Reputation Effects under Interdependent Values

Harry Di PEI*

This Draft: August 12, 2018.

Abstract: I study reputation effects when players have persistent private information that directly affects 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 reputation models, the uninformed players' payoffs depend on the state. This *interdependence of values* introduces new challenges to reputation building, as the informed player faces a trade-off between establishing a reputation for commitment and signaling favorable information about the state. I derive predictions on the informed player's payoff and behavior that uniformly apply across *all* the Nash equilibria. When the stage-game payoffs satisfy a *monotone-supermodularity* condition, I show that the informed player can overcome this new challenge 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 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 payoff, equilibrium behavior
JEL Codes: C73, D82, D83

1 Introduction

Economists have long recognized that 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 (1989,1992) 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 enable researchers to make sharp predictions in many decentralized markets where there is no mediator helping participants to coordinate on a particular equilibrium.

However, previous works on robust reputation effects restrict attention to private value environments. This excludes situations where reputation building agents have persistent private information that directly affects

*Department of Economics, Northwestern University. Email: harrydp@northwestern.edu. I am indebted to Daron Acemoglu, Drew Fudenberg, Juuso Toikka and Alex Wolitzky for guidance and support. I thank Jie Bai, Vivek Bhattacharya, Alessandro Bonatti, Gonzalo Cisternas, Dan Clark, Joyee Deb, Miaomiao Dong, Jetlir Duraj, Mehmet Ekmekci, Glenn Ellison, Chishio Furukawa, Bob Gibbons, Ben Golub, Yingni Guo, Kevin He, Christian Hellwig, Yuhta Ishii, Annie Liang, Ernest Liu, Shuo Liu, Yusuke Narita, Ben Roth, Bruno Strulovici, Jean Tirole, Muhamet Yildiz and my seminar participants for helpful comments. Errors are mine.

their opponents' payoffs. For example in the pharmaceutical, cable TV and passenger airline industries, incumbents can benefit from a reputation for fighting potential entrants, but are also better informed about the market demand curve.[1] The latter directly affects the entrants' profits from entry. In the markets for food, electronics and custom software, merchants can benefit from a reputation for providing good customer service, but can also have private information about their product quality (Banerjee and Duflo 2000, Bai 2016), which can affect consumers' willingness to pay. This coexistence of *reputation for playing certain actions* and *payoff relevant private information* introduces new economic forces, as the reputation building agent's behavior not only shows his resolve to maintain his reputation but also signals his payoff relevant private information. The dynamic interactions between the two can affect the value of a good reputation and the agent's incentives to maintain his reputation. Understanding these effects 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 studies reputation building in *interdependent value* environments. In my model, a patient long-run player 1 (e.g. incumbent, seller) interacts with a sequence of myopic short-run player 2s (e.g. entrants, buyers). The key modeling innovation is that player 1 has perfectly persistent private information about two aspects: (1) the state of the world (e.g. market demand, product quality) that can directly affect both players' stage-game payoffs; (2) whether he is strategic or committed. The strategic long-run player maximizes his discounted average payoff and will be referred to by the state he observes. The committed long-run player mechanically follows some state-contingent stationary strategies. Every player 2 perfectly observes all the actions taken in the past but cannot observe their predecessors' payoffs.

My results address the predictions on the long-run player's payoff and behavior that uniformly apply across all the Nash equilibria of this repeated game.[2] I show that in general, interdependent values introduce new challenges to reputation building as the long-run player faces a trade-off between maintaining reputation for commitment and signalling the state. My second contribution is to identify an important class of interdependent value games in which the long-run player can secure high payoffs by building reputations. Under a condition on the state distribution, he will behave consistently over time and maintain his reputation in all equilibria.

To illustrate the new economic forces brought by interdependent values, consider an example of an incumbent firm 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 whether the incumbent is fighting (F) or accommodating (A), but also on the state θ, interpreted as the price elasticities of demand and is either high

[1] As incumbents have been in the market for a long time, they can better estimate many aspects of the market demand curve using their sales data, which include the market size, the price elasticities of demand, the effectiveness of advertising, etc. See Ellison and Ellison (2011), Seamans (2013), Sweeting, Roberts and Gedge (2016) for empirical evidence of such private information.

[2] To address the objection that Nash equilibrium is too permissive in the study of repeated games, I adopt the following *double standard* throughout the paper. For the counterexamples, I use stringent solution concepts such as sequential equilibrium to make the conclusion more convincing. For the positive results, I adopt permissive solution concepts in order to make the predictions more robust.

(H) or low (L). This is modeled as the following entry deterrence game:

$\theta = \text{High}$	Out	Enter
Fight	$2, 0$	$0, \mathbf{-1}$
Accommodate	$3, 0$	$1,\ 2$

$\theta = \text{Low}$	Out	Enter
Fight	$2-\eta, 0$	$-\eta, \mathbf{1}$
Accommodate	$3, 0$	$1, 2$

where $\eta \in \mathbb{R}$ is a parameter. When $\theta = H$ is common knowledge (call it the *private value benchmark*), Fudenberg and Levine (1989) establish the following *commitment payoff theorem*: if with positive probability the incumbent is non-strategic and mechanically fights in every period, then a patient strategic incumbent can secure his commitment payoff from fighting (equal to 2) 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.

The above logic no longer applies when θ is the incumbent's private information. This is because an entrant's best reply to F depends on θ (it is O when $\theta = H$ and E when $\theta = L$), which is signalled through the incumbent's past actions. In situations where F is more likely to be played in state L,[3] the strategic incumbent will be facing a trade-off between maintaining his reputation for fighting and signalling that the state is high. Therefore, it is unclear whether he can still secure his commitment payoff. Furthermore, obtaining sharp predictions on the incumbent's equilibrium behavior faces additional challenges as he is repeatedly signalling the state. This could lead to various self-fulfilling beliefs and multiple possible behaviors. Even the commitment payoff theorem cannot imply that he will maintain his reputation in every equilibrium, as a strategy that can secure himself a high payoff is not necessarily his optimal strategy.

My first result, Theorem 1, establishes the generality of the above trade-off by characterizing the set of type distributions under which the commitment payoff theorem applies *regardless* of the long-run player's payoff function. The primary motivation for this exercise is to evaluate whether the *economic mechanism* behind the private value reputation results remains valid in interdependent value environments. According to my theorem, a sufficient and (almost) necessary condition is that the prior likelihood ratio between certain strategic types and the commitment type be below some finite cutoff.[4] This implies that the canonical reputation building logic will fail as long as the short-run players' best reply to the commitment action depends on the state and the probability with which the long-run player is committed is relatively small.

My proof of Theorem 1 constructs the long-run player's strategy that secures his commitment payoff (sufficiency part) as well as equilibria that result in low payoffs (necessity part). Interestingly, the long-run player

[3] This is a serious concern as player 1's action today can affect players' future equilibrium play. Equilibria in which player 2 attaches higher probability to state L after observing F are constructed in Appendix G for all signs of η.

[4] To be more precise, a strategic type is *bad* if player 2's best reply to the commitment action under his state is different from that under the target state. My conditions require the probability of each bad strategic type to be small compared to the probability of the commitment type. My conditions are *almost necessary* as they leave out a degenerate set of beliefs.

cannot secure his mixed commitment payoff by replicating the mixed commitment strategy. This is because playing certain actions in the support of a mixed commitment action can trigger negative inferences about the state in the sense that it will *increase* some of those likelihood ratios. This implies that the timing of player 1's actions matters for his long-term payoff, which is a novel feature of interdependent value environments, making the existing techniques in Fudenberg and Levine (1992), Sorin (1999), Gossner (2011), etc. inapplicable. I overcome this challenge using martingale techniques and the central limit theorem to construct a *non-stationary strategy* under which player 1 can achieve three goals simultaneously: (1) avoiding negative inferences about the state; (2) matching the frequencies 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.

Motivated by Theorem 1 and the leading applications of reputation models, my Theorems 2 and 3 establish robust reputation effects when players' stage-game payoffs satisfy a regularity condition called *monotone-supermodularity* (or MSM) and player 2's action choice is binary. MSM requires that the states and 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 (*monotonicity*); (2) the action profile and the state are complements in players' stage-game payoff functions (*supermodularity*).[5] In the entry deterrence example, once we rank the states and actions according to $H \succ L$, $F \succ A$ and $O \succ E$, MSM translates into $\eta > 0$. This is the case when θ is the price elasticity of demand, the market size, the effectiveness of advertising, etc. MSM is also satisfied in product choice games when providing good service is costly to the seller and it is less costly when the seller's quality is higher. This fits into the custom software industry as both the effort cost of making a timely delivery and the quality of the software's design are correlated with the talent of the firm's engineers. It also applies to the restaurant industry as the effort cost of cooking delicious dishes is decreasing with the quality of raw materials.

I establish robust predictions on player 1's equilibrium payoff and behavior when she can build a reputation for playing her highest action. When the commitment payoff from the highest action profile is one of the equilibrium payoffs in the benchmark game without commitment types (i.e. when the high states are relatively more likely compared to the low states, or the *optimistic prior* case), I show 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 occurs with probability greater than $1/2$, 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 a strictly higher payoff by establishing reputations for playing other pure commitment actions.[6]

[5] My results only require part of the supermodularity requirement, instead of full supermodularity. See Assumption 2 for details. When player 2 has more than two actions, the results require additional conditions, see Online Appendix D.

[6] This conclusion extends to mixed commitment actions if (1) the short-run players' best reply to the long-run player's highest action depends on the state; (2) the long-run player strictly prefers the highest action profile to the lowest action profile in every state. In the

In the complementary scenario (the *pessimistic prior* case), I show that when the total probability of commitment types is small, (1) player 1 can secure his highest equilibrium payoff in the benchmark game without commitment types; (2) his on-path behavior is the same across all the Nash equilibria of this infinitely repeated signalling game.[7] According to this unique equilibrium behavior, there exists a cutoff state (in the example, state L) such that the strategic player 1 plays his highest action in every period if the state is above this cutoff, plays his lowest action in every period if the state is below this cutoff, and mixes between playing his highest action in every period and playing his lowest action in every period at the cutoff state. That is to say, he will behave consistently over time and maintain his reputation in every equilibrium.

The intuition behind this behavioral uniqueness result is the following *disciplinary effect*: (1) player 1 can obtain a high continuation payoff by playing his highest action thanks to the commitment type; (2) but it is impossible for him to receive a high continuation payoff after he has failed to do so, as player 2's posterior belief will attach higher probability to the low states compared to her prior. This contrasts to the private value benchmark of Fudenberg and Levine (1989, 1992) and the optimistic prior case, in which a patient player 1 can still receive a high continuation payoff after deviating from his commitment action. As a result, behaving inconsistently is strictly optimal in many equilibria, leading to multiple on-path behaviors for the long-run player as well as strict incentives to abandon reputation.

Conceptually, the above comparison suggests that interdependent values can contribute to the *sustainability* of reputation. This channel is novel compared to the ones proposed in the existing literature, such as impermanent commitment types (Mailath and Samuelson 2001, Ekmekci, et al. 2012), competition between long-run players (Hörner 2002), imperfect observation about the long-run player's past actions (Ekmekci 2011), etc.[8]

In terms of the proofs of Theorems 2 and 3, a challenge 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 and the receiver's action choice is binary, then the sender's equilibrium action must be non-decreasing in the state. In a repeated signalling game with MSM stage game payoffs, it implies that in equilibria where playing the highest action in every period is optimal for player 1 in a given state (call them *regular equilibria*), then he must be playing the highest action with probability 1 in every higher state. That is to say in every regular equilibrium, player 2's posterior about the state will never decrease if player 1

example, it implies that the incumbent prefers (F, O) to (A, E) in every state, which translates into $\eta \in (0, 1)$.

[7]Theorem 3 states the result when all the commitment actions are pure. Theorem 3' in Appendix D.2 generalizes Theorem 3 by incorporating mixed strategy commitment types. It shows that when the total probability of commitment types is small enough, player 1's on-path behaviors across different equilibria are arbitrarily close in terms of the distributions over (player 1's) action paths. When the total probability of commitment types vanishes to 0, the maximal distance between those equilibrium distributions converges to 0 and the probability with which player 1 behaves consistently converges to 1 in every equilibrium.

[8]In 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.

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 state, and it is possible that at some on-path histories, playing the highest action will lead to negative inferences about the state. Examples of such sequential equilibria are constructed in Appendix G.6. To deal with this, I establish a belief lower bound result that in every irregular equilibrium, player 2's belief about the state can never fall below a hyperplane as long as player 1 has always played his highest action.

