha_1^+, \theta)\) to simplify notation.
When $BR$ is a $k$-dimensional hyperplane that contains all the intersections between $\alpha_1^*$ and the coordinates. In general, when $BR_2(\alpha_1^*, \theta)$ may not be a singleton, $\Lambda(\alpha_1^*, \theta)$ is defined as $\mathbb{R}^{k(\alpha_1^*, \theta)} \setminus \text{co}(\mathbb{R}^{k(\alpha_1^*, \theta)} \setminus \Lambda(\alpha_1^*, \theta))$, where $\text{co}(\cdot)$ denotes the convex hull. I show in Lemma B.2 (Online Appendix B) that $\bar{\lambda} \in \Lambda(\alpha_1^*, \theta)$ if and only if there exists $\phi \in (0, +\infty)^{k(\alpha_1^*, \theta)}$ such that:

$$
\bar{\lambda} \in \left\{ \bar{\lambda} \mid \sum_{\tilde{\theta} \in \Theta_{(\alpha_1^*, \theta)}} \bar{\lambda}(\tilde{\theta})/\phi(\tilde{\theta}) < 1 \right\} \subseteq \Lambda(\alpha_1^*, \theta).
$$

Figure 1 depicts the three sets in an example with two bad strategic types. I summarize some geometric properties of these sets for future reference. First, despite $\Lambda(\alpha_1^*, \theta)$ can be unbounded, both $\Lambda(\alpha_1^*, \theta)$ and $\Lambda(\alpha_1^*, \theta)$ are bounded sets. Furthermore, they are convex polyhedrons with characterizations independent of both player 1’s payoff function and the probabilities of commitment types other than $\alpha_1^*$. Second, as suggested by the notation, $\Lambda(\alpha_1^*, \theta) \subset \Lambda(\alpha_1^*, \theta) \subset \Lambda(\alpha_1^*, \theta)$. Third, if there is only one bad strategic type, i.e. $k(\alpha_1^*, \theta) = 1$ and $\Lambda(\alpha_1^*, \theta) \neq \emptyset$, then there exists a scalar $\psi^* \in (0, +\infty)$ such that:

$$
\Lambda(\alpha_1^*, \theta) = \Lambda(\alpha_1^*, \theta) = \Lambda(\alpha_1^*, \theta) = \{ \bar{\lambda} \in \mathbb{R} \mid 0 \leq \bar{\lambda} < \psi^* \}. \quad (3.5)
$$

When $k(\alpha_1^*, \theta) \geq 2$, however, these three sets can be different, as I show in Figure 1.

---

**Definition 2** (Strong Saturation Set). The strong saturation set for $(\alpha_1^*, \theta) \in \Omega^m \times \Theta$ is:

$$
\Lambda(\alpha_1^*, \theta) \equiv \left\{ \bar{\lambda} \mid \sum_{\tilde{\theta} \in \Theta_{(\alpha_1^*, \theta)}} \bar{\lambda}(\tilde{\theta})/\phi(\tilde{\theta}) < 1 \right\} \quad \text{if } \Lambda(\alpha_1^*, \theta) \neq \emptyset,

\left\{ \emptyset \right\} \quad \text{if } \Lambda(\alpha_1^*, \theta) = \emptyset \quad (3.4)
$$

Intuitively, the strong saturation set contains every non-negative vector that lies below the $k(\alpha_1^*, \theta) - 1$ dimensional hyperplane that contains all the intersections between $\Lambda(\alpha_1^*, \theta)$ and the coordinates. In general, $\Lambda(\alpha_1^*, \theta)$ may not be a singleton, $\Lambda(\alpha_1^*, \theta)$ is defined as $\mathbb{R}^{k(\alpha_1^*, \theta)} \setminus \text{co}(\mathbb{R}^{k(\alpha_1^*, \theta)} \setminus \Lambda(\alpha_1^*, \theta))$, where $\text{co}(\cdot)$ denotes the convex hull. I show in Lemma B.2 (Online Appendix B) that $\bar{\lambda} \in \Lambda(\alpha_1^*, \theta)$ if and only if there exists $\phi \in (0, +\infty)^{k(\alpha_1^*, \theta)}$ such that:

$$
\bar{\lambda} \in \left\{ \bar{\lambda} \mid \sum_{\tilde{\theta} \in \Theta_{(\alpha_1^*, \theta)}} \bar{\lambda}(\tilde{\theta})/\phi(\tilde{\theta}) < 1 \right\} \subseteq \Lambda(\alpha_1^*, \theta).
$$

Figure 1 depicts the three sets in an example with two bad strategic types. I summarize some geometric properties of these sets for future reference. First, despite $\Lambda(\alpha_1^*, \theta)$ can be unbounded, both $\Lambda(\alpha_1^*, \theta)$ and $\Lambda(\alpha_1^*, \theta)$ are bounded sets. Furthermore, they are convex polyhedrons with characterizations independent of both player 1’s payoff function and the probabilities of commitment types other than $\alpha_1^*$. Second, as suggested by the notation, $\Lambda(\alpha_1^*, \theta) \subset \Lambda(\alpha_1^*, \theta) \subset \Lambda(\alpha_1^*, \theta)$. Third, if there is only one bad strategic type, i.e. $k(\alpha_1^*, \theta) = 1$ and $\Lambda(\alpha_1^*, \theta) \neq \emptyset$, then there exists a scalar $\psi^* \in (0, +\infty)$ such that:

$$
\Lambda(\alpha_1^*, \theta) = \Lambda(\alpha_1^*, \theta) = \Lambda(\alpha_1^*, \theta) = \{ \bar{\lambda} \in \mathbb{R} \mid 0 \leq \bar{\lambda} < \psi^* \}. \quad (3.5)
$$

When $k(\alpha_1^*, \theta) \geq 2$, however, these three sets can be different, as I show in Figure 1.
3.2 Statement of Result

My first result characterizes the set of \((\mu, \phi)\) under which the commitment payoff bound is fully robust i.e. it applies to every \(u_1\). Let \(\mu_t\) be player 2’s belief in period \(t\). Let \(\lambda\) and \(\lambda_t\) be the likelihood ratio vectors induced by \(\mu\) and \(\mu_t\), respectively. For a set \(X \subset \mathbb{R}^n\), recall that \(\text{co}(X)\) is its convex hull and let \(\text{cl}(X)\) be its closure.

**Theorem 1.** For every \((\alpha^*_1, \theta) \in \Omega^m \times \Theta\) with \(\alpha^*_1\) being pure,

1. If \(\lambda \in \Lambda(\alpha^*_1, \theta)\), then \(\liminf_{\delta \to 1} V_\theta(\delta, \mu, \phi) \geq v_\theta(\alpha^*_1)\) for every \(u_1\).

2. If \(\lambda \notin \text{cl}\left(\Lambda(\alpha^*_1, \theta)\right)\) and \(\text{BR}_2(\alpha^*_1, \phi_{\alpha^*_1})\) is a singleton, then there exists \(u_1\) such that \(\limsup_{\delta \to 1} V_\theta(\delta, \mu, \phi) < v_\theta(\alpha^*_1)\).

For every \((\alpha^*_1, \theta) \in \Omega^m \times \Theta\) with \(\alpha^*_1\) being mixed,

3. If \(\lambda \in \Lambda(\alpha^*_1, \theta)\), then \(\liminf_{\delta \to 1} V_\theta(\delta, \mu, \phi) \geq v_\theta(\alpha^*_1)\) for every \(u_1\).

4. If \(\lambda \notin \text{cl}\left(\Lambda(\alpha^*_1, \theta)\right)\), \(\text{BR}_2(\alpha^*_1, \phi_{\alpha^*_1})\) is a singleton and \(\alpha^*_j \notin \text{co}\left(\Omega^m \setminus \{\alpha^*_1\}\right)\), then there exists \(u_1\) such that \(\limsup_{\delta \to 1} V_\theta(\delta, \mu, \phi) < v_\theta(\alpha^*_1)\).

According to Theorem 1, full robustness requires that the likelihood ratio between every bad strategic type and the relevant commitment type be below some cutoff, while it does not depend on the probabilities of the other strategic types and commitment types. Intuitively, this is because type \(\theta\) needs to come up with a history-dependent action plan under which the likelihood ratio vector will remain low forever along every dimension. When the commitment payoff bound is fully robust, such action plans should exist regardless of player 2’s belief about the other strategic types’ strategies. This includes the adverse belief in which all the good strategic types separate from, while all the bad strategic types pool with, the commitment type.

However, unlike the private value benchmark, player 1 cannot guarantee his mixed commitment payoff by replicating the mixed commitment strategy. This is because playing some actions in the support of the mixed commitment strategy can increase some likelihood ratios, after which player 2’s belief about the persistent state becomes pessimistic and player 1 cannot guarantee a high continuation payoff. Moreover, as \(\Lambda(\alpha^*_1, \theta) \subset \Lambda(\alpha^*_1, \theta)\), overcoming the lack-of-commitment problem and securing the commitment payoff requires more demanding conditions when the commitment strategy is mixed. This implies that small trembles by a pure commitment type can lead to a large decrease in player 1’s guaranteed equilibrium payoff. This highlights another distinction between private and interdependent values, which I formalize in Online Appendix A.

This theorem has two interpretations. First, it evaluates the robustness of reputation effects in private value reputation games against a richer set of perturbations. Starting from Fudenberg and Levine (1989) in which \(\theta\) is common knowledge and there is a positive chance of a commitment type, one can allow the short-run players
to entertain the possibility that their opponent is another strategic type who may have private information about their preferences. My result implies that the fully-robust commitment payoff bound extends when these interdependent value perturbations are relatively less likely compared to the commitment type, and vice versa.

Second, it points out the limitations of reputation effects in repeated incomplete information games with interdependent values and unrestricted payoffs. According to this view, the modeler is perturbing a repeated game with interdependent values with commitment types. Therefore, every commitment type is arbitrarily unlikely compared to any strategic type. As a result, my conditions fail whenever $k(\alpha^*_1, \theta) > 0$. This motivates the study of games with specific payoff structures in Section 4, which allows one to further explore the robust implications of reputation effects in interdependent value environments.

The proof of Theorem 1 appears in Appendices A, B and Online Appendix A. I make several remarks on the conditions before explaining the proof. First, Theorem 1 left out two degenerate sets of beliefs, which are the boundaries of $\Lambda(\alpha^*_1, \theta)$ and $\Delta(\alpha^*_1, \theta)$. In these knife-edge cases, the attainability of the commitment payoff bound depends on the presence of other mixed strategy commitment types and their correlations with the state. Second, the assumption that $\text{BR}_2(\alpha^*_1, \phi_{\alpha^*_1})$ is a singleton in statements 2 and 4 is satisfied under generic parameter values, and is only required for the proof when $\Lambda(\alpha^*_1, \theta) = \{\varnothing\}$, which is used to rule out pathological cases where $a^*_2 \in \text{BR}_2(\alpha^*_1, \phi_{\alpha^*_1})$ but $\{a^*_2\} \neq \text{BR}_2(\alpha^*_1, \phi_{\alpha^*_1})$. An example on this issue is presented in Appendix B. Third, according to the separating hyperplane theorem, the requirement that $\alpha^*_1 \notin \text{co}(\Omega^m \setminus \{\alpha^*_1\})$ guarantees the existence of a payoff function $u_1(\theta, \cdot, \cdot)$ under which type $\theta$’s commitment payoff from any alternative commitment action in $\Omega^m$ is strictly below $v_\theta(\alpha^*_1)$. This convex independence assumption cannot be dispensed, as no restrictions are made on $\mu(\Omega^m \setminus \{\alpha^*_1\})$ and $\{\phi_{\alpha_1}\}_{\alpha_1 \neq \alpha^*_1}$. Therefore, commitment types other than $\alpha^*_1$ are allowed to occur with arbitrarily high probability and can have arbitrary correlations with the state.

### 3.3 Proof Ideas of Statements 1 & 3

I start with the case in which $\alpha^*_1$ is pure and then move on to those in which $\alpha^*_1$ is mixed.

**Pure Commitment Payoff:** Since $\alpha^*_1$ is pure, $\lambda_t(\tilde{\theta})$ will not increase if player 2 observes $a^*_1$ for every $\tilde{\theta} \in \Theta^b_{a^*_1, \theta}$. Therefore, $\lambda_t(\tilde{\theta}) \leq \lambda(\tilde{\theta})$ for every $t \in \mathbb{N}$ if player 1 imitates the commitment type. By definition, if $\lambda_t \in \Lambda(\alpha^*_1, \theta)$ and $a^*_2$ is not a strict best reply (call period $t$ a bad period), then the strategic types must be playing actions other than $a^*_1$ in period $t$ with probability bounded from below, after which they will be separated from the commitment type. As in Fudenberg and Levine (1989), the number of bad periods is uniformly bounded from above, which implies that player 1 can secure his commitment payoff as $\delta \to 1$. 

13
Mixed Commitment Payoff when \( k(\alpha_1^*, \theta) = 1 \): Let \( \Theta^{b}_{(\alpha_1^*, \theta)} \equiv \{ \tilde{\theta} \} \). Recall from equation (3.5) in subsection 3.1 that when \( \Lambda(\alpha_1^*, \theta) \neq \{ \emptyset \} \), there exists \( \psi^* > 0 \) such that \( \Lambda(\alpha_1^*, \theta) = \{ \lambda | 0 < \lambda < \psi^* \} \). The main difference from the pure commitment action case is that \( \lambda_t \) can increase after player 2 observes some actions in the support of \( \alpha_1^* \). As a result, type \( \theta \) cannot secure his commitment payoff by replicating \( \alpha_1^* \) since he may end up playing actions that are more likely to be played by type \( \tilde{\theta} \), in which case \( \lambda_t \) will exceed \( \psi^* \).

The key step in my proof shows that for every equilibrium strategy of the short-run players, one can construct a non-stationary strategy for the long-run player under which the following three goals are achieved simultaneously: (1) To avoid negative inferences about the state, i.e. \( \lambda_t < \psi^* \) for every \( t \in \mathbb{N} \). (2) In expectation, the short-run players believe that actions within a small neighborhood of \( \alpha_1^* \) will be played for all but a bounded number of periods. (3) Every \( a_1 \in A_1 \) will be played with occupation measure close to \( \alpha_1^*(a_1) \).

To understand why one can make such a construction, note that \( \{ \lambda_t \}_{t \in \mathbb{N}} \) is a non-negative supermartingale conditional on \( \alpha_1^* \). Since \( \lambda_0 < \psi^* \), the probability measure over histories (induced by \( \alpha_1^* \)) in which \( \lambda_t \) never exceeds \( \psi^* \) is bounded from below by the Doob’s Upcrossing Inequality\(^{[3]}\). When \( \delta \) is close to 1, the Lindeberg-Feller Central Limit Theorem (Chung 1974) ensures that the set of player 1’s action paths, in which the discounted time average frequency of every \( a_1 \) being close to \( \alpha_1^*(a_1) \), occurs with probability close to 1 under the measure induced by \( \alpha_1^* \). Each of the previous steps defines a subset of histories, and the intersection between them occurs with probability bounded from below. Then I derive a uniform upper bound on the expected sum of relative entropy between \( \alpha_1^* \) and player 2’s predicted action conditional on only observing histories at the intersection. According to Gossner (2011), the unconditional expected sum is bounded from above by a positive number that does not explode as \( \delta \to 1 \). Given that the intersection between the two sets has probability bounded from below, the Markov Inequality implies that the conditional expected sum is also bounded from above. Therefore, the expected number of periods that player 2’s predicted action is far away from \( \alpha_1^* \) is bounded from above.

**Mixed Commitment Payoff when \( k(\alpha_1^*, \theta) \geq 2 \):** Let \( S_t \equiv \sum_{\tilde{\theta} \in \Theta^{b}_{(\alpha_1^*, \theta)}} \lambda_t(\tilde{\theta})/\psi(\tilde{\theta}) \), which is a non-negative supermartingale conditional on \( \alpha_1^* \). The assumption that \( \lambda \in \Lambda(\alpha_1^*, \theta) \) implies that \( S_0 < 1 \). Doob’s Upcrossing Inequality provides a lower bound on the probability measure over histories under which \( S_t \) is always strictly below 1, i.e. \( \lambda_t \in \Lambda(\alpha_1^*, \theta) \) for every \( t \in \mathbb{N} \). The proof then follows from the \( k(\alpha_1^*, \theta) = 1 \) case.

To illustrate why \( \lambda \in \Lambda(\alpha_1^*, \theta) \) is insufficient when \( k(\alpha_1^*, \theta) \geq 2 \) and \( \alpha_1^* \) is mixed, I present an example in Appendix G.8 where \( \lambda \in \Lambda(\alpha_1^*, \theta) \) but type \( \theta \)’s equilibrium payoff is bounded below his commitment payoff.

\(^{[12]}\)There is a remaining step after this to deal with correlations between action and state, with details shown in Appendix A.2 Part II.

\(^{[13]}\)In private value reputation games with noisy monitoring, Fudenberg and Levine (1992) use the upcrossing inequality to bound the number of bad periods when player 1 imitates the commitment strategy. In contrast, I use the upcrossing inequality to show that player 1 can cherry-pick actions in the support of his mixed commitment strategy in order to prevent \( \lambda(\tilde{\theta}) \) from exceeding \( \psi(\tilde{\theta}) \) while simultaneously making his opponents believe that actions close to \( \alpha_1^* \) will be played in all but a bounded number of periods.
The idea is to construct equilibrium strategies for the bad strategic types, under which playing every action in the support of $\alpha_1^*$ will increase the likelihood ratio along some dimensions. As a result, player 2’s belief in period 1 is bounded away from $\Xi(\alpha_1^*, \theta)$ regardless of the action played in period 0.

### 3.4 Proof Ideas of Statements 2 & 4

To prove statement 2, let $\alpha_1^*$ be the Dirac measure on $a_1^* \in A_1$. Let player 1’s payoff be given by:

$$u_1(\tilde{\theta}, a_1, a_2) = 1\{\tilde{\theta} = \theta, a_1 = a_1^*, a_2 = a_2^*\}. \quad (3.6)$$

I construct an equilibrium in which type $\theta$ obtains a payoff strictly bounded below 1 even when $\delta$ is arbitrarily close to 1. The key idea is to let the bad strategic types pool with the commitment type (with high probability) and the good ones separate from the commitment type. As a result, type $\theta$ cannot simultaneously build a reputation for commitment while separating away from the bad strategic types. However, the proof is complicated by the presence of other commitment types that are playing mixed strategies. To understand this issue, consider an example where $\Theta = \{\theta, \tilde{\theta}\}$ with $\tilde{\theta} \in \Theta^b(\alpha_1^*, \theta)$, $\Omega^m = \{a_1^*, \alpha_1\}$ with $\alpha_1$ non-trivially mixed, attaching positive probability to $a_1^*$ and $\{a_2^*\} = \text{BR}_2(a_1^*, \phi_{a_1^*}) = \text{BR}_2(\alpha_1, \phi_{a_1})$.

The naive construction in which type $\tilde{\theta}$ plays $a_1^*$ all the time does not work, as type $\theta$ can then obtain a payoff arbitrarily close to 1 by playing $a_1 \in \text{supp}(\alpha_1) \setminus \{a_1^*\}$ in period 0 and $a_1^*$ in every subsequent period.

To circumvent this problem, I construct a sequential equilibrium in which type $\theta$’s action is deterministic on the equilibrium path. Type $\tilde{\theta}$ plays $a_1^*$ in every period with probability $p \in (0, 1)$ and plays strategy $\sigma(\alpha_1)$ with probability $1 - p$, with $p$ being large enough that $\lambda_1$ is bounded away from $\Xi(\alpha_1^*, \theta)$ after observing $a_1^*$ in period 0. The strategy $\sigma(\alpha_1)$ is described as follows: At histories that are consistent with type $\theta$’s equilibrium strategy, play $\alpha_1$; at histories that are inconsistent, play a completely mixed action $\tilde{\alpha}_1(\alpha_1)$ which attaches strictly higher probability to $a_1^*$ than to any element in $\Omega^m \setminus \{a_1^*\}$.

To verify incentive compatibility, I keep track of the likelihood ratio between the fraction of type $\tilde{\theta}$ who plays $\sigma(\alpha_1)$ and the commitment type $\alpha_1$. If type $\theta$ has never deviated before, then this ratio remains constant. If type $\theta$ has deviated before, then this ratio increases every time $a_1^*$ is observed. Therefore, once type $\theta$ has deviated from his equilibrium play, he will face a trade-off between obtaining a high stage-game payoff (by playing $a_1^*$) and reducing the likelihood ratio. This uniformly bounds his continuation payoff after any deviation from above, which is strictly below 1. Type $\theta$’s on-path strategy is then constructed such that his payoff is strictly between 1 and his highest post-deviation continuation payoff. This can be achieved, for example, by using a public randomization device that prescribes $a_1^*$ with probability less than 1 in every period.

The proof of statement 4 involves several additional steps, with details shown in Online Appendix A. First,
the payoff function in equation (3.6) is replaced by one that is constructed via the separating hyperplane theorem, such that type $\theta$’s commitment payoff from every other action in $\Omega^m$ is strictly lower than his commitment payoff from playing $\alpha_1^*$. Second, in Online Appendix A.3, I show that there exists an integer $T$ (independent of $\delta$) and a $T$-period strategy for the strategic types other than $\theta$ such that the likelihood ratio vector in period $T$ is bounded away from $\Lambda(\alpha_1^*, \theta)$ regardless of player 1’s behavior in the first $T$ periods. Third, the continuation play after period $T$ modifies the construction in the proof of statement 2. The key step is to construct the bad strategic types’ strategies under which type $\theta$’s continuation payoff after any deviation is bounded below his commitment payoff from playing $\alpha_1^*$. The details are shown in Online Appendices A.5 and A.7.

4 Reputation Effects in Games with Monotone-Supermodular Payoffs

Motivated by the discussions in Section 3, I study stage games where players’ payoffs satisfy a monotone-supermodularity (or MSM) condition. I explore the robust predictions on the long-run player’s equilibrium payoff and on-path equilibrium behavior. All the results in this section apply even when the commitment types are arbitrarily unlikely compared to any strategic type.

4.1 Monotone-Supermodular Payoff Structure

Let $\Theta$, $A_1$ and $A_2$ be finite ordered sets. I use ‘$>$’, ‘$\gtrsim$’, ‘$<$’ and ‘$\lesssim$’ to denote the rankings between pairs of elements. The stage game has MSM payoffs if it satisfies the following pair of assumptions:

Assumption 1 (Monotonicity). $u_1(\theta, a_1, a_2)$ is strictly decreasing in $a_1$ and is strictly increasing in $a_2$.

Assumption 2 (Supermodularity). $u_1(\theta, a_1, a_2)$ has strictly increasing differences (or SID) in $(a_1, \theta)$ and increasing differences (or ID) in $(a_2, \theta)$. $u_2(\theta, a_1, a_2)$ has SID in $(a_1, a_2)$ and $(\theta, a_2)$.

I focus on games where player 2’s decision-making problem is binary, which have been a primary focus of the reputation literature, for example, Kreps and Wilson (1982), Milgrom and Roberts (1982), Mailath and Samuelson (2001), Ekmekci (2011), Liu (2011) and others.

Assumption 3. $|A_2| = 2$.

I will discuss how my assumptions fit into the applications to business transactions (or product choice game) and monopolistic competition (or entry deterrence game) in subsection 4.5.

14First, given Assumption 2, the case in which $u_1(\theta, a_1, a_2)$ is strictly increasing in $a_1$ and strictly decreasing in $a_2$ can be analyzed similarly by reversing the orders of the states and each player’s actions. Second, I only require $u_1$ to have ID in $(a_2, \theta)$ in order to accommodate the classic separable case, in which player 1’s return from player 2’s action does not depend on the state. Assumption 2 can be further relaxed, which can be seen in the conclusion (Assumption 2’) and Online Appendix G.

15The results extend to games with $|A_2| \geq 3$ under extra conditions on $u_1$. The details can be found in Online Appendix D.
Preliminary Analysis: Let $\overline{a}_i \equiv \max A_i$ and $\underline{a}_i \equiv \min A_i$, with $i \in \{1, 2\}$. For every $\pi \in \Delta(\Theta)$ and $\alpha_1 \in \Delta(A_1)$, let

$$D(\pi, \alpha_1) \equiv u_2(\pi, \alpha_1, \overline{a}_2) - u_2(\pi, \alpha_1, \underline{a}_2).$$

(4.1)

I classify the states into good, bad and negative by partitioning $\Theta$ into the following three sets:

$$\Theta_g \equiv \{ \theta | D(\theta, \overline{a}_1) \geq 0 \text{ and } u_1(\theta, \overline{a}_1, \overline{a}_2) > u_1(\theta, \underline{a}_1, \underline{a}_2) \},$$

$$\Theta_p \equiv \{ \theta \not\in \Theta_g | u_1(\theta, \overline{a}_1, \overline{a}_2) > u_1(\theta, \underline{a}_1, \underline{a}_2) \} \text{ and } \Theta_n \equiv \{ \theta | u_1(\theta, \overline{a}_1, \overline{a}_2) \leq u_1(\theta, \underline{a}_1, \underline{a}_2) \}.$$

Intuitively, $\Theta_g$ is the set of good states in which $\overline{a}_2$ is player 2’s best reply to $\overline{a}_1$ and player 1 strictly prefers the commitment outcome $(\overline{a}_1, \overline{a}_2)$ to his minmax outcome $(\underline{a}_1, \underline{a}_2)$. $\Theta_p$ is the set of bad states in which player 2 has no incentive to play $\overline{a}_2$ but player 1 strictly prefers $(\overline{a}_1, \overline{a}_2)$ to his minmax outcome. $\Theta_n$ is the set of negative states in which player 1 prefers his minmax outcome to the commitment outcome. Lemma 4.1 shows that every good state is higher than every bad state, and every bad state is higher than every negative state:

**Lemma 4.1.** If the stage game payoff satisfies Assumption 2 then:

1. For every $\theta_g \in \Theta_g$, $\theta_p \in \Theta_p$ and $\theta_n \in \Theta_n$, $\theta_g > \theta_p$, $\theta_p > \theta_n$ and $\theta_g > \theta_n$.

2. If $\Theta_p, \Theta_n \neq \{\varnothing\}$, then $D(\theta_n, \overline{a}_1) < 0$ for every $\theta_n \in \Theta_n$.

**Proof of Lemma 4.1:** For statement 1, since $D(\theta_g, \overline{a}_1) \geq 0$ and $D(\theta_p, \overline{a}_1) < 0$, SID of $u_2$ with respect to $(\theta, a_2)$ implies that $\theta_g > \theta_p$. Since $u_1(\theta_p, \overline{a}_1, \overline{a}_2) > u_1(\theta_p, \underline{a}_1, \underline{a}_2)$ and $u_1(\theta_n, \overline{a}_1, \overline{a}_2) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2)$, we know that $\theta_p > \theta_n$ due to the SID of $u_1$ in $(\theta, a_1)$ and ID of $u_1$ in $(\theta, a_2)$. If $\Theta_p \neq \{\varnothing\}$, then statement 1 is proved. If $\Theta_p = \{\varnothing\}$, then since $u_1(\theta_p, \overline{a}_1, \overline{a}_2) > u_1(\theta_p, \underline{a}_1, \underline{a}_2)$ and $u_1(\theta_n, \overline{a}_1, \overline{a}_2) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2)$, we have $\theta_g > \theta_n$. For statement 2, if $\Theta_p, \Theta_n \neq \{\varnothing\}$, then $\theta_n < \theta_p$. SID of $u_2$ with respect to $(\theta, a_2)$ implies that $D(\theta_n, \overline{a}_1) < D(\theta_p, \overline{a}_1) < 0$.

4.2 Statement of Results

My results in this section outline the robust implications on player 1’s payoff and behavior when he has the option to build a reputation for playing the highest action. For this purpose, I assume that $\overline{a}_1 \in \Omega^N$ and $D(\phi_{\pi_1}, \overline{a}_1) > 0$, i.e. there exists a commitment type that plays the highest action in every period, and player 2 has a strict incentive to play $\overline{a}_2$ conditional on knowing that she is facing commitment type $\pi_1$.

The qualitative features of equilibria depend on the relative likelihood between the strategic types who know that the state is good (call them good strategic types) and the ones who know that the state is bad (call them bad
strategic types). In particular, player 2’s prior belief is optimistic if:

$$\mu(\overline{a}_1)D(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta_y \cup \Theta_p} \mu(\theta)D(\theta, \overline{a}_1) > 0,$$  \hspace{1cm} (4.2)

and is pessimistic if:

$$\mu(\overline{a}_1)D(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta_y \cup \Theta_p} \mu(\theta)D(\theta, \overline{a}_1) \leq 0.$$  \hspace{1cm} (4.3)

Notice that formulas (4.2) and (4.3) allow type $\overline{a}_1$ to be arbitrarily unlikely compared to every strategic type. In my results, these inequalities can be replaced by

$$\sum_{\theta \in \Theta_y \cup \Theta_p} \mu(\theta)D(\theta, \overline{a}_1) \geq 0$$

and

$$\sum_{\theta \in \Theta_y \cup \Theta_p} \mu(\theta)D(\theta, \overline{a}_1) < 0,$$

respectively, when the total probability of commitment types is small enough. These expressions will be useful once we compare the reputation game to the benchmark game without commitment types.

**Equilibrium Payoff under Optimistic Priors:** The main result in the optimistic prior case is the commitment payoff bound for playing the highest action, which is stated as Theorem 2:

**Theorem 2.** If $\overline{a}_1 \in \Omega^m$, $D(\phi_{\overline{a}_1}, \overline{a}_1) > 0$ and $\mu$ satisfies (4.2), then for every $\theta \in \Theta$, we have:

$$\liminf_{\delta \to 1} V_{\theta}(\delta, \mu, \phi) \geq \max\{u_1(\theta, \overline{a}_1, \overline{a}_2), u_1(\theta, \overline{a}_1, \overline{a}_2)\}.$$

According to Theorem 2, a patient long-run player can overcome the lack-of-commitment problem and guarantee his payoff from $(\overline{a}_1, \overline{a}_2)$ in every state and in every equilibrium. It implies, for example, that a firm can secure high returns by maintaining a reputation for exerting high effort despite his customers’ skepticism about product quality; an incumbent who might have unfavorable information about the market demand curve (say, demand elasticities are low) can guarantee high profits by fighting entrants.

The key distinction between condition (4.2) and the distribution conditions in Theorem 1 is that the good strategic types can contribute to attaining the commitment payoff. As a result, the commitment type is allowed to be arbitrarily unlikely compared to every bad strategic type. Intuitively, this is driven by the following implication of MSM payoff: when player 1 has an incentive to pool with the commitment type $\overline{a}_1$ in a lower state, he will then play $\overline{a}_1$ with probability 1 at every on-path history in a higher state. According to Lemma 4.1, every good state is higher than every bad state. The above property implies that in equilibria where some bad strategic types pool with the commitment type, all the good strategic types will behave like the commitment type on the equilibrium path, which can help to guarantee the commitment payoff.

Nevertheless, there also exist repeated game equilibria in which the commitment strategy is not optimal for any bad strategic type. This undermines the implications of MSM and as a result, the good strategic types will
have a strict incentive not to play the commitment strategy in some of those equilibria. Therefore, reputation effects cannot provide accurate predictions on player 1’s equilibrium behavior in the optimistic prior case. Moreover, there can exist on-path histories at which every bad strategic type plays $\pi_1$ with strictly higher probability than every good strategic type, so player 2’s belief can become more pessimistic after observing $\pi_1$.

In order to establish the commitment payoff bound in those equilibria, I circumvent the aforementioned complications by showing that player 2’s posterior cannot become too pessimistic conditional on $a_1$ is always being played. This implies that in every period where $\pi_2$ is not player 2’s strict best reply, the strategic types must be separating from the commitment type with probability bounded from below. Hence, there can be at most a bounded number of such periods, which validates the commitment payoff bound in those equilibria.

One may also wonder whether player 1 can guarantee a strictly higher payoff by establishing a reputation for playing an alternative commitment action. In Online Appendix F, I adopt a notion of tightness introduced by Cripps, Schmidt and Thomas (1996) and show that when there are bad strategic types, i.e. $\Theta_p \neq \{\emptyset\}$, no type of player 1 can guarantee a strictly higher equilibrium payoff by establishing a reputation for playing another pure commitment action. Furthermore, if $\Theta_p \neq \{\emptyset\}$ and $\Theta_n = \{\emptyset\}$, then player 1 cannot guarantee a strictly higher equilibrium payoff by taking any other (pure or mixed) commitment actions.

**Equilibrium Payoff and Behavior under Pessimistic Priors:** When $\mu$ satisfies condition (4.3), there exists a unique pair of $(\theta_p^*, q(\mu)) \in \Theta_p \times (0,1]$ such that:

$$\mu(\pi_1)D(\phi_{a_1}, \pi_1) + q(\mu)\mu(\theta_p^*)D(\theta, \pi_1) + \sum_{\theta > \theta_p^*} \mu(\theta)D(\theta, \pi_1) = 0. \quad (4.5)$$

Since $\theta_p^* \in \Theta_p$, the definition of $\Theta_p$ and Assumption imply the existence of $r \in (0,1)$ such that:

$$ru_1(\theta_p^*, a_1, a_2) + (1 - r)u_1(\theta_p^*, a_1, a_2) = u_1(\theta_p^*, a_1, a_2). \quad (4.6)$$

Let

$$v^*_\theta = \begin{cases} 
    u_1(\theta, a_1, a_2) & \text{if } \theta \preceq \theta_p^* \\
    ru_1(\theta, a_1, a_2) + (1 - r)u_1(\theta, a_1, a_2) & \text{if } \theta > \theta_p^*.
\end{cases} \quad (4.7)$$

Let $H^\sigma(\theta) (\subset H)$ be the set of histories that occurs with strictly positive probability under strategy profile $\sigma$ conditional on player 1 being strategic type $\theta$. For the sake of exposition, I state the result when $\Omega^m = \{\pi_1\}$, which will be generalized to the case with multiple commitment types in Theorem 3’ (Appendix D.2) under the extra requirement that the total probability of commitment types, $\mu(\Omega^m)$, is small enough:
Theorem 3. If $\Omega^m = \{\overline{a}_1\}$ and $D(\phi_{\pi_1}, \overline{a}_1) > 0$, then for every $\mu$ satisfying equation (4.3), there exists $\delta \in (0, 1)$, such that for every $\delta > \delta$ and $\sigma \equiv (\sigma_\omega)_{\omega \in \Omega}, \sigma_2) \in \text{NE}(\delta, \mu, \phi)$.

1. For every $\theta > \theta_0^*$ and $h^t \in \mathcal{H}^\theta(\theta)$, type $\theta$ plays $\overline{a}_1$ at $h^t$.

For every $\theta < \theta_0^*$ and $h^t \in \mathcal{H}^\theta(\theta)$, type $\theta$ plays $\underline{a}_1$ at $h^t$.

In period 0, type $\theta_0^*$ plays $\overline{a}_1$ with probability $q(\mu)$ and $\underline{a}_1$ with probability $1 - q(\mu)$.

For every $h^t \in \mathcal{H}^\theta(\theta_0^*)$ with $t \geq 1$:

- If $h^t$ contains $\overline{a}_1$, then at $h^t$, type $\theta_0^*$ plays $\overline{a}_1$.

- If $h^t$ contains $\underline{a}_1$, then at $h^t$, type $\theta_0^*$ plays $\underline{a}_1$.

2. If $q(\mu) \neq 1$, then $V^\sigma_\theta(\delta) = v^*_\theta$ for every $\theta \in \Theta$.

According to Theorem 3, every strategic type’s equilibrium payoff and on-path behavior are (generically) the same across all Nash equilibria when player 1 is patient. Furthermore, his behavior and payoff are independent of the discount factor as long as it lies above some threshold $\delta$. On the equilibrium path, every type strictly above $\theta_0^*$ plays $\overline{a}_1$ in every period, every type strictly below $\theta_0^*$ plays $\underline{a}_1$ in every period, type $\theta_0^*$ randomizes (in period 0) between playing $\overline{a}_1$ in every period and playing $\underline{a}_1$ in every period. This suggests that the long-run player will behave consistently over time and maintain his reputation for commitment in every equilibrium. As implied by equation (4.5), player 2 is indifferent between $\overline{a}_2$ and $\underline{a}_2$ starting from period 1 conditional on player 1’s always having played $\overline{a}_1$ in the past. When the cutoff type $\theta_0^*$ plays a non-trivial mixed strategy, i.e. $q(\mu) \neq 1$\(^\text{16}\), his indifference condition in period 0 pins down every strategic type’s equilibrium payoff. In particular, it requires that the occupation measure of $\overline{a}_2$ be $r$ when $\overline{a}_1$ is played in every period and 0 when $\underline{a}_1$ is played in every period, as can be seen from equations (4.6) and (4.7).

Intuitively, the uniqueness of player 1’s on-path behavior is driven by the following disciplinary effect: he can obtain a high payoff by playing $\overline{a}_1$ in every period thanks to the commitment type, but it is impossible for him to receive a high payoff in the continuation game if he has ever failed to do so. To be more precise, I begin with the useful observation that under a pessimistic prior, always playing $\overline{a}_1$ must be optimal for some bad strategic types. Since players’ payoffs are MSM, the above statement implies that all the good strategic types will play $\overline{a}_1$ at every on-path history and in every equilibrium. Therefore, player 2’s belief about $\theta$ deteriorates whenever player 1 fails to play $\overline{a}_1$, after which his continuation payoff will be low. To see why playing the highest action for some time and then switching to a lower action is suboptimal for every strategic type, notice that (1) if his first deviation happened at an optimistic belief, then he could guarantee a strictly higher payoff

\(^{16}\)The condition that $q(\mu) \neq 1$ is satisfied for generic $\mu$. It is also satisfied if we fix the likelihood ratios between the strategic types and focus on cases where the total probability of commitment types is small.
by playing the highest action in every period thanks to the commitment type; (2) if his first deviation occurred after period 0 when player 2’s belief is pessimistic, then he could strictly save the cost of playing $\bar{a}_1$ by playing $\bar{a}_1$ from period 0. The probabilities with which the cutoff type $\theta^*_p$ mixes in period 0 can be uniquely pinned down due to the substitutability between his return from playing $\bar{a}_1$ and the equilibrium probability with which he plays $\bar{a}_1$. In particular, if type $\theta^*_p$ plays $\bar{a}_1$ with higher probability, then it reduces player 2’s incentive to play $\bar{a}_2$ after observing $\bar{a}_1$ and hence reduces type $\theta^*_p$’s return from playing $\bar{a}_1$.

In contrast, player 1 exhibits multiple on-path behaviors in Fudenberg and Levine (1989) and behaving inconsistently is strictly optimal in some sequential equilibria. This is because deviating from the commitment action only signals that player 1 is strategic, but cannot preclude him from obtaining a high payoff in the continuation game according to Fudenberg et al. (1990). As a result, he may have an incentive to separate from the commitment type in any given period, depending on which equilibrium players coordinate on. Similarly in the optimistic prior case, deviating from the commitment action can still lead to an optimistic posterior about the state, after which player 1’s continuation payoff can still be high, leading to multiple equilibrium behaviors.

For an overview of the extension to multiple commitment types (Theorem 3’ in Appendix D.2): If there are only pure strategy commitment types and type $\bar{a}_1$ is the only commitment type under which $\bar{a}_2$ is optimal, then all the conclusions in Theorem 3 apply without any further qualifications. If there are only pure commitment types, but there are commitment types other than $\bar{a}_1$ under which player 2 has a strict incentive to play $\bar{a}_2$, then as long as $\mu(\Omega^m)$ is small enough, player 1’s equilibrium payoff and behavior are the same across all equilibria. Every strategic type’s equilibrium behavior is the same as described in Theorem 3 except for the cutoff type $\theta^*_p$, who can play actions other than $\bar{a}_1$ and $\bar{a}_1$ with positive probability. If there are mixed commitment types and $\mu(\Omega^m)$ is small enough, then there exists a cutoff type $\theta^*_p \in \Theta_p$ (which is the same across all equilibria) such that all types above $\theta^*_p$ play $\bar{a}_1$ all the time, all types below $\theta^*_p$ play $\bar{a}_1$ all the time. Type $\theta^*_p$’s various on-path behaviors in different equilibria will coincide with (ex ante) probability of at least $1 - \epsilon$, and moreover, he will either play $\bar{a}_1$ all the time or $\bar{a}_1$ all the time with probability of at least $1 - \epsilon$, with $\epsilon$ vanishing as $\mu(\Omega^m) \to 0$.

I conclude this subsection by adding two caveats. To begin with, my behavior uniqueness result requires that the long-run player be patient. I show by counterexample in Appendix G.7 that he can have multiple possible equilibrium behaviors when $\delta$ is low. Intuitively, this is because an impatient bad strategic type has no incentive to pay the cost of imitating the commitment type. As a result, the disciplinary effect will disappear.

Next, neither Theorem 3 nor its extension (Theorem 3’) can imply the uniqueness of Nash equilibrium or Nash

---

17 The behavioral uniqueness conclusion will also fail in repeated incomplete information games where the state only affects player 1’s payoff, regardless of how pessimistic the prior belief is. To see this, consider for example that player 1 has persistent private information about his discount factor (Ghosh and Ray 1996) or his cost of taking a higher action (Schmidt 1993b). In these cases, the strategic types who have low discount factors or high costs either have no incentive to pool with the commitment type, in which case the disciplinary effect only works temporarily; or if they play the commitment strategy in equilibrium, then they are equivalent to the commitment type in player 2’s best-response problem, in which case they are no longer ‘bad’.
equilibrium outcome. This is because first, Nash equilibrium places no restriction on players’ behaviors off the equilibrium path. Second, player 2’s behavior on the equilibrium path is not necessarily unique. To see this, assume for example, \( \Omega^m = \{ \bar{a}_1 \} \). Since player 2 is indifferent starting from period 1 conditional on \( \bar{a}_1 \) always being played, her behavior is only restricted by two sets of constraints. The first is, type \( \theta_p^* \)’s indifference condition in period 0. The second constraint is, type \( \theta_p^* \)’s incentives to play \( \bar{a}_1 \) in period \( t \in \mathbb{N} \). The first one only pins down the occupation measure of \( \bar{a}_2 \) conditional on \( \bar{a}_1 \) being played in every period, and the second one only requires that \( \bar{a}_2 \) not be too front-loaded. Under these constraints, there are still multiple ways to allocate the play of \( \bar{a}_2 \) over time, leading to multiple equilibrium outcomes.

### 4.3 Proof Ideas of Theorems 2 and 3

The proofs of Theorems 2, 3 and 3’ can be found in Appendices C and D, and the counterexamples to my results in which each of my assumptions fails are in Appendices G.1, G.2 and G.3.

To recall the challenges, first, since values are interdependent and the commitment types are allowed to be arbitrarily unlikely compared to every strategic type, Theorem 1 suggests that the proofs need to exploit the properties of player 1’s payoff function. Therefore, the standard learning-based arguments in Fudenberg and Levine (1989, 1992), Sorin (1999), Gossner (2011) and others cannot be directly applied.

Second, a repeated supermodular game is not supermodular, as player 1’s action today can affect future equilibrium play. Consequently, the monotone selection result on static supermodular games (see Topkis 1998) is not applicable. Similar issues have been highlighted in complete information extensive form games (Echenique 2004) and 1-shot signalling games (Liu and Pei 2017). For an illustrative example, consider the following 1-shot signalling game where the sender is the row player and the receiver is the column player:

<table>
<thead>
<tr>
<th>( \theta = H )</th>
<th>( l )</th>
<th>( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( U )</td>
<td>4, 8</td>
<td>0, 0</td>
</tr>
<tr>
<td>( D )</td>
<td>2, 4</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \theta = L )</th>
<th>( l )</th>
<th>( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( U )</td>
<td>(-2, -2)</td>
<td>( 2, 0 )</td>
</tr>
<tr>
<td>( D )</td>
<td>0, (-4)</td>
<td>( 5, 1 )</td>
</tr>
</tbody>
</table>

If we rank the states and players’ actions according to \( H \succ L, U \succ D \) and \( l \succ r \), one can verify that both players’ payoffs are strict supermodular functions of the triple \( (\theta, a_1, a_2) \). However, there exists a sequential equilibrium in which the sender plays \( D \) in state \( H \) and \( U \) in state \( L \). The receiver plays \( l \) after she observes \( D \) and \( r \) after she observes \( U \). Therefore, the sender’s equilibrium action can be strictly decreasing in the state, despite all the complementarities between players’ actions and the state.

The game studied in this paper is trickier than 1-shot signalling games, as the sender (or player 1) is repeatedly signalling his private information. The presence of intertemporal incentives provides a rationale for many different behaviors and belief-updating processes that cannot be rationalized in 1-shot interactions.
example, even when the stage game has MSM payoffs, there can still exist equilibria in the repeated signalling game where at some on-path histories, player 1 plays $\tilde{a}_1$ with higher probability in a lower state compared to a higher state. As a result, player 1’s reputation could deteriorate even when he plays the highest action.

4.3.1 Proof Sketch in the Entry Deterrence Game

I illustrate the logic of the proof using the entry deterrence game in the introduction. Recall that players’ stage game payoffs are given by:

<table>
<thead>
<tr>
<th>$\theta = H$</th>
<th>$O$</th>
<th>$E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F$</td>
<td>2, 0</td>
<td>0, −1</td>
</tr>
<tr>
<td>$A$</td>
<td>3, 0</td>
<td>1, 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\theta = L$</th>
<th>$O$</th>
<th>$E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F$</td>
<td>$2 - \eta, 0$</td>
<td>$-\eta, 1$</td>
</tr>
<tr>
<td>$A$</td>
<td>3, 0</td>
<td>1, 2</td>
</tr>
</tbody>
</table>

Let $H \succ L$, $F \succ A$ and $O \succ E$. One can check that Assumptions [1] and [3] are satisfied. I focus on the case where $\eta \in (0, 1)$, which satisfies Assumption [2] and moreover, $L \in \Theta_p$. I make several simplifying assumptions which are relaxed in the Appendix. First, player 2 can only observe player 1’s past actions, i.e. $h_t = \{a_{1,s}\}_{s=0}^{t-1}$. Second, there is only one commitment type, i.e. $\Omega^m \equiv \{F\}$. Third, let $\phi_F$ be the Dirac measure on state $H$.

Two Classes of Equilibria: I classify the set of equilibria into two classes, depending on whether or not playing $F$ in every period is type $L$’s best reply. Formally, let $h_F^t$ be the period $t$ history at which all past actions were $F$. For any given equilibrium $\sigma \equiv (\{\sigma_\omega\}_{\omega \in \Omega}, \sigma_2)$, $\sigma$ is called a regular equilibrium if playing $F$ at every history in $\{h_F^t\}_{t=0}^{\infty}$ is type $L$’s best reply to $\sigma_2$. Otherwise, $\sigma$ is called an irregular equilibrium.

Regular Equilibria: I use a monotone selection result on 1-shot signalling games (Liu and Pei 2017):

- If a 1-shot signalling game has MSM payoffs and the receiver’s action choice is binary, then the sender’s action is non-decreasing in the state in every Nash equilibrium.

This result implies that in the repeated signalling game studied in this section, if playing the highest action in every period is player 1’s best reply in a lower state, then he will play the highest action with probability 1 at every on-path history in a higher state (see Lemma [C.1] for a formal statement). In the context of the entry deterrence game, if an equilibrium is regular, then playing $F$ in every period is type $L$’s best reply. Since $H \succ L$, the result implies that type $H$ will play $F$ with probability 1 at $h_F^t$ for every $t \in \mathbb{N}$.

Irregular Equilibria: I establish two properties of irregular equilibria. First, at every history $h_F^t$ where player 2’s belief attaches higher probability to type $H$ than to type $L$, either $O$ is her strict best reply, or the strategic types will be separated from the commitment type at $h_F^t$ with significant probability. Next, I show that when $\delta$
is large enough, player 2’s posterior belief will attach higher probability to type $H$ than to type $L$ at every $h^t_F$.

Let $q_t$ be the ex ante probability that player 1 is type $L$ and he has played $F$ from period 0 to $t - 1$, and let $p_t$ be the ex ante probability that player 1 is type $H$ and he has played $F$ from period 0 to $t - 1$.

**Claim 1.** For every $t \in \mathbb{N}$, if $p_t \geq q_t$ but $O$ is not a strict best reply at $h^t_F$, then:

\[
(p_t + q_t) - (p_{t+1} + q_{t+1}) \geq \mu(F)/2. \tag{4.8}
\]

**Proof of Claim 1:** Player 2 does not have a strict incentive to play $O$ at $h^t_F$ if and only if: $\mu(F) + p_{t+1} - (p_t - p_{t+1}) - q_{t+1} - 2(q_t - q_{t+1}) \leq 0$, which implies that $\mu(F) + 2p_{t+1} + 2q_{t+1} \leq p_t + 2q_t + q_{t+1} \leq p_t + 3q_t \leq 2p_t + 2q_t$, where the last inequality makes use of the assumption that $p_t \geq q_t$. If we rearrange the terms, the result is inequality (4.8).

**Claim 2.** If $\delta$ is large enough, then in every irregular equilibrium, $p_t \geq q_t$ for all $t \geq 0$.

Claim 2 establishes an important property of irregular equilibria, namely, despite the fact that playing the highest action could lead to negative inferences about the state, player 2’s belief about the strategic types cannot become too pessimistic. Intuitively, this is because type $L$’s continuation payoff must be low if he separates from the commitment type in the last period with a pessimistic belief, while he can guarantee himself a high payoff by continuing to play $F$. This contradicts his incentive to separate in that last period.

**Proof of Claim 2:** Suppose towards a contradiction that $p_t < q_t$ for some $t \in \mathbb{N}$. Given that playing $F$ in every period is not type $L$’s best reply, there exists $T \in \mathbb{N}$ such that type $L$ has a strict incentive to play $A$ at $h^T_F$. That is to say, $p_s \geq q_s = 0$ for every $s > T$. Let $t^* \in \mathbb{N}$ be the largest integer $t$ such that $p_t < q_t$. The definition of $t^*$ implies that (1) player 2’s belief at history $(h^t_F, A)$ attaches probability strictly more than $1/2$ to type $L$, (2) type $L$ is supposed to play $A$ with strictly positive probability at $h^T_F$.

Let us examine type $L$’s incentives at $h^t_F$. If he plays $A$, then his continuation payoff at $(h^t_F, A)$ is 1. This is because player 2’s belief is a martingale, so there exists an action path played with positive probability by type $L$ such that at every history along this path, player 2 attaches probability strictly more than $1/2$ to state $L$, which implies that she has a strict incentive to play $E$, and type $L$’s stage game payoff is at most 1.

If he plays $F$ at $h^t_F$ and in all subsequent periods, then according to Claim 1, there exists at most $\overline{T} \equiv \lceil 2/\mu(F) \rceil$ periods in which $O$ is not player 2’s strict best reply. This is because by definition, $p_s \geq q_s$ for all $s > t^*$. Therefore, type $L$’s guaranteed continuation payoff is close to $2 - \eta$ when $\delta$ is large. This is strictly larger than 1. Comparing his continuation payoffs by playing $A$ versus playing $F$ reveals a contradiction.

---

This is no longer true when player 2 can condition her actions on her predecessors’ actions and the realizations of public randomization devices, in which case it can only imply that type $L$ has a strict incentive to play $A$ at some on-path histories where he has always played $F$ before. These complications will be discussed in Remark II and will be treated formally in Appendix C.
Optimistic Prior Belief: When the prior belief is optimistic, i.e. \( \mu(F) + \mu(H) > \mu(L) \), I establish the commitment payoff theorem for the two classes of equilibria separately. For regular equilibria, since type \( H \) behaves in the same way as the commitment type \( F \), one can directly apply statement 1 of Theorem 1, and obtain the commitment payoff bound for playing \( F \). For irregular equilibria, Claims 1 and 2 imply that conditional on playing \( F \) in every period, there exist at most \( \overline{T} \) periods in which \( O \) is not player 2’s strict best reply. Therefore, type \( H \) can guarantee a payoff close to 2 and type \( L \) can guarantee payoff close to \( 2 - \eta \).

Pessimistic Prior Belief: When the prior belief is pessimistic, i.e. \( \mu(F) + \mu(H) \leq \mu(L) \), we know that \( p_0 = \mu(H) < \mu(L) = q_0 \). According to Claim 2, there is no irregular equilibria. So every equilibrium is regular, and therefore, type \( H \) will play \( F \) with probability 1 at every \( h_F^1 \).

Next, I pin down the probability with which type \( L \) plays \( F \) at every \( h_F^t \). I start by introducing a measure of optimism for player 2’s belief at \( h_F^1 \) by letting

\[
X_t \equiv \mu(F)D(H, F) + p_tD(H, F) + q_tD(L, F).
\]

(4.9)

Note that \( \{X_t\}_{t=0}^\infty \) is a non-decreasing sequence as \( D(H, F) > 0, D(L, F) < 0, p_t \) is constant and \( q_t \) is non-increasing. The pessimistic prior assumption translates into \( X_0 \leq 0 \). The key step is to show that:

Claim 3. If \( \delta \) is large enough, then \( X_t = 0 \) for all \( t \geq 1 \).

Proof of Claim 3: Suppose towards a contradiction that \( X_t < 0 \) for some \( t \geq 1 \), then let us examine type \( L \)’s incentives at \( h_F^t \). Since \( X_t < 0 \), type \( L \) will play \( F \) with positive probability at \( h_F^t \). If he plays \( F \) at \( h_F^t \), then his continuation payoff at \( h_F^t \) is 1. If he plays \( A \) at \( h_F^t \), then his continuation payoff at \( (h_F^t, A) \) is 1, but he can receive a strictly higher stage game payoff in period \( t - 1 \). This leads to a contradiction.

Suppose towards a contradiction that \( X_t > 0 \) for some \( t \geq 1 \), then let \( t^* \) be the smallest \( t \) such that \( X_t > 0 \). Since \( X_s \leq 0 \) for every \( s < t^* \), we know that type \( L \) will play \( A \) with positive probability at \( h_F^{t^*} \). In what follows, I examine type \( L \)’s incentives at \( h_F^{t^*} \). If he plays \( A \), then his continuation payoff at \( (h_F^{t^*}, A) \) is 1. If he plays \( F \) forever, then I will show below that \( O \) is player 2’s strict best reply at \( h_F^{t^*} \) for every \( s \geq t^* \). Once this is shown, we know that type \( L \)’s guaranteed continuation payoff at \( h_F^{t^*} \) is \( 2 - \eta \), which is strictly greater than 1 and leads to a contradiction.

I complete the proof by showing that \( O \) is player 2’s strict best reply at \( h_F^{s} \) for every \( s \geq t^* \). Suppose

---

10When there are other commitment types playing mixed strategies, \( X_t \) is close to albeit not necessarily equal to 0. Nevertheless, the variation of \( X_t \) across different equilibria vanishes as the total probability of commitment types goes to 0. When there are no mixed commitment types under which player 2 has a strict incentive to play \( \overline{a}_2 \), the sequence \( \{X_t\}_{t=0}^\infty \) is generically unique.
towards a contradiction that player 2 does not have a strict incentive to play $O$ at $h_F^t$ for some $s \geq t^*$, then:

$$\mu(F)D(H,F) + p_s D(H,F) + q_{s+1} D(L,F) + (q_s - q_{s+1}) D(L,A) \leq 0, \quad (4.10)$$

$$\Rightarrow q_s - q_{s+1} \geq \frac{X_s}{D(L,F) - D(L,A)} \geq \frac{X_{t^*}}{D(L,F) - D(L,A)} \equiv Y. \quad (4.11)$$

Hence, there exist at most $\lceil q_0/Y \rceil$ such periods, which is a finite number. Let period $\bar{t}$ be the last of such periods. Let us examine type $L$’s incentive at $h_F^{\bar{t}}$. On one hand, he plays $A$ with positive probability at this history in equilibrium, which results in a continuation payoff close to 1. On the other hand, his continuation payoff from playing $F$ in every period is $2 - \eta$, which results in a contradiction.

**Remark I:** When $L \in \Theta_n$, i.e. $\eta \geq 1$, Claim 1 as well as the conclusion on regular equilibria will remain intact. What needs to be modified is Claim 2: despite the fact that $p_t$ can be less than $q_t$ for some $t \in \mathbb{N}$ in some equilibria (think about for example, when the prior attaches very high probability to state $L$ such that $p_0 < q_0$), type $H$ can still guarantee a payoff close to 2 in every equilibrium.

To see this, in every irregular equilibrium where $p_t < q_t$ for some $t$, let $t^*$ be the largest of such $t$ and let us examine type $L$’s incentives in period 0. For this to be an equilibrium, he must prefer playing $F$ from period 0 to $t^* - 1$ and then $A$ in period $t^*$, compared to playing $A$ forever starting from period 0. By adopting the first strategy, his continuation payoff is 1 after period $t^* + 1$, his stage game payoff from period 0 to $t^* - 1$ is no more than 1 if $O$ is played, and is no more than $-\eta$ if $E$ is played. By adopting the second strategy, he can guarantee himself a payoff of at least 1. For the first strategy to be better than the second, the occupation measure with which $E$ is played from period 0 to $t^* - 1$ needs to be arbitrarily close to 0 as $\delta \to 1$. That is to say, if type $H$ plays $F$ in every period, then the discounted average payoff he loses from period 0 to $t^* - 1$ (relative to 2 in each period) vanishes as $\delta \to 1$. According to Claim 1, his guaranteed continuation payoff after period $t^*$ is close to 2. Summing up, his guaranteed payoff in period 0 is at least 2 in the $\delta \to 1$ limit.

**Remark II:** In Appendices C and D, I extend the above idea and provide full proofs to Theorems 2 and 3, which incur two additional complications. First, there can be arbitrarily many strategic types, and in particular, good, bad and negative types could co-exist. Second, player 2’s actions can be conditioned on the past realizations of public randomization devices as well as on her predecessors’ actions, which could open up new equilibrium possibilities and therefore, can potentially undermine the robust predictions on payoff and behavior.

In terms of the proof, the main difference occurs in the analysis of irregular equilibria, as there may not exist a last history at which the probability of the bad strategic types is greater than the probability of the good
strategic types. This is because the predecessor-successor relationship is incomplete on the set of histories where player 1 has always played $a_1$ once $\{a_2,s,\xi_s\}_{s\leq t-1}$ is also included in $h^t$.

My proof overcomes this difficulty by showing that every time a switch from a pessimistic to an optimistic belief happens, the bad strategic types must be separating from the commitment type with ex ante probability bounded from below. This implies that such switches can only happen finitely many times conditional on every positive probability event. On the other hand, the bad strategic types only have incentives to separate at those switching histories when their continuation payoffs from imitating the commitment type are low, which implies that another such switch needs to happen again in the future. This implies that such switches must happen infinitely many times if it happens at least once, leading to a contradiction.

4.4 Implications for Equilibrium Refinement

In this subsection, I revisit a classic application of reputation results by studying how they refine equilibrium payoffs and behaviors in repeated incomplete information games with long-run and short-run players. To do this, I study a benchmark repeated Bayesian game without commitment types, that is, player 1 is strategic with probability 1. I compare player 1’s equilibrium payoff and behavior in the benchmark game and in the reputation game. This comparison addresses the question of which repeated game equilibria are more plausible when the long-run player can build reputations.

Let $\Theta$ be the set of states, which is also the set of player 1’s types in the benchmark game. Let $\hat{\mu} \in \Delta(\Theta)$ be player 2’s prior belief. Let $\mathcal{V}(\delta, \hat{\mu}) \subset \mathbb{R}^{\Theta}$ be player 1’s equilibrium payoff set with $v \equiv (v_\theta)_{\theta \in \Theta}$ a typical element. I start with the optimistic prior case in which:

$$\sum_{\theta \in \Theta_y \cup \Theta_p} \hat{\mu}(\theta) D(\theta, \pi_1) \geq 0. \quad (4.12)$$
To see the relationship between equations (4.12) and (4.2), notice that in the original reputation game where \( \mu(\Omega_m) \) is small, (4.2) implies (4.12). Moreover, when we perturb a benchmark game that satisfies (4.12), the reputation game will satisfy (4.2) given that \( D(\phi_{\pi_1}, a_1) > 0 \). Let

\[
\bar{\delta} \equiv \max_{a_2 \in \Delta(A_2)} \left\{ \frac{u_1(\bar{\vartheta}, a_1, a_2) - u_1(\bar{\vartheta}, a_1, a_2)}{u_1(\bar{\vartheta}, a_1, a_2) - u_1(\bar{\vartheta}, a_1, a_2) + u_1(\bar{\vartheta}, a_1, a_2) - u_1(\bar{\vartheta}, a_1, a_2)} \right\},
\]

the result is stated as Proposition 4.1:

**Proposition 4.1.** If \( a_1 \) is player 1’s pure Stackelberg action in state \( \bar{\vartheta} \), then for every \( v \in \mathcal{V}(\delta, \hat{\mu}) \), we have \( v_{\bar{\vartheta}} \leq u_1(\bar{\vartheta}, a_1, a_2) \). Furthermore, if \( \hat{\mu} \) satisfies (4.12) and \( \delta \geq \bar{\delta} \), then

\[
\sup_{v \in \mathcal{V}(\delta, \hat{\mu})} v_{\theta} \in \left[ (1 - \delta)u_1(\bar{\vartheta}, a_1, a_2) + \delta u_1(\bar{\vartheta}, a_1, a_2), u_1(\bar{\vartheta}, a_1, a_2) \right].
\]

The proof can be found in Appendix E. To summarize the role of reputation in refining equilibrium payoffs, first, it rules out equilibria with bad payoffs (for example, those with payoff \( u_1(\theta, a_1, a_2) \)) and selects equilibria that deliver every strategic type \( \theta \) a payoff no less than his highest equilibrium payoff in a complete information repeated game where \( \theta \) is common knowledge (Fudenberg, Kreps and Maskin 1990).

Second, according to Proposition 4.1, reputation effects select the highest equilibrium payoff for type \( \bar{\vartheta} \) in the benchmark incomplete information game. However, types lower than \( \bar{\vartheta} \) can obtain payoff strictly higher than \( u_1(\theta, a_1, a_2) \) in the benchmark game even in the \( \delta \to 1 \) limit. Figure 2 in page 27 depicts player 1’s limiting equilibrium payoff set in the entry deterrence game, with more details coming in Online Appendix C.

Next, I analyze the pessimistic prior case in which

\[
\sum_{\theta \in \Theta_1 \cup \Theta_p} \hat{\mu}(\theta) D(\theta, a_1) < 0. \tag{4.13}
\]

The above inequality translates equation (4.3) into the benchmark game without commitment types, given that \( \mu(\Omega_m) \) is small enough. Recall the definition of \( v^*_\theta \) in expression (4.7), I state the result as Proposition 4.2:

**Proposition 4.2.** For every \( \hat{\mu} \) satisfying (4.13), there exists \( \hat{\delta} \in (0, 1) \) such that for every \( \delta > \hat{\delta} \) and \( \theta \in \Theta \), we have:

\[
\sup_{v \in \mathcal{V}(\delta, \hat{\mu})} v_{\theta} = v^*_\theta. \tag{4.14}
\]

The proof is detailed in Appendix F. Proposition 4.2 implies that reputation effects select the highest equilibrium payoff for every strategic type in the benchmark incomplete information game. Moreover, since the unique equilibrium play for player 1 in Theorem 3 constitutes an equilibrium in the benchmark game as well,
reputation effects also lead to the selection of a unique on-path behavior for the long-run player.

Unlike Proposition 4.1, Proposition 4.2 does not require $\pi_1$ to be player 1’s pure Stackelberg action in state $\theta$. This is because under a pessimistic prior belief, playing actions other than $\pi_1$ cannot induce player 2 to play $\pi_2$, while under an optimistic prior belief, such possibilities cannot be ruled out unless we assume that $\pi_1$ is type $\theta$’s pure Stackelberg action. Moreover, the equilibrium selection result applies to all strategic types under a pessimistic belief but only applies to the highest type under an optimistic belief, as the pessimistic prior belief condition implies tight upper bounds on every bad strategic type’s equilibrium payoff.

4.5 Related Applications

I discuss two applications of reputation models as well as how they fit into my MSM assumptions: the product choice game which highlights the lack-of-commitment problem in business transactions (Mailath and Samuelson 2001, Liu 2011, Ekmekci 2011) and the entry deterrence game that studies predatory pricing behaviors in monopolistic competition (Kreps and Wilson 1982, Milgrom and Roberts 1982).

Limit Pricing & Predation with Unknown Price Elasticities: Player 1 is an incumbent choosing between a low price (interpreted as limit pricing or predation) and a normal price, every player 2 is an entrant choosing between out and enter. The incumbent has private information about the demand elasticities $\theta \in \mathbb{R}_+$, which measures the increase in his product’s demand when he lowers the price. The payoff matrix is given by:

<table>
<thead>
<tr>
<th>State is $\theta$</th>
<th>Out</th>
<th>Enter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Price</td>
<td>$p_L(Q_M + \theta), 0$</td>
<td>$p_L(Q_D + \gamma \theta), \Pi_L(\theta) - f$</td>
</tr>
<tr>
<td>Normal Price</td>
<td>$p_N Q_M, 0$</td>
<td>$p_N Q_D, \Pi_N - f$</td>
</tr>
</tbody>
</table>

where $p_L$ and $p_N$ are the low and normal prices, $f$ is the sunk cost of entry, $Q_M$ and $Q_D$ are the incumbent’s monopoly and duopoly demands under a normal price, $\Pi_L$ and $\Pi_N$ are the entrant’s profits when the incumbent’s price is low and normal, $\gamma \in (0, 1)$ is a parameter measuring the effect of price elasticity on the incumbent’s demand in duopoly markets relative to monopoly markets. This parameter is less than 1 as the entrant captures part of the market, which offsets some of the demand increase (of the incumbent’s product) from a price cut.

In this example, Assumptions 1 and 2 require that (1) setting a low price is costly for the incumbent and he strictly prefers the entrant to stay out; (2) the entrant’s profit from entering the market is lower when the incumbent sets a low price and when the demand elasticity is higher; (3) it is less costly for the incumbent to set a low price when the demand elasticity is higher. The first and third requirements are natural. The second one is reasonable, since lowering prices leaves the entrant a smaller market share, and this effect is more pronounced when the demand elasticity is higher.
Among other entry deterrence games, my assumptions also apply when the entrant faces uncertainty about the market size or the elasticity of substitution between her product and the incumbent’s. It is also valid when the incumbent uses non-pricing strategies to deter entry, such as choosing the intensity of advertising in the pharmaceutical industry where advertising has positive spillovers to the entrant’s product (Ellison and Ellison 2011). However, my supermodularity assumption fails in the entry deterrence model of Harrington (1986), in which the incumbent’s and the entrant’s production costs are positively correlated and the entrant does not know her own production costs before entering the market.

Product Choice Games: Consider an example of a software firm (player 1) and a sequence of clients (player 2). Every client chooses between the custom software ($C$) and the standardized software ($S$). In response to his client’s request, the firm either exerts high effort ($H$) which can ensure a timely delivery and reduce the cost overruns, or exerts low effort ($L$). A client’s willingness to pay depends not only on the delivery time and the expected cost overruns, but also on the quality of the software, which can be either good ($G$) or bad ($B$), and is the firm’s private information. Here, quality is interpreted as the hidden running risks, the software’s adaptability to future generations of operation systems, etc. Therefore, compared to delivery time and cost overruns, quality is much harder to observe directly, so it is reasonable to assume that future clients learn about quality mainly through the firm’s past behaviors. This is modeled as the following product choice game:

<table>
<thead>
<tr>
<th>$\theta =$ Good</th>
<th>Custom</th>
<th>Standardized</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Effort</td>
<td>1, 3</td>
<td>$-1, 2$</td>
</tr>
<tr>
<td>Low Effort</td>
<td>2, 0</td>
<td>0, 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\theta =$ Bad</th>
<th>Custom</th>
<th>Standardized</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Effort</td>
<td>$1 - \eta, 0$</td>
<td>$-1 - \eta, 1$</td>
</tr>
<tr>
<td>Low Effort</td>
<td>$2, -2$</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

MSM requires that (1) exerting high effort is costly for the firm but it can result in more profit when the client purchases the custom software; (2) clients are more inclined to buy the custom software if it can be delivered on-time and its quality is high; (3) firms that produce higher quality software face lower effort costs. The first and second requirements are natural. The third one is reasonable since both the cost of making timely deliveries and the software’s quality are positively correlated with the talent of the firm’s employees. Indeed, Banerjee and Duflo (2000) provide empirical evidence in the Indian software industry, showing that firms enhance their reputations for competence via making timely deliveries and reducing cost overruns.

5 Concluding Remarks

A central theme of my analysis is that reputation building can be challenging when the uninformed players’ learning is confounded. Even though the informed player can convince his opponents about his future actions, he may still fail to teach them how to best reply since their payoffs depend on the state. Similar in spirit
is a contemporary work of Deb and Ishii (2017), which revisits the commitment payoff theorem when the uninformed players face uncertainty about the monitoring structure\textsuperscript{20}. Their paper is complementary to mine, with the main difference being: the state can be identified through some exogenous public signals in their model (see their Assumption 2.3), while it can only be learned through the informed player’s actions in my model. Under their identification assumption, they show that the strategic long-run player cannot guarantee his Stackelberg payoff when there are only stationary commitment types. They also construct a countably-infinite set of non-stationary commitment types, under which the informed player can guarantee his Stackelberg payoff. Their informational assumption fits in environments where informative signals about the state arrive frequently (or, in every period), as for example, when the state is the performance of vehicles, mobile phones, etc.

In contrast, my informational assumption fits into applications where exogenous signals are unlikely to arrive for a long time, or the variations of their realizations are mostly driven by noise orthogonal to the state. For example, when the state is the adaptability of a software to future generations of operating systems, the resilience of an architectural design to earthquakes, the long-run health impact of a certain type of food, the demand elasticity in markets with high sunk costs, the effectiveness of advertising in the NBA Finals, the amount of connection traffic in the hub of a major airline, and the like. My assumption has been adopted in many repeated Bayesian game models with interdependent values, such as Hart (1985), Aumann and Maschler (1995), Hörner and Lovo (2009), Kaya (2009), Hörner et al. (2011), Roddie (2012), Pęski and Toikka (2017).