This paper 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 the effects of reputations for commitment on an informed player's guaranteed payoff.[9] A separate strand of works on dynamic signalling games, including Bar-Isaac (2003), Lee and Liu (2013), Pei (2015), etc. examine the effects of persistent private information about payoff-relevant variables (such as talent, quality, value of outside options) 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. I evaluate the robustness of the private value reputation mechanism against interdependent value perturbations. In games with MSM payoffs, I derive robust predictions on the informed player's payoff and behavior that apply across all equilibria.

In the study of repeated Bayesian games with interdependent values,[10] my reputation results can be interpreted as an equilibrium refinement, just as Fudenberg and Levine (1989) did for the repeated complete information games in Fudenberg, Kreps and Maskin (1990). By allowing the informed long-run player to be non-strategic and mechanically playing state-contingent stationary strategies, Theorem 2 shows that reputation effects can sharpen the predictions on a patient player's equilibrium payoff. Theorem 3 advances this research agenda one step further by showing that reputations can also lead to sharp predictions on a patient player's equilibrium behavior, which requires values to be interdependent.

In terms of the applications, my result offers a robust explanation to the classic observation of Bain (1949) that "*...established sellers persistently forego high prices for fear of attracting new entry to the industry...*". This

[9] The commitment payoff theorem has been extended to environments with imperfect monitoring (Fudenberg and Levine 1992, Gossner 2011), frequent interactions (Faingold 2013), long-lived uninformed players (Schmidt 1993a, Cripps, Dekel and Pesendorfer 2005, Atakan and Ekmekci 2012), weaker solution concepts (Watson 1993), etc. Another set of papers characterize Markov equilibria (in infinite horizon games) or sequential equilibria (in finite horizon games) in private value reputation games with a stationary commitment type, which includes Kreps and Wilson (1982), Milgrom and Roberts (1982), Schmidt (1993b), Phelan (2006), Ekmekci (2011), Liu (2011), Liu and Skrzypacz (2014), etc. See Mailath and Samuelson (2006) for an overview of this literature.

[10] This is currently a challenging area and not much is known except for zero-sum games (Aumann and Maschler 1995, Pęski and Toikka 2017), undiscounted games (Hart 1985), belief-free equilibrium payoff sets in games with two equally patient long-run players (Hörner and Lovo 2009, Hörner, Lovo and Tomala 2011).

will only happen in some particular equilibria under private values when the incumbent has no private information about market demand, but will occur in every equilibrium when the incumbent has private information about market demand and the potential entrants are optimistic about their prospects of entry. 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 to punish the firm after observing low effort. This will only happen in some particular equilibria under private values, but will occur in every equilibrium when the firm has persistent private information about quality and the consumers are initially pessimistic about it.

2 The Model

Time is discrete, indexed by $t = 0, 1, 2....$ A long-run player 1 (he) with discount factor $\delta \in (0, 1)$ interacts with an infinite sequence of myopic short-run player 2s (she or they), arriving one in each period and each plays the game only once. In period t, players simultaneously choose their actions $(a_{1,t}, a_{2,t}) \in A_1 \times A_2$. Players have access to a public randomization device, with $\xi_t \in \Xi$ the realization in period t.

Player 1 has perfectly persistent private information about two aspects: a payoff relevant state of the world $\theta \in \Theta$ as well as whether he is *strategic* or *committed*. The strategic player 1 can flexibly choose his action in every period. The committed player 1 mechanically follows some state-contingent *commitment plan* $\gamma : \Theta \to \Delta(A_1)$. Let Γ be the set of commitment plans that the committed long-run player could possibly follow. My model rules out non-stationary commitment strategies, the presence of which will be discussed in section 5.

Player 2's full support prior belief is μ, which is a joint distribution of θ and player 1's *characteristics*, i.e. whether he is strategic or committed, and if committed which plan in Γ is he following. The two dimensions can be arbitrarily correlated. I assume A_1, A_2, Θ and Γ are finite sets with $|A_1|, |A_2| \geq 2$.

For every $\theta \in \Theta$, I say that player 1 is '*(strategic) type* θ' if he is strategic and knew that the state is θ. Let $\mu(\theta)$ be the probability of the event that player 1 is type θ. Let

$$\Omega \equiv \{\alpha_1 \in \Delta(A_1) | \text{ there exist } \gamma \in \Gamma \text{ and } \theta \in \Theta \text{ such that } \gamma(\theta) = \alpha_1\}, \tag{2.1}$$

be the *set of commitment actions*. For every $\alpha_1 \in \Omega$, I say that player 1 is '*(commitment) type* α_1' if he is committed and is playing α_1 in every period. Let $\mu(\alpha_1)$ be the probability that player 1 is type α_1. Let $\phi_{\alpha_1} \in \Delta(\Theta)$ be the distribution of state *conditional on* player 1 being type α_1, which can be derived from μ.

Players' past action choices 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 $\sigma_\theta : \mathcal{H} \to \Delta(A_1)$ be type θ's strategy. Let $\sigma_2 : \mathcal{H} \to \Delta(A_2)$ be player 2's strategy. Let $\sigma \equiv \Big((\sigma_\theta)_{\theta \in \Theta}, \sigma_2\Big)$ be a typical strategy profile and let Σ be the set of strategy profiles.

Player $i \in \{1,2\}$'s stage game payoff in period t is $u_i(\theta, a_{1,t}, a_{2,t})$. Both u_1 and u_2 are naturally extended to the domain $\Delta(\Theta) \times \Delta(A_1) \times \Delta(A_2)$. Unlike the canonical reputation models in Fudenberg and Levine (1989,1992), my model has *interdependent values* as player 2's payoff depends on θ, which is player 1's private information. Strategic type θ maximizes $\sum_{t=0}^{\infty}(1-\delta)\delta^t u_1(\theta, a_{1,t}, a_{2,t})$. Every player 2 maximizes her 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 a_1 and θ are independently distributed with marginal distributions $\alpha_1 \in \Delta(A_1)$ and $\pi \in \Delta(\Theta)$. For every $(\alpha_1, \theta) \in \Omega \times \Theta$, let

$$v_\theta(\alpha_1) \equiv \min_{a_2 \in \text{BR}_2(\alpha_1, \theta)} u_1(\theta, \alpha_1, a_2), \tag{2.2}$$

be type θ's (complete information) *commitment payoff* from playing α_1. If α_1 is pure, then $v_\theta(\alpha_1)$ is a *pure commitment payoff*. If α_1 is non-trivially mixed, then $v_\theta(\alpha_1)$ is a *mixed commitment payoff*.

The solution concept is Bayes Nash equilibrium (or *equilibrium* for short). Since Θ, Γ, A_1 and A_2 are finite sets and the game is continuous at infinity, an equilibrium exists. Let $\text{NE}(\delta, \mu) \subset \Sigma$ be the set of equilibria under parameter configuration (δ, μ). Let $V_\theta^\sigma(\delta)$ be type θ's discounted average payoff under strategy profile σ and discount factor δ. Let $\underline{V}_\theta(\delta, \mu) \equiv \inf_{\sigma \in \text{NE}(\delta,\mu)} V_\theta^\sigma(\delta)$ be type θ's worst equilibrium payoff under (δ, μ).

The goal of this paper is to address the predictions on player 1's payoff and behavior that uniformly apply across *all equilibria* of this infinitely repeated game. Specifically, I am interested in the following set of questions. First, will the *economic mechanism* behind the private value reputation results remain valid in interdependent value environments? Formally, for every u_2 and $(\alpha_1, \theta) \in \Omega \times \Theta$, when does

$$\underbrace{\liminf_{\delta \to 1} \underline{V}_\theta(\delta, \mu)}_{\text{patient player 1's guaranteed equilibrium payoff in state } \theta} \quad \geq \quad \underbrace{v_\theta(\alpha_1)}_{\text{player 1's complete information commitment payoff in state } \theta} \tag{2.3}$$

apply across all u_1? Second, can we find good lower bounds for player 1's *guaranteed equilibrium payoff*, namely $\liminf_{\delta \to 1} \underline{V}_\theta(\delta, \mu)$? Third, will player 1 behave consistently over time, play the same action in every period and maintain his reputation for commitment?

My first and second questions examine player 1's guaranteed equilibrium payoff when he can build reputations for commitment. When player 2's best reply to α_1 does not depend on the state, inequality (2.3) is implied by the results in Fudenberg and Levine (1989, 1992) and player 1 can guarantee his commitment payoff by playing α_1 in every period. This logic no longer applies in interdependent value environments, as convincing player 2 that α_1 will be played does not determine her best reply. In particular, playing α_1 in every period could signal a state other than θ under which player 2's best reply is different and will give player 1 a low payoff.

My third question advances this literature one step further by examining the robust predictions on player 1's equilibrium behavior. Nevertheless, delivering sharp behavioral predictions in infinitely-repeated signalling games is challenging as the conventional wisdom suggests that both infinitely-repeated games and signalling games have multiple equilibria. The commitment payoff theorem does not imply an affirmative answer to this question either, as a strategy that can secure player 1 a high payoff is not necessarily his optimal strategy. For example, in the entry deterrence game (section 1) where $\theta = H$ is common knowledge, there are many sequential equilibria in which the strategic incumbent will behave inconsistently and abandon his reputation.

3 Characterization Theorem in General Games

To evaluate the validity of the canonical reputation building mechanism in general interdependent value environments, I characterize, for given $(\alpha_1, \theta) \in \Omega \times \Theta$, the set of μ under which (2.3) applies regardless of u_1. My conditions require that the likelihood ratios between certain strategic types and the relevant commitment type to be below some cutoffs. My result implies that the economic mechanism behind the private value reputation results will fail whenever (1) the short-run players' best reply to the commitment action depends on the state, and (2) the probability of commitment type is small. This highlights the generality of the trade-off between maintaining reputations and signalling the state.

My proof constructs a strategy that secures the long-run player his commitment payoff (sufficiency part) and equilibria that result in low payoffs (necessity part). Interestingly, the long-run player cannot secure his mixed commitment payoff by replicating the mixed commitment strategy, meaning that the *timing* of his actions matters for his long-term payoff.

3.1 The Relevant Sets of Beliefs

First, I identify a sufficient statistic for my characterization, which I call the *likelihood ratio vector*. Then, I define two sets of beliefs that will later be related to the attainability of pure and mixed commitment payoffs. I focus on pairs of $(\alpha_1, \theta) \in \Omega \times \Theta$ such that $\text{BR}_2(\alpha_1, \theta)$ is a singleton, which is the case under generic $u_2(\theta, a_1, a_2)$. This assumption will be relaxed in Online Appendix B where I allow $\text{BR}_2(\alpha_1, \theta)$ to have multiple elements. For future reference, let a_2^* be the unique element in $\text{BR}_2(\alpha_1, \theta)$. For every $X \subset \mathbb{R}^n$, let $\text{co}(X)$ be its convex hull and $\text{cl}(\cdot)$ be its closure.

First, the set of *bad states* with respect to (α_1, θ) is given by:

$$\Theta^b_{(\alpha_1,\theta)} \equiv \left\{\tilde{\theta} \in \Theta \middle| a_2^* \notin \text{BR}_2(\alpha_1, \tilde{\theta})\right\}. \tag{3.1}$$

Intuitively, this is the set of states under which player 2's best reply to the commitment action is different from that under state θ. Let $k(\alpha_1, \theta) \equiv \left|\Theta^b_{(\alpha_1,\theta)}\right|$ be the number of bad states. By definition, $k(\alpha_1, \theta) = 0$ in private value models. If $\tilde{\theta} \in \Theta^b_{(\alpha_1,\theta)}$, then type $\tilde{\theta}$ is called a *bad strategic type*. For every belief $\tilde{\mu}$ with $\tilde{\mu}(\alpha_1) > 0$, let $\tilde{\lambda}(\tilde{\theta}) \equiv \tilde{\mu}(\tilde{\theta})/\tilde{\mu}(\alpha_1)$ be the *likelihood ratio* between strategic type $\tilde{\theta}$ and commitment type α_1. Let

$$\tilde{\lambda} \equiv \Big(\tilde{\lambda}(\tilde{\theta})\Big)_{\tilde{\theta}\in\Theta^b_{(\alpha_1,\theta)}} \in \mathbb{R}^{k(\alpha_1,\theta)}_+,$$

be the *likelihood ratio vector*. The *best response set* with respect to (α_1, θ) is defined as:

$$\overline{\Lambda}(\alpha_1,\theta) \equiv \Big\{\tilde{\lambda} \in \mathbb{R}^{k(\alpha_1,\theta)}_+ \Big| \{a_2^*\} = \arg\max_{a_2\in A_2} \big\{u_2(\phi_{\alpha_1},\alpha_1,a_2) + \sum_{\tilde{\theta}\in\Theta^b_{(\alpha_1,\theta)}} \tilde{\lambda}(\tilde{\theta})u_2(\tilde{\theta},\alpha_1,a_2)\big\}\Big\}. \tag{3.2}$$

Intuitively, a likelihood ratio vector belongs to the best response set if a_2^* is player 2's strict best reply to α_1 conditional on the *union* of the following set of events: (1) player 1 is strategic and the state is bad; (2) player 1 is committed and is playing α_1 in every period.