My work is also related to the papers on bad reputation, for example, Ely and Välimäki (2003), Ely, Fudenberg and Levine (2008). These papers study a class of private value reputation games with imperfect monitoring known as participation games. They show that a patient long-run player’s equilibrium payoff is low when the bad commitment types (i.e. ones that commit to play actions which discourage the short-run players from participating) are relatively more likely compared to the Stackelberg commitment type.

Although both my Theorem\textsuperscript{1} and their results underscore the possibilities of reputation failure, the economic forces behind them are very different. In their models, reputations are bad due to the tension between the long-run player’s forward-looking incentives and the short-run players’ participation incentives. In particular, a patient long-run player has a strong incentive to take actions that can generate good signals but harm the participating short-run players. This discourages the short-run players from participating, which prevents the long-run player from signalling and leads to a bad reputation. In contrast, reputation failure occurs in my model as the short-run players’ learning is confounded. This is because the long-run player’s actions signal the payoff-relevant state. In different equilibria, these signals are interpreted in different ways, which will affect
the short-run players’ best reply to the commitment action. If the bad strategic types are believed to be pooling with the commitment type with high probability, then the strategic long-run player cannot simultaneously build a reputation for commitment while separating from the bad strategic types.

**Extensions:** I conclude by discussing several extensions of my results. First, players move sequentially in the stage game rather than simultaneously in some applications, such as a firm that chooses its service standards after consumers decide which product to purchase; an incumbent sets prices before or after observing the entrant’s entering decision. My results are robust when the long-run player moves first. When the short-run players move first, my results are valid when every commitment type’s strategy is independent of the short-run players’ actions. This requirement is not redundant, as the short-run players cannot learn the long-run player’s reaction following an unchosen $a_2$.

Along this line, my analysis can be applied to the following repeated bargaining problem, which models conflict resolution between employers and employees, firms and clients and other contexts. In every period, a long-run player bargains with a short-run player. The short-run player makes a take-it-or-leave-it offer, which is either soft or tough, and the long-run player either accepts the offer or chooses to resolve the dispute via arbitration. The long-run player has persistent private information about both parties’ payoffs from arbitration, which can be interpreted as the quality of his supporting evidence. The short-run players observe the long-run player’s bargaining postures in the previous periods and update their beliefs about their payoffs from arbitration.

In this context, my results provide accurate predictions on the long-run player’s payoff and characterize his unique Nash equilibrium behavior when his (ex ante) expected payoff from arbitration is below a cutoff.

In some other applications where the uninformed players move first, the informed player cannot take actions at certain information sets. For example, the firm cannot exert effort when its client refuses to purchase, the incumbent cannot fight if the entrant stays out. My results in Section 4 apply to these scenarios as long as the informed long-run player can make an action choice in period $t$ if $a_{2,t} \neq \pi_2$. This condition allows for entry deterrence games but rules out participation games defined in Ely, Fudenberg and Levine (2008).

Second, my results are robust against the presence of non-stationary commitment types given that (1) all the non-stationary commitment types are pure, (2) different commitment strategies behave differently on the equilibrium path. When there are non-stationary commitment types playing mixed strategies, the attainability of the commitment payoff bound also depends on the probabilities of these non-stationary commitment types and their correlations with the state. To see this, consider the entry deterrence game. If there is a commitment

---

21Lee and Liu (2013) study a similar game without commitment types, but the short-run players observe their realized payoffs in addition to the long-run player’s past actions. Their model applies to litigation, where the court’s decisions are publicly available. My model applies to arbitration, as arbitration hearings are usually confidential and the final decisions are not publicly accessible.
type who plays $F$ in every period and another one who plays strategy $\tilde{\sigma}_1$, which is defined as:

$$
\tilde{\sigma}_1(h^t) \equiv \begin{cases} 
\frac{1}{2}F + \frac{1}{2}A & \text{if } t = 0 \\
F & \text{otherwise.}
\end{cases}
$$

Conditional on commitment type $\tilde{\sigma}_1$, state $L$ occurs with certainty. If the probability of type $\tilde{\sigma}_1$ is three times larger than that of type $F$, then Theorems 2 and 3 fail. This is because player 2 has no incentive to play $O$ even conditional on the event that $F$ will be played in every future period and player 1 is committed.

Third, Assumption 2 can be replaced by the following weaker condition as players’ incentives remain unchanged under affine transformations on player 1’s state contingent payoffs.

- **Assumption 2’**: There exists $f : \Theta \to (0, +\infty)$ such that $\tilde{u}_1(\theta, a_1, a_2) \equiv f(\theta)u_1(\theta, a_1, a_2)$ has SID in $(a_1, \theta)$ and ID in $(a_2, \theta)$. $u_2$ has SID in $(\theta, a_2)$ and $(a_1, a_2)$.

To see how this generalization expands the applicability of Theorems 2 and 3 consider for example a repeated prisoner’s dilemma game between a patient long-run player (player 1) and a sequence of short-run players (player 2s) in which players are reciprocal altruistic. As in Levine (1998), every player maximizes a weighted average of his monetary payoff and his opponent’s monetary payoff, with the weight on his opponent be a strictly increasing function of his belief about his opponent’s level of altruism. This can be applied to a number of situations in development economics, for example, a foreign firm, NGO or missionary (player 1) trying to cooperate with different local villagers (player 2s) in different periods. When player 1’s level of altruism is his private information, this game violates Assumption 2 as his cost from playing a higher action (cooperate) and his benefit from player 2’s higher action (cooperate) are both decreasing with his level of altruism. I show in Online Appendix G that the game satisfies Assumption 2’ under an open set of parameters. I also provide a full characterization of Assumption 2’ based on the primitives of the model.

**References**


33


A Proof of Theorem 1, Statements 1 & 3

I use \( a^*_1 \in A_1 \) to denote the Dirac measure on \( a^*_1 \), so when \( a^*_1 \) is pure, I will replace \( a^*_1 \) with \( a^*_1 \). Recall that \( \text{BR}_2(a^*_1, \theta) = \{ a^*_2 \} \) (or \( \text{BR}_2(a^*_1, \theta) = \{ a^*_2 \} \)). Since \( \Lambda(a^*_1, \theta) = \{ \emptyset \} \) (or \( \Lambda(a^*_1, \theta) = \{ \emptyset \} \)) if \( \text{BR}_2(a^*_1, \phi_{a^*_1}) \neq \{ a^*_2 \} \) (or \( \text{BR}_2(a^*_1, \phi_{a^*_1}) \neq \{ a^*_2 \} \)), in which case statement 1 (or statement 3) is void. Therefore, it is without loss of generality to assume that \( \text{BR}_2(a^*_1, \phi_{a^*_1}) = \{ a^*_2 \} \) (or \( \text{BR}_2(a^*_1, \phi_{a^*_1}) = \{ a^*_2 \} \)).

A.1 Proof of Statement 1

When \( \Omega^m = \{ a^*_1 \} \) and \( \lambda \in \Lambda(a^*_1, \theta) \), for every \( \mu \) with \( \tilde{\mu}(\tilde{\theta}) \in [0, \mu(\tilde{\theta})] \) for all \( \tilde{\theta} \in \Theta \), we have:

\[
\{ a^*_2 \} = \arg \max_{a^*_2 \in A_2} \left\{ \mu(a^*_1)u_2(\phi_{a^*_1}, a^*_1, a_2) + \sum_{\tilde{\theta} \in \Theta} \tilde{\mu}(\tilde{\theta})u_2(\tilde{\theta}, a^*_1, a_2) \right\}.
\]

Let \( h^*_t \) be the period \( t \) public history such that \( a^*_1 \) is always played. Let \( q_t(\omega) \) be the (ex ante) probability that the history is \( h^*_t \) and player 1’s type is \( \omega \in \Omega \). By definition, \( q_t(a^*_1) = \mu(a^*_1) \) for all \( t \). Player 2’s maximization problem at \( h^*_t \) is:

\[
\text{max}_{a^*_2 \in A_2} \left\{ \mu(a^*_1)u_2(\phi_{a^*_1}, a^*_1, a_2) + \sum_{\tilde{\theta} \in \Theta} q_{t+1}(\tilde{\theta})u_2(\tilde{\theta}, a^*_1, a_2) \right\}
\]

where \( \alpha_{1,t}(\tilde{\theta}) \in \Delta(A_1 \setminus \{ a^*_1 \}) \) is the distribution of type \( \tilde{\theta} \)’s action at \( h^*_t \) conditional on it is not \( a^*_1 \).

Fixing \( \mu(a^*_1) \) and given the fact that \( \lambda \in \Lambda(a^*_1, \theta) \), there exists \( \rho > 0 \) such that \( a^*_2 \) is player 2’s strict best reply if

\[
\sum_{\tilde{\theta} \in \Theta} q_{t+1}(\tilde{\theta}) > \sum_{\tilde{\theta} \in \Theta} q_t(\tilde{\theta}) - \rho.
\]

Let \( T = \lceil 1/\rho \rceil \), which is independent of \( \delta \). There exist at most \( T \) periods in which \( a^*_2 \) fails to be a strict best reply conditional on \( a^*_1 \) has always been played. Therefore, type \( \theta \)’s payoff is bounded from below by:

\[
(1 - \delta^T) \min_{a^*_1 \in A} u_1(\theta, a) + \delta^Tu_2(a^*_1),
\]

which converges to \( v_{\theta}(a^*_1) \) as \( \delta \rightarrow 1 \).

When there are other commitment types, let \( \bar{p} = \max_{a^*_1 \in A^m \setminus \{ a^*_1 \}} \alpha_1(a^*_1) \), which is strictly below 1. There exists \( T \in \mathbb{N} \), such that for every \( t \geq T \), \( a^*_2 \) is player 2’s strict best reply at \( h^*_t \) if:

\[
\sum_{\tilde{\theta} \in \Theta} q_{t+1}(\tilde{\theta}) \geq \sum_{\tilde{\theta} \in \Theta} q_t(\tilde{\theta}) - \rho/2.
\]

Consider the subgame starting from history \( h^*_T \), we obtain the commitment payoff bound.

A.2 Proof of Statement 3

Notation: For every \( \alpha_1 \in \Omega^m \setminus \{ a^*_1 \} \), \( \theta \in \Theta \) and \( \tilde{\mu} \in \Delta(\Omega) \) with \( \tilde{\mu}(a^*_1) \neq 0 \), let

\[
\tilde{\lambda}(\alpha_1) \equiv \tilde{\mu}(\alpha_1)/\tilde{\mu}(a^*_1) \quad \text{and} \quad \tilde{\lambda}(\theta) \equiv \tilde{\mu}(\theta)/\tilde{\mu}(a^*_1)
\]

Abusing notation, let \( \tilde{\lambda} = \left( (\tilde{\lambda}(\theta))_{\alpha_1 \in \Omega^m \setminus \{ a^*_1 \}}, (\tilde{\lambda}(\theta))_{\theta \in \Theta} \right) \) be the (expanded) likelihood ratio vector. Let \( n \equiv |A_1| \) and \( m \equiv |\Omega| - 1 \). For convenience, I write \( \Omega \setminus \{ a^*_1 \} \equiv \{ \omega_1, ..., \omega_m \} \) and \( \tilde{\lambda} = (\tilde{\lambda}_1, ..., \tilde{\lambda}_m) \). The proof is divided into two parts.
A.2.1 Part I

Let $\Sigma_2$ be the set of player 2's strategies with $\sigma_2$ a typical element. Let $\text{NE}_2(\mu, \phi) \subset \Sigma_2$ be the set of player 2's Nash equilibrium strategies for some $\delta \in (0, 1)$, i.e.

$$\text{NE}_2(\mu, \phi) = \left\{ \sigma_2 \mid \exists \delta \in (0, 1) \text{ such that } (\sigma_1, \sigma_2) \in \text{NE}(\delta, \mu, \phi) \right\}.$$  

For every $\sigma_\omega : \mathcal{H} \rightarrow \Delta(A_1)$ and player 2’s strategy $\sigma_2$, let $\mathcal{P}^{(\sigma_\omega, \sigma_2)}$ be the probability measure over $\mathcal{H}$ induced by $(\sigma_\omega, \sigma_2)$, let $\mathcal{H}^{(\sigma_\omega, \sigma_2)}$ be the set of histories that occur with positive probability under $\mathcal{P}^{(\sigma_\omega, \sigma_2)}$ and let $\mathbb{E}^{(\sigma_\omega, \sigma_2)}$ be its expectation operator. Abusing notation, I use $\alpha_1^*$ to denote the strategy of always playing $\alpha_1^*$.

For every $\psi \equiv (\psi_1, ... \psi_m) \in \mathbb{R}_+^m$ and $\chi \geq 0$, let

$$\Delta(\psi, \chi) \equiv \{ \hat{\lambda} \mid \sum_{i=1}^m \hat{\lambda}_i/\psi_i = \chi \}$$

Let $\lambda$ be the likelihood ratio vector induced by player 2’s prior belief $\mu$. Let $\lambda(h^t) \equiv (\lambda_1(h^t), ..., \lambda_m(h^t))$ be the likelihood ratio vector following history $h^t$. For every infinite history $h^\infty$, let $h^\infty_\omega$ be its projection on $a_{1,t}$. Let $\alpha_1(\cdot|h^t)$ be player 2’s conditional expectation over player 1’s next period action at history $h^t$. I show the following Proposition:

**Proposition A.1.** For every $\chi > 0$, $\lambda \in \Delta(\psi, \chi)$, $\sigma_2 \in \text{NE}_2(\mu, \phi)$ and $\epsilon > 0$, there exist $\delta \in (0, 1)$ and $T \in \mathbb{N}$ such that for every $\delta > \delta$, there exists $\sigma_\omega : \mathcal{H} \rightarrow \Delta(A_1)$ that satisfies:

1. $\lambda(h^t) \in \bigcup_{\chi \in [0, \chi+\epsilon]} \Delta(\psi, \chi)$ for every $h^t \in \mathcal{H}^{(\sigma_\omega, \sigma_2)}$.

2. For every $h^\infty \in \mathcal{H}^{(\sigma_\omega, \sigma_2)}$ and every $a_1 \in A_1$,

$$\left| \sum_{t=0}^{\infty} (1 - \delta)^t \mathbb{1}\{h^\infty_t = a_1\} - \alpha_1^*(a_1) \right| < \frac{\epsilon}{2(2\chi + \epsilon)} \quad \text{(A.1)}$$

3. 

$$\mathbb{E}^{(\sigma_\omega, \sigma_2)} \left[ \# \left\{ t \mid d(\alpha_1^*(\cdot|h^t)) > \epsilon^2/2 \right\} \right] < T \quad \text{(A.2)}$$

Intuitively, Proposition [A.1] demonstrates the existence of a strategy for player 1 (for every equilibrium strategy of player 2) such that the following three goals can be achieved simultaneously: first, inducing favorable beliefs about the state; second, the occupation measure of actions is closely matched to $\alpha_1^*$; and third, the expected number of periods in which player 2’s believed action differs significantly from $\alpha_1^*$ is uniformly bounded from above by an integer independent of $\delta$. My proof follows three steps, which is the same as the description in subsection 4.3.

**Step 1:** Let $A_1^* \equiv \text{supp}(\alpha_1^*)$. Recall that $\mathcal{P}^{(\alpha_1^*, \sigma_2)}$ is the probability measure over $\mathcal{H}$ induced by the commitment type that always plays $\alpha_1^*$.

Let $\chi(h^t) \equiv \sum_{i=1}^m \lambda_i(h^t)/\psi_i$. Since $\lambda \in \Delta(\psi, \chi)$, we have $\chi(h^0) = \chi$. Using the observation that $\{\lambda_i(h^t), \mathcal{P}^{(\alpha_1^*, \sigma_2)}, \mathcal{F}_t\}_{t \in \mathbb{N}}$ is a non-negative supermartingale for every $i \in \{1, 2, ..., m\}$, where $\mathcal{F}_t \equiv \mathcal{F}_{\tau_1} \vee \mathcal{F}_{\tau_2} \vee \mathcal{F}_{\tau_3}$ is the filtration induced by the public history$^{22}$ we know that $\{\chi_t, \mathcal{P}^{(\alpha_1^*, \sigma_2)}, \mathcal{F}_t\}_{t \in \mathbb{N}}$ is also a non-negative supermartingale. For every $a < b$, let $U(a, b)$ be the number of upcrossings from $a$ to $b$. According to the Doob’s Upcrossing Inequality (Chung 1974),

$$\mathcal{P}^{(\alpha_1^*, \sigma_2)} \left\{ U(\chi, \chi + \frac{\epsilon}{2}) \geq 1 \right\} \leq \frac{2\chi}{2\chi + \epsilon}. \quad \text{(A.3)}$$

$^{22}$When $\alpha_1^*$ has full support, $\{\lambda_i(h^t), \mathcal{P}^{(\alpha_1^*, \sigma_2)}, \mathcal{F}_t\}_{t \in \mathbb{N}}$ is a martingale. However, when $A_1^* \neq A_1$ and type $\omega_i$ plays action $a_1^* \notin A_1^*$ with positive probability, then the expected value of $\lambda_i(h^t)$ can strictly reduce.
Let \( \mathcal{H}^\infty \) be the set of infinite histories that \( \chi(h^t) \) is always below \( \chi + \frac{\epsilon}{2} \). According to \((A.3)\), it occurs with probability at least \( \frac{\epsilon}{2X+\epsilon} \).

**Step 2:** In this step, I show that for large enough \( \delta \), there exists a subset of \( \mathcal{H}^\infty \), which occurs with probability bounded from below by a positive number, such that the occupation measure over \( A_1 \) induced by every history in this set is \( \epsilon \)-close to \( \alpha_1^* \). For every \( a_1 \in A_1^* \), let \( \{X_t\} \) be a sequence of i.i.d. random variables such that:

\[
X_t = \begin{cases} 
1 & \text{when } a_{1,t} = a_1 \\
0 & \text{otherwise} .
\end{cases}
\]

Under \( \mathcal{D}^{(\alpha_1^*,\sigma_2)} \), \( X_t = 1 \) with probability \( \alpha_1^*(a_1) \), so \( X_t \) has mean \( \alpha_1^*(a_1) \) and variance \( \sigma^2 \equiv \alpha_1^*(a_1)(1 - \alpha_1^*(a_1)) \). Recall that \( n = |A_1| \). I start with the following Lemma:

**Lemma A.1.** For any \( \varepsilon > 0 \), there exists \( \tilde{\delta} \in (0, 1) \), such that for all \( \delta \in (\tilde{\delta}, 1) \),

\[
\limsup_{\delta \to 1} \mathcal{D}^{(\alpha_1^*,\sigma_2)} \left( \left| \sum_{t=0}^{\infty} (1 - \delta)^t X_t - \alpha_1^*(a_1) \right| \geq \varepsilon \right) \leq \frac{\varepsilon}{n} . \tag{A.4}
\]

**Proof of Lemma A.1:** For every \( n \in \mathbb{N} \), let \( \hat{X}_n \equiv \delta^n(X_n - \alpha_1^*(a_1)) \). Define a triangular sequence of random variables \( \{X_{k,n}\}_{0 \leq n \leq k, k,n \in \mathbb{N}} \), such that \( X_{k,n} \equiv \xi_k \hat{X}_n \), where

\[
\xi_k \equiv \sqrt{\frac{1}{\sigma^2} \frac{1}{1 - \delta^{2k}}} .
\]

Let \( Z_k \equiv \sum_{n=1}^{k} X_{k,n} = \xi_k \sum_{k=1}^{n} \hat{X}_n \). By the Lindeberg-Feller Central Limit Theorem (Chung 1974), \( Z_k \) converges in law to \( \mathcal{N}(0, 1) \). By construction,

\[
\frac{\sum_{n=1}^{k} \hat{X}_n}{1 + \delta + \ldots + \delta^{k-1}} = \sigma \sqrt{\frac{1}{1 - \delta^{2k}} \frac{1 - \delta}{1 - \delta^2}} Z_k ,
\]

the RHS of this expression converges (in distribution) to a normal distribution with mean 0 and variance

\[
\sigma^2 \frac{1}{1 - \delta^2} \frac{(1 - \delta)^2}{(1 - \delta^2)^2} .
\]

The variance term converges to \( O\left((1 - \delta)\right) \) as \( k \to \infty \). Using Theorem 7.4.1 in Chung (1974), we have:

\[
\sup_{x \in \mathbb{R}} |F_k(x) - \Phi(x)| \leq A_0 \sum_{n=1}^{k} |X_{k,n}|^3 \sim A_1 (1 - \delta)^{2} ,
\]

where \( A_0 \) and \( A_1 \) are constants, \( F_k \) is the empirical distribution of \( Z_k \) and \( \Phi(\cdot) \) is the standard normal distribution. Both the variance and the approximation error goes to 0 as \( \delta \to 1 \).

Using the properties of normal distribution, we know that for every \( \varepsilon > 0 \), there exists \( \tilde{\delta} \in (0, 1) \) such that for every \( \delta > \tilde{\delta} \), there exists \( K \in \mathbb{N} \), such that for all \( k > K \),

\[
\mathcal{D}^{(\alpha_1^*,\sigma_2)} \left( \left| \sum_{i=1}^{k} \hat{X}_n \right| \geq \varepsilon \right) < \frac{\varepsilon}{n} .
\]

Taking the \( k \to +\infty \) limit, one can obtain the conclusion in Lemma \([A.1]\) \( \square \).
Step 3: According to Lemma [A.1], for every \( a_1 \in A_1 \) and \( \epsilon > 0 \), there exists \( \delta \in (0, 1) \), such that for all \( \delta > \delta \), there exists \( H_{\epsilon,a_1}(\delta) \subset \mathcal{H}^\infty \), such that

\[
P(\alpha_1^*, \sigma_2)(H_{\epsilon,a_1}(\delta)) \geq 1 - \epsilon / n, \tag{A.5}
\]

and for every \( h^\infty \in H_{\epsilon,a_1}(\delta) \), the occupation measure of \( a_1 \) is \( \epsilon \)-close to \( \alpha_1^*(a_1) \). Let \( H_\epsilon(\delta) \equiv \bigcap_{a_1 \in A_1} H_{\epsilon,a_1}(\delta) \). According to \( (A.5) \), we have:

\[
P(\alpha_1^*, \sigma_2)(H_\epsilon(\delta)) \geq 1 - \epsilon. \tag{A.6}
\]

Take \( \epsilon \equiv \frac{\epsilon}{2(2\chi + \epsilon)} \) and let

\[
\hat{H}^\infty \equiv \hat{H}^\infty \cap H_\epsilon(\delta), \tag{A.7}
\]

we have:

\[
P(\alpha_1^*, \sigma_2)(\hat{H}^\infty) \geq \frac{\epsilon}{2(2\chi + \epsilon)} \tag{A.8}
\]

According to Gossner (2011), we have

\[
\mathbb{E}(\alpha_1^*, \sigma_2) \left[ \sum_{\tau=0}^{+\infty} d(\alpha^*||\alpha(\cdot|h^\tau)) \right] \leq - \log \mu(\alpha_1^*). \tag{A.9}
\]

The Markov Inequality implies that:

\[
\mathbb{E}(\alpha_1^*, \sigma_2) \left[ \sum_{\tau=0}^{+\infty} d(\alpha^*||\alpha(\cdot|h^\tau)) \right] \bigg|_{\hat{H}^\infty} \leq - \frac{2(2\chi + \epsilon) \log \mu(\alpha_1^*)}{\epsilon} \tag{A.10}
\]

Let \( P^* \) be the probability measure over \( \mathcal{H}^\infty \) such that for every \( \mathcal{H}_{0}^\infty \subset \mathcal{H}^\infty \),

\[
P^*(\mathcal{H}_{0}^\infty) \equiv \frac{P(\alpha_1^*, \sigma_2)(\mathcal{H}_{0}^\infty \cap \hat{H}^\infty)}{P(\alpha_1^*, \sigma_2)(\hat{H}^\infty)}. \tag{A.11}
\]

Let \( \sigma_{\omega} : \mathcal{H} \rightarrow \Delta(A_1) \) be player 1’s strategy that induces \( P^* \). The expected number of periods in which \( d(\alpha_1^*||\alpha(\cdot|h^\tau)) > \epsilon^2 / 2 \) is bounded from above by:

\[
T \equiv \left[ - \frac{4(2\chi + \epsilon) \log \mu(\alpha_1^*)}{\epsilon^3} \right], \tag{A.12}
\]

which is an integer independent of \( \delta \). The three steps together establish Proposition [A.1].

A.2.2 Part II

Proposition [A.1] and \( \lambda \in \Lambda(\alpha_1^*, \theta) \) do not imply that player 1 can guarantee himself his commitment payoff. This is because player 2 may not have an incentive to play \( a_2^* \) despite \( \lambda \in \Lambda(\alpha_1^*, \theta) \) and the average action is close to \( \alpha_1^* \). My proof overcomes this challenge using two observations, which are the two steps of my proof.

1. If \( \lambda \in \Lambda(\alpha_1^*, \theta) \), is small in all but at most one entry and player 1’s average action is close to \( \alpha_1^* \), then player 2 has a strict incentive to play \( a_2^* \) regardless of the correlation. Let \( \Lambda^0 \) be the set of beliefs that has the above feature.

2. If \( \lambda \in \Lambda(\alpha_1^*, \theta) \) and player 1’s average action is close to \( \alpha_1^* \) but player 2 does not have a strict incentive to play \( a_2^* \), then different types of player 1’s actions must be sufficiently different. This implies that there is significant learning about player 1’s type after observing his action.

I show that for every \( \lambda \in \Lambda(\alpha_1^*, \theta) \), there exists an integer \( T \) and a strategy such that if player 1 picks his action according to this strategy in periods with the above feature, then after at most \( T \) such periods, player 2’s belief about his type will be in \( \Lambda^0 \), which concludes the proof.
Recall that $m \equiv |\Omega|^n + |\Theta| - 1$. Let $\psi \equiv \{\psi_i\}_{i=1}^m \in \mathbb{R}^m_+$ be defined as:

- If $\omega_i \in \Theta^b(\alpha^*_1, \theta)$, then $\psi_i$ equals to the intercept of $\Lambda(\alpha^*_1, \theta)$ on dimension $\omega_i$.
- Otherwise, $\psi_i > 0$ is chosen to be large enough such that
  \[\sum_{i=1}^m \lambda_i / \psi_i < 1.\] (A.12)

Such $\psi$ exists given that $\lambda \in \Lambda(\alpha^*_1, \theta)$. Let $\overline{\psi} \equiv \max\{\psi_i | i = 1, 2, \ldots, m\}$. Recall that Part I has established the existence of a strategy for player 1 under which:

- Player 2’s belief always satisfies (A.12), or more precisely, bounded from above by some $\chi < 1$.
- The occupation measure over $A_1$ at every on-path infinite history is $\epsilon$-close to $\alpha^*_1$.
- In expectation, there exists at most $T$ periods in which player 2’s believed action differs significantly from $\alpha^*_1$, where $T$ is independent of $\delta$.

**Step 1:** For every $\xi > 0$, a likelihood ratio vector $\lambda$ is of *size $\xi$* if there exists $\tilde{\psi} \equiv (\tilde{\psi}_1, \ldots, \tilde{\psi}_m) \in \mathbb{R}^m_+$ such that: $\tilde{\psi}_i \in (0, \psi_i)$ for all $i$ and

\[\lambda \in \{\tilde{\lambda} \in \mathbb{R}^m_+ | \sum_{i=1}^m \tilde{\lambda}_i / \tilde{\psi}_i < 1\} \subset \{\tilde{\lambda} \in \mathbb{R}^m_+ | \#\{i | \tilde{\lambda}_i \leq \xi\} \geq m - 1\}.\] (A.13)

Intuitively, $\lambda$ is of size $\xi$ if there exists a downward sloping hyperplane such that all likelihood ratio vectors below this hyperplane have at least $m - 1$ entries that are no larger than $\xi$. Therefore, for every $\xi > \xi' > 0$, if $\lambda$ is of size $\xi'$, then it is also of size $\xi$. Proposition A.2 establishes the commitment payoff bound when $\lambda$ is of size $\xi$ for $\xi$ small enough.

**Proposition A.2.** There exists $\xi > 0$, such that if $\lambda$ is of size $\xi$, then

\[\liminf_{\delta \to 1} V_\delta(\mu, \delta, \phi) \geq u_1(\theta, \alpha^*_1, a^*_2).\]

In the proof, I show that using the strategy constructed in Proposition A.1 we can ensure that $a^*_2$ is player 2’s strict best reply at every $h^t$ where $d(\alpha^*_1||\alpha_1(\cdot | h^t)) < \epsilon^2/2$. This implies Proposition A.2

**Proof of Proposition A.2:** Let $\alpha_1(\cdot | h^t, \omega_i) \in \Delta(A_1)$ be the equilibrium action of type $\omega_i$ at history $h^t$. Let

\[B_{i, \alpha_1}(h^t) \equiv \lambda_i(h^t)\left(\alpha^*_1(a_1) - \alpha_1(a_1 | h^t, \omega_i)\right).\] (A.14)

Recall that

\[\alpha_1(\cdot | h^t) \equiv \frac{\alpha^*_1 + \sum_{i=1}^m \lambda_i(h^t) \alpha_1(\cdot | h^t, \omega_i)}{1 + \sum_{i=1}^m \lambda_i(h^t)}\]

is the average action anticipated by player 2. For any $\lambda \in \Lambda(\alpha^*_1, \theta)$ and $\epsilon > 0$, there exists $\epsilon > 0$ such that at every likelihood ratio vector $\tilde{\lambda}$ satisfying:

\[\sum_{i=1}^m \tilde{\lambda}_i / \tilde{\psi}_i < \frac{1}{2} \left(1 + \sum_{i=1}^m \lambda_i / \psi_i\right),\] (A.15)

$a^*_2$ is player 2’s strict best reply to every $\{\alpha_1(\cdot | h^t, \omega_i)\}_{i=1}^m$ satisfying the following two conditions.
• $|B_{i,a_1}(h^t)| < \varepsilon$ for all $i$ and $a_1$.

• $\|\alpha^*_i - \alpha_1(\cdot|h^t)\| \leq \varepsilon$.

This is because when the prior belief satisfies (A.15), $\alpha^*_2$ is player 2’s strict best reply when all types of player 1 are playing $\alpha^*_1$. When $\varepsilon$ and $\varepsilon$ are both small enough, an $\varepsilon$-deviation of the average action together with an $\varepsilon$ correlation between types and actions cannot overturn this strictness.

According to the Pinsker’s Inequality, $\|\alpha^*_1 - \alpha_1(\cdot|h^t)\| \leq \varepsilon$ is implied by $d(\alpha^*_1||\alpha_1(\cdot|h^t)) \leq \varepsilon^2/2$. Pick $\varepsilon$ and $\xi$ small enough such that:

$$\varepsilon < \frac{\varepsilon}{2(1 + \bar{\psi})}$$

(A.16)

and

$$\xi < \frac{\varepsilon}{(m - 1)(1 + \varepsilon)}.$$  

(A.17)

Suppose $\lambda_i(h^t) \leq \xi$ for all $i \geq 2$, since $\|\alpha^*_1 - \alpha_1(\cdot|h^t)\| \leq \varepsilon$, we have:

$$\frac{\|\lambda_1(\alpha^*_1 - \alpha_1(a_1|h^t, \omega_1)) + \sum_{i=2}^{m} \lambda_i(\alpha^*_i - \alpha_1(a_1|h^t, \omega_i))\|}{1 + \lambda_1 + \xi(m - 1)} \leq \varepsilon.$$  

The triangular inequality implies that:

$$\|\lambda_1(\alpha^*_1 - \alpha_1(a_1|h^t, \omega_1))\| \leq \sum_{i=2}^{m} \|\lambda_i(\alpha^*_i - \alpha_1(a_1|h^t, \omega_i))\| + \varepsilon (1 + \lambda_1 + \xi(m - 1)) \leq \xi(m - 1) + \varepsilon (1 + \bar{\psi} + \xi(m - 1)) \leq \varepsilon.$$  

(A.18)

where the last inequality uses (A.16) and (A.17). Inequality (A.18) implies that $|B_{1,a_1}(h^t)| \leq \varepsilon$, and therefore, when $\lambda$ is of size $\xi$, $a^*_2$ is player 2’s strict best reply at every history where $d(\alpha^*_1||\alpha_1(\cdot|h^t)) \leq \varepsilon^2/2$. This further implies that the commitment payoff bound is guaranteed.

**Step 2:** In this step, I use Proposition A.2 to show that the mixed commitment payoff is guaranteed for every $\lambda$ satisfying (A.12). Recall the definition of $B_{i,a_1}(h^t)$ in (A.14). According to Bayes Rule, if $a_1 \in A^*_i$ is observed at $h^t$, then

$$\lambda_i(h^t) - \lambda_i(h^t, a_1) = \frac{B_{i,a_1}(h^t)}{\alpha^*_i(a_1)} \text{ and } \sum_{a_1 \in A^*_i} \alpha^*_i(a_1) \left(\lambda_i(h^t) - \lambda_i(h^t, a_1)\right) \geq 0.$$  

Let

$$D(h^t, a_1) \equiv \left(\lambda_i(h^t) - \lambda_i(h^t, a_1)\right)_{i=1}^{m} \in \mathbb{R}^m.$$  

Suppose $B_{i,a_1}(h^t) \geq \varepsilon$ for some $i$ and $a_1 \in A^*_i$, then $\|D(h^t, a_1)\| \geq \varepsilon$ where $\|\cdot\|$ denotes the $L^2$-norm. Pick $\xi > 0$ small enough to meet the requirement in Proposition A.2. Define two sequences of sets, $\{\Lambda^k\}_{k=0}^{\infty}$ and $\{\tilde{\Lambda}^k\}_{k=1}^{\infty}$, which satisfy $\Lambda^k, \tilde{\Lambda}^k \subset \Lambda(\alpha^*_1, \theta)$ for all $k \in \mathbb{N}$, recursively as follows:

• Let $\Lambda^0$ be the set of likelihood ratio vectors that are of size $\xi$,

• For every $k \geq 1$, let $\tilde{\Lambda}^k$ be the set of likelihood ratio vectors in $\Lambda(\alpha^*_1, \theta)$ such that if $\lambda(h^t) \in \tilde{\Lambda}^k$, then either $\lambda(h^t) \in \Lambda^{k-1}$ or, for every $\{\alpha_1(\cdot|h^t, \omega_i)\}_{i=1}^{m}$ such that $\|D(h^t, a_1)\| \geq \varepsilon$ for some $a_1 \in A^*_i$, there exists $a^*_1 \in A^*_i$ such that $\lambda(h^t, a^*_1) \in \Lambda^{k-1}$.
Let $\Lambda^k$ be the set of likelihood ratio vectors in $\Lambda(\alpha^*_1, \theta)$ such that for every $\tilde{\lambda} \in \Lambda^k$, there exists $\tilde{\psi} \equiv (\tilde{\psi}_1, ..., \tilde{\psi}_m) \in \mathbb{R}^m_+$ such that: $\tilde{\psi}_i \in (0, \psi_i)$ for all $i$ and

$$\lambda \in \left\{ \tilde{\lambda} \in \mathbb{R}^m_+ \left| \sum_{i=1}^m \tilde{\lambda}_i/\tilde{\psi}_i < 1 \right\} \subset \left( \bigcup_{j=0}^{k-1} \Lambda^j \right) \bigcup \hat{\Lambda}^k.$$  \hspace{1cm} (A.19)

By construction, we know that:

$$\left\{ \tilde{\lambda} \in \mathbb{R}^m_+ \left| \sum_{i=1}^m \tilde{\lambda}_i/\tilde{\psi}_i < 1 \right\} \subset \bigcup_{j=0}^k \Lambda^j = \Lambda^k.$$  \hspace{1cm} (A.20)

Since $(0, ..., \psi_i - v, ..., 0) \in \Lambda^0$ for any $i \in \{1, 2, ..., m\}$ and $v > 0$, so $\co(\Lambda^0) = \Lambda(\alpha^*_1, \theta)$. By definition, $\{\Lambda^k\}_{k \in \mathbb{N}}$ is an increasing sequence with $\Lambda^k \subset \Lambda(\alpha^*_1, \theta) = \co(\Lambda^k)$ for any $k \in \mathbb{N}$, i.e. it is bounded from above by a compact set. Therefore $\lim_{k \to \infty} \bigcup_{j=0}^k \Lambda^j \equiv \Lambda^\infty$ exists and is a subset of $\co\left( \Lambda(\alpha^*_1, \theta) \right)$. The next Lemma shows that $\co(\Lambda^\infty)$ coincides with $\co(\Lambda(\alpha^*_1, \theta))$.