Nevertheless, the prior likelihood ratio vector belongs to the best response set is not sufficient for our purpose as player 2s' belief is updated over time and their posterior could fall outside of $\overline{\Lambda}(\alpha_1,\theta)$. Such concerns motivate us to find the *largest subset* of $\overline{\Lambda}(\alpha_1,\theta)$ with the following *property* $(*)$:

(*) If the prior likelihood ratio vector belongs to this subset, then for *every* feasible belief updating process, player 1 has a strategy under which the posterior likelihood ratio belongs to $\overline{\Lambda}(\alpha_1,\theta)$ in every period.

By definition, this largest subset will depend on the set of feasible belief updating processes, which in turn will depend on whether the commitment action α_1 is pure or mixed.

When α_1 is pure, none of the likelihood ratios can increase as long as player 1 is playing α_1. The largest subset of $\overline{\Lambda}(\alpha_1,\theta)$ with property (*) is given by:

$$\Lambda(\alpha_1,\theta) \equiv \Big\{\widetilde{\lambda}\Big|\lambda' \in \overline{\Lambda}(\alpha_1,\theta) \text{ for every } \mathbf{0} \ll \lambda' \ll \tilde{\lambda}\Big\}, \tag{3.3}$$

where '$\ll$' denotes weak dominance in product order on $\mathbb{R}^{k(\alpha_1,\theta)}$ and $\mathbf{0}$ is the null vector in $\mathbb{R}^{k(\alpha_1,\theta)}$. Examples of $\overline{\Lambda}(\alpha_1,\theta)$ and $\Lambda(\alpha_1,\theta)$ when there are two bad states are shown in Figure 1.

When α_1 is mixed, the set of feasible belief updating processes becomes richer. A new concern arises which is embodied in the updating process described by the dashed lines in Figure 1 (right panel). This can occur, for example, when $\alpha_1 = \frac{1}{2}a_1' + \frac{1}{2}a_1''$, the prior likelihood ratio vector is $\widetilde{\lambda} \in \Lambda(\alpha_1,\theta)$ and the set of bad states is $\{\theta_1,\theta_2\}$. When player 2 believes that type θ_1 plays a_1' and type θ_2 plays a_1'', no matter which action player 1 plays in the support of α_1, the posterior likelihood ratio vector will be bounded away from $\overline{\Lambda}(\alpha_1,\theta)$. This

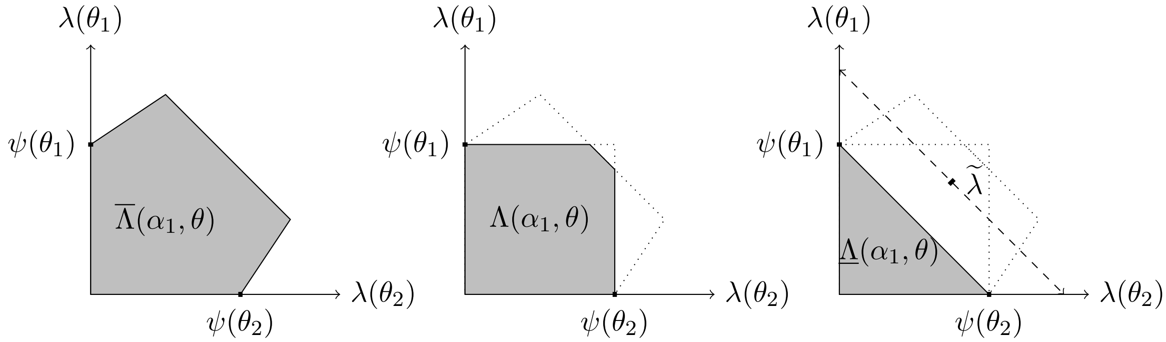

Figure 1: $k(\alpha_1,\theta) = 2$ with $\overline{\Lambda}(\alpha_1,\theta)$ in the left, $\Lambda(\alpha_1,\theta)$ in the middle and $\underline{\Lambda}(\alpha_1,\theta)$ in the right.

implies that the largest subset with property (*) is even smaller, which is given by:

$$\underline{\Lambda}(\alpha_1,\theta) \equiv \mathbb{R}_+^{k(\alpha_1,\theta)} \Big\backslash \text{co}\Big(\mathbb{R}_+^{k(\alpha_1,\theta)} \backslash \Lambda(\alpha_1,\theta)\Big). \tag{3.4}$$

An example of $\underline{\Lambda}(\alpha_1,\theta)$ is shown in the right panel of Figure 1. When $\text{BR}_2(\alpha_1,\theta)$ is a singleton and $\overline{\Lambda}(\alpha_1,\theta)$ is non-empty, $\underline{\Lambda}(\alpha_1,\theta)$ is characterized by a linear inequality:

$$\underline{\Lambda}(\alpha_1,\theta) = \Big\{\tilde{\lambda} \in \mathbb{R}_+^{k(\alpha_1,\theta)} \Big| \sum_{\tilde{\theta} \in \Theta^b_{(\alpha_1,\theta)}} \tilde{\lambda}(\tilde{\theta})/\psi(\tilde{\theta}) < 1\Big\}, \tag{3.5}$$

where $\psi(\tilde{\theta})$ is the intercept of $\Lambda(\alpha_1,\theta)$ on the $\lambda(\tilde{\theta})$-coordinate.

For future reference, I summarize some geometric properties of these sets. First, despite $\overline{\Lambda}(\alpha_1,\theta)$ is not necessarily bounded, both $\Lambda(\alpha_1,\theta)$ and $\underline{\Lambda}(\alpha_1,\theta)$ are bounded. This is because $\psi(\tilde{\theta})$ is a positive real number for every bad state. Second, $\Lambda(\alpha_1,\theta)$ and $\underline{\Lambda}(\alpha_1,\theta)$ are convex polyhedrons that are independent of u_1. Third,

$$\underline{\Lambda}(\alpha_1,\theta) \subset \Lambda(\alpha_1,\theta) \subset \overline{\Lambda}(\alpha_1,\theta). \tag{3.6}$$

Fourth, if $k(\alpha_1,\theta) = 1$ and $\Lambda(\alpha_1,\theta) \neq \{\varnothing\}$, then there exists a scalar $\psi^* \in (0,+\infty)$ such that:

$$\underline{\Lambda}(\alpha_1,\theta) = \Lambda(\alpha_1,\theta) = \overline{\Lambda}(\alpha_1,\theta) = \{\tilde{\lambda} \in \mathbb{R} | 0 \leq \tilde{\lambda} < \psi^*\}. \tag{3.7}$$

However, the three sets can be different from each other when $k(\alpha_1,\theta) \geq 2$, as illustrated in Figure 1. The condition under which these three sets coincide is characterized in Online Appendix A.

3.2 Statement of Result

My first result characterizes the set of μ under which the commitment payoff bound, namely inequality (2.3), applies regardless of u_1. Let μ_t be player 2's belief in period t. Let λ and λ_t be the likelihood ratio vectors induced by the prior μ and μ_t, respectively.

Theorem 1. *For every* $(\alpha_1, \theta) \in \Omega \times \Theta$ *with* α_1 *being a pure action,*

1. *If* $\lambda \in \Lambda(\alpha_1, \theta)$*, then* $\liminf_{\delta \to 1} \underline{V}_\theta(\delta, \mu) \geq v_\theta(\alpha_1)$ *for every* u_1.
2. *If* $\lambda \notin cl\Big(\Lambda(\alpha_1, \theta)\Big)$ *and* $BR_2(\alpha_1, \phi_{\alpha_1})$ *is a singleton, then there exists* u_1 *such that* $\limsup_{\delta \to 1} \underline{V}_\theta(\delta, \mu) < v_\theta(\alpha_1)$.

For every $(\alpha_1, \theta) \in \Omega \times \Theta$ *with* α_1 *being a non-trivially mixed action,*

3. *If* $\lambda \in \underline{\Lambda}(\alpha_1, \theta)$*, then* $\liminf_{\delta \to 1} \underline{V}_\theta(\delta, \mu) \geq v_\theta(\alpha_1)$ *for every* u_1.
4. *If* $\lambda \notin cl\Big(\underline{\Lambda}(\alpha_1, \theta)\Big)$*,* $BR_2(\alpha_1, \phi_{\alpha_1})$ *is a singleton and* $\alpha_1 \notin co\Big(\Omega \Big\backslash \{\alpha_1\}\Big)$*, then there exists* u_1 *such that* $\limsup_{\delta \to 1} \underline{V}_\theta(\delta, \mu) < v_\theta(\alpha_1)$.

According to Theorem 1, the private value reputation building logic applies if and only if the likelihood ratio between every bad strategic type and the relevant commitment type be below some finite cutoff. Moreover, it does not depend on the probabilities of the good strategic types and commitment types playing actions other than α_1. Intuitively, this is because type θ needs to come up with a strategy under which the likelihood ratio vector will remain low along every dimension in every period and this strategy should exist regardless of player 2's belief about the other types' strategies. This includes the adverse self-fulfilling belief in which all the good strategic types separate from, while all the bad strategic types pool with, the commitment type. This explains why the probabilities of bad strategic types matter while those of the good strategic types do not. Since all the commitment strategies are stationary, commitment types other than α_1 will be statistically distinguished from type α_1 in the long-run, which explains why their probabilities do not matter either.

Theorem 1 has two interpretations. First, it evaluates the robustness of reputation effects in private value reputation games against a richer set of perturbations. Starting from a private value reputation game where θ 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 private value reputation result extends when these interdependent value perturbations are relatively less likely compared to the commitment type, and vice versa. In cases where $\underline{\Lambda}(\alpha_1, \theta) \subsetneq \Lambda(\alpha_1, \theta)$, securing the commitment payoff requires more demanding conditions when the commitment action is mixed. This implies that small trembles to a pure commitment action can lead to a large decrease in player 1's guaranteed equilibrium payoff. This highlights another distinction between private and interdependent values, which will be formalized in Online Appendix A.

Second, Theorem 1 also points out the failure of the canonical reputation building mechanism in repeated incomplete information games with non-trivial interdependent values. That is, starting from a repeated incomplete information game with interdependent values, and then perturb it by introducing a small probability of commitment types. According to this view, every commitment type is *arbitrarily unlikely* compared to every strategic type. Since $\Lambda(\alpha_1, \theta)$ is bounded and the prior has full support, the likelihood ratio does not belong to $\Lambda(\alpha_1, \theta)$ whenever $k(\alpha_1, \theta) > 0$. This motivates the study of games with regularity conditions on the underlying payoff structure (for example stage games with monotone-supermodular payoffs in section 4), which enables us to make progress in addressing the robust predictions in interdependent value reputation games.

I conclude this subsection by commenting on the technical conditions in Theorem 1. First, my characterization excludes two degenerate sets of beliefs, which are the boundaries of $\Lambda(\alpha_1, \theta)$ and $\underline{\Lambda}(\alpha_1, \theta)$, respectively. In these knife-edge cases, the attainability of the commitment payoff depends on the presence of other mixed commitment types as well as their correlations with the state. Second, the assumption in statements 2 and 4 that $\text{BR}_2(\alpha_1, \phi_{\alpha_1})$ being singleton is satisfied under generic parameter values, and is only required for the proof when $\Lambda(\alpha_1, \theta) = \{\varnothing\}$. This is 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 of this issue will be shown in Appendix B. Third, according to the separating hyperplane theorem, the requirement that $\alpha_1 \notin \text{co}\big(\Omega \backslash \{\alpha_1\}\big)$ guarantees the existence of a stage-game payoff function $u_1(\theta, \cdot, \cdot)$ under which type θ's commitment payoff from any alternative commitment action in Ω is strictly below $v_\theta(\alpha_1)$. This convex independence assumption cannot be dispensed as no restrictions are made on the probabilities of other commitment types. That is to say, commitment types other than α_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 where α_1 is pure and then move on cases where α_1 is mixed. Unlike the private value benchmark, player 1 *cannot* guarantee his mixed commitment payoff by replicating the mixed commitment action. This is because playing some actions in the support of the mixed commitment action can *increase* the likelihood ratios, which is a novel feature of interdependent value environments.