**Lemma A.2.** $\co(\Lambda^\infty) = \co(\Lambda(\alpha^*_1, \theta))$

**Proof of Lemma A.2:** Since $\Lambda^k \subset \Lambda(\alpha^*_1, \theta)$ for every $k \in \mathbb{N}$, it is obvious that $\co(\Lambda^\infty) \subset \co(\Lambda(\alpha^*_1, \theta))$. Suppose towards a contradiction that

$$\co(\Lambda^\infty) \subset \co(\Lambda(\alpha^*_1, \theta)).$$  \hspace{1cm} (A.21)

Let $\hat{\Lambda} \subset \Lambda(\alpha^*_1, \theta)$ be such that if $\lambda(h^t) \in \hat{\Lambda}$, then either $\lambda(h^t) \not\in \Lambda^\infty$ or:

- For every $\{\lambda(1 | h^t, \omega_i)\}_{i=1}^m$ such that $||D(h^t, a_1)|| \geq \varepsilon$ for some $a_1 \in A^*_1$, there exists $a^*_1 \in A^*_1$ such that $\lambda(h^t, a^*_1) \in \infty$.

- Let $\hat{\Lambda}$ be the set of likelihood ratio vectors in $\Lambda(\alpha^*_1, \theta)$ such that for every $\tilde{\lambda} \in \hat{\Lambda}$, there exists $\tilde{\psi} \equiv (\tilde{\psi}_1, ..., \tilde{\psi}_m) \in \mathbb{R}^m_+$ such that: $\tilde{\psi}_i \in (0, \psi_i)$ for all $i$ and

$$\lambda \in \left\{ \tilde{\lambda} \in \mathbb{R}^m_+ \left| \sum_{i=1}^m \tilde{\lambda}_i/\tilde{\psi}_i < 1 \right\} \subset \left( \Lambda^\infty \bigcup \hat{\Lambda} \right).$$  \hspace{1cm} (A.22)

Since $\Lambda^\infty$ is defined as the limit of the above operator, so in order for (A.21) to be true, it has to be the case that $\hat{\Lambda} = \Lambda^\infty$, or $\Xi \cap \hat{\Lambda} = \{\emptyset\}$ where

$$\Xi \equiv \co(\Lambda(\alpha^*_1, \theta)) \setminus \co(\Lambda^\infty).$$  \hspace{1cm} (A.23)

One can check that $\Xi$ is convex and has non-empty interior. For every $\rho > 0$, there exists $x \in \Xi$, $\theta \in (0, \pi/2)$ and a halfspace: $H(\chi) \equiv \left\{ \tilde{\lambda} \left| \sum_{i=1}^m \tilde{\lambda}_i/\chi_i \leq \chi \right\} \right.$ with $\phi > 0$ satisfying:

1. $\sum_{i=1}^m x_i/\psi_i = \chi$.
2. $\partial B(x, r) \cap H(\chi) \cap \Lambda(\alpha^*_1, \theta) \subset \Lambda^\infty$ for every $r \geq \rho$.
3. For every $r \geq \rho$ and $y \in \partial B(x, r) \cap \Lambda(\alpha^*_1, \theta)$, either $y \in \Lambda^\infty$ or $d(y, H(\chi)) > r \sin \theta$, where $d(\cdot, \cdot)$ denotes the Hausdorff distance.
The second and third property used the non-convexity of \( \clo(\Lambda^\infty) \). Suppose \( \lambda(h^t) = x \) for some \( h^t \) and there exists \( a_1 \in A_1^* \) such that \( \|D(h^t, a_1)\| \geq \epsilon 
\)

- Either \( \lambda(h^t, a_1) \in \Lambda^\infty \), in which case \( x \in \bar{\Lambda} \) but \( x \in \Xi \), leading to a contradiction.

- Or \( \lambda(h^t, a_1) \notin \Lambda^\infty \). Requirement 3 implies that \( d(\lambda(h^t, a_1), H(\chi)) > \epsilon \sin \theta \). On the other hand,

\[
\sum_{a'_1 \in A_1^*} \alpha_1^*(a'_1) \lambda_i(h^t, a'_1) \leq \lambda_i(h^t)
\]  

(A.24)

for every \( i \). Requirement 1 then implies that \( \sum_{a'_1 \in A_1^*} \alpha_1^*(a'_1) \lambda_i(h^t, a'_1) \in H(\chi) \), which is to say:

\[
\sum_{a'_1 \in A_1^*} \alpha_1^*(a'_1) \sum_{i=1}^m \lambda_i(h^t, a'_1)/\psi_i \leq \chi.
\]  

(A.25)

According to Requirement 2, \( \lambda(h^t, a_1) \notin H(\chi) \), i.e. \( \sum_{i=1}^m \lambda_i(h^t, a_1)/\psi_i > \chi + \epsilon \kappa \) for some constant \( \kappa > 0 \). Take

\[
\rho \equiv \frac{1}{2} \min_{a_1 \in A_1^*} \{\alpha_1^*(a_1)\} \epsilon \kappa,
\]

(A.24) implies the existence of \( a_1^* \in A_1^* \setminus \{a_1\} \) such that \( \lambda(h^t, a_1^*) \in H(\chi) \cap B(x, \rho) \). Requirement 2 then implies that \( x = \lambda(h^t) \in \bar{\Lambda} \). Since \( x \in \Xi \), this leads to a contradiction.

Therefore, (A.21) cannot be true, which validates the conclusion of Lemma A.2.

Lemma A.2 implies that for every \( \lambda \in \bar{\Delta}(a_1^*, \theta) \), there exists an integer \( K \in \mathbb{N} \) independent of \( \delta \) such that \( \lambda \in \Lambda^K \). Statement 3 can then be shown by induction on \( K \). According to Proposition A.2, the statement holds when \( K = 0 \). Suppose the statement applies to every \( K \leq K^* - 1 \), let us consider the case when \( K = K^* \). According to the construction of \( \Lambda^K \), there exists a strategy for player 1 such that whenever \( a_2^* \) is not player 2’s best reply despite \( d(a_1^*, \cdot | h^t) < \epsilon^2/2 \), then the posterior belief after observing \( a_{1,1} \) is in \( \bar{\Lambda}^{K^*-1} \), under which the commitment payoff bound is attained by the induction hypothesis.

## B Proof of Theorem 1, Statement 2

In this Appendix, I prove statement 2. The proof of statement 4 involves some additional technical complication, which is relegated to Online Appendix B. The key intuition behind the distinction of pure and mixed commitment strategies in the construction of low payoff equilibria is summarized in Proposition B.3 and Proposition B.6 in Online Appendix B.

In this section, I replace \( \alpha_1^* \) with \( a_1^* \). Let \( \Pi(a_1^*, \theta) \), \( \Pi(a_1^*, \theta) \) and \( \bar{\Pi}(a_1^*, \theta) \) be the exteriors of \( \bar{\Lambda}(a_1^*, \theta) \), \( \Lambda(a_1^*, \theta) \) and \( \bar{\Lambda}(a_1^*, \theta) \), respectively. I start with the following Lemma, which clarifies the restriction that \( \text{BR}_2(a_1^*, \phi_{a_1^*}) \) being a singleton.

**Lemma B.1.** For every \( \lambda \in \Pi(a_1^*, \theta) \), there exist \( 0 < \lambda' < \lambda \) and \( a_2' \neq a_2^* \) such that \( \lambda' \in \Pi(a_1^*, \theta) \) and

\[
\sum_{\tilde{\theta} \in \Theta_{a_1^*, \theta}} \lambda'(...) > 0
\]  

(B.1)

if either one of the following three conditions hold:

1. \( \Lambda(a_1^*, \theta) \neq \emptyset \).

2. \( \Lambda(a_1^*, \theta) = \{\emptyset\} \) and \( \text{BR}_2(a_1^*, \phi_{a_1^*}) \) is a singleton.
3. \( \Lambda(a_1^*, \theta) = \{ \varnothing \} \) and \( a_2^* \notin \text{BR}_2(a_1^*, \phi_{a_1}^*) \).

**Proof of Lemma B.1:** When \( \Lambda(a_1^*, \theta) \neq \{ \varnothing \} \), by definition of \( \Pi(a_1^*, \theta) \), there exists \( 0 < \lambda' < \lambda \) and \( a_2' \neq a_2^* \) such that:

\[
\left( u_2(\phi_{a_1}^*, a_1^*, a_2') - u_2(\phi_{a_1}^*, a_1^*, a_2^*) \right) + \sum_{\tilde{\theta} \in \Theta_{(a_1^*, \theta)}} \lambda'(\tilde{\theta}) \left( u_2(\tilde{\theta}, a_1^*, a_2') - u_2(\tilde{\theta}, a_1^*, a_2^*) \right) > 0. \tag{B.2}
\]

But \( \Lambda(a_1^*, \theta) \neq \{ \varnothing \} \) implies that \( a_2^* = \text{BR}_2(a_1^*, \phi_{a_1}^*) \), so (B.2) implies (B.1).

When \( \Lambda(a_1^*, \theta) = \{ \varnothing \} \), if \( \text{BR}_2(a_1^*, \phi_{a_1}^*) \) is a singleton, then \( \text{BR}_2(a_1^*, \phi_{a_1}^*) \neq \{ a_2^* \} \). Therefore, under condition 2 or 3, \( a_2' \notin \text{BR}_2(a_1^*, \phi_{a_1}^*) \), which implies the existence of \( \theta' \neq \theta \) and \( a_2' \neq a_2^* \) such that \( u_2(\theta', a_1^*, a_2') > u_2(\theta', a_1^*, a_2^*) \). By definition, \( \theta' \in \Theta_{(a_1^*, \theta)} \). Let

\[
\lambda'(\tilde{\theta}) = \begin{cases} 
\frac{\lambda(\tilde{\theta})}{\lambda(\tilde{\theta})} & \text{if } \tilde{\theta} = \theta' \\
0 & \text{otherwise}
\end{cases}
\]

\( \lambda' \) satisfies (B.1) since \( \mu(\omega) > 0 \) for every \( \omega \in \Omega \).

**Remark:** Lemma B.1 leaves out the case in which \( \Lambda(a_1^*, \theta) = \{ \varnothing \} \) and \( a_2^* \in \text{BR}_2(a_1^*, \phi_{a_1}^*) \). In this pathological case, whether player 1 can guarantee his commitment payoff or not depends on the presence of other commitment types. For example, when \( \Theta = \{ \theta, \theta' \} \), \( A_1 = \{ a_1^*, a_1^1 \} \), \( A_2 = \{ a_2^*, a_2^2 \} \) and \( \Theta' = \{ a_1^*, (1-\varepsilon) a_1^1 + \varepsilon a_1^2 \} \) with \( \phi_{a_1}^1(\theta) = 1 \) and \( \phi_{a_1}^2(\theta') = 1 \). Suppose \( \{ a_2^* \} = \text{BR}_2(a_1^*, \theta) = \text{BR}_2(a_1^*, \theta'') \) and \( \{ a_2^*, a_2'' \} = \text{BR}_2(a_1^*, \theta') \). Then type \( \theta \) can guarantee himself payoff \( u_1(\theta, a_1^*, a_2^*) \) by always playing \( a_1^* \) even though \( \lambda \in \Pi(a_1^*, \theta) \) since \( a_1^* \) is always player 2’s strictly best reply given the presence of commitment type playing \( (1-\varepsilon) a_1^1 + \varepsilon a_1^2 \).

**Overview of Two Phase Construction:** Let player 1’s payoff function be:

\[
u_0(a_1^*) = 1. \text{ The sequential equilibrium I construct has a ‘normal phase’ and an ‘abnormal phase’. Type } \theta \text{'s equilibrium action is pure at every history occurring with positive probability under } (\sigma_\theta, \sigma_2). \text{ Play starts from the normal phase and remains in it as long as all past actions equal to type } \theta \text{’s equilibrium actions. Otherwise, play switches to the abnormal phase and stays there forever.}

Let \( A_1 = \{ a_1^1, ..., a_1^{n-1} \}. \) I show there exists a constant \( q \in (0, 1) \) (independent of \( \delta \)) such that:

- After a bounded number of periods (uniform for all \( \delta \)), type \( \theta \) obtains expected payoff \( 1 - q \) in every period in the normal phase, i.e. his payoff is approximately \( 1 - q \) when \( \delta \rightarrow 1 \).

- Type \( \theta \)’s continuation payoff is bounded below \( 1 - 2q \) in the beginning of the abnormal phase.

**Strategies in the Normal Phase:** Let \( \Theta(a_1^*, \theta) = \Theta_{(a_1^*, \theta)} \), which are the set of good strategic types.

- ‘Mechanical’ Strategic Types: Every strategic type in \( \Theta(a_1^*, \theta) \setminus \{ \theta \} \) plays \( \alpha_1 \in \Theta(a_1^*, \theta) \setminus \{ a_1^* \} \) forever, with \( \alpha_1 \) being arbitrarily chosen.\(^{21}\) For every strategic type \( \tilde{\theta} \in \Theta_{(a_1^*, \theta)} \), he plays \( \alpha_1 \) forever with probability \( x(\tilde{\theta}) \in [0, 1] \) such that conditional on player 2 knowing that

\(^{21}\)If \( \Omega^m = \{ a_1^* \} \), then all types in \( \Theta(a_1^*, \theta) \setminus \{ \theta \} \) play some arbitrarily chosen \( a_1' \neq a_1^* \).
the likelihood ratio vector induced by her belief equals to $\lambda'$, with $\lambda'$ being defined in Lemma B.1.

In what follows, I treat the strategic types that are always playing $\alpha_1$ as the commitment type that is playing $\alpha_1$. Formally, let

$$\tilde{\Omega}^m = \begin{cases} \{a_1^*\} & \text{if } |\Omega^m| = 1 \\ \Omega^m \setminus \{a_1^*\} & \text{otherwise} \end{cases} .$$

Let $l \equiv |\tilde{\Omega}^m|$. By construction, we have $l \geq 1$. Let $\phi_{\alpha_1} \in \Delta(\Theta)$ be the adjusted distribution conditional on player 1 being either commitment type $\alpha_1$ or strategic type $\theta \in \Theta(a_1^*, \emptyset) \setminus \{\emptyset\}$ that always plays $\alpha_1$.

- **Other Bad Strategic Types:** Conditional on not always playing $\alpha_1$, type $\tilde{\theta} \in \Theta^b_{(a_1^*, \emptyset)}$ plays $a_1^*$ forever with probability $p \in [0, 1)$, with $p$ chosen such that there exists $a_2' \neq a_2^*$ with

$$u_2(\phi_{a_1^*}, a_1^*, a_2') + \tilde{p} \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*, \emptyset)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2') > u_2(\phi_{a_1^*}, a_1^*, a_2^*) + \tilde{p} \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*, \emptyset)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2^*) \quad (B.4)$$

for all $\tilde{p} \in [p, 1]$. According to the construction of $\lambda'$, Lemma B.1 also implies that

$$\sum_{\tilde{\theta} \in \Theta^b_{(a_1^*, \emptyset)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2') > \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*, \emptyset)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2^*) . \quad (B.5)$$

For every $\tilde{\theta} \in \Theta^b_{(a_1^*, \emptyset)}$, type $\tilde{\theta}$ plays $a_1^*$ forever with probability $p$, he plays $\alpha_1 \in \tilde{\Omega}^m$ in the normal phase with probability $\frac{1-p}{\tilde{p}}$.

Call the bad strategic type(s) who always play $\alpha_1 \in \tilde{\Omega}^m \cup \{a_1^*\}$ in the normal phase type $\theta(\alpha_1)$. Let $\mu_t(\theta(\alpha_1))$ be the total probability of such type in period $t$. By construction, throughout the normal phase, if $\mu_t(\alpha_1) = 0$, then $\mu_t(\theta(\alpha_1)) = 0$; if $\mu_t(\alpha_1) \neq 0$, then $\mu_t(\theta(\alpha_1))/\mu_t(\alpha_1) = \mu_0(\theta(\alpha_1))/\mu_0(\alpha_1)$.

Next, I describe type $\theta$’s normal phase strategy:

1. **Preparation Sub-Phase:** This phase lasts from period 0 to $n - 1$. Type $\theta$ plays $a_1^*$ in period $i$ for all $i \in \{0, 1, \ldots, n - 1\}$. This is to rule out all pure strategy commitment types.

2. **Value Delivery Sub-Phase:** This phase starts from period $n$. Type $\theta$ plays either $a_1^*$ or some $a_i^* \neq a_1^*$, depending on the realization of $\xi_t$. The probability that $a_1^*$ being prescribed is $q$.

I claim that type $\theta$’s expected payoff is close to $1 - q$ if he plays type $\theta$’s equilibrium strategy when $\delta$ is sufficiently close to 1. This is because in the normal phase:

- After period $n$, player 2 attaches probability 0 to all pure strategy commitment types.

- Starting from period $n$, whenever player 2 observes player 1 playing his equilibrium action, there exists $q > 1$ such that:

$$\mu_{t+1}(\theta)/\left(\mu_{t+1}(\alpha_1) + \mu_{t+1}(\theta(\alpha_1))\right) \geq q \mu_t(\theta)/\left(\mu_t(\alpha_1) + \mu_t(\theta(\alpha_1))\right) . \quad (B.6)$$

for every $\alpha_1 \in \tilde{\Omega}^m$ satisfying $\mu_t(\alpha_1) \neq 0$.

So there exists $T \in \mathbb{N}$ independent of $\delta$ such that in period $t \geq T$, $a_2^*$ is player 2’s strict best reply conditional on $\xi_t$ prescribing $a_1^*$ and play remains in the normal phase. Therefore, type $\theta$’s expected payoff at every normal phase information set must be within the following interval:

$$\left[ (1 - \delta^T)q + \delta^T(1 - q), (1 - \delta^T) + \delta^T(1 - q) \right] .$$

Both the lower and the upper bound of this interval will converge to $1 - q$ as $\delta \to 1$. 

46
**Strategies in the Abnormal Phase:** In the abnormal phase, player 2 has ruled out the possibility that player 1 is type \( \theta \). Type \( \theta(a_1^*) \)’s strategy also remains the same. For every \( \alpha_1 \in \tilde{\Omega}^m \), type \( \theta(\alpha_1) \) plays:

\[
\tilde{\alpha}_1(\alpha_1) \equiv (1 - \frac{\eta}{2})a_1^* + \frac{\eta}{2}\tilde{\alpha}_1(\alpha_1)
\]

where:

\[
\tilde{\alpha}_1(\alpha_1)[a_1] = \begin{cases} 0 & \text{when } a_1 = a_1^* \\ \alpha_1(a_1)/(1 - \alpha_1(a_1^*)) & \text{otherwise}. \end{cases}
\]

I choose \( \eta > 0 \) such that \( \max_{\alpha_1 \in \tilde{\Omega}^m} \alpha_1(a_1^*) < 1 - \eta \), and for every \( \alpha_1' \in \Delta(A_1) \) satisfying \( \alpha_1'(a_1^*) \geq 1 - \eta \), we have:

\[
\sum_{\tilde{\theta} \in \Theta_{\tilde{\alpha}_1}(\alpha_1^*)} x'(\tilde{\theta})u_2(\tilde{\theta}, \alpha_1', a_2') > \sum_{\tilde{\theta} \in \Theta_{\tilde{\alpha}_1}(\alpha_1^*)} x'(\tilde{\theta})u_2(\tilde{\theta}, \alpha_1', a_2^*).
\]

Such \( \eta \) exists according to (B.5).

Next, I verify that type \( \theta \) has no incentive to trigger the abnormal phase. Instead of explicitly constructing his abnormal phase strategy, I compute an upper bound on his payoff in the beginning of the abnormal phase.

\[
\beta(\alpha_1) \equiv \mu(\theta(\alpha_1))/\mu_1(\alpha_1).
\]

Since \( \max_{\alpha_1 \in \tilde{\Omega}^m} \alpha_1(a_1^*) < 1 - \eta \), so if \( a_1^* \) is observed in period \( t \),

\[
\beta_{t+1}(\alpha_1) \geq \frac{1 - \eta/2}{1 - \eta} \beta_t(\alpha_1),
\]

for every \( \alpha_1 \in \tilde{\Omega}^m \). Let \( \gamma \equiv 1 - \min_{\alpha_1 \in \tilde{\Omega}^m} \alpha_1(a_1^*) \). If \( a_1 \neq a_1^* \) is observed in period \( t \), by definition of \( \tilde{\alpha}_1(\alpha_1) \),

\[
\beta_{t+1}(\alpha_1) \geq \frac{\eta}{2\gamma} \beta_t(\alpha_1).
\]

Let \( \tilde{k} \equiv \log \frac{2\gamma}{\eta} / \log \frac{1 - \eta/2}{1 - \eta} \). For every \( \alpha_1 \in \tilde{\Omega}^m \), let \( \bar{\beta}(\alpha_1) \) be the smallest \( \beta \in \mathbb{R}_+ \) such that:

\[
w_2(\tilde{\phi}_{\alpha_1}, \alpha_1, a_2') + \beta \sum_{\tilde{\theta} \in \Theta_{\tilde{\alpha}_1}(\alpha_1^*)} x'(\tilde{\theta})u_2(\tilde{\theta}, \tilde{\alpha}_1(\alpha_1), a_2') \geq w_2(\tilde{\phi}_{\alpha_1}, \alpha_1, a_2^*), \quad \forall \alpha_1 \in \tilde{\Omega}^m.
\]

The choice of \( \eta \) and (B.5) ensures the existence of such \( \bar{\beta}(\alpha_1) \). Let \( \bar{\beta} = 2 \max_{\alpha_1 \in \tilde{\Omega}^m} \bar{\beta}(\alpha_1) \) and \( \beta \equiv \min_{\alpha_1 \in \tilde{\Omega}^m} \frac{\mu(\theta(\alpha_1))}{\mu_1(\alpha_1)} \). Let \( T_1 \equiv \log \frac{2\gamma}{\beta} / \log \frac{1 - \eta/2}{1 - \eta} \). In the beginning of the abnormal phase (regardless of when it is triggered), \( \beta_t(\alpha_1) \geq \beta \) for all \( \alpha_1 \in \tilde{\Omega}^m \). After player 2 observing \( a_1^* \) for \( T_1 \) consecutive periods, \( a_2^* \) is being strictly dominated by \( a_2' \) until he observes some \( a_1' \neq a_1^* \). Every time player 1 plays any \( a_1' \neq a_1^* \), he can trigger outcome \((a_1^*, a_2^*)\) for at most \( \tilde{k} \) consecutive periods before \( a_2^* \) is being strictly dominated by \( a_2' \) again. Therefore, type \( \theta \)'s payoff in the abnormal phase is at most:

\[
(1 - \delta^{T_1}) + \delta^{T_1} \left\{ (1 - \delta^{T_1} - 1) + \delta^{T_1}(1 - \delta^{T_1} - 1) - \delta^{2T_1}(1 - \delta^{T_1} - 1) + \ldots \right\}
\]

The term in the curly bracket converges to \( \frac{\tilde{k}}{1 + \tilde{k}} \) as \( \delta \to 1 \). Let \( q \equiv \frac{\tilde{k}}{2(\tilde{k} + 1)} \), type \( \theta \)'s payoff in beginning of the abnormal phase cannot exceed \( 1 - 2q \).

**Remark:** My construction in the abnormal phase is reminiscent of Jehiel and Samuelson (2012), in which the short-run players mistakenly believe that the strategic long-run player is using a stationary strategy. In the abnormal phase of my construction, player 2’s belief attaches positive probability only to types that are playing stationary strategies. This leads to a similar reputation manipulation procedure: Type \( \theta \) faces a trade-off between playing \( a_1^* \) at the expense of his reputation and playing other actions to build-up his reputation in the abnormal phase. My construction ensures that the speed of reputation building is bounded from above while the speed of reputation deterioration is bounded from below. When player 1’s reputation is sufficiently bad, player 2 has a strict incentive to play \( a_2' \), which punishes type \( \theta \) for at least one period.
C Proof of Theorem 2

I prove Theorem 2 for all games satisfying Assumptions 1-3 while allowing \( a_{2,t} \) to depend on \( h^t \equiv \{ a_{1,s}, a_{2,s}, \xi_s \}_{s \leq t-1} \). To avoid cumbersome notation, I focus on the case where \( \xi_t \) has a finite number of realizations and there are no other commitment types, i.e. \( \Omega^m = \{ \alpha_1 \} \). This is without loss of generality since when player 1 always plays \( \pi_1 \), the probability of other commitment types becomes negligible relative to type \( \pi_1 \) after a bounded number of periods, and those periods have negligible payoff consequences as \( \delta \to 1 \).

C.1 Several Useful Constants

I start with defining several useful constants which depend only on \( \mu, u_1 \) and \( u_2 \), while making no reference to \( \sigma \) and \( \delta \). Let \( M \equiv \max_{\theta, a_1, a_2} |u_1(\theta, a_1, a_2)| \) and

\[
K \equiv \max_{\theta \in \Theta} \left\{ u_1(\theta, \alpha_1, \alpha_2) - u_1(\theta, \bar{\alpha}_1, \alpha_2) \right\} / \min_{\theta \in \Theta} \left\{ u_1(\theta, \alpha_1, \alpha_2) - u_1(\theta, \bar{\alpha}_1, \bar{\alpha}_2) \right\}.
\]

Since \( D(\phi_{\pi_1}, \pi_1) > 0 \), expression (4.2) implies the existence of \( \kappa \in (0, 1) \) such that:

\[
\kappa \mu(\alpha_1) D(\phi_{\pi_1}, \alpha_1) + \sum_{\theta \in \Theta} \mu(\theta) D(\theta, \alpha_1) > 0.
\]

For any \( \kappa \in (0, 1) \), let

\[
\rho_0(\kappa) \equiv \frac{(1 - \kappa) \mu(\alpha_1) D(\phi_{\pi_1}, \alpha_1)}{2 \max_{\theta, a_1, a_2} |D(\theta, a_1)|} > 0 \tag{C.1}
\]

and

\[
\bar{T}_0(\kappa) \equiv \left\lfloor 1 / \rho_0(\kappa) \right\rfloor. \tag{C.2}
\]

Let

\[
\rho_1(\kappa) \equiv \frac{\kappa \mu(\alpha_1) D(\phi_{\pi_1}, \alpha_1)}{\max_{\theta, a_1, a_2} |D(\theta, a_1)|}, \tag{C.3}
\]

and

\[
\bar{T}_1(\kappa) \equiv \left\lfloor 1 / \rho_1(\kappa) \right\rfloor. \tag{C.4}
\]

Let \( \delta \in (0, 1) \) be close enough to 1 such that for every \( \delta \in [\delta, 1) \) and \( \theta_p \in \Theta_p \),

\[
(1 - \delta \bar{T}_0(\delta)) u_1(\theta_p, \alpha_1, \bar{\alpha}_2) + \delta \bar{T}_0(\delta) u_1(\theta_p, \alpha_1, \bar{\alpha}_2) > \frac{1}{2} \left( u_1(\theta_p, \alpha_1, \bar{\alpha}_2) + u_1(\theta_p, \alpha_1, \bar{\alpha}_2) \right). \tag{C.5}
\]

C.2 Random History & Random Path

Let \( h^t \equiv (a^t, r^t) \), with \( a^t \equiv (a_{1,s})_{s \leq t-1} \) and \( r^t \equiv (a_{2,s}, \xi_s)_{s \leq t-1} \). Let \( a^t_\pi \equiv (\alpha_1, ..., \bar{\alpha}_1) \). I call \( h^t \) a public history, \( r^t \) a random history and \( r^\infty \) a random path.

Let \( H \) and \( R \) be the set of public histories and random histories, respectively, with \( \succ, \succcurlyeq, \prec \) and \( \preccurlyeq \) naturally defined. Recall that a strategy profile \( \sigma \) consists of \( (\sigma_\omega)_{\omega \in \Omega} \) with \( \sigma_\omega : H \to \Delta(A_1) \) and \( \sigma_2 : H \to \Delta(A_2) \). Let \( P^\sigma(\omega) \) be the probability measure over public histories induced by \( (\sigma_\omega, \sigma_2) \). Let \( P^\sigma \equiv \sum_{\omega \in \Omega} \mu(\omega) P^\sigma(\omega) \).

Let \( V^\sigma(h^t) \equiv (V^\sigma_\theta(h^t))_{\theta \in \Theta} \in \mathbb{R}^{\Theta} \) be the continuation payoff vector for strategic types at \( h^t \).

Let \( \mathcal{H}^\sigma \subset H \) be the set of histories \( h^t \) such that \( P^\sigma(h^t) > 0 \), and let \( \mathcal{H}^\sigma(\omega) \subset \mathcal{H} \) be the set of histories \( h^t \) such that \( P^\sigma(\omega)(h^t) > 0 \). Let

\[
\mathcal{R}^\sigma_\pi \equiv \left\{ r^\infty \left| (a^t, r^t) \in \mathcal{H}^\sigma \text{ for all } t \text{ and } r^t \prec r^\infty \right. \right\}
\]

be the set of random paths consistent with player 1 always playing \( \bar{\alpha}_1 \). For every \( h^t = (a^t, r^t) \), let \( \bar{\sigma}_1[h^t] : H \to \mathcal{A}_1 \) be a continuation strategy at \( h^t \) satisfying \( \bar{\sigma}_1[h^t](h^s) = \bar{\alpha}_1 \) for all \( h^s \prec h^t \) with \( h^s = (a^t, \bar{\alpha}_1, ..., \bar{\alpha}_1, r^s) \).
\( \mathcal{H}^\sigma \). Let \( \sigma_1[h^t] : \mathcal{H} \to A_1 \) be a continuation strategy that satisfies \( \sigma_1[h^t](h^s) = a_1 \) for all \( h^s \supseteq h^t \) with \( h^s = (a^t, a_{i1}, \ldots, a_{ir}, r^s) \in \mathcal{H}^\sigma \). For every \( \theta \in \Theta \), let

\[
\mathcal{R}_\theta'(t) \equiv \left\{ r^t \mid \sigma_1[a^t, r^t] \text{ is type } \theta \text{’s best reply to } \sigma_2 \right\} \text{ and } \mathcal{R}_\theta^\sigma(t) \equiv \left\{ r^t \mid \sigma_1[a^t, r^t] \text{ is type } \theta \text{’s best reply to } \sigma_2 \right\}.
\]

### C.3 Beliefs & Best Response Sets

Let \( \mu(a^t, r^t) \in \Delta(\Omega) \) be player 2’s posterior belief at \( (a^t, r^t) \) and specifically, let \( \mu^*(r^t) \equiv \mu(a^t, r^t) \). Let

\[
\mathcal{B}_\kappa \equiv \left\{ \hat{\mu} \in \Delta(\Omega) \mid \mu^*(r^t) \leq \hat{\mu}(\theta) \right\} \geq 0 \right\}.
\]

By definition, only \( \{\hat{\mu}(\theta)\}_{\theta \in \Theta} \) matters for whether \( \hat{\mu} \) belongs to \( \mathcal{B}_\kappa \) or not. Moreover, \( \mathcal{B}_{\kappa'} \subseteq \mathcal{B}_{\kappa} \) for every \( \kappa, \kappa' \in [0,1] \) with \( \kappa' < \kappa \).

Let \( q^*(r^t)(\omega) \) be the (ex ante) probability that player 1’s type being \( \omega \) and his past actions being \( a^t \) conditional on \( r^t \). Let \( q^*(r^t) \equiv \mathbb{R}_+^\sigma \) be the corresponding vector of probabilities. For any \( \delta \) and \( \sigma \in \Delta(\mu, \delta) \), Bayes Rule implies that:

- For any \( a^t \) and \( r^t, r^t \supseteq r^{t-1} \) satisfying \( (a^t, r^t), (a^t, r^t) \in \mathcal{H}^\sigma \), we have \( \mu(a^t, r^t) = \mu(a^t, r^t) \).
- For any \( r^t, r^t \supseteq r^{t-1} \) with \( (a^t, r^t), (a^t, r^t) \in \mathcal{H}^\sigma \), we have \( q^*(r^t) = q^*(r^t) \).

This is because player 1’s action in period \( t - 1 \) depends on \( r^t \) only through \( r^{t-1} \), so is player 2’s belief at every on-path history. Since the commitment type always plays \( \bar{a}_1 \), we have \( q^*(r^t)(\bar{a}_1) = \mu_0(\bar{a}_1) \).

For future reference, I introduce two set of random histories based on player 2’s posterior beliefs. Let

\[
\mathcal{R}_g^\sigma \equiv \left\{ r^t \mid (a^t, r^t) \in \mathcal{H}^\sigma \text{ and } \mu^*(r^t)(\Theta) = 0 \right\}.
\]

and let

\[
\mathcal{R}_g^\sigma \equiv \left\{ r^t \mid \exists r^T \supseteq r^t \text{ such that } r^T \in \mathcal{R}_g^\sigma \right\}.
\]

### C.4 A Few Useful Observations

I present four Lemmas, which are useful preliminary results towards the final proof. Recall that \( \sigma_\theta : \mathcal{H} \to \Delta(A_1) \) is type \( \theta \)’s strategy. The first one shows the implications of MSM on player 1’s equilibrium strategy:

**Lemma C.1.** Suppose \( \sigma \in \Delta(\delta, \mu, \phi) \), \( \theta \supseteq \bar{\theta} \) and \( h^t_\theta = (a^t, r^t) \in \mathcal{H}^\sigma(\theta) \cap \mathcal{H}^\sigma(\bar{\theta}) \).