Pure Commitment Payoff: Since α_1 is pure, for every $\tilde{\theta} \in \Theta^b_{(\alpha_1, \theta)}$, $\lambda_t(\tilde{\theta})$ will not increase as long as player 1 plays α_1 . 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 in period t, the strategic types are playing actions other than α_1 with probability bounded from below, after which they will be separated from the commitment type. Similar to Fudenberg and Levine (1989), the number of bad periods is uniformly bounded from above, which implies that player 1 can approximately secure his commitment payoff as $\delta \to 1$.

Mixed Commitment Payoff when $k(\alpha_1, \theta) = 1$: Let $\Theta^b_{(\alpha_1,\theta)} \equiv \{\tilde{\theta}\}$. Recall that when $\underline{\Lambda}(\alpha_1, \theta) \neq \{\varnothing\}$, there exists $\psi^* > 0$ such that $\underline{\Lambda}(\alpha_1, \theta) = \{\tilde{\lambda} | 0 \leq \tilde{\lambda} < \psi^*\}$. The main difference from the pure commitment action case is that λ_t can *increase* after player 2 observes some actions in the support of α_1. As a result, type θ *cannot* secure his commitment payoff by replicating α_1 since he may end up playing actions that are more likely to be played by type $\tilde{\theta}$, in which case λ_t will exceed ψ^*.

The key step in my proof shows that for every equilibrium strategy of player 2, one can construct a *non-stationary strategy* for player 1 under which the following three objectives are achieved simultaneously: (1) Avoid negative inferences about the state, i.e. $\lambda_t < \psi^*$ for every $t \in \mathbb{N}$; (2) Every $a_1 \in A_1$ will be played with discounted average frequency close to $\alpha_1(a_1)$; (3) In expectation, the short-run players believe that actions within a small neighborhood of α_1 will be played for all but a bounded number of periods.[11]

To understand the ideas behind my construction, note that $\{\lambda_t\}_{t \in \mathbb{N}}$ is a non-negative supermartingale conditional on α_1. Since $\lambda_0 < \psi^*$, the probability measure over histories (induced by α_1) in which λ_t never exceeds ψ^* is bounded from below by the Doob's Upcrossing Inequality.[12] When δ is close to 1, the Lindeberg-Feller Central Limit Theorem ensures that the set of player 1's action paths, in which the discounted time average frequency of every $a_1 \in A_1$ being close to $\alpha_1(a_1)$, occurs with probability close to 1 under the measure induced by α_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 an upper bound on the expected sum of relative entropies between α_1 and player 2's predicted action conditional on the aforementioned 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 occurs with 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 being far away from α_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 α_1. The assumption that $\lambda \in \underline{\Lambda}(\alpha_1, \theta)$ implies that $S_0 < 1$. The Doob's Upcrossing Inequality provides a lower bound on the probability measure over histories under which S_t is always

[11] There is a remaining step to deal with correlations between the actions and the state, with details shown in Part II of Appendix A.2.

[12] My proof uses the upcrossing inequality for a different purpose compared to Fudenberg and Levine (1992). In private value environments, they establish an upper bound on the number of periods in which player 2's prediction about player 1's action differs significantly from the commitment action if player 1 plays the commitment action in every period. 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 negative belief updating about a payoff relevant state that cannot be statistically identified via public signals.

strictly below 1, i.e. $\lambda_t \in \underline{\Lambda}(\alpha_1, \theta)$ 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 α_1 is mixed, I present a counterexample in Appendix G.8 where $\lambda \in \Lambda(\alpha_1, \theta)$ but type θ's equilibrium payoff is bounded below his commitment payoff. The idea is to construct equilibrium strategies for the bad strategic types, under which playing every action in the support of α_1 will increase the likelihood ratio along some dimensions. As a result, player 2's belief in period 1 will be bounded away from $\overline{\Lambda}(\alpha_1, \theta)$ regardless of the action being played in period 0.

3.4 Proof Ideas of Statements 2 & 4

To prove statement 2, let α_1 be the Dirac measure on $a_1^* \in A_1$. Let player 1's stage-game payoff be:

$$u_1(\hat{\theta}, a_1, a_2) = \mathbf{1}\{\hat{\theta} = \theta, a_1 = a_1^*, a_2 = a_2^*\}. \tag{3.8}$$

I construct equilibria in which type θ's payoff is bounded below 1 in the $\delta \to 1$ limit. The idea is to let the bad strategic types pool with commitment type a_1^* with high probability (albeit not equal to 1) while the good ones separate from type a_1^*. As a result, type θ cannot simultaneously build a reputation for commitment while separating away from the bad strategic types.

The key challenge comes from the presence of other commitment types that are playing mixed strategies. To understand this issue, consider an example where $\Theta = \{\theta, \tilde{\theta}\}$ with $\tilde{\theta} \in \Theta^b_{(a_1^*, \theta)}$, $\Omega = \{a_1^*, \alpha_1'\}$ with α_1' non-trivially mixed, attaching positive probability to a_1^* and $\{a_2^*\} = \mathrm{BR}_2(a_1^*, \phi_{a_1^*}) = \mathrm{BR}_2(\alpha_1', \phi_{\alpha_1'})$. The naive construction in which type $\tilde{\theta}$ plays a_1^* in every period does not work, as type θ can then obtain a payoff arbitrarily close to 1 by playing $a_1 \in \mathrm{supp}(\alpha_1') \backslash \{a_1^*\}$ in period 0 and a_1^* in every subsequent period.

To overcome this challenge, I construct sequential equilibria in which all the bad strategic types are playing non-stationary strategies. In the example, type $\tilde{\theta}$ plays a_1^* in every period with probability $p \in (0, 1)$ and plays non-stationary strategy $\sigma(\alpha_1')$ with probability $1 - p$, with p being large enough such that λ_1 is bounded away from $\overline{\Lambda}(a_1^*, \theta)$ after observing a_1^* in period 0. Strategy $\sigma(\alpha_1')$ is described as follows:

- Play α_1' at histories that are consistent with type θ's equilibrium strategy.

- Otherwise, play a completely mixed action $\widehat{\alpha}_1'$ that attaches higher probability to a_1^* compared to α_1'.

To verify incentive compatibility, I keep track of the likelihood ratio between type $\tilde{\theta}$ who plays $\sigma(\alpha_1')$ and the commitment type α_1. If type θ has never deviated before, then this ratio remains constant. If type θ has deviated before, then this ratio increases every time a_1^* is observed. Therefore, once type θ has deviated from his equilibrium play, he will be constantly facing 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 θ'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 stage-game payoff function in (3.8) is replaced by one that is constructed via the separating hyperplane theorem, such that type θ's commitment payoff from every other action in Ω is strictly lower than his commitment payoff from playing α_1. Second, in Online Appendix A.3, I show that there exists an integer T (independent of δ) and a T-period strategy for the strategic types other than θ such that the likelihood ratio vector in period T is bounded away from $\overline{\Lambda}(\alpha_1, \theta)$ 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 θ's continuation payoff after any deviation is bounded below his commitment payoff from playing α_1. The details are shown in Online Appendices A.5 and A.7.

4 Reputation Effects in Monotone-Supermodular Games

Motivated by Theorem 1 and the leading applications of reputation effects such as the entry deterrence game and the product choice game, I focus on stage-game payoffs that satisfy a *monotone-supermodularity* condition (or MSM). Theorems 2 and 3 derive results on player 1's guaranteed payoff and on-path behavior that are valid across all equilibria. Different from those in Theorem 1, these robust predictions apply even when the total probability of commitment types is arbitrarily small compared to the probability of any strategic type.

4.1 Monotone-Supermodular Payoff Structure

Let Θ, A_1 and A_2 be finite ordered sets, with '$\succ$', '$\succsim$', '$\prec$' and '$\precsim$' be the ranking among elements. The stage game has MSM payoffs if u_1, u_2 satisfy a *monotonicity* condition and a *supermodularity* condition.

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 (in short, SID) in* (a_1, a_2) *and* θ*.* $u_2(\theta, a_1, a_2)$ *has strictly increasing differences in* (θ, a_1) *and* a_2*.*[13]

Theoretically, MSM is related to a result on the monotonicity of the sender's equilibrium strategy with respect to the state in *one-shot signalling games* (Liu and Pei 2017). Economically, MSM has three implications. First, player 1 faces a *lack-of-commitment problem*, i.e. he would like to commit to a higher a_1 as it gives player

[13]First, 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, Assumption 2 will be relaxed in section 5 (see Assumption 2') to address applications such as reciprocal altruism in repeated prisoner's dilemma where u_1 has decreasing differences in θ and a_2.

2 more incentives to play a higher a_2, but he is tempted to save cost by reducing a_1. This holds in the product choice game and the entry deterrence game but fails in coordination games (battle of sexes) and zero sum games (matching pennies) where player 1's ordinal preference over a_1 depends on a_2 and vice versa. Second, player 1 always wants to signal that θ is high. This is a natural assumption when θ is product quality and sellers wish to signal high quality, or when θ is the demand curve and incumbents want to convince entrants that the market conditions are adverse to entry. It rules out zero-sum games in which senders want to signal the opposite state and common interest games in which senders want to signal the true state. Third, there are complementarities between the state and the action profile in the long-run player's payoff. This is reasonable in settings such as, fighting entrants is less costly in markets where the price elasticity is higher, exerting effort is less costly when the seller's quality is higher, etc. I will discuss the relevant applications in subsection 4.4.

For illustration purposes, I study games where player 2's action choice is binary, which have been a primary focus of the reputation literature (Mailath and Samuelson 2001, Phelan 2006, Ekmekci 2011, Liu 2011). Extensions to games with $|A_2| \geq 3$ under extra conditions can be found in Online Appendix D.

Assumption 3. $|A_2| = 2$.

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 $\alpha_1 \in \Delta(A_1)$, the following expression measures player 2's incentives to play $\overline{a}_2$ given her beliefs about θ and a_1.

$$\mathcal{D}(\pi, \alpha_1) \equiv u_2(\pi, \alpha_1, \overline{a}_2) - u_2(\pi, \alpha_1, \underline{a}_2), \tag{4.1}$$

For future reference, I classify the states into *good*, *positive* and *negative* by partitioning Θ into the following three sets:

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

$$\Theta_p \equiv \Big\{\theta \notin \Theta_g \Big| u_1(\theta, \overline{a}_1, \overline{a}_2) > u_1(\theta, \underline{a}_1, \underline{a}_2)\Big\} \text{ and } \Theta_n \equiv \Big\{\theta \Big| u_1(\theta, \overline{a}_1, \overline{a}_2) \leq u_1(\theta, \underline{a}_1, \underline{a}_2)\Big\}.$$

Intuitively, Θ_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 $(\underline{a}_1, \underline{a}_2)$. Θ_p is the set of *positive 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. Θ_n is the set of *negative bad states* in which player 1 prefers his minmax outcome to the commitment outcome. I show that every good state is higher than every positive state, and every positive 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 $\theta_n \in \Theta_n$, we have $\theta_g \succ \theta_p$, $\theta_p \succ \theta_n$ and $\theta_g \succ \theta_n$.*

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

PROOF OF LEMMA 4.1: Since $\mathcal{D}(\theta_g, \overline{a}_1) \geq 0$ and $\mathcal{D}(\theta_p, \overline{a}_1) < 0$, SID of u_2 in θ and a_2 implies that $\theta_g \succ \theta_p$. 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 $\theta_p \succ \theta_n$ due to the SID of u_1 in θ and (a_1, a_2). If $\Theta_p \neq \{\varnothing\}$, then statement 1 is proved. If $\Theta_p = \{\varnothing\}$, then since $u_1(\theta_g, \overline{a}_1, \overline{a}_2) > u_1(\theta_g, \underline{a}_1, \underline{a}_2)$ and $u_1(\theta_n, \overline{a}_1, \overline{a}_2) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2)$, we have $\theta_g \succ \theta_n$. If $\Theta_p, \Theta_n \neq \{\varnothing\}$, then $\theta_n \prec \theta_p$. SID of u_2 in θ and a_2 implies that $\mathcal{D}(\theta_n, \overline{a}_1) < \mathcal{D}(\theta_p, \overline{a}_1) < 0$. □

4.2 Statement of Results

I study player 1's guaranteed payoff and on-path behavior when he can build a reputation for playing his *highest action* $\overline{a}_1$. For this purpose, I assume that $\overline{a}_1 \in \Omega$ and $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$, i.e. the highest action is one of the commitment actions and building a reputation for playing the highest action is valuable. I will state Theorems 2 and 3 in this subsection. In subsection 4.3, I examine a benchmark repeated game with the same state distribution but *without* commitment types in order to motivate the study of reputations for playing $\overline{a}_1$ as well as the conditions on the type distribution in the statement of the theorems. I will discuss the related economic applications in subsection 4.4. I will sketch the proofs of these theorems in subsection 4.5 using an illustrative example. The full proofs can be found in Appendices C and D. Counterexamples when each of the assumptions fails are provided in Appendices G.1, G.2 and G.3.