- if \( r^t \in \mathcal{R}_\theta(\bar{\theta}) \), then \( \sigma_\theta(a^s, r^s)(\bar{a}_1) = 1 \) for every \( (a^s, r^s) \in \mathcal{H}(\sigma^*(h^t), \sigma_\theta(\bar{\theta})) \) with \( r^s \supseteq r^t \).
- if \( r^t \in \mathcal{R}_\theta^\sigma(\theta) \), then \( \sigma_\theta(a^s, r^s)(\bar{a}_1) = 1 \) for every \( (a^s, r^s) \in \mathcal{H}(\sigma^*(h^t), \sigma_\theta(\theta)) \) with \( (a^s, r^s) \supseteq (a^t, r^t) \).

**Proof of Lemma C.1:** I only prove first part, since the second part can be shown similarly by switching signs. Without loss of generality, I focus on history \( h^0 \). For notation simplicity, let \( \sigma_1[h^0] = \sigma_1 \). For every \( \sigma_\omega \) and \( \sigma_2 \), let \( P(\sigma_\omega, \sigma_2) : A_1 \times A_2 \to [0,1] \) be defined as:

\[
P(\sigma_\omega, \sigma_2)(a_1, a_2) \equiv \sum_{t=0}^{+\infty} (1 - \delta)^t p_t(\sigma_\omega, \sigma_2)(a_1, a_2)
\]

where \( p_t(\sigma_\omega, \sigma_2)(a_1, a_2) \) is the probability of \( (a_1, a_2) \) occurring in period \( t \) under \( (\sigma_\omega, \sigma_2) \). Let \( P_t(\sigma_1, \sigma_2) \in \Delta(A_2) \) be \( P(\sigma_1, \sigma_2) \)’s marginal distribution on \( A_1 \).
Suppose towards a contradiction that $\bar{\pi}_1$ is type $\tilde{\theta}$’s best reply and there exists $\sigma_\theta$ with $P_1^{(\sigma_\theta,\sigma_2)}(\bar{\pi}_1) < 1$ such that $\sigma_\theta$ is type $\theta$’s best reply, then type $\tilde{\theta}$ and $\theta$’s incentive constraints require that:

$$\sum_{a_2 \in A_2} \left( P_2^{(\sigma_1,\sigma_2)}(a_2) - P_2^{(\sigma_\theta,\sigma_2)}(a_2) \right) u_1(\tilde{\theta}, \bar{a}_1, a_2) \geq \sum_{a_2 \in A_2, a_1 \neq \bar{\pi}_1} P^{(\sigma_\theta,\sigma_2)}(a_1, a_2) \left( u_1(\tilde{\theta}, a_1, a_2) - u_1(\bar{\theta}, \bar{a}_1, a_2) \right),$$

and

$$\sum_{a_2 \in A_2} \left( P_2^{(\sigma_1,\sigma_2)}(a_2) - P_2^{(\sigma_\theta,\sigma_2)}(a_2) \right) u_1(\theta, \bar{a}_1, a_2) \leq \sum_{a_2 \in A_2, a_1 \neq \bar{\pi}_1} P^{(\sigma_\theta,\sigma_2)}(a_1, a_2) \left( u_1(\theta, a_1, a_2) - u_1(\bar{\theta}, \bar{a}_1, a_2) \right).$$

Since $P_1^{(\sigma_\theta,\sigma_2)}(\bar{\pi}_1) < 1$ and $u_1$ has SID in $\theta$ and $a_1$, we have:

$$\sum_{a_2 \in A_2, a_1 \neq \bar{\pi}_1} P^{(\sigma_\theta,\sigma_2)}(a_1, a_2) \left( u_1(\tilde{\theta}, a_1, a_2) - u_1(\bar{\theta}, \bar{a}_1, a_2) \right) > \sum_{a_2 \in A_2, a_1 \neq \bar{\pi}_1} P^{(\sigma_\theta,\sigma_2)}(a_1, a_2) \left( u_1(\theta, a_1, a_2) - u_1(\bar{\theta}, \bar{a}_1, a_2) \right)$$

which implies that:

$$\sum_{a_2 \in A_2} \left( P_2^{(\sigma_\theta,\sigma_2)}(a_2) - P_2^{(\sigma_\theta,\sigma_2)}(a_2) \right) \left( u_1(\theta, \bar{a}_1, a_2) - u_1(\bar{\theta}, \bar{a}_1, a_2) \right) > 0. \quad (C.9)$$

On the other hand, since $u_1$ is strictly decreasing in $a_1$, we have:

$$\sum_{a_2 \in A_2, a_1 \neq \bar{\pi}_1} P^{(\sigma_\theta,\sigma_2)}(a_1, a_2) \left( u_1(\tilde{\theta}, a_1, a_2) - u_1(\bar{\theta}, \bar{a}_1, a_2) \right) > 0$$

Type $\tilde{\theta}$’s incentive constraint implies that:

$$\sum_{a_2 \in A_2} \left( P_2^{(\sigma_1,\sigma_2)}(a_2) - P_2^{(\sigma_\theta,\sigma_2)}(a_2) \right) u_1(\tilde{\theta}, \bar{a}_1, a_2) > 0. \quad (C.10)$$

Since $P_2^{(\sigma_\theta,\sigma_2)}$ and $P_2^{(\sigma_1,\sigma_2)}$ are both probability distributions, $\sum_{a_2 \in A_2} \left( P_2^{(\sigma_\theta,\sigma_2)}(a_2) - P_2^{(\sigma_1,\sigma_2)}(a_2) \right) = 0$.

Since $u_1(\theta, \bar{a}_1, a_2) - u_1(\tilde{\theta}, \bar{a}_1, a_2)$ is weakly increasing in $a_2$, (C.9) implies that $P_2^{(\sigma_\theta,\sigma_2)}(\bar{\pi}_2) - P_2^{(\sigma_1,\sigma_2)}(\bar{\pi}_2) > 0$. Since $\theta(\bar{\theta}, \bar{a}_1, a_2)$ is strictly increasing in $a_2$, (C.10) implies that $P_2^{(\sigma_\theta,\sigma_2)}(\bar{\pi}_2) - P_2^{(\sigma_1,\sigma_2)}(\bar{\pi}_2) < 0$, leading to a contradiction.

The next Lemma places a uniform upper bound on the number of ‘bad periods’ in which $\bar{\pi}_2$ is not player 2’s best reply despite $\bar{\pi}_1$ has always been played and $\mu^*(r^t) \in B_\kappa$.

**Lemma C.2.** If $\mu^*(r^t) \in B_\kappa$ and $\bar{\pi}_2$ is not a strict best reply at $(a^*_t, r^t)$, then:

$$\sum_{\theta \in \Theta} \left( q^*(r^t)(\theta) - q^*(r^{t+1})(\theta) \right) \geq \mu^*(r^t). \quad (C.11)$$

**Proof of Lemma C.2:** If $\mu^*(r^t) \in B_\kappa$, then

$$\kappa \mu(\bar{\pi}_1) \mathcal{D}(\phi_1, a_1) + \sum_{\theta \in \Theta} q^*(r^t)(\theta) \mathcal{D}(\theta, a_1) \geq 0.$$

Suppose $\bar{\pi}_2$ is not a strict best reply at $(a^*_t, r^t)$, then,

$$\mu(\bar{\pi}_1) \mathcal{D}(\phi_1, a_1) + \sum_{\theta \in \Theta} q^*(r^{t+1})(\theta) \mathcal{D}(\theta, a_1) + \sum_{\theta \in \Theta} \left( q^*(r^t)(\theta) - q^*(r^{t+1})(\theta) \right) \mathcal{D}(\theta, a_1) \leq 0,$$

24 According to Bayes Rule, $\mu^*(r^t)(\theta) \geq q^*(r^t)(\theta)$ for all $\theta \in \Theta$ and $\frac{q^*(r^t)(\theta)}{q^*(r^t)}$ is independent of $\theta$ as long as $q^*(r^t)(\theta) \neq 0$. 

50
or equivalently, 
\[
\kappa \mu(\overline{\pi}_1) D(\phi_{\pi_1}, \overline{\pi}_1) + \sum_{\theta \in \Theta} q^*(r^t)(\theta) D(\theta, \overline{\pi}_1) + \sum_{\theta \in \Theta} (1 - \kappa) \mu(\overline{\pi}_1) D(\phi_{\pi_1}, \overline{\pi}_1) \geq 0 \\
+ \sum_{\theta \in \Theta} \left( q^*(r^{t+1})(\theta) - q^*(r^t)(\theta) \right) D(\theta, \overline{\pi}_1) + \sum_{\theta \in \Theta} \left( q^*(r^t)(\theta) - q^*(r^{t+1})(\theta) \right) D(\theta, \overline{\pi}_1) \leq 0.
\]

According to (C.1), we have:
\[
\sum_{\theta \in \Theta} \left( q^*(r^t)(\theta) - q^*(r^{t+1})(\theta) \right) D(\theta, \overline{\pi}_1) \leq 0,
\]

Lemma C.2 implies that for any \( \sigma \in \text{NE}(\delta, \mu, \phi) \) and along any \( r^\infty \in \mathcal{R}^\sigma_0 \), the number of \( r^t \) such that \( \mu^*(r^t) \in \mathcal{B}_0 \) but \( \overline{\sigma}_2 \) is not a strict best reply is at most \( T_0(\kappa) \). The next Lemma obtains an upper bound for player 1’s drop-out payoff at any unfavorable belief.

**Lemma C.3.** For every \( \sigma \in \text{NE}(\delta, \mu, \phi) \) and \( h^t \in \mathcal{H}^\sigma \) with
\[
\mu(h^t)(\overline{\pi}_1) D(\phi_{\pi_1}, \overline{\pi}_1) + \sum_{\theta} \mu(h^t)(\theta) D(\theta, \overline{\pi}_1) < 0.
\]  
(C.12)

Let \( \overline{\theta} \equiv \min \{ \text{supp}(\mu(h^t)) \} \), then:
\[
V_{\overline{\theta}}(h^t) = u_1(\overline{\theta}, a_1, a_2).
\]

**Proof of Lemma C.3:** Let
\[
\Theta^* \equiv \left\{ \hat{\theta} \in \Theta_2 \cup \Theta_2 \mid \mu(h^t)(\hat{\theta}) > 0 \right\}.
\]

According to (C.12), \( \Theta^* \neq \{ \emptyset \} \). The rest of the proof is done via induction on \( |\Theta^*| \). When \( |\Theta^*| = 1 \), there exists a pure strategy \( \sigma^*_2 : \mathcal{H} \rightarrow A_1 \) in the support of \( \sigma_{\overline{\theta}} \) such that (C.12) holds for all \( h^s \) satisfying \( h^s \in \mathcal{H}^{(\delta^s_{\overline{\theta}}, \sigma^*_2)} \) and \( h^s \supseteq h^t \). At every such \( h^s \), \( a_2 \) is player 2’s strict best reply. When playing \( \sigma^*_2 \), type \( \overline{\theta} \)'s stage game payoff is no more than \( u_1(\overline{\theta}, a_1, a_2) \) in every period.

Suppose towards a contradiction that the conclusion holds when \( |\Theta^*| \leq k - 1 \) but fails when \( |\Theta^*| = k \), then there exists \( h^s \in \mathcal{H}^\sigma(\overline{\theta}) \) with \( h^s \supseteq h^t \) such that

- \( \mu(h^s) \notin \mathcal{B}_k \) for all \( h^s \supseteq h^s \supseteq h^t \).
- \( V_{\overline{\theta}}(h^s) > u_1(\overline{\theta}, a_1, a_2) \).
- For all \( A_1 \) such that \( \mu(h^s, a_1) \notin \mathcal{B}_k \), \( \sigma_{\overline{\theta}}(h^s)(a_1) = 0 \).\(^{25}\)

According to the martingale property of beliefs, there exists \( a_1 \) such that \( (h^s, a_1) \in \mathcal{H}^\sigma \) and \( \mu(h^s, a_1) \) satisfies (C.12). Since \( \mu(h^s, a_1)(\overline{\theta}) = 0 \), there exists \( \hat{\theta} \in \Theta^* \setminus \{ \emptyset \} \) such that \( (h^s, a_1) \in \mathcal{H}^\sigma(\hat{\theta}) \). Our induction hypothesis suggests that:
\[
V_{\hat{\theta}}(h^s) = u_1(\hat{\theta}, a_1, a_2).
\]

The incentive constraints of type \( \overline{\theta} \) and type \( \hat{\theta} \) at \( h^s \) require the existence of \( (\alpha_{1, \tau}, \alpha_{2, \tau})_{\tau=0}^\infty \) with \( \alpha_{i, \tau} \in \Delta(A_i) \) such that:
\[
\mathbb{E} \left[ \sum_{\tau=0}^\infty (1-\delta)^\tau \left( u_1(\overline{\theta}, \alpha_{1, \tau}, \alpha_{2, \tau}) - u_1(\overline{\theta}, a_1, a_2) \right) \right] > 0 \geq \mathbb{E} \left[ \sum_{\tau=0}^\infty (1-\delta)^\tau \left( u_1(\hat{\theta}, \alpha_{1, \tau}, \alpha_{2, \tau}) - u_1(\hat{\theta}, a_1, a_2) \right) \right],
\]

\(^{25}\)I omit \( (a_2, \xi_\overline{\theta}) \) in the expression for histories since they play no role in the posterior belief on \( \Omega \) at every on-path history.
where \( \mathbb{E}[-] \) is taken over probability measure \( \mathcal{P}^\sigma \). However, the supermodularity condition implies that,
\[
    u_1(\theta, \alpha_1, \alpha_2) - u_1(\tilde{\theta}, \alpha_1, \alpha_2) \leq u_1(\tilde{\theta}, \alpha_1, \alpha_2) - u_1(\hat{\theta}, \alpha_1, \alpha_2),
\]
leading to a contradiction.

The next Lemma outlines an important implication of \( r^t \notin \hat{\mathcal{R}}^\sigma_g \).

**Lemma C.4.** If \( r^t \notin \hat{\mathcal{R}}^\sigma_g \) and \( (a^t, r^t) \in \mathcal{H}^\sigma \), then there exists
\[
    \theta \in (\Theta_p \cup \Theta_n) \cap \text{supp}(\mu^*(r^t))
\]
such that \( r^t \in \mathcal{R}^\sigma(\theta) \).

**Proof of Lemma C.4:** Suppose towards a contradiction that \( r^t \notin \hat{\mathcal{R}}^\sigma_g \) but no such \( \theta \) exists. Let
\[
    \theta_1 \equiv \max \left\{ (\Theta_p \cup \Theta_n) \cap \text{supp}(\mu^*(r^t)) \right\}{26}
\]
Let \((a^t_1, r^t_1) \gtrless (a^t_2, r^t_2)\) be the history at which type \( \theta_1 \) has a strict incentive not to play \( a_1 \) with \((a^t_1, r^t_1) \in \mathcal{H}^\sigma \). For any \((a^t_1+1, r^t_1+1) \gtrless (a^t_2, r^t_1)\) with \((a^t_1, r^t_1+1) \in \mathcal{H}^\sigma \), on one hand, we have \( \mu^*(r^t_1+1)(\theta_1) = 0 \). On the other hand, the fact that \( r^t \notin \hat{\mathcal{R}}^\sigma_g \) implies that \( \mu^*(r^t_1+1)(\Theta_n \cup \Theta_p) > 0 \).

Let
\[
    \theta_2 \equiv \max \left\{ (\Theta_p \cup \Theta_n) \cap \text{supp}(\mu^*(r^{t+1})) \right\}.
\]
Examine type \( \theta_1 \) and \( \theta_2 \)'s incentive constraints at \((a^t_1, r^{t+1})\). According to Lemma [C.1], there exists \( r^{t+1} \succ r^t \) such that type \( \theta_2 \) has a strict incentive not to play \( a_1 \) at \((a^t_2, r^{t+1}) \in \mathcal{H}^\sigma \).

Therefore, we can iterate this process and obtain \( r^{t+1} \succ r^{t+1} \) ... Since
\[
    \left| \text{supp}(\mu^*(r^{t+1})) \right| \leq \left| \text{supp}(\mu^*(r^{t+1})) \right| - 1,
\]
for any \( k \in \mathbb{N} \), there exists \( m \leq |\Theta_p \cup \Theta_n| \) such that \((a^t_m, r^{t_m}) \in \mathcal{H}^\sigma \), \( r^{t_m} \succ r^t \) and \( \mu^*(r^{t_m})(\Theta_n \cup \Theta_p) = 0 \), which contradicts \( r^t \notin \hat{\mathcal{R}}^\sigma_g \).

**C.5 Positive Types**

In this part, I show the following Proposition:

**Proposition C.1.** If \( \Theta_n = \{ \emptyset \} \) and \( \mu \in \mathcal{B}_\kappa \), then for every \( \theta \), we have:
\[
    V_\theta(a^0, r^0) \geq u_1(\theta, \bar{a}_1, \bar{a}_2) - 2M(1 - \delta T_0(\kappa)).
\]

Despite Proposition [C.1], it is stated in terms of the prior belief, the conclusion applies to all \( r^t \) and \( \theta \in \Theta_p \cup \Theta_n \) when \( \mu^*(r^t) \in \mathcal{B}_\kappa \) and \((a^t, r^t) \in \mathcal{H}^\sigma(\theta) \cup_{\theta, \kappa} \mathcal{H}^\sigma(\theta_n) \). The proof is decomposed into Lemma [C.5] and Lemma [C.6], which together imply Proposition [C.1]. Let \( \sigma \in \text{NE}(\delta, \mu, \phi) \).

**Lemma C.5.** If \( \mu^*(r^t) \in \mathcal{B}_\kappa \) for all \( r^t \in \hat{\mathcal{R}}^\sigma_g \), then for any \( r^\infty \in \mathcal{R}^\sigma \),
\[
    \left| \left\{ t \in \mathbb{N} \mid r^\infty \succ r^t \text{ and } \bar{a}_2 \text{ is not a strict best reply at } (a^t, r^t) \right\} \right| \leq T_0(\kappa). \tag{C.13}
\]

When \( r^t \notin \hat{\mathcal{R}}^\sigma_g \), then the intersection of \( \Theta_p \cup \Theta_n \) and \( \text{supp}(\mu^*(r^t)) \) cannot be empty by definition.
PROOF OF LEMMA C.5: Pick any \( r^\infty \in \mathcal{R}_g^\sigma \), if \( r^0 \notin \hat{\mathcal{R}}_g^\sigma \), then let \( t^* = -1 \). Otherwise, let

\[
t^* \equiv \max \left\{ t \in \mathbb{N} \cup \{+\infty\} \mid r^t \in \hat{\mathcal{R}}_g^\sigma \text{ and } r^\infty > r^t \right\}.
\]

Using the argument in Lemma C.2 for any \( t \leq t^* \), if \( \hat{\sigma}_2 \) is not a strict best reply at \( (a^*_1, r^t) \), inequality (C.11) holds.

Next, I show that \( \mu^*(r^{t^*+1}) \in \mathcal{B}_R \). If \( t^* = -1 \), this is a direct implication of (4.2). If \( t^* \geq 0 \), then there exists \( r^{t^*+1} > r^{t^*} \) such that \( r^{t^*+1} \in \hat{\mathcal{R}}_g^\sigma \). Let \( r^{t^*+1} > r^\infty \), we have \( q^*(r^{t^*+1}) = q^*(r^{t^*+1}) \). Moreover, since \( \mu^*(r^t) \in \mathcal{B}_R \) for every \( r^t \in \hat{\mathcal{R}}_g^\sigma \), we have \( \mu^*(r^{t^*+1}) = \mu^*(r^{t^*+1}) \in \mathcal{B}_R \).

Since \( r^{t^*+1} \notin \hat{\mathcal{R}}_g^\sigma \), Lemma C.4 implies the existence of

\[
\theta \in (\Theta_p \cup \Theta_n) \cap \text{supp}(\mu^*(r^{t^*+1}))
\]

such that \( r^{t^*+1} \in \overline{\mathcal{R}}(\theta) \). Since \( \theta_R \geq \theta \) for all \( \theta_R \in \Theta_R \), Lemma C.1 implies that for every \( \theta_R \) and \( r^\infty > r^t \geq r^{t^*+1} \), we have \( \sigma_{\theta_R}(a^*_1, r^t) = 1 \), and therefore, \( q^*(r^t)(\theta) = q^*(r^{t^*+1})(\theta_R) \). This implies that \( \mu^*(r^t) \in \mathcal{B}_R \) for every \( r^\infty > r^t \geq r^{t^*+1} \).

To sum up, for every \( t \in \mathbb{N} \), if \( \hat{\sigma}_2 \) is not a strict best reply at \( (a^*_1, r^t) \), then:

\[
\sum_{\theta \in \Theta} \left( q^*(r^t)(\theta) - q^*(r^{t^*+1})(\theta) \right) \geq \rho(\kappa),
\]

from which we obtain (C.13).

The next result shows that the condition in Lemma C.5 holds in every equilibrium when \( \delta \) is large enough.

**Lemma C.6.** If \( \delta > \overline{\delta} \), then \( \mu^*(r^t) \in \mathcal{B}_0 \) for every \( r^t \in \hat{\mathcal{R}}_g^\sigma \) with \( \mu^*(r^t)(\Theta_R) = 0 \).

**Proof of Lemma C.6:** For any given \( \delta > \overline{\delta} \), according to (C.5), there exists \( \kappa^* \in (0, 1) \) such that:

\[
(1 - \delta \mathcal{T}_0(\kappa^*)u_1(\theta_p, \overline{a}_1, a_2) + \delta \mathcal{T}_0(\kappa^*)u_1(\theta_p, \overline{a}_1, a_2) > \frac{1}{2} \left( u_1(\theta_p, \overline{a}_1, a_2) + u_1(\theta_p, \overline{a}_1, a_2) \right).
\]

Suppose towards a contradiction that there exist \( r^{t_1} \) and \( r^{T_1} \) such that:

- \( r^{T_1} \geq r^{t_1}, r^{T_1} \in \mathcal{R}_g^\sigma \) and \( \mu^*(r^{t_1}) \notin \mathcal{B}_0 \).

Since \( \mu^*(r^{T_1}) \in \mathcal{B}_0 \), let \( t^*_1 \) be the largest \( t \in \mathbb{N} \) such that \( \mu^*(r^t) \notin \mathcal{B}_0 \) for \( r^{T_1} \geq r^t \gg r^{t^*_1} \). Then there exists \( a_1 \neq \overline{a}_1 \) and \( r^{t^*_1+1} \geq r^{t^*_1} \) such that \( \mu((a^*_1, a_1), r^{t^*_1+1}) \notin \mathcal{B}_0 \) and \( ((a^*_1, a_1), r^{t^*_1+1}) \in \mathcal{H}_R^\sigma \). This also implies the existence of \( \theta_p \in \Theta_p \cap \text{supp}(\mu((a^*_1, a_1), r^{t^*_1+1})) \).

According to Lemma C.3, type \( \theta_p \)'s continuation payoff at \( (a^*_1, r^{t^*_1}) \) by playing \( a_1 \) is at most

\[
(1 - \delta)u_1(\theta_p, \overline{a}_1, a_2) + \delta u_1(\theta_p, \overline{a}_1, a_2).
\]

His incentive constraint at \( (a^*_1, r^{t^*_1}) \) requires that his expected payoff from \( \sigma_1 \) is weakly lower than (C.15), i.e. there exists \( r^{t^*_1+1} \geq r^{t^*_1} \) satisfying \( (a^*_1, r^{t^*_1+1}) \in \mathcal{H}_R^\sigma \) and type \( \theta_p \)'s continuation payoff at \( (a^*_1, r^{t^*_1+1}) \) is no more than:

\[
\frac{1}{2} \left( u_1(\theta_p, \overline{a}_1, a_2) + u_1(\theta_p, \overline{a}_1, a_2) \right).
\]

If \( \mu^*(r^t) \in \mathcal{B}_0 \) for every \( r^t \in \hat{\mathcal{R}}_g^\sigma \cap \{r^t \geq r^{t^*_1}\} \), then according to Lemma C.5, his continuation payoff at \( (a^*_1, r^{t^*_1}) \) by playing \( \overline{a}_1 \) is at least:

\[
(1 - \delta \mathcal{T}_0(\kappa^*)u_1(\theta_p, \overline{a}_1, a_2) + \delta \mathcal{T}_0(\kappa^*)u_1(\theta_p, \overline{a}_1, a_2),
\]

which is strictly larger than (C.16) by the definition of \( \kappa^* \) in (C.14), leading to a contradiction.

Suppose on the other hand, there exists \( r^{t^*_2} \gg r^{t^*_1} \) such that:

53
\[ r^{t_2} \in \mathcal{R}_y^\kappa \text{ while } \mu^*(r^{t_2}) \notin B_{\kappa^*}. \]

There exists \( r^{T_2} \succ r^{t_2} \) such that \( r^{T_2} \in \mathcal{R}_y^\kappa \) and \( r^{T_2} \succ r^{t_2} \). Again, we can find \( r^{t_2} \) such that \( t^*_2 \) be the largest \( t \in [t_2, T_2] \) such that \( \mu^*(r^t) \notin B_0 \) for \( r^{T_2} \succ r^t \). Then there exists \( a_1 \neq \overline{a}_1 \) and \( r^{t_2+1} \succ r^{t_2} \) such that \( \mu((a^{t_2}_2, a_1), r^{t_2+1}) \notin B_0 \) and \((a^{t_2}_2, a_1), r^{t_2+1}) \in \mathcal{H}^\sigma\).

Repeating this argument by iterating the above process, for every \( k \geq 1 \), in order to satisfy player 1’s incentive constraint to play \( a_1 \neq \overline{a}_1 \) at \((a^{t_k}_2, r^{t_k})\), we can find the triple \((r^{t_k+1}, r^{t_k+1}, r^{T_{k+1}})\), i.e. this process cannot stop after finite rounds of iteration. Since \( \mu^*(r^{t_k}) \notin B_{\kappa^*} \) but \( \mu^*(r^{t_k+1}) \in B_0 \) as well as \( r^{t_k+1} \succ r^{t_k+1} \), we have:

\[
\sum_{\theta \in \Theta} q^*(r^k)(\theta) - q^*(r^{k+1})(\theta) \geq \sum_{\theta \in \Theta} q^*(r^k)(\theta) - q^*(r^{k+1})(\theta) \geq \rho_1(\kappa^*) \tag{C.17}
\]

for every \( k \geq 2 \). (C.17) and (C.4) together suggest that this iteration process cannot last for more than \( T_1(\kappa^*) \) rounds, which is an integer independent of \( \delta \), leading to a contradiction.

The next Lemma is not needed for the proof of Proposition C.1 but will be useful for future reference.

**Lemma C.7.** For any \( \delta \geq \overline{\delta} \) and any \( \sigma \in NE(\delta, \mu, \phi) \), for every \( r^t \) such that \((a^t_1, r^t) \in \mathcal{H}^\sigma \), \( \mu^*(r^t)(\Theta_\sigma) = 0 \), \( r^t \notin \mathcal{R}_y^\kappa \) and

\[
\mu(\overline{a}_1)D(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta} q^*(r^t)(\theta)D(\theta, \overline{a}_1) > 0, \tag{C.18}
\]

\( \overline{a}_2 \) is player 2’s strict best reply at every \((a^*_2, r^s) \succeq (a^*_2, r^t) \) with \((a^*_2, r^s) \in \mathcal{H}^\sigma \).

**Proof of Lemma C.7:** Since \( \mu^*(r^t)(\Theta_\sigma) = 0 \) and \( r^t \notin \mathcal{R}_y^\kappa \), Lemma C.4 implies the existence of \( \sigma_\theta(a^*_2, r^s)(\overline{a}_1) = 1 \) for every \((a^*_2, r^s) \in \mathcal{H}^\sigma(\theta) \) with \( r^s \succ r^t \). From (C.18), we know that \( \overline{a}_2 \) is not a strict best reply only if there exists type \( \theta_p \in \Theta_p \) who plays \( a_1 \neq \overline{a}_1 \) with positive probability. In particular, (C.18) implies the existence of \( \pi \in (0, 1) \) such that \( \pi \mu(\pi_1)D(\phi_{\pi_1}, \pi_1) + \sum_{\theta \in \Theta} q^*(r^t)(\theta)D(\theta, \pi_1) > 0. \)

According to (C.11), we have:

\[
\sum_{\theta \in \Theta_p} \left( q^*(r^s)(\theta) - q^*(r^{s+1})(\theta) \right) \geq \rho_0(\pi) \]

whenever \( \overline{a}_2 \) is not a strict best reply at \((a^*_2, r^s) \succeq (a^*_2, r^t) \). Therefore, there can be at most \( T_0(\pi) \) such periods. Hence, there exists \( r^N \) with \((a^*_2, r^N) \in \mathcal{H}^\sigma \) such that:

- \( \overline{a}_2 \) is not a strict best reply at \((a^*_2, r^N) \).
- \( \overline{a}_2 \) is a strict best reply for all \((a^*_2, r^s) \succ (a^*_2, r^N) \) with \((a^*_2, r^s) \in \mathcal{H}^\sigma \).

Then there exists \( \theta_p \in \Theta_p \) that plays \( a_1 \neq \overline{a}_1 \) in equilibrium at \((a^*_2, r^N) \), his continuation payoff by always playing \( \pi_1 \) is at least \((1 - \delta)u_1(\theta_p, a_1, \overline{a}_2) + \delta u_1(\theta_p, \pi_1, \overline{a}_2) \) while his equilibrium continuation payoff from playing \( a_1 \) is at most \((1 - \delta)u_1(\theta_p, a_1, a_2) + \delta u_1(\theta_p, \pi_1, a_2) \) according to Lemma C.3. The latter is strictly less than the former when \( \delta > \overline{\delta} \), leading to a contradiction. \[ \square \]

\[27\]The reasons why we cannot directly apply Lemma C.2 are, first, stronger conclusion is required for Lemma C.7 and second, \( \kappa \) can be arbitrarily close to 1, while \( \kappa \) is uniformly bounded below 1 for any given \( \mu \).
C.6 Incorporating Negative Types

Next, we extend the proof by allowing for $\Theta_n \neq \{\emptyset\}$. Lemmas C.5 and C.6 imply the following result in this general environment:

**Proposition C.2.** For any $\delta > \delta$ and $\sigma \in \text{NE}(\delta, \mu, \phi)$, there do not exist $\theta_p \in \Theta_p$, $r^{t+1} > r^t$ and $a_1 \neq a_1$ that simultaneously satisfy:

1. $r^{t+1} \in \mathcal{R}_g^\sigma$.
2. $((a^*_s, a_1), r^{t+1}) \in \mathcal{H}^\sigma(\theta_p)$.
3. $V^\theta_p(((a^*_s, a_1), r^{t+1})) = u_1(\theta_p, a_1, a_2)$ for all $r^{t+1} > r^t$.

**Proof of Proposition C.2:** Suppose towards a contradiction that such $\theta_p \in \Theta_p$, $r^t$, $r^{t+1}$ and $a_1$ exist. From requirement 3, we know that $r^t \in \mathcal{R}_g^\sigma(\theta_p)$, according to Lemma 4.1, $\theta_n \prec \theta_p$ for all $\theta_n \in \Theta_n$. The second part of Lemma C.6 then implies that $\mu^*(r^{t+1})(\Theta_n) = 0$ for all $r^{t+1} > r^t$ with $(a^*_s, r^{t+1}) \in \mathcal{H}^\sigma$.

If $\mu^*(r^{t+1}) \in \mathcal{B}_n$, then requirement 2 and Proposition C.1 result in a contradiction when examining type $\theta_p$’s incentive at $(a^*_s, r^t)$ to play $a_1$ as opposed to $a_1$. If $\mu^*(r^{t+1}) \notin \mathcal{B}_n$, since $\delta > \delta$ and $r^{t+1} \in \mathcal{R}_g^\sigma$, we obtain a contradiction from Lemma C.6.

The rest of the proof is decomposed into several steps by considering any $\sigma \in \text{NE}(\mu, \delta)$ when $\delta$ is large enough. First,\(^{25}\)

\[
\mu(\pi_1)D(\phi_{\pi_1}, \pi_1) + \sum_{\theta \in \Theta} q^*(r^t)(\theta)D(\theta, \pi_1) \geq 0 \tag{C.19}
\]

for all $t \geq 1$ and $r^t$ satisfying $(a^*_s, r^t) \in \mathcal{H}^\sigma$. This is because otherwise, according to Lemma C.3, there exists $\theta \in \text{supp}(\mu^*(r^t))$ such that $V^\theta_p(a^*_s, r^t) = u_1(\theta, a_1, a_2)$. But then, at $(a^*_s, r^{t-1})$ with $r^{t-1} < r^t$, he could obtain strictly higher payoff by playing $a_1$ instead of $a_1$, leading to a contradiction.

Next comes the following Lemma:

**Lemma C.8.** $V^\theta_p(a^*_s, r^t) \geq u_1(\theta, a_1, a_2) - 2M(K + 1)(1 - \delta)$ for every $\theta$ and $r^t \notin \mathcal{R}_g^\sigma$ satisfying:

- $(a^*_s, r^t) \in \mathcal{H}^\sigma$.
- Either $t = 0$ or $t \geq 1$ but there exists $r^t$ such that $r^t, r^t > r^{t-1}$, $(a^*_s, r^t) \in \mathcal{H}^\sigma$ and $\hat{r}^t \in \mathcal{R}_g^\sigma$.

**Proof of Lemma C.8:** If $\mu^*(r^t) \in \mathcal{B}_n$ and $r^t \notin \mathcal{R}_g^\sigma$, then Lemmas C.1 and C.4 suggest that $\mu^*(r^t) \in \mathcal{B}_n$ for all $r^t \succ r^t$ and the conclusion is straightforward from Lemma C.2.

Therefore, for the rest of the proof, I assume that $\mu^*(r^t) \notin \mathcal{B}_n$. I consider two cases. First, when $\mu^*(r^t)(\Theta_n) > 0$, then according to (C.17),\(^{26}\)

\[
\mu(\pi_1)D(\phi_{\pi_1}, \pi_1) + \sum_{\theta \in \Theta_n \cup \Theta_g} q^*(r^t)(\theta)D(\theta, \pi_1) > 0.
\]

Since $r^t \notin \mathcal{R}_g^\sigma$, according to Lemma C.4, there exists $\theta \in \Theta_n \cup \Theta_g$ with $(a^*_s, r^t) \in \mathcal{H}^\sigma(\theta)$ such that $r^t \in \mathcal{R}_g^\sigma(\theta)$. According to Lemma C.1, for all $\theta_g \in \Theta_g$ with $(a^*_s, r^t) \in \mathcal{H}^\sigma(\theta_g)$ and every $(a^*_s, r^s) \in \mathcal{H}^\sigma(\theta)$ with $r^s \succ r^t$, we have $\sigma_{\theta_g}(a^*_s, r^s) = 1$.

This implies that for every $h^s = (a^*_s, r^s) \succ (a^*_s, r^t)$ with $a^s \neq a^*_s$ and $h^s \in \mathcal{H}^\sigma$, we have $\mu(h^s)(\Theta_g) = 0$, so for every $\theta$,

\[
V^\theta_p(h^s) = u_1(\theta, a_1, a_2). \tag{C.20}
\]

Let $\tau : \mathcal{R}_g^\sigma \to \mathbb{N} \cup \{+\infty\}$ be such that for $r^t \prec r^{t+1} \prec r^\infty$, we have:

\(^{25}\)Inequality (C.19) trivially applies to $r^0$ due to (4.2).