Optimistic and Pessimistic Priors: Due to the presence of interdependent values and persistent private information, players' equilibrium payoffs and behaviors will depend on the distribution of the payoff relevant state θ. For player 1's guaranteed payoff and on-path behavior, the critical aspect is the relative likelihood between the states in Θ_g and those in Θ_p. In particular, I say that player 2's prior belief is *optimistic* if:

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta_g \cup \Theta_p} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) > 0, \tag{4.2}$$

and her prior belief is *pessimistic* in the complementary scenario where:

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta_g \cup \Theta_p} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0.\text{[14]} \tag{4.3}$$

Notice that (4.2) and (4.3) allow the commitment types to be arbitrarily unlikely compared to every strategic type. When the total probability of commitment types $\mu(\Omega)$ is small enough, the above inequalities are implied by $\sum_{\theta \in \Theta_g \cup \Theta_p} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$ and $\sum_{\theta \in \Theta_g \cup \Theta_p} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0$, respectively.

[14] In the knife-edge case where equality holds, the guaranteed payoff in Theorem 2 applies and the unique prediction on player 1's behavior applies when there exists no $\alpha_1 \in \Omega \backslash \{\overline{a}_1\}$ such that $\mathcal{D}(\phi_{\alpha_1}, \alpha_1) > 0$.

4.2.1 Reputation Effects on Guaranteed Payoffs

I start from introducing a payoff vector for every pessimistic μ which will be related to player 1's guaranteed payoff. For every μ satisfying (4.3), there exists a unique pair of $\Big(\theta_p^*(\mu), q(\mu)\Big) \in \Theta_p \times (0,1]$ such that:

$$\mu(\overline{a}_1)\mathcal{D}\big(\phi_{\overline{a}_1}, \overline{a}_1\big) + q(\mu)\mu(\theta_p^*(\mu))\mathcal{D}\Big(\theta_p^*(\mu), \overline{a}_1\Big) + \sum_{\theta \succ \theta_p^*(\mu)} \mu(\theta)\mathcal{D}\big(\theta, \overline{a}_1\big) = 0. \tag{4.4}$$

Since $u_1(\theta_p, \overline{a}_1, \overline{a}_2) > u_1(\theta_p, \underline{a}_1, \underline{a}_2) > u_1(\theta_p, \overline{a}_1, \underline{a}_2)$ for every $\theta_p \in \Theta_p$, there exists $r(\mu) \in (0,1)$ such that:

$$r(\mu)u_1(\theta_p^*(\mu), \overline{a}_1, \overline{a}_2) + \Big(1 - r(\mu)\Big)u_1(\theta_p^*(\mu), \overline{a}_1, \underline{a}_2) = u_1(\theta_p^*(\mu), \underline{a}_1, \underline{a}_2). \tag{4.5}$$

Let

$$v_\theta^*(\mu) \equiv \begin{cases} u_1(\theta, \underline{a}_1, \underline{a}_2) & \text{if } \theta \precsim \theta_p^*(\mu) \\ r(\mu)u_1(\theta, \overline{a}_1, \overline{a}_2) + \Big(1 - r(\mu)\Big)u_1(\theta, \overline{a}_1, \underline{a}_2) & \text{if } \theta \succ \theta_p^*(\mu). \end{cases} \tag{4.6}$$

By definition, $\Big(v_\theta^*(\mu)\Big)_{\theta \in \Theta}$ depends on μ only through the cutoff state $\theta_p^*(\mu)$. I will show in the next subsection that $v_\theta^*(\mu)$ is type θ's *highest equilibrium payoff* in the benchmark game with the same state distribution but without commitment types. Theorem 2 establishes the patient long-run player's guaranteed payoff.

Theorem 2. *Suppose $\overline{a}_1 \in \Omega$ and $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$.*

1. *If μ satisfies (4.2), then $\liminf_{\delta \to 1} \underline{V}_\theta(\delta, \mu) \geq u_1(\theta, \overline{a}_1, \overline{a}_2)$ for every $\theta \in \Theta$.*

2. *If μ satisfies (4.3), then $\liminf_{\delta \to 1} \underline{V}_\theta(\delta, \mu) \geq v_\theta^*(\mu)$ for every $\theta \in \Theta$.*

The proof is in Appendix C. According to Theorem 2, the patient long-run player can secure high returns from building a reputation for playing his highest action in an important class of games with non-trivial interdependent values. The exact notion of *high returns* depends on the state distribution, which will be related to the no-commitment-type benchmark in Propositions 4.1, 4.2 and 4.3. Mapping Theorem 2 back to the economic applications, it implies that a seller can secure high payoffs by maintaining a reputation for exerting high effort despite his customers' skepticism about product durability/quality; an incumbent who might have unfavorable information about the market demand curve (e.g. demand elasticities are low) can guarantee high profits by building a reputation for pricing aggressively.

Different from the distribution conditions in Theorem 1, (4.2) allows the total probability of commitment types to be arbitrarily small compared to that of every bad strategic type. Intuitively, this is because the strategic types in Θ_g can also contribute to attaining the commitment payoff. This is driven by the following implication of MSM stage-game payoffs in this *repeated signalling game*, which is derived from a companion result on *one-shot signalling games* with MSM payoffs (Liu and Pei 2017).

(*) In equilibria where some strategic type in Θ_p has an incentive to play $\overline{a}_1$ in every period, every strategic type in Θ_g will play $\overline{a}_1$ with probability 1 at every on-path history.

In those equilibria, player 2's belief about θ will not decline as long as $\overline{a}_1$ has always been played in the past.

Nevertheless, there also exist equilibria in which playing $\overline{a}_1$ in every period is not optimal for any strategic type. This undermines the implications of MSM, and in particular, every strategic type could have incentives to separate from type $\overline{a}_1$ at certain on-path histories. Moreover, playing $\overline{a}_1$ can also lead to negative inferences about θ, with an example shown in Appendix G.6. Such equilibria exist as in repeated signalling games, player 1's action choice today can affect the future equilibrium play. In particular, there exist self-fulfilling beliefs that playing $\overline{a}_1$ in the current period will result in a lower frequency of $\overline{a}_1$ being played in the future. At histories where this type of belief prevails, the bad strategic types will play $\overline{a}_1$ with strictly higher probability compared to the good ones and as a result, player 2 will interpret $\overline{a}_1$ as a negative signal about θ.

To circumvent the aforementioned complications, I establish a lower bound on player 2's posterior belief about θ that uniformly applies across all the on-path histories where $\overline{a}_1$ has been played in every period. This implies that in every period where $\overline{a}_2$ is not player 2's strict best reply, the strategic types must be separating from commitment type $\overline{a}_1$ with probability bounded from below. As a result, there exist at most a bounded number of such periods, which establishes a patient long-run player's guaranteed payoff in those equilibria.

4.2.2 Reputation Effects on Equilibrium Behavior

When the prior belief about θ is pessimistic and the total probability of commitment types is small enough, reputation effects can also lead to sharp predictions on player 1's on-path behavior in addition to his payoffs. For the ease of exposition, I focus on the case where all commitment actions are pure in the main text. The result will be generalized in Appendix D by incorporating mixed strategy commitment types (Theorem 3').

Formally, let $h_1^t \equiv \{a_{1,0}, ..., a_{1,t-1}\}$ be player 1's t-period *action path*. Let $\mathcal{H}_1^t$ be the set of h_1^t. Let $\mathcal{H}_1 \equiv \cup_{t=0}^{\infty} \mathcal{H}_1^t$ be the set of player 1's action paths. I say that $h_1^t = \{a_{1,0}, ..., a_{1,t-1}\}$ is *consistent* if $a_{1,0} = ... = a_{1,t-1}$, i.e. player 1 has played the same action in every period. For every σ and θ, let $\mathcal{P}_1^{\sigma}(\theta)$ be the probability measure over $\mathcal{H}_1$ induced by $(\sigma_\theta, \sigma_2)$. Deriving sharp predictions on player 1's on-path behavior also requires the total probability of commitment types to be small enough, with one of the sufficient conditions given by:[15]

$$\mu(\Omega)\mathcal{D}(\overline{\theta}, \overline{a}_1) + \sum_{\theta \in \Theta_g \cup \Theta_p} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0. \tag{4.7}$$

[15] In Appendix G.5, I show by counterexample that player 1 can have multiple equilibrium behaviors and his on-path play can be inconsistent over time when μ satisfies (4.3) but the total probability of commitment types is large enough to violate (4.7).

To motivate (4.7) and to see how it is related to the pessimistic prior condition in (4.3), note that first, (4.7) is more demanding compared to (4.3) when $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$; second, for every μ satisfying (4.3), it also satisfies (4.7) when the total probability of commitment types $\mu(\Omega)$ is small enough.

Theorem 3. *If $\overline{a}_1 \in A_1$, $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$, all actions in Ω are pure and μ satisfies (4.7), then there exists $\overline{\delta} \in (0,1)$ such that for every $\delta > \overline{\delta}$, $\theta \in \Theta$ and $\sigma, \sigma' \in NE(\delta, \mu)$, we have $\mathcal{P}_1^{\sigma}(\theta) = \mathcal{P}_1^{\sigma'}(\theta)$. Moreover, $\mathcal{P}_1^{\sigma}(\theta)$ only attaches positive probability to consistent action paths.*

Theorem 3 says that when the prior is pessimistic and the total probability of commitment types is small, the patient long-run player's on-path behavior is the same across all equilibria, according to which he will behave consistently over time and maintain his reputation for commitment. To describe this unique behavior, let

$$\Omega^g \equiv \{\alpha_1 \in \Omega | \mathcal{D}(\phi_{\alpha_1}, \alpha_1) > 0\} \tag{4.8}$$

be the set of *good commitment actions*, i.e. ones that can induce player 2 to play $\overline{a}_2$. In every equilibrium, an action path occurs with positive probability if it is consistent and player 1's action belongs to $\Omega^g \cup \{\underline{a}_1\}$. The probability with which each type plays each action is pinned down by two conditions (1) for every $\theta \succ \theta'$, every action path played by type θ is no lower than every action path played by type θ'; (2) after observing $a_1 \in \Omega^g \backslash \{\underline{a}_1\}$ in period 0, player 2 is indifferent between $\overline{a}_2$ and $\underline{a}_2$ on the equilibrium path starting from period 1. For example, in the simplest case where $\Omega^g = \{\overline{a}_1\}$, every strategic type strictly higher than $\theta_p^*(\mu)$ plays $\overline{a}_1$ in every period, every strategic type strictly lower than $\theta_p^*(\mu)$ plays $\underline{a}_1$ in every period, type $\theta_p^*(\mu)$ plays $\overline{a}_1$ in every period with probability $q(\mu)$ and plays $\underline{a}_1$ in every period with probability $1 - q(\mu)$.

This uniqueness and consistency of equilibrium behavior contrasts to the private value benchmark of Fudenberg and Levine (1989), in which the patient long-run player has multiple equilibrium behaviors and behaving inconsistently can be strictly optimal. Intuitively, this is because interdependent values introduce a novel *disciplinary effect*, namely, the patient long-run player can obtain a high payoff by playing $\overline{a}_1$ consistently thanks to the presence of the commitment type, but due to the existence of bad states (i.e. states in Θ_p), his continuation payoff after behaving inconsistently will be low. My proof uses the observation that under a pessimistic prior, playing $\overline{a}_1$ in every period is optimal for some types in Θ_p, which comes from the lower bound on player 2's posterior derived in the proof of Theorem 2. When stage-game payoffs are MSM, this implies that all the good strategic types will play $\overline{a}_1$ in every period on the equilibrium path. Therefore, player 1's reputation will be bad once he fails to play $\overline{a}_1$, after which his continuation payoff will drop to its minmax.

This disciplinary effect is absent in Fudenberg and Levine (1989), as behaving inconsistently only signals that player 1 is strategic, but cannot preclude him from obtaining a high continuation payoff according to the folk theorem result in Fudenberg, Kreps and Maskin (1990). Therefore, he may have an incentive to separate from the commitment type at any given history, depending on the continuation equilibrium players coordinate on in the future. Similarly when μ is optimistic, separating from the commitment type can still lead to an optimistic posterior about the state, after which player 1's continuation payoff can be high. This leads to multiple possible behaviors with an example shown in Appendix G.6. Player 1's on-path behavior also fails to be unique in private value incomplete information games, for example when the long-run player has persistent private information about his production cost (Schmidt 1993b) or his discount factor (Ghosh and Ray 1996). This is because the bad types who have high costs or low discount factors either have no incentive to pool with the commitment type, in which case the disciplinary effect only works temporarily; or they pool with the commitment type but then they are equivalent to the latter in player 2's best-response problem. Therefore, attaching a high probability to those strategic types cannot discipline player 1 in the long-run.[16]

As a caveat, Theorem 3 does not imply the uniqueness of equilibrium or equilibrium outcome. This is because first, Bayes Nash equilibrium has no restriction on players' off-path behaviors. Second, player 2's behavior on the equilibrium path is not unique. To see this, suppose $\Omega = \{\overline{a}_1, \underline{a}_1\}$, as player 2 is indifferent after period 1 given that $\overline{a}_1$ has always been played, the dynamics of her behavior only face two constraints: (1) type θ_p^*'s indifference condition in period 0; (2) type θ_p^*'s incentives to play $\overline{a}_1$ in period $t \in \mathbb{N}$. The first one only pins down the occupation measure of $\overline{a}_2$ conditional on $\overline{a}_1$ is played in every period. The second one only requires that $\overline{a}_2$ not be too front-loaded. Nevertheless, there are still many ways to allocate the play of $\overline{a}_2$ over time that can meet both requirements, leading to multiple equilibrium outcomes.

For an overview of the extension to incorporate mixed commitment types, if all the actions in Ω^g are pure, then player 1's on-path behavior is still the same across all equilibria. However, after observing $a_1 \in \Omega^g \backslash \{\underline{a}_1\}$, player 2 won't be indifferent between $\overline{a}_2$ and $\underline{a}_2$ starting from period 1. Instead, her on-path behavior will be unique under generic μ, according to which she will have a strict incentive to play $\underline{a}_2$ in the initial N periods, followed by one period in which she is indifferent, and after that, she will have a strict incentive to play $\overline{a}_2$. The integer N depends on δ as well as the commitment action being played in period 0.

When Ω^g contains mixed actions, then the distributions over player 1's action paths are arbitrarily close across different equilibria. In particular, there exists a cutoff state $\theta_p^* \in \Theta_p$ such that all strategic types above θ_p^* play $\overline{a}_1$ in every period, all strategic types below θ_p^* play $\underline{a}_1$ in every period. Type θ_p^*'s on-path behaviors across different equilibria will coincide with (ex ante) probability at least $1-\epsilon$. Moreover, the action paths of playing $\overline{a}_1$ in every period and playing $\underline{a}_1$ in every period will occur with probability at least $1-\epsilon$ in every equilibrium, with ϵ vanishes to 0 as $\mu(\Omega) \to 0$. To summarize the general lesson in words, when the prior is pessimistic and the total probability of commitment types is small, player 1's on-path behavior is *almost unique* and attaches

[16]Indeed, the aforementioned papers restrict attention to a subset of equilibria, such as weak Markov equilibia in Schmidt (1993b) and renegotiation proof equilibria in Ghosh and Ray (1996).

arbitrarily high probability to consistent action paths.

Remarks: I conclude this subsection by commenting on the assumptions in Theorem 3. To begin with, the condition that $\mu(\Omega)$ being small is only required when $\Omega^g\backslash\{\underline{a}_1\}$ is not a singleton. In Appendix G.5, I provide an example where player 1 has multiple equilibrium behaviors when μ satisfies (4.3) but not (4.7). This is driven by the presence of other good commitment actions. Intuitively, if commitment types that are playing actions in $\Omega^g\backslash\{\overline{a}_1, \underline{a}_1\}$ occur with high probability, it will allow the bad strategic types to pool with those types and obtain high payoffs. This weakens those bad types' incentives to imitate commitment type $\overline{a}_1$ and softens the punishment for behaving inconsistently, which undermines the disciplinary effect.

Next, my behavior uniqueness result requires player 1 to be patient. I show by counterexample in Appendix G.7 that player 1 has multiple possible behaviors when δ is intermediate. Intuitively, this is because the bad strategic types have no incentive to pay the cost of playing $\overline{a}_1$ when he is impatient. This lowers their rewards from imitating the commitment type, which softens the punishment when a good strategic type plays $\overline{a}_1$ for a finite number of periods and then deviates to a lower action.

4.3 Benchmark Game without Commitment Types

In this subsection, I analyze a benchmark repeated incomplete information game *without* commitment types and then compare it to the reputation game with the same state distribution and a small probability of commitment types. The objective is to motivate the questions, conditions and results in subsection 4.2. I also address the implications of reputation effects on refining equilibria in repeated incomplete information games.

Let NE(δ, π) be the set of equilibria in the benchmark game with state distribution $\pi \in \Delta(\Theta)$. The analogs of the optimistic and pessimistic prior belief conditions in the benchmark game are given by:

$$\sum_{\theta \in \Theta_g \cup \Theta_p} \pi(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0, \tag{4.9}$$

and

$$\sum_{\theta \in \Theta_g \cup \Theta_p} \pi(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0. \tag{4.10}$$

To see how they are related to (4.2), (4.3) and (4.7), take a benchmark game with state distribution π and perturb it with a small probability of commitment types. Let μ be the type distribution in the *perturbed reputation game* and suppose the probability of commitment types $\mu(\Omega)$ is sufficiently small, $\overline{a}_1$ is one of the commitment actions and $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$. If π satisfies (4.9), then μ satisfies (4.2). If π satisfies (4.10), then μ satisfies (4.3) and (4.7). Proposition 4.1 relates whether a prior belief is optimistic or pessimistic to the attainability of player 1's complete information commitment payoff in *some* equilibria of the benchmark game.

Proposition 4.1. *There exists* $\{\sigma(\delta)\}_{\delta\in(0,1)}$ *with* $\sigma(\delta)\in NE(\delta,\pi)$ *such that*

$$\liminf_{\delta\to 1} V_\theta^{\sigma(\delta)}(\delta)\geq u_1(\theta,\overline{a}_1,\overline{a}_2) \text{ for every } \theta\in\Theta \tag{4.11}$$

if and only if π *satisfies (4.9).*

Optimistic Prior: According to Proposition 4.1 and the first statement of Theorem 2, as long as the payoff from the highest action profile is attainable in *some* equilibria of the benchmark game, it can be guaranteed in *all* equilibria of the perturbed reputation game. In terms of equilibrium refinement, reputation effects rule out equilibria with bad payoffs, for example those with payoff $\{u_1(\theta,\underline{a}_1,\underline{a}_2)\}_{\theta\in\Theta}$, and select ones that deliver every strategic type a payoff no less than his highest equilibrium payoff in a *repeated complete information game* where θ is common knowledge (Fudenberg, Kreps and Maskin 1990).

The remaining questions are: (1) Is $u_1(\theta,\overline{a}_1,\overline{a}_2)$ type θ's highest payoff in the benchmark *repeated incomplete information game*? (2) Can player 1 guarantee strictly higher payoffs by building up reputations for playing alternative commitment actions? In terms of the first question, $u_1(\overline{\theta},\overline{a}_1,\overline{a}_2)$ is type $\overline{\theta}$'s highest equilibrium payoff in the benchmark game when $\overline{a}_1$ is player 1's pure Stackelberg action. Let

$$\overline{\delta}\equiv\max_{\alpha_2\in\Delta(A_2)}\Big\{\frac{u_1(\overline{\theta},\underline{a}_1,\alpha_2)-u_1(\overline{\theta},\overline{a}_1,\alpha_2)}{u_1(\overline{\theta},\underline{a}_1,\alpha_2)-u_1(\overline{\theta},\overline{a}_1,\alpha_2)+u_1(\overline{\theta},\overline{a}_1,\overline{a}_2)-u_1(\overline{\theta},\underline{a}_1,\underline{a}_2)}\Big\}.$$

The above observation is stated as Proposition 4.2:

Proposition 4.2. *If* π *satisfies (4.9),* $\overline{a}_1$ *is player* 1*'s pure Stackelberg action in state* $\overline{\theta}$ *and* $\delta\geq\overline{\delta}$*, then:*

$$\sup_{\sigma\in NE(\delta,\pi)} V_{\overline{\theta}}^{\sigma}\in\Big[(1-\delta)u_1(\overline{\theta},\overline{a}_1,\underline{a}_2)+\delta u_1(\overline{\theta},\overline{a}_1,\overline{a}_2),\quad u_1(\overline{\theta},\overline{a}_1,\overline{a}_2)\Big].$$

Proposition 4.2 shows that by building reputations for playing the highest action, the highest type can secure his highest equilibrium payoff in the benchmark game without commitment types. Nevertheless, types strictly lower than $\overline{\theta}$ can obtain higher payoffs compared to their payoff lower bounds in Theorem 2. This is because they can extract information rent from imitating higher types. In Online Appendix C, I characterize player 1's limiting equilibrium payoff set in the entry deterrence game of section 1 and the resulting sets are depicted in Figure 2. The takeaway lesson is: introducing a small probability of commitment types rules out low-payoff equilibria and the highest equilibrium payoff for every type coincides with that in the benchmark game.

To answer the second question and to further motivate the study of player 1's guaranteed payoff from playing $\overline{a}_1$, I examine whether he can ensure himself a strictly higher payoff by establishing reputations for

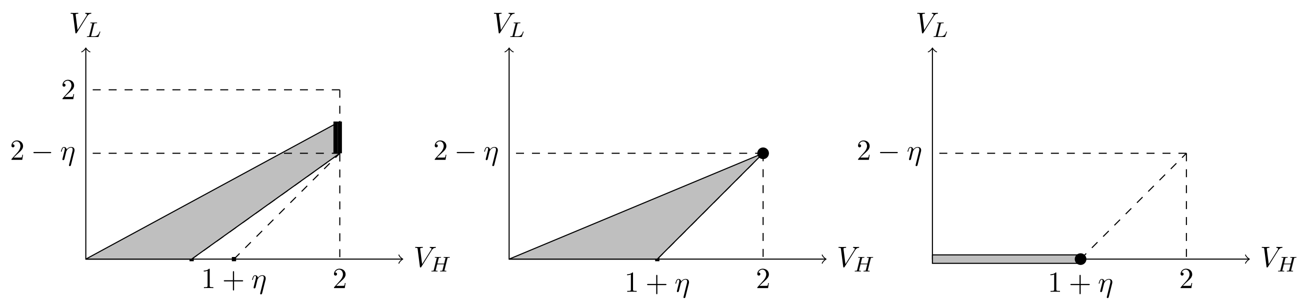

Figure 2: Limiting equilibrium payoff set in the interdependent value entry deterrence game (section 1) without commitment types (in gray) and the selected payoffs under reputation effects (in black) when $\eta \in (0,1)$. The horizontal axis: Player 1's discounted average payoff in state H. The vertical axis: Player 1's discounted average payoff in state L. Left panel: $\pi(H) > 1/2$. Middle panel: $\pi(H) = 1/2$. Right panel: $\pi(H) < 1/2$.

playing alternative commitment actions. In Online Appendix F, I adopt a notion of tightness introduced by Cripps, Schmidt and Thomas (1996) and show that when $\Theta_p \neq \{\varnothing\}$, no type of player 1 can guarantee a strictly higher payoff by establishing reputations for playing other pure commitment actions. If $\Theta_p \neq \{\varnothing\}$ and $\Theta_n = \{\varnothing\}$, in another word it is common knowledge that player 1 prefers the commitment outcome to his minmax outcome, then he cannot guarantee a strictly higher payoff by establishing reputations for playing other commitment actions, no matter whether they are pure or mixed.