\(^{26}\)To see this, consider three cases. If $\Theta_n = \{\emptyset\}$, then this inequality is obvious. If $\Theta_n \neq \{\emptyset\}$, then $D(\theta_n, \pi_1) \leq 0$ for all $\theta_n \in \Theta_n$ according to Lemma 4.1. When $D(\theta_n, \pi_1) < 0$ for all $\theta_n$, then the inequality follows from (C.19). When $D(\theta_n, \pi_1) = 0$ for some $\theta_n \in \Theta_n$, then $D(\theta_n, \pi_1) = 0$ for all $\theta_p \in \Theta_p$. The inequality then follows from $D(\theta_n, \pi_1) > 0$ for all $\theta_p \in \Theta_n$ as well as $\hat{\theta} \in \Theta_g$. 

55
\( \mu^*(r^\tau)(\Theta_n) > 0 \) while \( \mu^*(r^{\tau+1})(\Theta_n) = 0 \).

Let
\[
\bar{\theta}_n \equiv \max \left\{ \supp(\mu^*(r^\tau)) \cap \Theta_n \right\}.
\]

The second part of Lemma C.1 and C.20 together imply that \( \mu^*(r^\tau)(\bar{\theta}_n) > 0 \). Let us examine type \( \bar{\theta}_n \)'s incentive at \( (a^*_s, r^t) \) to play his equilibrium strategy as opposed to play \( a^*_1 \) all the time. This requires that:
\[
E \left[ \sum_{s=t}^{\tau-1} (1-\delta)^{s-t} u_1(\bar{\theta}_n, a_1, \alpha_{2,s}) + (\delta^{s-t} - \delta^{s+1-t}) u_1(\bar{\theta}_n, a_1, \tau, \alpha_{2,\tau}) + \delta^{s+1-t} u_1(\bar{\theta}_n, a_1, a_2) \right] \geq u_1(\bar{\theta}_n, a_1, a_2).
\]
where \( E[\cdot] \) is taken over \( \mathcal{P}^\sigma \) and \( \alpha_{2,s} \in \Delta(A_2) \) is player 2's action in period \( s \).

Using the fact that \( u_1(\bar{\theta}_n, a_1, a_2) \geq u_1(\bar{\theta}_n, a_1, a_2) \), the above inequality implies that:
\[
E \left[ \sum_{s=t}^{\tau-1} (1-\delta)^{s-t} \left( u_1(\bar{\theta}_n, a_1, \alpha_{2,s}) - u_1(\bar{\theta}_n, a_1, a_2) \right) + (\delta^{s-t} - \delta^{s+1-t}) \left( u_1(\bar{\theta}_n, a_1, \alpha_{2,\tau}) - u_1(\bar{\theta}_n, a_1, a_2) \right) \right] \leq 0.
\]

According to the definitions of \( K \) and \( M \), we know that for all \( \theta \),
\[
E \left[ \sum_{s=t}^{\tau} (1-\delta)^{s-t} \left( u_1(\theta_n, \bar{a}_1, \alpha_{2,s}) - u_1(\theta_n, \bar{a}_1, a_2) \right) \right] \leq 2M(K+1)(1-\delta). \tag{C.21}
\]

For every \( r^\infty \in \mathcal{R}_\sigma^\theta \), since \( r^t \notin \hat{\mathcal{R}}_\sigma^\theta \), we have:
\[
\mu(\bar{\alpha}_1) D(\phi_{\alpha_1, \bar{a}_1}) + \sum_{\theta \in \Theta} q^*(r^t(r^\infty+1))(\theta) D(\theta, \bar{a}_1) \geq \mu(\bar{\alpha}_1) D(\phi_{\alpha_1, \bar{a}_1}) + \sum_{\theta \in \Theta \cup \Theta_g} q^*(r^t) D(\theta, \bar{a}_1) \geq \mu(\bar{\alpha}_1) D(\phi_{\alpha_1, \bar{a}_1}) + \sum_{\theta \in \Theta} q^*(r^t) D(\theta, \bar{a}_1) \geq 0.
\]

According to Lemma C.7, we know that \( V_\theta(a^*_s, r^1(r^\infty+1)) = u_1(\theta, \bar{a}_1, a_2) \) for all \( \theta \in \Theta_g \cup \Theta_p \) and \( r^\infty \in \mathcal{R}_\sigma^\theta \). This together with (C.21) gives the conclusion.

Second, when \( \mu^*(r^\tau)(\Theta_n) = 0 \). If \( t = 0 \), the conclusion directly follows from Proposition C.1. If \( t \geq 1 \) and there exists \( \hat{r}^t \) such that \( r^t, \hat{r}^t > r^{t-1}, (a^*_s, \hat{r}^t) \in \mathcal{H}_\sigma \) and \( \hat{r}^t \in \hat{\mathcal{R}}_\sigma^\theta \). Then, since
\[
\mu^*(r^t) = \mu^*(\hat{r}^t),
\]
we have \( \mu^*(r^t)(\Theta_n) = 0 \). Since \( \hat{r}^t \in \hat{\mathcal{R}}_\sigma^\theta \), according to Lemma C.6 \( \mu^*(r^t) = \mu^*(r^t) \in \mathcal{B}_\kappa \). The conclusion then follows from Lemma C.7.

The next Lemma puts an upper bound on type \( \theta_n \in \Theta_n \)'s continuation payoff at \( (a^*_s, r^t) \) with \( r^t \notin \hat{\mathcal{R}}_\sigma^\theta \).

**Lemma C.9.** For every \( \theta_n \in \Theta_n \) such that \( \bar{a}_2 \notin \mathcal{B}_2(\bar{a}_1, \theta_n) \) and \( r^t \notin \hat{\mathcal{R}}_\sigma^\theta \) with \( (a^*_s, r^t) \in \mathcal{H}_{\theta_n} \) and \( \mu^*(r^t) \notin \mathcal{B}_\kappa \), we have:
\[
V_{\theta_n}(a^*_s, r^t) \leq u_1(\theta_n, a_1, a_2) + 2(1-\delta)M. \tag{C.22}
\]

The proof is contained in the proof for the first case in Lemma C.8. Let
\[
A(\delta) \equiv 2M(K+1)(1-\delta), \quad B(\delta) \equiv 2M(1-\delta T^\alpha(\kappa))
\]
and
\[
C(\delta) \equiv 2MK|\Theta_n|(1-\delta).
\]
Notice that when \( \delta \to 1 \), all three functions converge to 0. The next Lemma puts a uniform upper bound on player 1's payoff when \( r^t \in \hat{\mathcal{R}}_\sigma^\theta \).

**Lemma C.10.** When \( \delta > \bar{\delta} \) and \( \sigma \in \mathcal{N}(\delta, \mu, \phi) \), for every \( r^t \in \hat{\mathcal{R}}_\sigma^\theta \),
\[
V_\theta(a^*_s, r^t) \geq u_1(\theta, \bar{a}_1, \bar{a}_2) - \left( A(\delta) + B(\delta) \right) - 2\bar{T}_1(\kappa) \left( A(\delta) + B(\delta) + C(\delta) \right). \tag{C.23}
\]

\(^{30}\)One can further tighten this bound. However, (C.23) is sufficient for the purpose of proving Theorem 2.
for all $\theta$ such that $(a_s^t, r^s) \in H^\sigma(\theta)$.

**Proof of Lemma C.10:** The non-trivial part of the proof deals with situations where $\mu^*(r^t) \notin B_n$. Since $r^t \in \hat{R}_n$, Lemma C.6 implies that $\mu^*(r^t)(\Theta_n) \neq 0$. Without loss of generality, assume $\Theta_n \subseteq \text{supp}(\mu^*(r^t))$. Let me introduce $|\Theta_n| + 1$ integer valued random variables on the space $\mathcal{R}_n$.

- $\tau : \mathcal{R}_n \to \mathbb{N} \cup \{+\infty\}$ be the first period $s$ along random path $r^\infty$ such that either one of the following two conditions is met.

1. $\mu^*(r^{s+1}) \in B_{n/2}$ for $r^{s+1} \succ r^s$ with $(a^t_s, r^{s+1}) \in H^\sigma$.
2. $r^s \notin \hat{R}_n$.

In the first case, there exists $a_1 \neq \bar{a}_1$ and $r^{t+1} \succ r^t$ such that

- $((a^s_t, a_1), r^{t+1}) \in H^\sigma(\tilde{\theta})$ for some $\tilde{\theta} \in \Theta_p \cup \Theta_n$.
- $\mu((a^s_t, a_1), r^{t+1}) \notin B_0$.

Lemma C.3 implies the existence of $\theta \in \Theta_p \cup \Theta_n$ with $((a^s_t, a_1), r^{t+1}) \in H^\sigma(\theta)$ such that $V_\theta((a^s_t, a_1), r^{t+1}) = u_1(\theta, \bar{a}_1, \bar{a}_2)$.

Suppose towards a contradiction that $\theta \in \Theta_p$, then Lemma C.1 implies that $\mu^*(r^{t+1})(\Theta_n) = 0$. Since $\mu^*(r^{t+1}) \in B_{n/2}$, Proposition C.1 implies that type $\theta$’s continuation payoff by always playing $\bar{a}_1$ is at least

$$(1 - \delta T_0(n/2))u_1(\theta, \bar{a}_1, \bar{a}_2) + \delta T_0(n/2)u_1(\theta, \bar{a}_1, \bar{a}_2),$$

which is strictly larger than his payoff from playing $a_1$, which is at most $2M(1 - \delta) + u_1(\theta, \bar{a}_1, a_2)$, leading to a contradiction.

Hence, there exists $\theta_n \in \Theta_n$ such that $V_{\theta_n}((a^s_t, a_1), r^{t+1}) = u_1(\theta_n, \bar{a}_1, \bar{a}_2)$, which implies that $V_{\theta_n}(a^s_t, r^t) \leq u_1(\theta_n, \bar{a}_1, \bar{a}_2) + 2(1 - \delta)M$.

In the second case, Lemma C.9 implies that $V_{\theta_n}(a^s_t, r^t) \leq u_1(\theta_n, \bar{a}_1, \bar{a}_2) + 2(1 - \delta)M$ for all $\theta_n \in \Theta_n$ with $r^s \in H^\sigma(\theta_n)$.

- For every $\theta_n \in \Theta_n$, let $\tau_{\theta_n} : \mathcal{R}^s_n \to \mathbb{N} \cup \{+\infty\}$ be the first period $s$ along random path $r^\infty$ such that either one of the following three conditions is met.

1. $\mu^*(r^{s+1}) \in B_{n/2}$ for $r^{s+1} \succ r^s$ with $(a^t_s, r^{s+1}) \in H^\sigma$.
2. $r^s \notin \hat{R}_n$.
3. $\mu^*(r^{s+1})(\theta_n) = 0$ for $r^{s+1} \succ r^s$ with $(a^t_s, r^{s+1}) \in H^\sigma$.

By definition, $\tau \geq \tau_{\theta_n}$, so $\tau \geq \max_{\theta_n \in \Theta_n}\{\tau_{\theta_n}\}$. Next, I show that

$$\tau = \max_{\theta_n \in \Theta_n}\{\tau_{\theta_n}\}. \quad (C.24)$$

Suppose on the contrary that $\tau > \max_{\theta_n \in \Theta_n}\{\tau_{\theta_n}\}$ for some $r^\infty \in \mathcal{R}^\sigma_n$. Then there exists $(a^t_s, r^s) \succ (a^t_s, r^t)$ such that $r^s \in \hat{R}_n$, $\mu^*(r^t) \notin B_n$ and $\mu^*(r^s)(\Theta_n) = 0$, which contradicts Lemma C.6 when $\delta > \bar{\delta}$.

Next, I show by induction over $|\Theta_n|$ that

$$\mathbb{E}\left[\sum_{s=t}^\tau (1 - \delta)^{\tau-s} \left(u_1(\theta, \bar{a}_1, \bar{a}_2) - u_1(\theta, \bar{a}_1, \bar{a}_2)\right)\right] \leq 2MK|\Theta_n|(1 - \delta), \quad (C.25)$$

57
for all $\theta \in \Theta$ and
\[ V_{\hat{\theta}}(a_2^{\tau_{\theta_n}}, r^{\tau_{\theta_n}}) \leq u_1(\theta_n, a_1, a_2) + 2(1 - \delta)M, \tag{C.26} \]
for
\[ \hat{\theta} \equiv \min \{ \Theta_n \cap \text{supp}(\mu^*(r^{\tau_{\theta_n} + 1})) \} \]
with $\theta_n, \hat{\theta}_n \in \Theta_n$, where $E[\cdot]$ is taken over $P^\sigma$ and $\hat{\alpha}_{2,s} \in \Delta(A_2)$ is player 2’s (mixed) action at $(a_s^*, r^\tau)$.

When $|\Theta_n| = 1$, let $\theta_n$ be its unique element. Consider player 1’s pure strategy of playing $\bar{a}_1$ until $r^\tau$ and then play $a_1$ forever. This is one of type $\theta_n$’s best responses according to (C.24), which results in payoff at most:
\[ E\left[ \sum_{s=t}^{\tau-1} (1 - \delta)^{s-t}u_1(\theta_n, \bar{a}_1, \hat{\alpha}_{2,s}) + \delta^{\tau-t} \left( u_1(\theta_n, a_1, a_2) + 2(1 - \delta)M \right) \right]. \]

The above expression cannot be smaller than $u_1(\theta_n, a_1, a_2)$, which is the payoff he can guarantee by always playing $a_1$. Since $u_1(\theta_n, a_1, a_2) \geq u_1(\theta_n, \bar{a}_1, \bar{a}_2)$, and from the definition of $K$, we get for all $\theta$,
\[ E\left[ \sum_{s=t}^{\tau-1} (1 - \delta)^{s-t} \left( u_1(\theta_n, \bar{a}_1, \bar{a}_2) - u_1(\theta, \bar{a}_1, \hat{\alpha}_{2,s}) \right) \right] \leq 2MK(1 - \delta). \]

We can then obtain (C.26) for free since $\tau = \tau_{\theta_n}$ and type $\theta_n$’s continuation value at $(a_s^*, r^\tau)$ is at most $u_1(\theta_n, a_1, a_2) + 2(1 - \delta)M$ by Lemma C.3.

Suppose the conclusion holds for all $|\Theta_n| \leq k - 1$, consider when $|\Theta_n| = k$ and let $\theta_n \equiv \min \Theta_n$. If $(a_s^*, r^\tau) \notin \mathcal{H}^\sigma(\theta_n)$, then there exists $(a_s^{\tau_{\theta_n}}, r^{\tau_{\theta_n}}) \in \mathcal{H}^\sigma(\theta_n)$ at which type $\theta_n$ plays $\bar{a}_1$ with probability 0. I put an upper bound on type $\theta_n$’s continuation payoff at $(a_s^{\tau_{\theta_n}}, r^{\tau_{\theta_n}})$ by examining type $\hat{\theta}_n \in \Theta_n \setminus \{ \theta_n \}$’s incentive to play $\bar{a}_1$ at $(a_s^{\tau_{\theta_n}}, r^{\tau_{\theta_n}})$, where
\[ \hat{\theta} \equiv \min \{ \Theta_n \cap \text{supp}(\mu^*(r^{\tau_{\theta_n} + 1})) \} \]
This requires that:
\[ E\left[ \sum_{s=0}^{\infty} (1 - \delta)^s u_1(\theta_n, a_1, a_2) \right] \leq u_1(\hat{\theta}_n, a_1, a_2) + 2(1 - \delta)M, \]
by induction hypothesis

where $\{(\alpha_{1,s}, \alpha_{2,s})\}_{s \in \mathbb{N}}$ is the equilibrium continuation play following $(a_s^{\tau_{\theta_n}}, r^{\tau_{\theta_n}})$. By definition, $\hat{\theta}_n > \theta_n$, so the supermodularity condition implies that:
\[ u_1(\theta_n, a_1, a_2) - u_1(\hat{\theta}_n, a_1, a_2) \geq u_1(\theta_n, a_1, a_2) - u_1(\hat{\theta}_n, a_1, a_2). \]

Therefore, we have:
\[ V_{\theta_n}(a_s^{\tau_{\theta_n}}, r^{\tau_{\theta_n}}) = E\left[ \sum_{s=0}^{\infty} (1 - \delta)^s u_1(\theta_n, a_1, a_2) \right] \leq E\left[ \sum_{s=0}^{\infty} (1 - \delta)^s \left( u_1(\hat{\theta}_n, a_1, a_2) - u_1(\theta_n, a_1, a_2) \right) \right] \leq u_1(\theta_n, a_1, a_2) + 2(1 - \delta)M. \]

Back to type $\theta_n$’s incentive constraint. Since it is optimal for him to play $\bar{a}_1$ until $r^{\tau_{\theta_n}}$ and then play $a_1$ forever, doing so must give him a higher payoff than playing $a_1$ forever starting from $r^\tau$, which gives:
\[ E\left[ \sum_{s=t}^{\tau_{\theta_n} - 1} (1 - \delta)^{s-t} u_1(\theta_n, \bar{a}_1, \hat{\alpha}_{2,s}) + \delta^{\tau_{\theta_n}} \left( u_1(\theta_n, a_1, a_2) + 2(1 - \delta)M \right) \right] \geq u_1(\theta_n, a_1, a_2). \]
This implies that:

$$E\left[\sum_{s=t}^{\tau_{\theta_n}-1} (1 - \delta)^{s-t} \left( u_1(\theta_n, \bar{a}_1, \bar{a}_2) - u_1(\theta_n, \bar{a}_1, \hat{\alpha}_{2,s}) \right) \right] \leq 2M(1 - \delta),$$

which also implies that for every $\theta \in \Theta$,

$$E\left[\sum_{s=t}^{\tau_{\theta_n}-1} (1 - \delta)^{s-t} \left( u_1(\theta, \bar{a}_1, \bar{a}_2) - u_1(\theta, \bar{a}_1, \hat{\alpha}_{2,s}) \right) \right] \leq 2MK(1 - \delta). \quad (C.27)$$

When $\tau > \tau_{\theta_n}$, the induction hypothesis implies that:

$$E\left[\sum_{s=t}^{\tau-1} (1 - \delta)^{s-\tau_0} \left( u_1(\theta, \bar{a}_1, \bar{a}_2) - u_1(\theta, \bar{a}_1, \hat{\alpha}_{2,s}) \right) \right] \leq 2MK(k - 1)(1 - \delta). \quad (C.28)$$

According to (C.27) and (C.28),

$$E\left[\sum_{s=t}^{\tau} (1 - \delta)^{\tau-\tau_0} \left( u_1(\theta, \bar{a}_1, \bar{a}_2) - u_1(\theta, \bar{a}_1, \hat{\alpha}_{2,s}) \right) \right] \leq 2MKk(1 - \delta),$$

which shows (C.25) when $|\Theta_n| = k$. \[C.26\] can be obtained directly from the induction hypothesis.

Next, I examine player 1’s continuation payoff at on-path histories following $(a^{\tau+1}_*, r^{\tau+1}) \in H^\sigma$. I consider three cases:

1. If $r^{\tau+1} \notin \hat{R}^\sigma_g$, by Lemma \[C.8\] then for every $\theta$,
   $$V_\theta(a^{\tau+1}_*, r^{\tau+1}) \geq u_1(\theta, \bar{a}_1, \bar{a}_2) - A(\delta).$$

2. If $r^{\tau+1} \in \hat{R}^\sigma_g$ and $\mu^*(r^s) \in B_\kappa$ for all $r^s$ satisfying $r^s \succ r^{\tau+1}$ and $r^s \in \hat{R}^\sigma_g$, then for every $\theta$,
   $$V_\theta(a^{\tau+1}_*, r^{\tau+1}) \geq u_1(\theta, \bar{a}_1, \bar{a}_2) - B(\delta).$$

3. If there exists $r^s$ such that $\mu^*(r^s) \notin B_\kappa$ with $r^s \succ r^{\tau+1}$ and $r^s \in \hat{R}^\sigma_g$, then repeat the procedure in the beginning of this proof by defining random variables
   - $\tau' : R^\sigma_s \to \{n \in \mathbb{N} \cup \{+\infty\} | n \geq s\}$
   - $\tau_{\theta_n} : R^\sigma_s \to \{n \in \mathbb{N} \cup \{+\infty\} | n \geq s\}$

   similarly as we have defined $\tau$ and $\tau_{\theta_n}$, and then examine continuation payoffs at $r^{\tau'+1}$...

Since $\mu^*(r^{\tau+1}) \in B_{\kappa/2}$ but $\mu^*(r^s) \notin B_\kappa$, then

$$\sum_{\theta \in \Theta} \left(q^*(r^{\tau+1})(\theta) - q^*(r^s)(\theta)\right) \geq \frac{\rho_1(\kappa)}{2}. \quad (C.29)$$

Therefore, such iterations can last for at most $2T_1(\kappa)$ rounds.

Next, I establish the payoff lower bound in case 3. I introduce a new concept called ‘trees’. Let

$$\mathcal{R}_b^\sigma \equiv \left\{r^t | \mu^*(r^t) \notin B_\kappa \text{ and } r^t \in \hat{R}^\sigma_g \right\}$$

Define set $\mathcal{R}^\sigma(k) \subset \mathcal{R}$ for all $k \in \mathbb{N}$ recursively as follows. Let

$$\mathcal{R}_b^\sigma(1) \equiv \left\{r^t | r^t \in \mathcal{R}_b^\sigma \text{ and there exists no } r^s \prec r^t \text{ such that } r^s \in \mathcal{R}_b^\sigma \right\}.$$

For every $r^t \in \mathcal{R}_b^\sigma(1)$, let $\tau[r^t] : \mathcal{R}_b^\sigma \to \mathbb{N} \cup \{+\infty\}$ as the first period $s > t$ (starting from $r^t$) such that either one of the following two conditions is met:

59
1. \( \mu^s(r^{s+1}) \in B_{k/2} \) for \( r^{s+1} \succ r^s \) with \( (a^{s+1}_{s+1}, r^{s+1}) \in \mathcal{H}^s \).

2. \( r^s \notin \mathcal{R}_g^s \).

I call

\[ T(r^t) \equiv \left\{ r^s \bigg| r^{r^t} \succ r^s \succ r^t \right\} \]

a ‘tree’ with root \( r^t \). For any \( k \geq 2 \), let

\( \mathcal{R}^s(k) \equiv \left\{ r^t \bigg| r^t \in \mathcal{R}^s_0, r^t > r^{r^t} \right\} \) for some \( r^s \in \mathcal{R}^s(k-1) \) and there exists no \( r^s < r^t \) that satisfy these two conditions.

Let \( T \) be the largest integer such that \( \mathcal{R}^s(T) \neq \{ \emptyset \} \). According to (C.29), we know that \( T \leq 2T_1(n) \).

Similarly, we can define trees with roots in \( \mathcal{R}(k) \) for every \( k \leq T \).

In what follows, I show that for every \( \theta \) and every \( r^t \in \mathcal{R}^s(k) \),

\[ V_\theta(a^t_s, r^t) \geq u_1(\theta, \pi_1, \pi_2) - (T + 1 - k) \left( A(\delta) + B(\delta) + C(\delta) \right). \]

(C.30)

The proof is done by inducting on \( k \) from backwards. When \( k = T \), player 1’s continuation value at \( (a^{r^t}_s, r^{r^t}) \) is at least \( u_1(\theta, \pi_1, \pi_2) - A(\delta) - B(\delta) \) according to Lemma C.2 and Lemma C.8. His continuation value at \( r^t \) is at least:

\[ u_1(\theta, \pi_1, \pi_2) - A(\delta) - B(\delta) - C(\delta). \]

Suppose the conclusion holds for all \( k \geq n + 1 \), then when \( k = n \), type \( \theta \)’s continuation payoff at \( (a^t_s, r^t) \) is at least:

\[ \mathbb{E}\left[ 1 - \delta^{r^t} - r^{r^t} - t \right] V_\theta(a^{r^t}_s, r^{r^t}) - C(\delta) \]

Pick any \( (a^{r^t}_s, r^{r^t}) \), consider the set of random paths \( r^\infty \) that it is consistent with, let this set be \( \mathcal{R}^\infty(a^{r^t}_s, r^{r^t}) \). Partition it into two subsets:

- \( \mathcal{R}_+^\infty(a^{r^t}_s, r^{r^t}) \) consists of \( r^\infty \) such that for all \( s \geq r^t + 1 \) and \( r^s \prec r^\infty \), we have \( r^s \notin \mathcal{R}_g^s \).

- \( \mathcal{R}_-^\infty(a^{r^t}_s, r^{r^t}) \) consists of \( r^\infty \) such that there exists \( s \geq r^t + 1 \) and \( r^s \prec r^\infty \) at which \( r^s \in \mathcal{R}(n + 1) \).

Conditional on \( r^\infty \in \mathcal{R}_+^\infty(a^{r^t}_s, r^{r^t}) \), we have:

\[ V_\theta(a^{r^t}_s, r^{r^t}) \geq u_1(\theta, \pi_1, \pi_2) - A(\delta) - B(\delta). \]

Conditional on \( r^\infty \in \mathcal{R}_-^\infty(a^{r^t}_s, r^{r^t}) \), type \( \theta \)’s continuation payoff is no less than

\[ V_\theta(a^t_s, r^t) \geq u_1(\theta, \pi_1, \pi_2) - (T - n) \left( A(\delta) + B(\delta) + C(\delta) \right) \]

after reaching \( r^s \in \mathcal{R}(n) \) according to the induction hypothesis. Moreover, since can he lose at most \( A(\delta) + B(\delta) \) before reaching \( r^s \) (according to Lemmas C.2 and C.8), we have:

\[ V_\theta(a^{r^t}_s, r^{r^t}) \geq u_1(\theta, \pi_1, \pi_2) - (T + 1 - n) \left( A(\delta) + B(\delta) + C(\delta) \right). \]

which obtains (C.30). (C.23) is implied by (C.30) since player 1’s loss is bounded above by \( A(\delta) + B(\delta) \) from \( r^0 \) to any \( r^t \in \mathcal{R}(0) \).

Theorem 2 is then implied by the conclusions of Lemma C.8, C.9 and C.10

---

60
D  Proof of Theorem 3 & Extensions

In this Appendix, I show Theorem [3] I also generalize the result by allowing for multiple commitment types. Recall the definitions of $\mathcal{H}^\sigma$ and $\mathcal{H}^\sigma(\omega)$ in the previous section.

D.1  Proof of Theorem 3

Step 1: Let

$$X(h^t) \equiv \mu(\vec{\omega}_1) \mathcal{D}(\phi_{\vec{\omega}_1}, \vec{\omega}_1) + \sum_{\theta \in \Theta_p \cup \Theta_g} q(h^t)(\theta) \mathcal{D}(\theta, \vec{\omega}_1).$$  (D.1)

According to (4.3), $X(h^0) < 0$. Moreover, at every $h^t \in \mathcal{H}^\sigma$ with $X(h^t) < 0$, player 2 has a strict incentive to play $a_2$. Applying Lemma [C.3] there exists $\theta_p \in \Theta_p$ with $h^t \in \mathcal{H}(\theta_p)$ such that type $\theta_p$’s continuation value at $h^t$ is $u_1(\theta_p, a_1, a_2)$, which further implies that always playing $a_1$ is his best reply. According to Lemma [C.1] and using the fact that $X(h^0) < 0$, every $\theta_n \in \Theta_n$ plays $a_1$ for sure at every $h^t \in \mathcal{H}(\theta_n)$.

Step 2: Let us examine the equilibrium behaviors of the types in $\Theta_p \cup \Theta_g$. I claim that for every $h^1 = (\vec{\omega}_1, r^1) \in \mathcal{H}^\sigma$,

$$\sum_{\theta \in \Theta_p \cup \Theta_g} q(h^1)(\theta) \mathcal{D}(\theta, \vec{\omega}_1) < 0.\tag{D.2}$$

Suppose towards a contradiction that $\sum_{\theta \in \Theta_p \cup \Theta_g} q(h^1)(\theta) \mathcal{D}(\theta, \vec{\omega}_1) \geq 0$, then $X(h^1) \geq \mu(\vec{\omega}_1) \mathcal{D}(\phi_{\vec{\omega}_1}, \vec{\omega}_1)$. According to Proposition [C.1] there exists $K \in \mathbb{R}_+$ independent of $\delta$ such that type $\theta$’s continuation payoff is at least $u_1(\theta, \vec{\omega}_1, a_2) - (1 - \delta)K$ at every $h^t \in \mathcal{H}^\sigma$. When $\delta$ is large enough, this contradicts the conclusion in the previous step that there exists $\theta_p \in \Theta_p$ such that type $\theta_p$’s continuation value at $h^0$ is $u_1(\theta_p, a_1, a_2)$, as he can profitably deviate by playing $a_1$ in period 0.

Step 3: According to (D.2), we have $\mu^*(r^1) \notin B_0$. So Lemma [C.6] implies that $r^1 \notin \hat{\mathcal{R}}^\sigma_g$. According to Lemma [C.1] type $\theta_g$ plays $a_1$ at every $h^t \in \mathcal{H}(\theta_g)$ with $t \geq 1$ for every $\theta_g \in \Theta_g$.

Next, I show that $r^0 \notin \hat{\mathcal{R}}^\sigma_g$. Suppose towards a contradiction that $r^0 \in \hat{\mathcal{R}}^\sigma_g$, then there exists $h^T = (a_T^r, r^T) \in \mathcal{H}^\sigma$ such that $\mu(h^T)(\Theta_p \cup \Theta_n) = 0$. If $T \geq 2$, it contradicts our previous conclusion that $r^1 \notin \hat{\mathcal{R}}^\sigma_g$. If $T = 1$, then it contradicts (D.2). Therefore, we have $r^0 \notin \hat{\mathcal{R}}^\sigma_g$ and we have shown that type $\theta_g$ plays $a_1$ at every $h^t \in \mathcal{H}(\theta_g)$ with $t \geq 0$ for every $\theta_g \in \Theta_g$.

Step 4: In the last step, I pin down the strategies of type $\theta_p$ by showing that $X(h^t) = 0$ for every $h^t = (a^t_1, r^t) \in \mathcal{H}^\sigma$ with $t \geq 1$. First, I show that $X(h^1) = 0$. The argument at other histories follows similarly.

Suppose first that $X(h^1) > 0$, then according to Lemma [C.7] type $\theta_p$’s continuation payoff at $(a_T^t, r^t)$ is $u_1(\theta_p, a_1, a_2)$ by always playing $a_1$, while his continuation payoff at $(a_T^t, a_1, r^t)$ is $u_1(\theta_p, a_1, a_2)$, leading to a contradiction. Suppose next that $X(h^1) < 0$, similar to the previous paragraph, there exists type $\theta_p \in \Theta_p$ with $h^1 \in \mathcal{H}(\theta_p)$ such that his incentive constraint is violated. Similarly, one can show that $X(h^t) = 0$ for every $t \geq 1, h^t = (a^t_1, r^t) \in \mathcal{H}^\sigma$. Hence, we have established the uniqueness of player 1’s equilibrium play.

D.2  Generalizations to Multiple Commitment Types

Next, I generalize Theorem [3] by accommodating multiple commitment types. For every $\theta \in \Theta$, let $\lambda(\theta)$ be the likelihood ratio between strategic type $\theta$ and the lowest strategic type $\hat{\theta} \equiv \min \Theta$ and let $\lambda \equiv \{\lambda(\theta)\}_{\theta \in \Theta}$ be the likelihood ratio vector between strategic types. I use this likelihood ratio vector to characterize the sufficient conditions for behavioral uniqueness as the result under multiple commitment type requires that the
total probability of commitment types being small enough. The upper bound of this probability depends on the distribution of strategic types. Let

$$\Omega^\phi \equiv \{\alpha_1 \in \Omega^m \cup \{\pi_1\}|D(\alpha_1, \phi_{\alpha_1}) > 0\}$$

which are the set of commitment types under which player 2 has a strict incentive to play \(\pi_2\).

For every \(t \geq 1\) and \(h^t \in \mathcal{H}^t\), let \(h^t_1 \equiv \{a_{1,0}, \ldots, a_{1,t-1}\}\) be the projection of \(h^t\) on \(x_{s=0}^{t-1} A_{1,s}\). Let \(\mathcal{H}^t_1\) be the set of \(h^t_1\). Let \(\mathcal{H}^t_1 \equiv \{(\pi_1, \ldots, \pi_t), (\alpha_1, \ldots, \alpha_t)\}\). For every probability measure \(\mathcal{P}\) over \(\mathcal{H}\), let \(\mathcal{P}_{1,t}\) be its projection on \(\mathcal{H}^t_1\). Recall that \(\mathcal{P}^\sigma(\theta)\) is the probability measure over \(\mathcal{H}\) under strategy profile \(\sigma\) conditional on player 1 being strategic type \(\theta\). For every \(\gamma \geq 0\) and two Nash equilibria \(\sigma\) and \(\sigma'\), strategic type \(\theta\)'s on-path behavior is \(\gamma\)-close between these equilibria if for every \(t \geq 1\),

$$D_B\left(\mathcal{P}_{1,t}^\sigma(\theta), \mathcal{P}_{1,t}^{\sigma'}(\theta)\right) \leq \gamma,$$

where \(D_B(p, q)\) denotes the Bhattacharyya distance between distributions \(p\) and \(q\).\(^{31}\) If \(\gamma = 0\), then type \(\theta^*_p\)'s on-path behavior is the same between these equilibria. Intuitively, the above distance measures the difference between the \(ex\ ante\) distributions over player 1's action paths across different equilibria. The generalization of Theorem 3 to multiple commitment types is stated below:

**Theorem 3'**. Suppose \(\pi_1 \in \Omega^m\) and \(D(\phi_{\pi_1}, \pi_1) > 0\), then for every \(\lambda \in [0, +\infty)^{|\Theta|}\) satisfying:

$$\sum_{\theta \in \Theta_p \cup \Theta_g} \lambda(\theta) D(\theta, \pi_1) < 0,$$

(D.3)

1. There exist \(\tau > 0\) and a function \(\gamma : (0, \tau) \to \mathbb{R}_+\) satisfying \(\lim_{\epsilon \downarrow 0} \gamma(\epsilon) = 0\), such that for every \(\mu \in \Delta(\Omega)\) with \(\{\mu(\theta)/\mu(\phi)\}_{\theta \in \Theta} = \lambda \) and \(\mu(\Omega^m) < \tau\), there exist \(\delta \in (0, 1)\) and \(\theta^*_p \in \Theta_p\) such that for every \(\sigma \in NE(\delta, \mu, \phi)\) with \(\delta > \delta\):

   - For every \(\theta > \theta^*_p\) and \(h^t \in \mathcal{H}^t(\theta)\), type \(\theta\) plays \(\pi_1\) at \(h^t\).
   - For every \(\theta < \theta^*_p\) and \(h^t \in \mathcal{H}^t(\theta)\), type \(\theta\) plays \(\alpha_1\) at \(h^t\).
   - As for type \(\theta^*_p\),
     - \(\mathcal{P}_{1,t}^\sigma(\theta^*_p)(\mathcal{H}^t_1) > 1 - \gamma(\tau)\) for every \(t \geq 1\).
     - For every \(\sigma \in NE(\delta, \mu, \phi)\), type \(\theta^*_p\)'s on-path behavior is \(\gamma(\tau)\)-close between \(\sigma\) and \(\sigma'\).
     - Furthermore, if there exists no mixed commitment type under which \(\pi_2\) is player 2's strict best reply, then type \(\theta^*_p\)'s on-path behavior is the same across all equilibria under generic parameters. Type \(\theta^*_p\) plays the same action in every period with \(ex\ ante\) probability 1.