Pessimistic Prior: In the complementary scenario where π is pessimistic, despite the long-run player cannot attain the payoff from the highest action profile, he is still facing a lack-of-commitment problem in the sense that there are equilibria that lead to his minmax payoff $\big\{u_1(\theta, \underline{a}_1, \underline{a}_2)\big\}_{\theta\in\Theta}$ and there are ones that result in strictly higher payoffs. Formally, let

$$v_\theta^*(\pi) \equiv \begin{cases} u_1(\theta, \underline{a}_1, \underline{a}_2) & \text{if } \theta \precsim \theta_p^*(\pi) \\ r(\pi)u_1(\theta, \overline{a}_1, \overline{a}_2) + \Big(1 - r(\pi)\Big)u_1(\theta, \overline{a}_1, \underline{a}_2) & \text{if } \theta \succ \theta_p^*(\pi), \end{cases} \tag{4.12}$$

where the cutoff state $\theta_p^*(\pi) \in \Theta_p$ is the unique element in Θ_p under which there exists $q \in (0,1]$ satisfying

$$q\pi(\theta_p^*(\pi))\mathcal{D}\Big(\theta_p^*(\pi), \overline{a}_1\Big) + \sum_{\theta \succ \theta_p^*(\pi)} \pi(\theta)\mathcal{D}\big(\theta, \overline{a}_1\big) = 0, \tag{4.13}$$

and $r(\pi) \in (0,1)$ is uniquely pinned down by:

$$r(\pi)u_1(\theta_p^*(\pi), \overline{a}_1, \overline{a}_2) + \Big(1 - r(\pi)\Big)u_1(\theta_p^*(\pi), \overline{a}_1, \underline{a}_2) = u_1(\theta_p^*(\pi), \underline{a}_1, \underline{a}_2). \tag{4.14}$$

When the total probability of commitment types is small enough, we have $\theta_p^*(\pi) = \theta_p^*(\mu)$ and $v_\theta^*(\pi) = v_\theta^*(\mu)$ for every $\theta \in \Theta$. Proposition 4.3 shows that $v_\theta^*(\pi)$ is type θ's highest payoff in the benchmark game:

Proposition 4.3. *There exists $\widehat{\delta} \in (0,1)$ such that for every $\delta > \widehat{\delta}$ and π satisfying (4.10), we have:*

$$\max_{\sigma \in NE(\delta,\pi)} V_\theta^\sigma(\delta) = v_\theta^*(\pi). \tag{4.15}$$

The proof is in Appendix E. Comparing this to the second statement of Theorem 2, player 1's highest payoff in the benchmark game can be secured in *all* equilibria of the perturbed reputation game. For an alternative interpretation, $\left(v_\theta^*(\pi)\right)_{\theta \in \Theta}$ is player 1 equilibrium payoff in the *informed principal game* where he can publicly commit to mixed actions after observing θ. Player 2 best responds to player 1's committed action based on her posterior belief about θ. In this sense, $\left(v_\theta^*(\pi)\right)_{\theta \in \Theta}$ is player 1's *commitment payoff* under a pessimistic prior π, which can be secured in all equilibria when he can build reputations.

4.4 Related Applications

I discuss the implications of my MSM conditions in two economic applications: the entry deterrence game that studies predatory pricing in monopolistic competition (Kreps and Wilson 1982, Milgrom and Roberts 1982) and the product choice game that highlights the lack-of-commitment problem in business transactions (Mailath and Samuelson 2001, Liu 2011, Ekmekci 2011).

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 per unit price. Players' stage-game payoffs are given by:

State is θ	Out	Enter
Low Price	$p_L(Q_M + \theta), 0$	$p_L(Q_D + \gamma\theta), \Pi_L(\theta) - f$
Normal Price	$p_N Q_M, 0$	$p_N Q_D, \Pi_N - f$

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, Π_L and Π_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 from the 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. As shown in Ellison and Ellison (2011), this is common in the pharmaceutical industry as advertising usually have positive spillovers to the entrant's product However, my supermodularity assumption fails in the entry deterrence problem studied in Harrington (1986), where the incumbent's and the entrant's production costs are positively correlated and the entrant does not know her own production cost 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. In this example, 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 the realized cost overruns, quality is much harder to observe, so it is reasonable to assume that future clients learn about quality mainly through the firm's past behaviors. This is modeled as a variant of the product choice game in Mailath and Samuelson (2001,2006):

$\theta =$ Good	Custom	Standardized
High Effort	$1, 3$	$-1, 2$
Low Effort	$2, 0$	$0, 1$

$\theta =$ Bad	Custom	Standardized
High Effort	$1-\eta, 0$	$-1-\eta, 1$
Low Effort	$2, -2$	$0, 0$

My MSM condition 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.

4.5 Proof Ideas of Theorems 2 and 3

I start from listing the challenges one needs to overcome to show Theorems 2 and 3. First, since values are interdependent and the commitment types can be arbitrarily unlikely compared to every strategic type, Theorem 1 suggests the necessity to exploit the properties of player 1's stage-game payoff function. Therefore, the standard arguments which are based purely on learning cannot be directly applied (for example the ones in Fudenberg and Levine 1989,1992, Sorin 1999, Gossner 2011).

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 e.g. 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:

$\theta = H$	l	r
U	$4, 8$	$0, 0$
D	$\mathbf{2, 4}$	$0, 0$

$\theta = L$	l	r
U	$-2, -2$	$\mathbf{2, 0}$
D	$0, -4$	$5, 1$

Suppose the states and players' actions are ranked 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 (θ, 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. That is to say, 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 persistent private information. The presence of intertemporal incentives provides a rationale for many different strategies and self-fulfilling beliefs that cannot be rationalized in 1-shot interactions. For 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 $\overline{a}_1$ 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.

4.5.1 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:

$\theta = H$	O	E
F	$2, 0$	$0, -1$
A	$3, 0$	$1, 2$

$\theta = L$	O	E
F	$2-\eta, 0$	$-\eta, 1$
A	$3, 0$	$1, 2$

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 two simplifying assumptions which will be 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 plan in which the committed long-run player plays F in every period when the state is H, and plays A in every period when the state is L. Translating this into the language of my model, $\Omega = \{F, A\}$, ϕ_F is the Dirac measure on state H and ϕ_A is the Dirac measure on state L.

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 $\sigma \equiv (\{\sigma_\theta\}_{\theta \in \{H,L\}}, \sigma_2)$, σ 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 σ_2. Otherwise, σ 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 equilibrium.

This result implies that in a repeated signalling game, 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 the 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 δ is large enough, player 2's posterior belief will attach higher probability to type H than to type L at every h_F^t. For some notation, 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_F^t, then:*

$$(p_t + q_t) - (p_{t+1} + q_{t+1}) \geq \mu(F)/2. \tag{4.16}$$

Proof of Claim 1: Player 2 does not have a strict incentive to play O at h_F^t 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 $p_t\geq q_t$. By rearranging the terms, one can obtain inequality (4.16). □

Claim 2. *If δ 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_F^T.[17] 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 $(h_F^{t^*},A)$ attaches probability strictly more than $1/2$ to type L, (2) type L is supposed to play A with strictly positive probability at $h_F^{t^*}$.

Let us examine type L's incentives at $h_F^{t^*}$. If he plays A, then his continuation payoff at $(h_F^{t^*},A)$ 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 $h_F^{t^*}$ 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 δ is large. This is strictly larger than 1. Comparing his continuation payoffs by playing A versus playing F reveals a contradiction. □

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 $2-\eta$.

[17] 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 played F in every period. These complications will be discussed in Remark II and a formal treatment is provided in Appendix C.

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^t.

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^t by letting

$$X_t \equiv \mu(F)\mathcal{D}(H,F)+p_t\mathcal{D}(H,F)+q_t\mathcal{D}(L,F). \tag{4.17}$$

Note that $\{X_t\}_{t=0}^{\infty}$ is a non-decreasing sequence as $\mathcal{D}(H,F)>0$, $\mathcal{D}(L,F)<0$, 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 δ is large enough, then $X_t=0$ for all $t\geq 1$.*[18]

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-1}. Since $X_t<0$, type L will play F with positive probability at h_F^{t-1}. If he plays F at h_F^{t-1}, then his continuation payoff at h_F^t is 1. If he plays A at h_F^{t-1}, then his continuation payoff at (h_F^{t-1},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 $h_F^{t^*-1}$. In what follows, I examine type L's incentives at $h_F^{t^*-1}$. If he plays A, then his continuation payoff at $(h_F^{t^*-1},A)$ is 1. If he plays F forever, then I will show below that O is player 2's strict best reply at h_F^s 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 towards a contradiction that player 2 does not have a strict incentive to play O at h_F^s for some $s\geq t^*$, then:

$$\mu(F)\mathcal{D}(H,F)+p_s\mathcal{D}(H,F)+q_{s+1}\mathcal{D}(L,F)+(q_s-q_{s+1})\mathcal{D}(L,A)\leq 0, \tag{4.18}$$

$$\Rightarrow \quad q_s-q_{s+1}\geq \frac{X_s}{\mathcal{D}(L,F)-\mathcal{D}(L,A)} \underbrace{\geq}_{\text{since } X_s\geq X_{t^*}} \underbrace{\frac{X_{t^*}}{\mathcal{D}(L,F)-\mathcal{D}(L,A)}}_{>0} \equiv Y. \tag{4.19}$$

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

[18]When 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.

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. $\square$

Remark: 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.

4.5.2 Overview of the Full Proof

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 many strategic types, and in particular, good, positive and negative states 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. This opens 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 played $\overline{a}_1$ in every period 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 equilibrium action path. 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. This implies that another switch needs to happen again in the future. Therefore, such switches must happen infinitely many times if it happens at least once, leading to a contradiction.

5 Concluding Remarks

Extensions: I discuss several extensions of my baseline model. First, players move sequentially 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 entrants' entry decisions. 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 past and update their beliefs about their payoffs from arbitration.[19] In this context, my results provide sharp predictions on the long-run player's payoff and characterize his unique 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 signal his toughness when 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 $a_{2,t} \neq \overline{a}_2$. This condition is satisfied in entry deterrence games but is violated in sequential-move product choice games,[20] or more generally, participation games defined in Ely, Fudenberg and Levine (2008).

Second, analogies of my results can be derived when there are *non-stationary* commitment types. The key difference is: the attainability of the commitment payoff from playing α_1 will depend not only on the probability of the commitment type playing α_1 in every period and its correlation with the state, but also on the

[19] Lee 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.

[20] Mailath and Samuelson (2015) provides an alternative interpretation of the product choice game. Instead of choosing whether to purchase the product or not, the consumer chooses between the customized and the standardized product, and the seller can exert high effort regardless of the consumer's action choice. My results can be applied to this game.

probabilities of those non-stationary commitment types and their correlations with the state. To illustrate this, consider the entry deterrence game in the introduction. Suppose there exists a commitment type who plays F in every period and another one who plays strategy $\widehat{\sigma}_1$, where

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

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

Nevertheless, as the state can only be learnt via the informed player's action choices, one *cannot* construct non-stationary commitment types in non-trivial interdependent value environments that can guarantee the informed player his commitment payoff regardless of the probabilities of other non-stationary commitment types. This is because for every non-stationary commitment plan $\sigma_1^* : \mathcal{H} \times \Theta \rightarrow \Delta(A_1)$, one can construct another commitment plan σ_1^{**} that

1. occurs with significantly higher probability compared to σ_1^*;
2. generates the same distribution over public histories as σ_1^*, in another word, σ_1^* and σ_1^{**} are observationally equivalent according to the uninformed players' perspective;
3. there exists a permutation $\tau : \Theta \rightarrow \Theta$ such that $\sigma_1^*(h^t, \theta) = \sigma_1^{**}(h^t, \tau(\theta))$ for every $(h^t, \theta) \in \mathcal{H} \times \Theta$, that is, the mapping from the states to the committed long-run player's stage-game actions is flipped.

This observation contrasts to the conclusion in Deb and Ishii (2018) that studies a reputation model in which the public signals can identify the state.

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 $f : \Theta \rightarrow (0, +\infty)$ such that $\widetilde{u}_1(\theta, a_1, a_2) \equiv f(\theta) u_1(\theta, a_1, a_2)$ has SID in θ and (a_1, a_2). u_2 has SID in a_2 and (θ, a_1).

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). Players are *reciprocal altruistic* in the sense that each of them maximizes a weighted average of his own monetary payoff and that of his opponent's, with the weight on the opponent's payoff being a strictly increasing function of his belief about the opponents' level of altruism (Levine 1998). 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.

Modeling Choices: I conclude by discussing several modeling assumptions and the connections to related works. A central theme of my analysis is that reputation building is challenging when the uninformed players' learning is confounded. Even though the informed player can convince his opponents about his future actions, he may not teach them how to best reply when their payoffs depend on the state. Conceptually this is related to the contemporary work of Deb and Ishii (2018) that revisits the commitment payoff theorem when the uninformed players are uncertain about the monitoring structure, captured by a perfectly persistent state.[21]

Neither model nests the other and the insights from the two papers are complementary. The main difference being: the state can be identified via the public signals in their model, while it can only be learnt via the informed player's actions in mine.[22] When the public signals can identify the state, they construct a set of non-stationary commitment types, under which the informed player can guarantee his commitment payoff regardless of the presence of other non-stationary commitment types. In contrast, I study a model in which the public signals *cannot* identify the state and all commitment types are playing state-contingent stationary strategies. My result highlights the new challenges to reputation building brought by interdependent values (Theorem 1). I also establish reputation effects on the long-run player's payoff and *behavior* in an interesting class of interdependent value games *without* state identification (Theorems 2 and 3).