   - Strategic type \(\theta\)'s equilibrium payoff is \(v^\theta_0\) for every \(\theta \in \Theta\).

2. If all commitment types are playing pure strategies, then there exists \(\tau > 0\), such that for every \(\mu \in \Delta(\Omega)\) with \(\{\mu(\theta)/\mu(\phi)\}_{\theta \in \Theta} = \lambda \) and \(\mu(\Omega^m) < \tau\), there exist \(\delta \in (0, 1)\) and \(\theta^*_p \in \Theta_p\) such that for every \(\sigma \in NE(\delta, \mu, \phi)\) with \(\delta > \delta\):

   - For every \(\theta > \theta^*_p\) and \(h^t \in \mathcal{H}^t(\theta)\), type \(\theta\) plays \(\pi_1\) at \(h^t\).
   - For every \(\theta < \theta^*_p\) and \(h^t \in \mathcal{H}^t(\theta)\), type \(\theta\) plays \(\pi_1\) at \(h^t\).

\(^{31}\)One can replace the Bhattacharyya distance with the Rényi divergence or Kullback-Leibler divergence in the following way: strategic type \(\theta\)'s on-path behavior is \(\gamma\)-close between \(\sigma\) and \(\sigma'\) if there exists a probability measure \(\mathcal{P}\) on \(\mathcal{H}\) such that for every \(t \geq 1\),

$$\max\left\{D\left(\mathcal{P}_{1,t}||\mathcal{P}_{1,t}^\sigma(\theta)\right), D\left(\mathcal{P}_{1,t}||\mathcal{P}_{1,t}^{\sigma'}(\theta')\right)\right\} \leq \gamma,$$

where \(D(||\cdot||)\) is either the Rényi divergence of order greater than 1 or the Kullback-Leibler divergence.
\begin{itemize}
  \item There exists \( a_1 \in \Delta(\Omega^g \cup \{\bar{a}_1, a_1\}) \) such that for every \( h^t \in \mathcal{H}^\sigma(\theta_p^r) \),
  \begin{itemize}
    \item Type \( \theta_p^r \) plays \( a_1 \) at \( h^0 \).
    \item If \( t \geq 1 \) and there exists \( a_1 \in \Omega^g \cup \{\bar{a}_1, a_1\} \) such that \( a_1 \in h^t \), then type \( \theta_p^r \) plays \( a_1 \) at \( h^t \).
  \end{itemize}
  \item Strategic type \( \theta \)'s equilibrium payoff is \( v_\theta^m \) for every \( \theta \in \Theta \).
\end{itemize}

Let me comment on the conditions in Theorem 3'. First, (D.3) is implied by (4.3) given that \( D(\phi_{\bar{a}_1, \bar{a}_1}) > 0 \). Second, when there are other commitment types under which player 2 has an incentive to play \( \bar{a}_2 \), then payoff and behavior uniqueness require the total probability of commitment types to be small enough. Intuitively, this is because the presence of multiple good commitment types gives the strategic player 1 many good reputations to choose from, which can lead to multiple behaviors and payoffs. An example is shown in Appendix G.5.

Next, I provide a sufficient condition on \( \tau \). Let
\[
Y(h^t) \equiv \mu(h^t)(\bar{a}_1)D(\phi_{\bar{a}_1, \bar{a}_1}) + \sum_{\alpha_1 \in \Omega^g} \mu(h^t)(\alpha_1)D(\phi_{\alpha_1, \alpha_1}) + \sum_{\theta \in \Theta_\mu \cup \Theta_g} \mu(h^t)(\theta)D(\theta, \bar{a}_1),
\]
(D.4)

which is an upper bound on player 2's incentive to play \( \bar{a}_2 \) at \( h^t \). I require \( \tau \) to be small enough such that
\[
\tau \max_{\theta \in \Theta} \{D(\bar{\theta}, \bar{a}_1)\} + (1 - \tau) \sum_{\theta \in \Theta_\mu \cup \Theta_g} \frac{\lambda(\theta)}{\sum_{\theta \in \Theta} \lambda(\theta)} D(\theta, \bar{a}_1) < 0.
\]
(D.5)

Such \( \tau \) exists since \( \sum_{\theta \in \Theta_\mu \cup \Theta_g} \lambda(\theta)D(\theta, \bar{a}_1) < 0 \). Inequality (D.5) implies that \( Y(h^0) < 0 \), which is also equivalent to (4.3) when \( \Omega^g = \{\emptyset\} \).

Third, when there are mixed strategy commitment types, the probabilities with which type \( \theta_p^r \) mixes may not be the same across all equilibria. Intuitively, this is because of two reasons. First, suppose player 2 has no incentive to play \( \bar{a}_2 \) under any mixed commitment type, then given that all strategic types either always plays \( \bar{a}_1 \) or always plays \( a_1 \), player 2's incentive to play \( \bar{a}_2 \) is increasing over time as more \( \bar{a}_1 \) has been observed. As a result, there will be \( T(\delta) \) periods in which player 2 has a strict incentive to play \( \bar{a}_2 \), followed by at most one period in which she is indifferent between \( \bar{a}_2 \) and \( a_2 \), followed by periods in which she has a strict incentive to play \( \bar{a}_2 \), with \( T(\delta) \) and the probabilities with which she mix between \( \bar{a}_2 \) and \( a_2 \) in period \( T(\delta) \) pinned down by type \( \theta_p^r \)'s indifference condition in period 0. Under degenerate parameter values in which there exists an integer \( T \) such that type \( \theta_p^r \) is just indifferent between always playing \( a_1 \) and always playing \( \bar{a}_1 \) when \( \bar{a}_2 \) will be played in the first \( T \) periods, his mixing probability between always playing \( \bar{a}_1 \) and always playing \( a_1 \) is not unique. Nevertheless, when the ex ante probability of \( \Omega^m \) is smaller than \( \epsilon \), his probability of mixing cannot vary by more than \( \gamma(\epsilon) \) even in this degenerate case, with \( \gamma(\cdot) \) diminishes as \( \epsilon \downarrow 0 \). Second, when there are good mixed strategy commitment types, the probability with which type \( \theta_p^r \) behaves inconsistently and builds a reputation for being a good mixed strategy commitment type cannot be uniquely pinned down by his equilibrium payoff. Nevertheless, the differences between these probabilities across different equilibria will vanish as the total probability of commitment types vanishes. Intuitively, this is because if type \( \theta_p^r \) imitates the mixed commitment type with significant probability, then player 2 will have a strict incentive to play \( a_2 \). This implies that as the probability of commitment type vanishes, the probability with which type \( \theta_p^r \) builds a mixed reputation also vanishes.

**D.3 Proof of Theorem 3’**

**Unique Equilibrium Behavior of Strategic Types when \( \theta \in \Theta_1 \cup \Theta_\gamma \).** This part of the proof is similar to the proof of Theorem 3 by replacing \( X(h^t) \) with \( Y(h^t) \). First, I show that every type \( \theta_n \in \Theta_n \) will play \( a_1 \) at every \( h^t \in \mathcal{H}(\theta_n) \) in every equilibrium \( \sigma \). This is similar to Step 1 in the proof of Theorem 3. Since \( Y(h^0) < 0 \) and at every \( h^t \in \mathcal{H}(\theta) \) with \( Y(h^t) < 0 \), player 2 has a strict incentive to play \( a_2 \). Applying Lemma C.3 there exists \( \theta_p \in \Theta_p \) with \( h^t \in \mathcal{H}(\theta_p) \) such that type \( \theta_p \)'s continuation value at \( h^t \) is \( u_1(\theta_p, a_1, a_2) \), and hence always playing \( a_1 \) is his best reply. Type \( \theta_n \)'s on-path behavior is pinned down by Lemma C.1.
Next, I establish (D.2). Suppose towards a contradiction that \( \sum_{q \in \Theta_p} q(h^t)(\theta)D(\theta, \hat{\sigma}_1) = 0 \), then \( Y(h^t) = \mu(\hat{\sigma}_1)D(\phi_{\hat{\sigma}_1}, \hat{\sigma}_1) \). According to Theorem 2 there exists \( K \in \mathbb{R}_+ \) independent of \( \delta \) such that type \( \theta^* \)'s continuation payoff is at least \( u_1(\theta, \sigma_1, a_2) - (1 - \delta) K \) at every \( h^t \in \mathcal{H}_\sigma^t \). When \( \delta \) is large enough, this contradicts the conclusion in the previous step that there exists \( \theta_p \in \Theta_p \) such that type \( \theta_p \)'s continuation value at \( h^t \) is \( u_1(\theta_p, a_1, a_2) \), as he can profitably deviate by playing \( \hat{\sigma}_1 \) in period 0. According to (D.2), we have \( \mu^*(r_1) \notin \mathcal{B}_0 \). Following the same procedure, one can show that \( r_1 \notin \mathcal{R}_\sigma^t \) and \( r_0 \notin \mathcal{R}_\sigma^t \). This implies that for every equilibrium \( \sigma \), type \( \theta_g \) plays \( \hat{\sigma}_1 \) at every \( h^t \in \mathcal{H}_\sigma^t(\theta_g) \) for every \( \theta_g \in \Theta_g \).

Consistency of Equilibrium Behavior and Generic Uniqueness of Equilibrium Payoff when \( \theta \in \Theta_p \): Let \( \Omega^{gm} \) be the set of mixed strategy commitment types under which player 2 has a strict incentive to play \( \sigma_2 \). I show that when \( \Omega^{gm} = \{\emptyset\} \), type \( \theta_p \) has to behave consistently over time for every \( \theta_p \in \Theta_p \). For every \( t \geq 1 \), let \( Z(h^t) = \mu(h^t)(\hat{\sigma}_1)D(\phi_{\hat{\sigma}_1}, \hat{\sigma}_1) + \sum_{\alpha_1 \in \hat{\Omega}^b} q(h^t)(\alpha_1)D(\phi_{\alpha_1}, \alpha_1) + \sum_{\theta \in \Theta_p \cup \Theta_g} q(h^t)(\theta)D(\theta, \hat{\sigma}_1) \) (D.6)

where \( \hat{\Omega}^b \equiv \{\alpha_1 \in \Omega^b \setminus \{\hat{\sigma}_1\} | D(\alpha_1, \phi_{\alpha_1}) < 0\} \). If \( \Omega^{gm} = \{\emptyset\} \), then \( \mu(h^t)(\Omega^g) = 0 \) for every \( h^t = (a^t_1, r^t) \in \mathcal{H}_\sigma^t \). According to (D.2), we have \( \hat{\sigma}_1 \) is consistent over time for every \( \theta_p \in \Theta_p \), whenever \( h^t > h^t \), we have \( Z(h^t) \geq Z(h^t) \).

Subcase 1: No Mixed Commitment Types Consider the case where there exists no \( \alpha_1 \in \hat{\Omega}^b \) such that \( \alpha_1 \notin \{A_1, \emptyset\} \). i.e. there are no mixed strategy commitment types that affect player 2's best reply. By definition, \( Z(h^t) = \mu(h^t)(\hat{\sigma}_1)D(\phi_{\hat{\sigma}_1}, \hat{\sigma}_1) \) for every \( t \geq 1 \). As shown in Theorem 3 we know that \( Z(h^t) = 0 \) for every \( h^t = (a^t_1, r^t) \in \mathcal{H}_\sigma^t \) and \( t \geq 1 \). When \( \Omega^g \neq \{\emptyset\} \), let \( \Omega^g \equiv \{a_1^1, ..., a_1^{n-1}\} \) with \( a_1^1 < a_2^1 < ... < a_1^n < a_1^n = \hat{\sigma}_1 \). There exists \( q : \Theta_p \rightarrow \Delta(\Omega^g \cup \{a_1^1\}) \) such that:

- **Monotonicity:** For every \( \theta_p \succ \theta_p' \) and \( a_1^j \in \Omega^g \cup \{a_1^1\} \). First, if \( q(\theta_p)(a_1^j) > 0 \), then \( q(\theta_p')(a_1^j) = 0 \) for every \( a_1^j > a_1^j \). Second, if \( q(\theta_p')(a_1^j) > 0 \), then \( q(\theta_p)(a_1^j) = 0 \) for every \( a_1^j < a_1^j \).

- **Indifference:** For every \( a_1^j \in \Omega^g \setminus \{a_1^1\} \), we have:

\[
\mu(a_1^j)D(\phi_{a_1^j}, a_1^j) + \sum_{\theta_p \in \Theta_p} \mu(\theta_p)q(\theta_p)(a_1^j)D(\theta_p, a_1^j) = 0. \tag{D.7}
\]

These two conditions uniquely pin down function \( q(\cdot) \), and therefore, the behavior of every type in \( \Theta_p \). In player 1’s unique equilibrium behavior, every strategic type always replicates his action in period 0.

Subcase 2: Presence of Mixed Commitment Types Consider the case where there are mixed strategy commitment types. Player 1’s action path \( a^t = (a_1, ..., a_{t-1}) \) (with \( t \geq 1 \)) is ‘inconsistent’ if there exists no \( a_1 \in \Omega^g \cup \{a_1^1\} \) such that \( a_1 = a_1, ..., a_{t-1} = a_1 \). Otherwise, it is consistent. A history is ‘consistent’ (or inconsistent) if the action path it contains is consistent (or inconsistent). Since \( \Omega^m = \{\emptyset\} \) and the types in \( \Theta_g \) are always playing \( \hat{\sigma}_1 \), so type \( \theta^* \)'s continuation value at every on-path inconsistent history must be \( u_1(\theta, a_1, \hat{\sigma}_2) \) for every \( \theta \in \Theta \).

I show that in every equilibrium, type \( \theta_p \)'s behavior must be consistent for every \( \theta_p \in \Theta_p \). Let

\[
W(h^t) = \mu(h^t)(\hat{\sigma}_1)D(\phi_{\hat{\sigma}_1}, \hat{\sigma}_1) + \sum_{\theta \in \Theta_p \cup \Theta_g} q(h^t)(\theta)D(\theta, \hat{\sigma}_1). \tag{D.8}
\]
For every consistent history $h^t$ where $a_1$ is the consistent action, we know that $W(h^t) \leq Z(h^t)$ since

$$
\sum_{\alpha_1 \in \Omega^b} q(h^t)(\alpha_1)D(\phi_{\alpha_1}, \alpha_1) \leq 0.
$$

As shown in the proof of Theorem 3, we know that $W(h^t) \geq 0$. Moreover, similar argument shows that:

1. If there exists $\alpha_1 \in \hat{\Omega}^b$ such that $\alpha_1(a_1) > 0$, then $W(h^t) > 0$.

2. If there exists no such $\alpha_1$, then $W(h^t) = 0$.

The consistency of type $\theta_p$'s behavior at the 2nd class of consistent histories directly follows from the argument in Theorem[3]. In what follows, I focus on the 1st class of consistent histories.

For every consistent history $h^t$ with $W(h^t) > 0$ and $\mu(h^t)(\Theta_p) \neq 0$, let $\hat{\theta}_p$ be the lowest type in the support of $\mu(h^t)$. According to Lemma[C.3], his expected payoff at any subsequent inconsistent history is $u_1(\hat{\theta}_p, a_1, a_2)$, i.e. playing $a_1$ all the time is his best reply. According to Lemma[C.1], if there exists $\theta_p \in \Theta_p$ playing inconsistently at $h^t$, then type $\hat{\theta}_p$ must be playing consistently at $h^t$ with probability 1.

Suppose type $\theta_p$ plays inconsistently with positive probability at $h^t$ with $Z(h^t) \leq 0$, then his continuation value at $h^t$ is $u_1(\theta_p, a_1, a_2)$. He strictly prefers to deviate and play $a_1$ forever at $h^t - 1 < h^t$ unless there exists $\hat{h}^T > h^t - 1$ such that $Z(h^T) \geq 0$ and type $\theta_p$ strictly prefers to play consistently from $h^t - 1$ to $\hat{h}^T$. This implies that every $\theta_p$ plays consistently with probability 1 from $h^t - 1$ to $\hat{h}^T$, i.e. for every $\hat{h}^T > h^t - 1$ in which type $\theta_p$ plays inconsistently with positive probability and $\hat{h}^T > h^t$, we have $Z(h^t) > Z(\hat{h}^T) \geq 0$. This implies that at $h^t$, type $\hat{\theta}_p$'s continuation payoff by playing consistently until $Z \geq 0$ is strictly higher than behaving inconsistently, leading to a contradiction.

If $h^t$, type $\hat{\theta}_p$ plays inconsistently with positive probability at $h^t$ with $Z(h^t) > 0$, then according to Lemma[C.7], his continuation value by playing consistently is at least $u_1(\hat{\theta}_p, a_1, a_2)$, which is no less than $u_1(\hat{\theta}_p, \bar{a}_1, \bar{a}_2)$, while his continuation value by playing inconsistently is at most $(1 - \delta) u_1(\hat{\theta}_p, a_1, \bar{a}_2) + \delta u_1(\hat{\theta}_p, a_1, \bar{a}_2)$, which is strictly less when $\delta$ is large enough, leading to a contradiction.

Consider generic $\mu$ such that there exist $\hat{\theta}_p \in \Theta_p$ and $q \in (0, 1)$ such that:

$$
\mu(\bar{a}_1)D(\phi_{\bar{a}_1}, \bar{a}_1) + q\mu(\hat{\theta}_p)D(\hat{\theta}_p, \bar{a}_1) + \sum_{\theta \neq \hat{\theta}_p} \mu(\theta)D(\theta, \bar{a}_1) = 0; \tag{D.9}
$$

as well as generic $\delta \in (0, 1)$ such that for every $a_1 \in \Omega^\theta \cup \{\bar{a}_1\}$, there exists no integer $T \in \mathbb{N}$ such that

$$
(1 - \delta^T) u_1(\hat{\theta}_p, a_1, a_2) + \delta^T u_1(\hat{\theta}_p, \bar{a}_1, \bar{a}_2) = u_1(\hat{\theta}_p, \bar{a}_1, \bar{a}_2). \tag{D.10}
$$

Hence, when $\mu(\Omega^m)$ is small enough such that:

$$
\sum_{\theta \neq \hat{\theta}_p} \mu(\theta)D(\theta, \bar{a}_1) + \sum_{\alpha_1 \in \Omega^b} \mu(\Omega^m)D(\phi_{\alpha_1}, \alpha_1) > 0 \tag{D.11}
$$

and

$$
(1 - q)\mu(\hat{\theta}_p)D(\hat{\theta}_p, \bar{a}_1) + \mu(\Omega^m) \max_{\alpha_1 \in \Omega^m} D(\phi_{\alpha_1}, \alpha_1) < 0, \tag{D.12}
$$

one can uniquely pin down the probability with which type $\hat{\theta}_p$ plays $\bar{a}_1$ all the time. To see this, there exists a unique integer $T$ such that:

$$
(1 - \delta^T) u_1(\hat{\theta}_p, \bar{a}_1, \bar{a}_2) + \delta^T u_1(\hat{\theta}_p, \bar{a}_1, \bar{a}_2) > u_1(\hat{\theta}_p, \bar{a}_1, \bar{a}_2) > (1 - \delta^{T+1}) u_1(\hat{\theta}_p, \bar{a}_1, \bar{a}_2) + \delta^{T+1} u_1(\hat{\theta}_p, \bar{a}_1, \bar{a}_2).
$$
The probability with which type $\theta_p^\star$ plays $\bar{a}_1$ all the time, denoted by $q^\star(\bar{a}_1) \in (0, 1)$, is chosen such that:

$$q^\star(\bar{a}_1)\mu(\theta_p^\star)\mathcal{D}(\theta_p^\star, \bar{a}_1) + \sum_{\theta \neq \theta_p^\star} \mu(\theta)\mathcal{D}(\theta, \bar{a}_1) + \sum_{\alpha_1 \in \Omega^m} \mu(\alpha_1) \alpha_1(\bar{a}_1)^T \mathcal{D}(\phi_{\alpha_1}, \alpha_1) = 0.$$

Similarly, the probability with which type $\theta_p^\star$ plays $a_1 \in \Omega^\theta$ all the time, denoted by $q^\star(a_1)$, is pinned down via:

$$q^\star(a_1)\mu(\theta_p^\star)\mathcal{D}(\theta_p^\star, \bar{a}_1) + \sum_{\alpha_1 \in \Omega^m} \mu(\alpha_1)\alpha_1(\bar{a}_1)\mathcal{D}(\phi_{\alpha_1}, \alpha_1) = 0,$$

where $T(a_1)$ is the unique integer satisfying:

$$(1 - \delta^{T(a_1)})u_1(\theta_p^\star, a_1, a_2) + \delta^{T(a_1)} u_1(\theta_p^\star, a_1, \bar{a}_2) > u_1(\theta_p^\star, a_1, a_2) + \delta^{T(a_1)+1} u_1(\theta_p^\star, a_1, a_2).$$

The argument above also pins down every type’s equilibrium payoff: type $\theta \preceq \theta_p^\star$ receives payoff $u_1(\theta, a_1, a_2).$ Every strategic type above $\theta_p^\star$’s equilibrium payoff is pinned down by the occupation measure with which $\bar{a}_2$ is played conditional on player 1 always plays $\bar{a}_1$, which itself is pinned down by type $\theta_p^\star$’s indifference condition.

### $\gamma$-closeness of on-path behavior:
Last, I claim that even when $\Omega^{gm} \neq \{\varnothing\}$, (1) All strategic types besides type $\theta_p^\star$ will either play $\bar{a}_1$ in every period or $a_1$ in every period, (2) strategic type $\theta_p^\star$ will either play $\bar{a}_1$ in every period or $a_1$ in every period with probability at least $1 - \gamma(\bar{a})$; (3) his on-path behavior across different equilibria are $\gamma(\bar{a})$-close, with $\lim_{\tau \to 0} \gamma(\bar{a}) = 0$.

Consider the expressions of $Y(h^t)$ in (D.4) and $Z(h^t)$ in (D.6) which provide upper and lower bounds, respectively, on player 2’s propensity to play $\bar{a}_2$ at $h^t$. When $\mu(\Omega^m) < \bar{a}$, previous arguments imply the existence of $\bar{a}(\bar{a})$ with $\lim_{\tau \to 0} \gamma(\bar{a}) = 0$, such that for every equilibrium,

$$Y(h^t), Z(h^t) \in [-\gamma(\bar{a}), \gamma(\bar{a})]$$

for every $h^t \in \mathcal{H}^{\tau}$ such that $s_1$ has always been played. When $\tau$ is sufficiently small, this implies the existence of $\theta_p^\star \in \Theta_p$ such that type $\theta_p^\star$ mixes between playing $\bar{a}_1$ in every period and playing $a_1$ in every period. This together with Lemma C.1 pins down every other strategic type’s equilibrium behavior aside from type $\theta_p^\star$. Moreover, it also implies that the ex ante probability with which type $\theta_p^\star$ plays $\bar{a}_1$ in every period or plays $a_1$ in every period cannot differ by $2\gamma(\bar{a})/\mu(\theta_p^\star)$ across different equilibria. Furthermore, when $\mu(\Omega^m)$ is small enough, player 2 will have a strict incentive to play $\bar{a}_2$ in period 0 as well as in period $t$ if $a_1$ has always been played in the past. This and type $\theta_p^\star$’s indifference condition pins down every type’s equilibrium payoff.

To show that the probability of type $\theta_p^\star$ behaving inconsistently vanishes with $\mu(\Omega^m)$, notice that first, there exists $s^* \in \mathbb{R}_+$ such that for every $s > s^*$, $\theta_p \in \Theta_p$ and $\alpha_1 \in \Omega^m$,

$$s\mathcal{D}(\theta_p, \bar{a}_1) + \mathcal{D}(\phi_{\alpha_1}, \alpha_1) < 0.$$  
(D.14)

Therefore, the probability with which every type $\theta_p \in \Theta_p$ playing time inconsistently must be below

$$s^*\tau \left\{ \min_{\theta_p \in \Theta_p} \left( 1 - \frac{\lambda(\theta_p)}{\sum_{\theta \in \Theta} \lambda(\theta)} \right) \right\}^{-1}. $$  
(D.15)

Expression (D.15) provides an upper bound for $\gamma(\bar{a})$, which vanishes as $\tau \downarrow 0$. When $\mu(\Omega^m)$ is sufficiently small, Lemma C.1 implies the existence of a cutoff type $\theta_p^\star$ such that all types strictly above $\theta_p^\star$ always plays $\bar{a}_1$ and all types strictly below $\theta_p^\star$ always plays $a_1$, and type $\theta_p^\star$ plays consistently with probability at least $1 - \gamma(\bar{a})$, concluding the proof.
Let \( \sigma = (\sigma_1, \sigma_2) \) be an equilibrium under \( (\tilde{\mu}, \delta) \). Recall the definitions of \( \mathcal{H} \) and \( \mathcal{H}^\sigma \). Since I will only be referring to on-path histories in this proof, I will replace \( \mathcal{H}^\sigma \) with \( \mathcal{H} \) from then on. I start with recursively defining the set of ‘high histories’. Let \( \mathcal{H}_0 \equiv \{ h^0 \} \) and

\[
\bar{h}_1(h^0) \equiv \max \left\{ \bigcup_{\theta \in \Theta} \text{supp}(\sigma_\theta(h^0)) \right\}.
\]

Let

\[
\mathcal{H}_1 \equiv \{ h^1 \} \text{ where there exists } h^0 \in \mathcal{H}_0 \text{ such that } h^1 \succ h^0 \text{ and } \bar{h}_1(h^0) \in h^1 \}.
\]

For every \( t \in \mathbb{N} \) and \( h^t \in \mathcal{H}_t \), let \( \Theta(h^t) \subset \Theta \) be the set of types that occur with positive probability at \( h^t \). Let

\[
\bar{h}_1(h^t) \equiv \max \left\{ \bigcup_{\theta \in \Theta(h^t)} \text{supp}(\sigma_\theta(h^t)) \right\}
\]

and

\[
\mathcal{H}_{t+1} \equiv \{ h^{t+1} \} \text{ where there exists } h^t \in \mathcal{H}_t \text{ such that } h^{t+1} \succ h^t \text{ and } \bar{h}_1(h^t) \in h^{t+1} \}.
\]

Let \( \mathcal{H} \equiv \bigcup_{t=0}^{\infty} \mathcal{H}_t \). For every \( \theta \in \Theta \), let \( \mathcal{H}(\theta) \) be a subset of \( \mathcal{H} \) such that \( h^t \in \mathcal{H}(\theta) \) if and only if:

1. For every \( h^s \succ h^t \) with \( h^s \in \mathcal{H} \), we have \( \theta \in \Theta(h^s) \).

2. If \( h^{t-1} < h^t \), then for every \( \tilde{\theta} \in \Theta(h^{t-1}) \), there exists \( h^s \in \mathcal{H} \) with \( h^s \succ h^{t-1} \) such that \( \tilde{\theta} \notin \Theta(h^s) \).

Let \( \mathcal{H}(\theta) \equiv \bigcup_{\theta \in \Theta} \mathcal{H}(\theta) \), which has the following properties:

1. \( \mathcal{H}(\theta) \subset \mathcal{H} \).

2. For every \( h^t, h^s \in \mathcal{H}(\theta) \), neither \( h^t \succ h^s \) nor \( h^t \prec h^s \).

In what follows, I show the following Lemma:

**Lemma E.1.** For every \( h^t \in \mathcal{H} \), if \( \theta = \max \Theta(h^t) \), then type \( \theta \)’s continuation payoff at \( h^t \) is no more than \( \max \{ u_1(\theta, \bar{a}_1, \bar{a}_2), u_1(\theta, a_1, a_2) \} \).

Lemma E.1 implies the conclusion in Proposition 4.1 as \( h^0 \in \mathcal{H} \) and \( \bar{\theta} = \max \Theta(h^0) \). A useful conclusion to show Lemma E.1 is the following observation:

**Lemma E.2.** For every \( h^t \in \mathcal{H} \), if \( \theta, \bar{\theta} \in \Theta(h^t) \) with \( \bar{\theta} \prec \theta \), then the difference between type \( \theta \)’s continuation payoff and type \( \bar{\theta} \)’s continuation payoff at \( h^t \) is no more than \( u_1(\theta, \bar{a}_1, \bar{a}_2) - u_1(\bar{\theta}, \bar{a}_1, \bar{a}_2) \).

**Proof of Lemma E.2:** Since \( u_1 \) has SID in \( \theta \) and \( (a_1, a_2) \), so for every \( \theta \succ \bar{\theta} \),

\[
(\bar{a}_1, \bar{a}_2) \in \arg \max_{(a_1, a_2)} \left\{ u_1(\theta, a_1, a_2) - u_1(\bar{\theta}, a_1, a_2) \right\}
\]

which yields the upper bound on the difference between type \( \theta \) and type \( \bar{\theta} \)’s continuation payoffs. \( \square \)

For every \( h^t \in \mathcal{H}(\bar{\theta}) \), at the subgame starting from \( h^t \), type \( \bar{\theta} \)’s stage game payoff is no more than \( u_1(\bar{\theta}, \bar{a}_1, \bar{a}_2) \) in every period and his continuation payoff at \( h^t \) cannot exceed \( u_1(\bar{\theta}, \bar{a}_1, \bar{a}_2) \). This is because
$\bar{a}_1$ is type $\bar{\theta}$'s Stackelberg action, so whenever player 1 plays an action $a_1 \prec \bar{a}_1$, $a_2$ is player 2's strict best reply. Lemma E.2 then implies that for every $\theta \in \Theta(h^t)$ with $\theta \succ \bar{\theta}$, type $\bar{\theta}$'s continuation payoff at $h^t$ cannot exceed $u_1(\theta, \bar{a}_1, \bar{a}_2)$.

In what follows, I prove Lemma E.1 by induction on $|\Theta(h^t)|$. When $|\Theta(h^t)| = 1$, i.e. there is only one type (call it type $\theta$) that can reach $h^t$ according to $\sigma$, then Lemma E.2 implies that type $\theta$'s payoff cannot exceed $u_1(\theta, \bar{a}_1, \bar{a}_2)$.

Suppose the conclusion in Lemma E.1 holds for every $|\Theta(h^t)| \leq n$, consider the case when $|\Theta(h^t)| = n+1$. Let $\theta \equiv \max \Theta(h^t)$. Let me introduce set $\overline{H}^B(h^t)$, which is a subset of $\overline{H}$. For every $h^s \succ h^t$ with $h^s \in \overline{H}$, $h^s \in \overline{H}^B(h^t)$ if and only if:

- $h^s \in \overline{H}(\theta)$,
- but $h^{s+1} \notin \overline{H}(\theta)$ for any $h^{s+1} \succeq h^s$ with $h^{s+1} \in \overline{H}$.

In another word, type $\theta$ has a strict incentive not to play $\bar{a}_1(h^s)$ at $h^s$. A useful property is:

- For every $h^\infty \in \overline{H}$ with $h^\infty \succ h^t$, either there exists $h^s \in \overline{H}^B(h^t)$ such that $h^s \prec h^\infty$, or there exists $h^s \in \overline{H}(\theta)$ such that $h^s \prec h^\infty$.

which means that play will eventually reach either a history in $\overline{H}^B(h^t)$ or $\overline{H}(\theta)$ if type $\theta$ keeps playing $\bar{a}_1(h^\tau)$ before that for every $t \leq \tau \leq s$.

In what follows, I examine type $\theta$'s continuation value at each kind of history.

1. For every $h^s \in \overline{H}^B(h^t)$, at $h^{s+1} \succ h^s$ and $h^{s+1} \in \overline{H}$, by definition,

   $|\Theta(h^{s+1})| \leq n$.

   Let $\tilde{\theta} \equiv \max \Theta(h^{s+1})$. By induction hypothesis, type $\tilde{\theta}$'s continuation payoff at $h^{s+1}$ is at most $u_1(\tilde{\theta}, \bar{a}_1, \bar{a}_2)$. This applies to every such $h^{s+1}$.

Type $\tilde{\theta}$'s continuation value at $h^s$ also cannot exceed $u_1(\tilde{\theta}, \bar{a}_1, \bar{a}_2)$ since he is playing $\bar{a}_1(h^s)$ with positive probability at $h^s$, and his stage game payoff from doing so is at most $u_1(\tilde{\theta}, \bar{a}_1, \bar{a}_2)$. Furthermore, his continuation value afterwards cannot exceed $u_1(\tilde{\theta}, \bar{a}_1, \bar{a}_2)$.

Lemma E.2 then implies that type $\theta$'s continuation value at $h^s$ is at most $u_1(\theta, \bar{a}_1, \bar{a}_2)$.

2. For every $h^s \in \overline{H}(\theta)$, always playing $\bar{a}_1(h^\tau)$ for all $h^\tau \succ h^s$ and $h^\tau \in \overline{H}$ is a best reply for type $\theta$. His stage game payoff from this strategy cannot exceed $u_1(\theta, \bar{a}_1, \bar{a}_2)$, which implies that his continuation value at $h^s$ also cannot exceed $u_1(\theta, \bar{a}_1, \bar{a}_2)$.

Starting from $h^t$ consider the strategy in which player 1 plays $\bar{a}_1(h^\tau)$ at every $h^\tau \succ h^t$ and $h^\tau \in \overline{H}$ until play reaches $h^s \in \overline{H}^B(h^t)$ or $h^s \in \overline{H}(\theta)$. By construction, this is type $\theta$'s best reply. Under this strategy, type $\theta$'s stage game payoff cannot exceed $u_1(\theta, \bar{a}_1, \bar{a}_2)$ before reaches $h^s$. Moreover, his continuation payoff after reaching $h^s$ is also bounded above by $u_1(\theta, \bar{a}_1, \bar{a}_2)$, which proves Lemma E.1 when $|\Theta(h^t)| = n+1$.

F Proof of Proposition 4.2

Throughout the proof, I normalize $u_1(\theta, a_1, a_2) = 0$ for every $\theta$. Let $x_\theta(a_1) \equiv -u_1(\theta, a_1, a_2)$ and $y_\theta(a_1) \equiv u_1(\theta, a_1, \bar{a}_2)$. Assumptions 1 and 2 imply that:

- $x_\theta(a_1) \geq 0$, with "=" holds only when $a_1 = \bar{a}_1$.
- $y_\theta(a_1) > 0$ for every $\theta \in \Theta$ and $a_1 \in A_1$. 

68
• $x_\theta(a_1)$ and $-y_\theta(a_1)$ are both strictly increasing in $a_1$.

• For every $\theta < \bar{\theta}$, $x_\theta(a_1) - x_\bar{\theta}(a_1)$ and $y_\theta(a_1) - y_\bar{\theta}(a_1)$ are both strictly increasing in $a_1$.

I start with defining ‘pessimistic belief path’ for every $\sigma \in \text{NE}(\delta, \hat{\mu})$. For every $a_1^\infty \equiv (a_{1,0}, a_{1,1}, ..., a_{1,t}, ...)$, we say $a_1^\infty \in \mathcal{A}_\sigma(\theta^*)$ if and only if for every $t \in \mathbb{N}$, there exists $r^t \in \mathcal{R}^t$ such that $(a_{1,0}, ..., a_{1,t-1}, r^t) \in \mathcal{H}^\sigma$ and $\sum_{\theta \geq \theta} \mu_t(\theta)D(\theta, \bar{\alpha}_1) < 0$, where $\mu_t$ is player $2$’s belief after observing $(a_{1,0}, ..., a_{1,t-1})$. For any $\theta \in \Theta$, if $(a_1^\infty, r^\infty) \in \mathcal{H}^\sigma(\theta)$ for some $a_1^\infty \in \mathcal{A}_\sigma(\theta)$, then $V_\theta^\sigma(\delta) = 0$ and always playing $\bar{\alpha}_1$ is type $\theta$’s best reply.

For every $\hat{\mu}$ satisfying (4.13), there exist a unique $\theta_p^* \in \Theta_p$ and a unique $q(\hat{\mu}) \in [0, 1)$ such that:

$$q(\hat{\mu}) \hat{m}(\theta_p^*)D(\theta_p^*, \bar{\alpha}_1) + \sum_{\theta > \theta_p^*} \hat{m}(\theta)D(\theta, \bar{\alpha}_1) = 0. \quad (F.1)$$

When $\hat{\mu}$ satisfies (4.13), then for every $\sigma \in \text{NE}(\delta, \hat{\mu})$ and $\theta \preceq \theta_p^*$, since $\sum_{\theta \geq \theta} \hat{m}(\theta)D(\theta, \bar{\alpha}_1) < 0$, using the martingale property of beliefs, we know that there exists $a_1^\infty \in \mathcal{A}_\sigma(\theta)$ such that $(a_1^\infty, r^\infty) \in \mathcal{H}^\sigma(\theta)$ for some $r^\infty$. This pins down the unique equilibrium payoff for all $\theta \preceq \theta_p^*$.

In what follows, I establish the upper bound in Proposition 4.2. For every $\theta > \theta_p^*$, in every action path $a_1^\infty = (a_{1,0}, a_{1,1}, ...)$ which type $\theta$ follows with strictly positive probability under $\sigma \in \text{NE}(\delta, \hat{\mu})$, it must be that:

$$V_\theta^\sigma(\delta) = \sum_{a_1, a_2} P^a_1(a_1, a_2)u_1(\theta, a_1, a_2)$$

and

$$0 = V_\theta^\sigma(\delta) \geq \sum_{a_1, a_2} P^a_1(a_1, a_2)u_1(\theta_p^*, a_1, a_2)$$

where

$$P^a_1(a_1, a_2) \equiv \sum_{t=0}^{\infty} (1 - \delta)\delta^t P^{(\sigma_1, \sigma_2)}(a_1, a_2)$$

with $\sigma_1$ playing $a_1^\infty$ on the equilibrium path. Therefore, $V_\theta^\sigma(\delta)$ must be weakly below the value of the following linear program:

$$\max_{\{\beta(1_1), \gamma(1_1)\}a_1 \in A_1} \left\{ \sum_{a_1 \in A_1} \beta(1_1)y_\theta(1_1) - \gamma(1_1)x_\theta(1_1) \right\}, \quad (F.2)$$

subject to

$$\sum_{a_1 \in A_1} \gamma(1_1) + \beta(1_1) = 1, \quad \gamma(1_1), \beta(1_1) \geq 0 \text{ for every } a_1 \in A_1,$$

and

$$\sum_{a_1 \in A_1} \beta(1_1)y_\theta(1_1) - \gamma(1_1)x_\theta(1_1) \leq 0. \quad (F.3)$$

Due to the linearity of this program, it is without loss of generality to focus on solutions where there exist $a_1^*$ and $a_1^{**}$ such that

$$\beta(1_1) > 0 \text{ if } a_1 = a_1^*, \quad \gamma(1_1) > 0 \text{ if } a_1 = a_1^{**}.$$  

According to (F.3), we have:

$$\beta(a_1^*)y_\theta^*(a_1^*) \leq (1 - \beta(a_1^*))x_\theta^*(a_1^{**}). \quad (F.4)$$

Plugging (F.4) into (F.2), the value of that expression cannot exceed:

$$\max_{a_1, a_1^{**} \in A_1} \left\{ \frac{y_\theta(a_1^*)x_\theta^*(a_1^{**}) - x_\theta(a_1^{**})y_\theta(a_1^*)}{x_\theta^*(a_1^*) + y_\theta(a_1^*)} \right\}. \quad (F.5)$$
Expression (F.5) is maximized when \( a_1^* = a_2^* = \bar{\alpha}_1 \), which gives an upper bound for \( V_\theta^u(\delta) \):

\[
V_\theta^u(\delta) \leq ru_1(\theta, \bar{\alpha}_1, \bar{\alpha}_2) + (1 - r)u_1(\theta, \bar{\alpha}_1, \bar{\alpha}_2),
\]

with \( r \in (0, 1) \) satisfying: \( ru_1(\theta^*_p, \bar{\alpha}_1, \bar{\alpha}_2) + (1 - r)u_1(\theta^*_p, \bar{\alpha}_1, \bar{\alpha}_2) = u_1(\theta^*_p, \bar{\alpha}_1, \bar{\alpha}_2) \). The upper bound in (F.6) is asymptotically achieved when \( \delta \to 1 \) in an equilibrium where:

- Type \( \theta \) always plays \( \bar{\alpha}_1 \) if \( \theta < \theta^*_p \), always plays \( \bar{\alpha}_1 \) if \( \theta > \theta^*_p \).
- Type \( \theta^*_p \) randomizes between always playing \( \bar{\alpha}_1 \) and always playing \( \bar{\alpha}_1 \) with prob \( q(\hat{\mu}) \) and \( 1 - q(\hat{\mu}) \).

### G Counterexamples

I present several counterexamples missing from the main text. For convenience, there is only one commitment type in every example besides the one in Appendix G.5. The intuition behind the examples still apply when there are multiple pure strategy commitment types. Abusing notation, I use \( \theta \) to denote the Dirac measure on \( \theta \) and \( a_i \) to denote the Dirac measure on \( a_i \) with \( i \in \{1, 2\} \).

#### G.1 Failure of Reputation Effects When Supermodularity is Violated

**Example 1:** To begin with, I construct equilibrium in the entry deterrence game of Harrington (1986) in which the supermodularity condition on \( u_1 \) is violated:\(^{32}\) Let the stage game payoff be:

<table>
<thead>
<tr>
<th>( \theta_1 )</th>
<th>O</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>1, 0</td>
<td>-1, -1</td>
</tr>
<tr>
<td>A</td>
<td>2, 0</td>
<td>0, 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \theta_0 )</th>
<th>O</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>5/2, 0</td>
<td>1/2, 1/2</td>
</tr>
<tr>
<td>A</td>
<td>3, 0</td>
<td>1, 3/2</td>
</tr>
</tbody>
</table>

One can verify that the monotonicity condition is satisfied. To see this game’s payoff fails the supermodularity assumption, let us rank the state and players’ actions via \( \theta_1 > \theta_0 \), \( F > A \) and \( O > I \). Player 1’s cost of fighting is 1 in state \( \theta_1 \) and is 1/2 in state \( \theta_0 \). Intuitively, when the incumbent’s and the entrant’s costs are positively correlated, the incumbent’s loss from fighting (by lowering prices) is higher when his cost is high, and the entrant’s profit from entry decreases with the cost and increases with the incumbent’s price.

When \( \Omega = \{F, \theta_1, \theta_0\} \) and \( \mu(F) \leq 2\mu(\theta_0) \), I construct an equilibrium with low payoffs in each of the following three cases, depending on the signs of:

\[
X \equiv \frac{\mu(\theta_0)}{2} + \left( \frac{\mu(F)\phi_F(\theta_0)}{2} - \mu(F)\phi_F(\theta_1) \right) - \mu(\theta_1)
\]

and

\[
Y \equiv \frac{\mu(F)\phi_F(\theta_0)}{2} - \mu(F)\phi_F(\theta_1).
\]

1. If \( X \leq 0 \), then type \( \theta_0 \) always plays \( F \), type \( \theta_1 \) mixes between always playing \( F \) and always playing \( A \), with the probability of playing \( F \) being \( 1 + X/\mu(\theta_1) \). Player 2 plays \( I \) for sure in period 0. Starting from period 1, she plays \( I \) for sure if \( A \) has been observed before and plays \( \frac{1}{4\delta}O + (1 - \frac{1}{4\delta})I \) otherwise. Despite the probability of type \( \theta_1 \) is large relative to that of type \( \theta_0 \), type \( \theta_1 \)'s equilibrium payoff is 0 and type \( \theta_0 \)'s equilibrium payoff is 3/2, both are lower than their commitment payoffs from playing \( F \).

2. If \( X > 0 \) and \( Y \leq 0 \), then type \( \theta_1 \) always plays \( A \), type \( \theta_0 \) mixes between always playing \( F \) and always playing \( A \), with the probability of playing \( F \) being \( -Y/\mu(\theta_0) \). Player 2 plays \( I \) for sure in period 0. Starting from period 1, she plays \( I \) for sure if \( A \) has been observed before and plays \( \frac{1}{4\delta}O + (1 - \frac{1}{4\delta})I \) otherwise. Type \( \theta_1 \)'s equilibrium payoff is 0 and type \( \theta_0 \)'s equilibrium payoff is 1.

3. If \( X > 0 \) and \( Y > 0 \), then both types always play \( A \). Player 2 plays \( I \) no matter what. Type \( \theta_1 \)'s equilibrium payoff is 0 and type \( \theta_0 \)'s equilibrium payoff is 1.

\(^{32}\)The case in which \( u_2 \) has decreasing differences between \( a_2 \) and \( \theta \) is similar once we reverse the order on the states.
Example 2: Next, I construct equilibrium in the entry deterrence game when the supermodularity condition on \( u_2 \) is violated. I focus on the case in which \( u_2 \) has decreasing differences between \( a_2 \) and \( a_1 \) \(^{33}\). Consider the following \( 2 \times 2 \times 2 \) game with payoffs given by:

<table>
<thead>
<tr>
<th>( \theta = \theta_1 )</th>
<th>( h )</th>
<th>( l )</th>
<th>( \theta = \theta_0 )</th>
<th>( h )</th>
<th>( l )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H )</td>
<td>1, -1</td>
<td>-1, 0</td>
<td>( H )</td>
<td>1 - ( \eta ), -2</td>
<td>-1 - ( \eta ), 0</td>
</tr>
<tr>
<td>( L )</td>
<td>2, 1</td>
<td>0, 0</td>
<td>( L )</td>
<td>2, -1</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

with \( \eta \in (0, 1) \). The states and players’ actions are ranked according to \( H > L, h > l \) and \( \theta_1 > \theta_0 \). Let \( \Omega = \{H, \theta_1, \theta_0\} \). Theorem 2 \(^{33}\) trivially applies as the commitment outcome \((H, l)\) gives every type his lowest feasible payoff. In what follows, I show the failure of Theorem 3 i.e. player 1 has multiple equilibrium behaviors. First, there exists an equilibrium in which \((L, h)\) is always played or \((L, l)\) is always played, depending on the prior belief. Second, consider the following equilibrium:

- In period 0, both strategic types play \( L \).
- From period 1 to \( T(\delta) \in \mathbb{N} \), type \( \theta_0 \) plays \( L \) and type \( \theta_1 \) plays \( H \). Player 2 plays \( h \) in period \( t(\geq 2) \) if and only if \( t \geq T(\delta) + 1 \) and player 1’s past play coincides with type \( \theta_1 \)’s equilibrium strategy. The integer \( T(\delta) \) is chosen such that:

\[
(1 - \delta^{T(\delta)}(-1) + 2\delta^{T(\delta)} > 0 > (1 - \delta^{T(\delta)})(-1 - \eta) + 2\delta^{T(\delta)},
\]

which exists when \( \delta \) is close enough to 1.

G.2 Failure of Reputation Effects When Monotonicity is Violated

I show that the monotonicity condition is indispensable for my reputation result. For this purpose, I consider two counterexamples in which Assumption 1 is violated in different ways.

Example 1: Consider the following \( 2 \times 2 \times 2 \) game:

<table>
<thead>
<tr>
<th>( \theta = \theta_1 )</th>
<th>( h )</th>
<th>( l )</th>
<th>( \theta = \theta_0 )</th>
<th>( h )</th>
<th>( l )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H )</td>
<td>3/2, 2</td>
<td>0, 0</td>
<td>( H )</td>
<td>-1, -1/2</td>
<td>1, 0</td>
</tr>
<tr>
<td>( L )</td>
<td>1, 1</td>
<td>0, 0</td>
<td>( L )</td>
<td>0, -1</td>
<td>5/2, 1/4</td>
</tr>
</tbody>
</table>

One can verify that this game satisfies the supermodularity assumption once we rank the states and actions according to \( \theta_1 > \theta_0, H > L \) and \( h > l \) \(^{34}\). However, the monotonicity assumption is violated as player 1’s ordinal preferences over \( a_1 \) and \( a_2 \) depend on the state.

Suppose \( \Omega = \{H, \theta_1, \theta_0\} \) with \( 4\mu(H) < \mu(\theta_0) \). Consider the following equilibrium in which player 2 plays a ‘tit-for-tat’ like strategy. Type \( \theta_1 \) plays \( L \) all the time and type \( \theta_0 \) plays \( H \) all the time. Starting from period 1, player 2 plays \( h \) in period \( t \geq 1 \) if \( L \) was played in period \( t - 1 \) and vice versa. Both types’ equilibrium payoffs are close to 1, which are strictly lower than their pure Stackelberg commitment payoffs, which are 3/2 and 5/2 respectively.

To verify that this is an equilibrium when \( \delta \) is high enough, notice that first, player 2’s incentive constraints are always satisfied. As for player 1, if \( \theta = \theta_1 \), deviating for one period gives him a stage game payoff at most 3/2 and in the next period his payoff is at most 0. If \( \delta > 1/2 \), then he has no incentive to deviate. if \( \theta = \theta_0 \),

\(^{33}\) The case in which \( u_2 \) has decreasing differences between \( a_2 \) and \( \theta \) is similar to the previous example. One only needs to reverse the order between the states.

\(^{34}\) In fact, the game’s payoffs even satisfy a stronger notion of complementarity, that is, both \( u_1 \) and \( u_2 \) are strictly supermodular functions of the triple \((\theta, a_1, a_2)\). The definition of supermodular function can be found in Topkis (1998).
deviating for one period gives him a stage game payoff at most $5/2$ and in the future, he will keep receiving payoff at most $0$ until he plays $H$ for one period. He has no incentive to deviate if and only if for every $t \in \mathbb{N}$,

$$(1 - \delta) \frac{5}{2} - (\delta^t - \delta^{t+1}) \leq 1 - \delta^{t+1}. \quad \text{(G.3)}$$

which is equivalent to:

$$\frac{5}{2} \leq 1 + \delta + \ldots + \delta^{t-1} + 2\delta^t.$$

The above inequality is satisfied for every integer $t \geq 1$ when $\delta > 0.9$. This is because when $t \geq 2$, the right hand side is at least $1 + 0.9 + 0.9^2$, which is greater than $5/2$. When $t = 1$, the right hand side equals to $2.8$, which is greater than $5/2$.

To see that player 1’s equilibrium behavior is not unique, consider another equilibrium in which type $\theta_1$ always plays $H$, type $\theta_0$ always plays $L$ and for every $t \in \mathbb{N}$, player 2 plays $h$ in period $t$ if $H$ is played in period $t - 1$, and plays $l$ in period $t$ if $L$ is played in period $t - 1$. This implies that the conclusion in Theorem 3 will fail in absence of the monotonicity assumption.

**Example 2:** Low payoff equilibria can be constructed when player 1’s ordinal preference over each player’s actions does not depend on the state, but the directions of monotonicity violate Assumption 1. For example, consider the following game:

<table>
<thead>
<tr>
<th>$\theta = \theta_1$</th>
<th>$h$</th>
<th>$l$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H$</td>
<td>2,2</td>
<td>0,0</td>
</tr>
<tr>
<td>$L$</td>
<td>1,1</td>
<td>-1/2,0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\theta = \theta_0$</th>
<th>$h$</th>
<th>$l$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H$</td>
<td>1/4, -1/2</td>
<td>1/8, 0</td>
</tr>
<tr>
<td>$L$</td>
<td>0, -1</td>
<td>-1/16, 1/4</td>
</tr>
</tbody>
</table>

Both players’ payoffs are supermodular functions of $(\theta, a_1, a_2)$. Player 1’s ordinal preferences over $a_1$ and $a_2$ are state independent but his payoff is strictly increasing in both $a_1$ and $a_2$, which is different from what Assumption 1 suggests. Rank the states and actions according to $\theta_1 \succ \theta_0$, $H \succ L$ and $h \succ l$.

Suppose $\Omega = \{H, \theta_1, \theta_0\}$ with $4\mu(H) < \mu(\theta_0)$. The following strategy profile is an equilibrium. Type $\theta_1$ plays $L$ all the time and type $\theta_0$ plays $H$ all the time. Starting from period 1, player 2 plays $h$ in period $t \geq 1$ if $L$ was played in period $t - 1$ and vice versa. Type $\theta_1$ and type $\theta_0$’s equilibrium payoffs are close to 1 and $1/8$, respectively as $\delta \to 1$. Their pure Stackelberg commitment payoffs are 2 and $1/4$, respectively, which are strictly higher. The verification of players’ incentive constraints is the same as the previous example.

Moreover, contrary to what Theorem 3 has suggested, player 1’s equilibrium behavior is not unique even when player 2’s prior belief is pessimistic, i.e.

$$2\mu(\theta_1) + \mu(H)(2\phi_H(\theta_1) - \frac{1}{2}\phi_H(\theta_0)) - \frac{1}{2}\mu(\theta_0) < 0. \quad \text{(G.4)}$$

This is because aside from the equilibrium constructed above, there also exists an equilibrium in which type $\theta_1$ always plays $H$, type $\theta_0$ mixes between always playing $H$ and always playing $L$ with probabilities such that player 2 becomes indifferent between $h$ and $l$ starting from period 1 conditional on $H$ has always been played. In equilibrium, player 2 plays $h$ in period $t \geq 1$ as long as player 1 has always played $H$ before, and switches to $l$ permanently otherwise.

**G.3 Failure of Reputation Effects When $|A_2| \geq 3$**

I present an example in which the reputation results in Theorems 2 and 3 fail when the stage game has MSM payoffs but player 2 has three or more actions. This motivates the additional conditions on the payoff structure in Online Appendix D. Consider the following $2 \times 2 \times 3$ game with payoffs:
\[ \begin{array}{c|ccc}
\theta = \theta_1 & l & m & r \\
\hline
H & 0,0 & 5/2,2 & 6,3 \\
L & \epsilon,0 & 5/2+\epsilon,-1 & 6+\epsilon,-2 \\
\end{array} \]

\[ \begin{array}{c|ccc}
\theta = \theta_0 & l & m & r \\
\hline
H & 0,0 & 2,-1 & 3,-2 \\
L & 2\epsilon,0 & 2+2\epsilon,-2 & 3+2\epsilon,-3 \\
\end{array} \]

where \( \epsilon > 0 \) is small enough. Let the rankings on actions and states be \( H > L, r > m > l \) and \( \theta_1 > \theta_0 \). One can check that the stage game payoffs are MSM.

Suppose \( \Omega = \{ \theta_1, \theta_0, H \} \) with \( \mu(\theta_0) = 2\eta, \mu(H) = \eta \) and \( \phi_H = \theta_1 \), with \( \eta \in (0, 1/3) \). Type \( \theta_1 \)'s commitment payoff from playing \( H \) is 6. However, consider the following equilibrium:

- Type \( \theta_0 \) plays \( H \) all the time. Type \( \theta_1 \) plays \( L \) from period 0 to \( T(\delta) \) and plays \( H \) afterwards, with \( 1 - \delta^{T(\delta)} \in (1/2 - \epsilon, 1/2 + \epsilon) \). Such \( T(\delta) \in \mathbb{N} \) exists when \( \delta > 1 - 2\epsilon \).

- Player 2 plays \( m \) starting from period 1 if player 1 has always played \( H \) in the past. She plays \( r \) from period 1 to \( T(\delta) \) and plays \( r \) afterwards if player 1’s past actions are consistent with type \( \theta_1 \)'s equilibrium strategy. She plays \( l \) at every off-path history.

Type \( \theta_1 \)'s equilibrium payoff is approximately \( 3 + \epsilon/2 \) as \( \delta \to 1 \), which is strictly less than his commitment payoff. To see that player 1 has multiple equilibrium behaviors under a pessimistic prior belief, i.e. \( \eta \in [1/4, 1/3] \), there exists another equilibrium in which all types of player 1 plays \( H \) at every on-path history. Player 2 plays \( m \) if all past actions were \( H \) and plays \( l \) otherwise.

### G.4 Time Inconsistent Equilibrium Plays in Private Value Reputation Game

I construct an equilibrium in the private value product choice game in which despite there is a commitment type that always exerts high effort, the strategic long-run player abandons his reputation early on in the relationship and \( L \) is played with significant probability. The game’s payoff matrix is given by:

\[ \begin{array}{c|cc}
- & C & S \\
\hline
H & 1,3 & -1,2 \\
L & 2,0 & 0,1 \\
\end{array} \]

Suppose there is a commitment type that always plays \( H \) (which is unlikely compared to the strategic type) and consider the following equilibrium when \( \delta > 1/2 \):

- The strategic type plays \( L \) for sure in period 0. He plays \( \frac{1}{2}H + \frac{1}{2}L \) starting from period 1.

- Player 2 plays \( S \) for sure in period 0. If \( H \) is observed in period 0, then she plays \( C \) for sure as long as \( H \) has always been played. She plays \( S \) for sure in all subsequent periods if \( L \) has been played before.

  If \( L \) is observed in period 0, \( C \) is played for sure in period 1. Starting from period 2, player 2 plays \( C \) for sure in period \( t \) if \( H \) was played in period \( t - 1 \), and \( (1 - \frac{1}{2\delta})C + \frac{1}{2\delta}S \) in period \( t \) if \( L \) was played in period \( t - 1 \).

It is straightforward to verify players’ incentive constraints. Intuitively, starting from period 1, every time player 1 shirks, he will be punished tomorrow as player 2 will play \( C \) with less probability. The probabilities with which he mixes between \( H \) and \( L \) are calibrated to provide player 2 the incentive to mix between \( C \) and \( S \). Despite the strategic type obtains equilibrium payoff \( \delta \), which is close to his pure Stackelberg commitment payoff given that he is sufficiently patient. However, the strategic long-run player’s equilibrium play is very different from the commitment type’s. Perhaps more surprisingly, (i) imitating the commitment type is a strictly dominated strategy, which yields payoff \(- (1 - \delta) + \delta \), strictly less than his equilibrium payoff; (ii) evaluating the occupation measure of every action ex ante, \( L \) is played with significant probability. On average, \( L \) is played with occupation measure strictly more than \( 1/2 \), which converges to \( 1/2 \) as \( \delta \to 1 \).
G.5 Low Probability of Commitment Type for Behavioral Uniqueness

The following example illustrates why $\mu(\Omega^m)$ being small is not redundant for Theorem 3’ when there exists $\alpha_1 \in \Omega^m \setminus \{\pi_1\}$ such that $\{\pi_2\} = BR_2(\alpha_1, \phi_{\alpha_1})$. Consider the following $3 \times 2 \times 2$ stage game:

<table>
<thead>
<tr>
<th>$\theta = 0_1$</th>
<th>$C$</th>
<th>$S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H$</td>
<td>1,2</td>
<td>$-2,0$</td>
</tr>
<tr>
<td>$M$</td>
<td>2,1</td>
<td>$-1,0$</td>
</tr>
<tr>
<td>$L$</td>
<td>3,-1</td>
<td>0,0</td>
</tr>
<tr>
<td>$\theta = 0_0$</td>
<td>$C$</td>
<td>$S$</td>
</tr>
<tr>
<td>$H$</td>
<td>$1/2, -1$</td>
<td>$-5/2,0$</td>
</tr>
<tr>
<td>$M$</td>
<td>$3/2, -2$</td>
<td>$-3/2,0$</td>
</tr>
<tr>
<td>$L$</td>
<td>3,-3</td>
<td>0,0</td>
</tr>
</tbody>
</table>

Let $\Omega \equiv \{H, M, 0_1, 0_0\}$ with $\mu(H) = \mu(0_1) = 1/20$, $\mu(0_0) = 3/10$ and $\mu(M) = 3/5$. Let $\phi_H = \phi_M$ be the Dirac measure on $0_1$. One can check that $M \in \Omega^0$ and $\mu$ satisfies (4.3). However, for every $\delta > 5/6$, one can construct the following class of equilibria indexed by $T \in \{1, 2, \ldots\}$:

- **Equilibrium $\sigma^T$:** Type $0_0$ plays $M$ forever. Type $0_1$ plays $M$ from period 0 to period $T$, and plays $H$ starting from period $T + 1$. Player 2 plays $C$ for sure from period 0 to $T + 1$ if player 1’s past actions were either all $H$ or all $M$. For period $t \geq T + 2$, player 2 plays $C$ for sure if player 1’s past actions were all $H$ or all $M$ from 0 to $T$ and all $H$ afterwards, he plays $\frac{35}{36}C + \frac{1}{36}S$ if player 1’s past actions were all $M$. Player 2 plays $S$ for sure at any other history.

One can verify players’ incentive constraints. In particular in period $T + 1$ conditional on player 1 has always played $M$ in the past, type $0_1$ is indifferent between playing $H$ and $M$ while type $0_0$ strictly prefers to play $M$. This class of equilibria can be constructed for an open set of beliefs.\(^{35}\) As we can see, player 1’s equilibrium behaviors are drastically different once we vary the index $T$, ranging from playing $M$ all the time to playing $H$ almost all the time. Moreover, the good strategic type, namely type $0_1$, have an incentive to play actions other than $H$ for a long period of time, contrary to what Theorems 3 and 3’ suggest.

G.6 Irregular Equilibria in Games with MSM Payoffs

I construct an equilibrium in the repeated product choice game with MSM payoffs in which at some on-path histories, player 1’s reputation deteriorates after playing the highest action.\(^{36}\) Recall that players’ stage game payoffs are given by:

<table>
<thead>
<tr>
<th>$\theta = 0_1$</th>
<th>$h$</th>
<th>$l$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H$</td>
<td>1,3</td>
<td>$-1,2$</td>
</tr>
<tr>
<td>$L$</td>
<td>2,0</td>
<td>0,1</td>
</tr>
<tr>
<td>$\theta = 0_0$</td>
<td>$h$</td>
<td>$l$</td>
</tr>
<tr>
<td>$H$</td>
<td>$1 - \eta, 0$</td>
<td>$-1 - \eta, 1$</td>
</tr>
<tr>
<td>$L$</td>
<td>$2, -2$</td>
<td>0,0</td>
</tr>
</tbody>
</table>

with $\eta \in (0, 1)$. Let $\Omega \equiv \{H, 0_1, 0_0\}$ with $\mu(H) = 0.06$, $\mu(0_0) = 0.04$, $\mu(0_1) = 0.9$ and $\phi_H$ is the Dirac measure on $0_1$. Consider the following strategy profile:

- In period 0, type $0_1$ plays $H$ with probability 2/45 and type $0_0$ plays $H$ with probability 1/4. Player 2 plays $l$.
- In period 1, if the history is $(L, l)$, then use the public randomization device. With probability $(1 - \delta) / \delta$, players play $(L, l)$ forever, with complementary probability, players play $(H, h)$ forever. If $(H, h)$ is prescribed and player 1 ever deviates to $L$, then player 2 plays $l$ at every subsequent history.
- In period 1, if the history is $(H, l)$, then both strategic types play $L$ and player 2 plays $h$. This is incentive compatible due to the presence of the commitment type.

\(^{35}\)Notice that under a generic prior belief, type $0_1$ needs to randomize between always playing $H$ and always playing $M$ in period $T + 1$. This can be achieved since he is indifferent by construction.

\(^{36}\)One can also verify that the constructed strategy profile is also part of a sequential equilibrium under its induced belief system.
• In period 2, if the history is \((H, l, H, h)\), then play \((H, h)\) forever on the equilibrium path. If player 2 ever observes player 1 plays \(L\), then she plays \(l\) in all subsequent periods.

• In period 2, if the history is \((H, l, L, h)\), then use the public randomization device:
  - With probability \((1 - \delta)/\delta\), play \((L, l)\) forever on the equilibrium path.
  - With probability \((1 - \frac{1 - \delta}{\delta}) - \frac{1 - \delta}{\delta}\), play \((H, h)\) forever on the equilibrium path. If player 2 ever observes player 1 plays \(L\), then she plays \(l\) in all subsequent periods.
  - With probability \((1 - \delta)/\delta^2\), type \(\theta_0\) plays \(L\) for sure and type \(\theta_1\) plays \(L\) with probability 1/4, and player 2 plays \(h\).

Following history \((H, l, L, h, H, h)\), play \((H, h)\) forever on the equilibrium path. If player 2 ever observes player 1 plays \(L\), then she plays \(l\) in all subsequent periods.

Following history \((H, l, L, h, L, h)\), use the public randomization device again. With probability \((1 - \delta)/\delta\), play \((L, l)\) forever. With complementary probability, play \((H, h)\) forever on the equilibrium path. If player 2 ever observes player 1 plays \(L\), then she plays \(l\) in all subsequent periods.

In period 0, player 2’s belief about \(\theta\) deteriorates after observing \(H\). This is true no matter whether we only count the strategic types (as strategic type \(\theta_0\) plays \(H\) with strictly higher probability) or also count the commitment type (probability of \(\theta_1\) decreases from 24/25 to 10/11).

**G.7 Multiple Equilibrium Behaviors when Player 1 is Impatient**

I present an example in which the game’s payoff satisfies Assumptions [1][3] to [3] player 2’s prior belief is pessimistic but player 1 has multiple equilibrium behaviors when \(\delta\) is not high enough. Consider the following product choice game:

<table>
<thead>
<tr>
<th>(\theta = \theta_1)</th>
<th>(C)</th>
<th>(S)</th>
<th>(\theta = \theta_0)</th>
<th>(C)</th>
<th>(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H)</td>
<td>1, 3</td>
<td>−1, 2</td>
<td>(H)</td>
<td>1 − (\eta), 0</td>
<td>−1 − (\eta), 1</td>
</tr>
<tr>
<td>(L)</td>
<td>2, 0</td>
<td>0, 1</td>
<td>(L)</td>
<td>2, −2</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

with \(\eta \in (0, 1)\), \(\Omega^m = \{H\}\) and \(\phi_H\) be the Dirac measure on \(\theta_1\). Player 2’s prior satisfies:

\[
\mu(\theta_0) = 0.7, \mu(\theta_1) + \mu(H) = 0.3 \text{ with } \mu(H) \in (0, 0.1).
\]

I construct a class of Nash equilibria when \(\delta \in (\frac{1}{2}, \frac{1 + \eta}{2})\), in which player 1’s on-path equilibrium behaviors are different across these equilibria:

- Type \(\theta_0\) always plays \(L\).
- Type \(\theta_1\) plays \(H\) in every period besides period \(t \in \{1, 2, \ldots\}\), in which he plays \(L\).
- Player 2 plays \(S\) in period 0 and period \(t\). In period \(s \neq 0, t\), she plays \(S\) if player 1 has played \(L\) before in any period besides \(t\); she plays \(C\) if player 1 has played \(H\) in every period or has only played \(L\) in period \(t\).

Intuitively, since player 1’s discount factor is low, strategic type \(\theta_0\) has no incentive to pool with the commitment type. Therefore, after playing \(H\) for one period, player 2’s belief becomes optimistic which leads to multiple equilibrium behaviors.

---

One can verify that these Nash equilibrium outcomes can also be implemented in sequential equilibrium.
G.8 Why $\lambda \in \Lambda(\alpha_1^*, \theta)$ is not sufficient when $\alpha_1^*$ is mixed?

I use a counterexample to show that $\lambda \in \Lambda(\alpha_1^*, \theta)$ is no longer sufficient to guarantee the commitment payoff bound when $\alpha_1^*$ is mixed. Players’ payoffs are given by:

\[
\begin{array}{c|ccc|c|ccc|c|ccc}
\theta_1 & l & m & r & \theta_2 & l & m & r & \theta_3 & l & m & r \\
\hline
H & 1, 3 & 0, 0 & 0, 0 & H & 0, 1/2 & 0, 3/2 & 0, 0 & H & 0, 1/2 & 0, 0 & 0, 3/2 \\
L & 2, -1 & 0, 0 & 0, 0 & L & 0, 1/2 & 0, 3/2 & 0, 0 & L & 0, 1/2 & 0, 0 & 0, 3/2 \\
D & 3, -1 & 1/2, 0 & 1/2, 0 & D & 0, 0 & 0, 0 & 0, 0 & D & 0, 0 & 0, 0 & 0, 0 \\
\end{array}
\]

Suppose $\Omega_m = \{\alpha_1^*\}$ with $\alpha_1^* = \frac{1}{2} H + \frac{1}{2} L$ and $\phi_{\alpha_1^*}$ is the Dirac measure on $\theta_1$, one can apply the definitions and obtain that $v_{\theta_1}(\alpha_1^*) = 3/2$ and $\Theta^b_{(\alpha_1^*, \theta_1)} = \{\theta_2, \theta_3\}$. If $\mu(\alpha_1^*) = 2\mu(\theta_2) = 2\mu(\theta_3) \equiv \rho$ for some $\rho \in (0, 1/2)$, then $\lambda = (1/2, 1/2) \in \Lambda(\alpha_1^*, \theta_1)$. In the following equilibrium, type $\theta_1$’s payoff is $1/2$ even when $\delta \to 1$.

- Type $\theta_1$ always plays $D$. In period 0, type $\theta_2$ plays $H$ and type $\theta_3$ plays $L$. Starting from period 1, both types play $\frac{1}{2} H + \frac{1}{2} L$. Player 2 plays $m$ in period 0. If she observes $H$ or $D$ in period 0, then she plays $m$ forever. If she observes $L$ in period 0, then she plays $r$ forever.

In the above equilibrium, either $\mu_t(\theta_2)/\mu_t(\alpha_1^*)$ or $\mu_t(\theta_3)/\mu_t(\alpha_1^*)$ will increase in period 0, regardless of player 1’s action in that period. As a result, player 2’s posterior belief in period 1 is outside $\overline{\Lambda}(\alpha_1^*, \theta_1)$ for sure. This provides him a rationale for not playing $l$ and gives type $\theta_1$ an incentive to play $D$ forever, making player 2’s belief self-fulfilling. This situation only arises when $\alpha_1^*$ is mixed and $k(\alpha_1^*, \theta) \geq 2$. 

76