In terms of the applications, their informational assumption fits in settings where informative signals about the state arrive frequently, as for example, when the state is the performance of vehicles, mobile phones, etc. In contrast, my informational assumption fits into applications where signals other than the informed player's actions are unlikely to arrive for a long time, or the variations of their realizations are mostly driven by noise orthogonal to the state. This includes for example, when the state is the resilience of an architectural design to

[21]Related ideas also appear in Wolitzky (2011), who studies reputational bargaining with non-stationary commitment types and shows the failure of the commitment payoff theorem. However, his negative result requires that the uninformed player being long-lived and the commitment types playing non-stationary strategies, none of which are needed for the counterexamples (see Appendix G) and negative results (statements 2 and 4 of Theorem 1) in my paper.

[22]Their Assumption 2.3 requires that for every pair of states $\theta, \theta' \in \Theta$, there exists $\alpha_1 \in \Delta(A_1)$ such that for every $\alpha_1' \in \Delta(A_1)$, the distribution over public signals under (θ, α_1) is different from that under (θ', α_1'). This is violated in my model as well as Aumann and Maschler (1995), Hörner and Lovo (2009), Kaya (2009), Roddie (2012), Pęski and Toikka (2017), etc. where the public signals cannot identify the state.

earthquakes, the long-run health impact of a certain type of food, the demand elasticity in markets with high sunk costs of entry, the adaptability of a software to future generations of operating systems, etc.

Ely and Välimäki (2003), Ely, Fudenberg and Levine (2008) study a class of private value reputation games with imperfect public monitoring called *participation games*. They show a bad reputation result, that a patient long-run player's equilibrium payoff is low when the *bad commitment types*, namely ones that discourage the short-run players from participating, are relatively more likely compared to the Stackelberg commitment type.

Although both my Theorem 1 and their results illustrate the possibilities of reputation failure, the economic forces behind them are different. Their bad reputation results are driven by the tension between the long-run player's forward-looking incentives and the short-run players' participation incentives. In particular, the long-run player has an incentive to take actions that can generate good signals but harm the participating short-run players. This discourages participation and prevents the long-run player from signalling his private information. In contrast, reputation failure occurs in my model as the short-run players' learning is confounded. Despite the long-run player can always choose actions to signal his type, the informational contents of his action choices are sensitive to equilibrium selection. 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 types.

Another point that is worth discussing is that in many applications, the underlying payoff environment and the characteristics of the long-run player can be changing over time. This has been taken into account in the reputation models of Tadelis (1999), Mailath and Samuelson (2001), Phelan (2006), Wiseman (2008), Ekmekci, Gossner and Wilson (2012), Board and Meyer-ter-Vehn (2013), etc.

My model abstracts away from such issues by focusing on settings where the long-run player's type (capturing both his characteristics and his knowledge about the payoff environment) is perfectly persistent. This is not to deny the importance of state-changes, but rather because the trade-off between maintaining reputations and signalling is most pronounced when the long-run player's type is perfectly persistent. As argued before, such a trade-off is the novel aspect of interdependent value environments and will be the focus of this paper. The effects and economic forces revealed in my analysis are also relevant in settings where the state is sufficiently persistent and changes infrequently, or the focus is within a certain time frame during which the probability of state-change is sufficiently low.

References

[1] Atakan, Alp and Mehmet Ekmekci (2012) "Reputation in Long-Run Relationships," *Review of Economic Studies*, 79(2), 751-780.

[2] Aumann, Robert and Michael Maschler (1995) *Repeated Games with Incomplete Information*, MIT Press.

[3] Bai, Jie (2016) "Melons as Lemons: Asymmetric Information, Consumer Learning and Seller Reputation," Working Paper, Harvard Kennedy School.

[4] Bain, Joe (1949) "A Note on Pricing in Monopoly and Oligopoly," *American Economic Review*, 39(2), 448-464.

[5] Banerjee, Abhijit and Esther Duflo (2000) "Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry," *Quarterly Journal of Economics*, 115(3), 989-1017.

[6] Bar-Isaac, Heski (2003) "Reputation and Survival: Learning in a Dynamic Signalling Model," *Review of Economic Studies*, 70(2), 231-251.

[7] Board, Simon and Mortiz Meyer-ter-Vehn (2013) "Reputation for Quality," *Econometrica*, 81(6), 2381-2462.

[8] Chung, Kai-Lai (1974) *A Course in Probability Theory*, Third Edition, Elsevier.

[9] Cripps, Martin, Eddie Dekel and Wolfgang Pesendorfer (2005) "Reputation with Equal Discounting in Repeated Games with Strictly Conflicting Interests," *Journal of Economic Theory*, 121, 259-272.

[10] Cripps, Martin, George Mailath and Larry Samuelson (2004) "Imperfect Monitoring and Impermanent Reputations," *Econometrica*, 72(2), 407-432.

[11] Cripps, Martin, Klaus Schmidt and Jonathan Thomas (1996) "Reputation in Perturbed Repeated Games," *Journal of Economic Theory*, 69(2), 387-410.

[12] Deb, Joyee and Yuhta Ishii (2019) "Reputation Building under Uncertain Monitoring," Working Paper, Yale SOM and ITAM.

[13] Echenique, Federico (2004) "Extensive-Form Games and Strategic Complementarities," *Games and Economic Behavior*, 46, 348-364.

[14] Ekmekci, Mehmet (2011) "Sustainable Reputations with Rating Systems," *Journal of Economic Theory*, 146(2), 479-503.

[15] Ekmekci, Mehmet, Olivier Gossner and Andrea Wilson (2012) "Impermanent Types and Permanent Reputations," *Journal of Economic Theory*, 147(1), 162-178.

[16] Ellison, Glenn and Sara Ellison (2011) "Strategic Entry Deterrence and the Behavior of Pharmaceutical Incumbets Prior to Patent Expiration," *American Economic Journal-Microeconomics*, 3(1), 1-36.

[17] Ely, Jeffrey, Drew Fudenberg and David Levine (2008) "When is Reputation Bad?" *Games and Economic Behavior*, 63, 498-526.

[18] Ely, Jeffrey and Juuso Välimäki (2003) "Bad Reputation," *Quarterly Journal of Economics*, 118(3), 785-814.

[19] Faingold, Eduardo (2013) "Reputation and the Flow of Information in Repeated Games," Working Paper, Yale University.

[20] Fudenberg, Drew, David Kreps and David Levine (1988) "On the Robustness of Equilibrium Refinements," *Journal of Economic Theory*, 44(2), 354-380.

[21] Fudenberg, Drew, David Kreps and Eric Maskin (1990) "Repeated Games with Long-Run and Short-Run Players," *Review of Economic Studies*, 57(4), 555-573.

[22] Fudenberg, Drew and David Levine (1989) "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57(4), 759-778.

[23] Fudenberg, Drew and David Levine (1992) "Maintaining a Reputation when Strategies are Imperfectly Observed," *Review of Economic Studies*, 59(3), 561-579.

[24] Ghosh, Parikshit and Debraj Ray (1996) "Cooperation and Community Interaction Without Information Flows," *Review of Economic Studies*, 63(3), 491-519.

[25] Gossner, Olivier (2011) "Simple Bounds on the Value of a Reputation," *Econometrica*, 79(5), 1627-1641.

[26] Harrington, Joseph (1986) "Limit Pricing when the Potential Entrant is Uncertain of Its Cost Function," *Econometrica*, 54(2), 429-437.

[27] Hart, Sergiu (1985) "Nonzero-Sum Two-Person Repeated Games with Incomplete Information," *Mathematics of Operations Research*, 10(1), 117-153.

[28] Hörner, Johannes (2002) "Reputation and Competition," *American Economic Review*, 92(3), 644-663.

[29] Hörner, Johannes and Stefano Lovo (2009) "Belief-Free Equilibria in Games with Incomplete Information," *Econometrica*, 77(2), 453-487.

[30] Hörner, Johannes, Stefano Lovo and Tritan Tomala (2011) "Belief-free Equilibria in Games with Incomplete Information: Characterization and Existence," *Journal of Economic Theory*, 146(5), 1770-1795.

[31] Jehiel, Philippe and Larry Samuelson (2012) "Reputation with Analogical Reasoning," *Quarterly Journal of Economics*, 127(4), 1927-1969.

[32] Kalai, Ehud and Ehud Lehrer (1993) "Rational Learning Leads to Nash Equilibrium," *Econometrica*, 61(5), 1019-1045.

[33] Kaya, Ayça (2009) "Repeated Singalling Games," *Games and Economic Behavior*, 66, 841-854.

[34] Klein, Benjamin and Keith Leffler (1981) "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89(4), 615-641.

[35] Kreps, David and Robert Wilson (1982) "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253-279.

[36] Lee, Jihong and Qingmin Liu (2013) "Gambling Reputation: Repeated Bargaining with Outside Options," *Econometrica*, 81(4), 1601-1672.

[37] Levine, David (1998) "Modeling Altruism and Spitefulness in Experiments," *Review of Economic Dynamics*, 1, 593-622.

[38] Liu, Qingmin (2011) "Information Acquisition and Reputation Dynamics," *Review of Economic Studies*, 78(4), 1400-1425.

[39] Liu, Qingmin and Andrzej Skrzypacz (2014) "Limited Records and Reputation Bubbles," *Journal of Economic Theory* 151, 2-29.

[40] Liu, Shuo and Harry Pei (2017) "Monotone Equilibria in Signalling Games," Working Paper, University of Zurich and MIT.

[41] Mailath, George and Larry Samuelson (2001) “Who Wants a Good Reputation?” *Review of Economic Studies*, 68(2), 415-441.

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

[43] Mailath, George and Larry Samuelson (2015) “Reputations in Repeated Games,” Handbook of Game Theory with Economic Applications, Elsevier.

[44] Milgrom, Paul and John Roberts (1982) “Predation, Reputation and Entry Deterence,” *Journal of Economic Theory*, 27, 280-312.

[45] Pei, Harry Di (2015) “Reputation with Strategic Information Disclosure,” Working Paper, MIT.

[46] Pei, Harry Di (2016) “Equilibrium Payoffs in Repeated Games with Interdependent Values,” Working Paper, MIT.

[47] Pęski, Marcin and Juuso Toikka (2017) “Value of Persistent Information,” *Econometrica*, 85(6), 1921-2948.

[48] Phelan, Christopher (2006) “Public Trust and Government Betrayal,” *Journal of Economic Theory*, 130(1), 27-43.

[49] Roddie, Charles (2012) “Signaling and Reputation in Repeated Games: Finite Games,” Working Paper, University of Cambridge.

[50] Schmidt, Klaus (1993a) “Reputation and Equilibrium Characterization in Repeated Games with Conflicting Interests,” *Econometrica*, 61(2), 325-351.

[51] Schmidt, Klaus (1993b) “Commitment through Incomplete Information in a Simple Repeated Bargaining Game,” *Journal of Economic Theory*, 60(1), 114-139.

[52] Seamans, Robert (2013) “Threat of Entry, Asymmetric Information and Pricing,” *Strategic Management Journal*, 34, 426-444.

[53] Sorin, Sylvain (1999) “Merging, Reputation, and Repeated Games with Incomplete Information,” *Games and Economic Behavior*, 29, 274-308.

[54] Sweeting, Andrew, James Roberts and Chris Gedge (2016) “A Model of Dynamic Limit Pricing with an Application to the Airline Industry,” Working Paper, University of Maryland and Duke University.

[55] Tadelis, Steven (1999) “What’s in a Name? Reputation as a Tradeable Asset,” *American Economic Review*, 89(3), 548-563.

[56] Topkis, Donald (1998) *Supermodularity and Complementarity*, Princeton University Press.

[57] Watson, Joel (1993) “A Reputation Refinement without Equilibrium,” *Econometrica*, 61(1), 199-205.

[58] Wiseman, Thomas (2008) “Reputation and Impermanent Types,” *Games and Economic Behavior*, 62(1), 190-210.

[59] Wolitzky, Alexander (2011) “Indeterminacy of Reputation Effects in Repeated Games with Contracts,” *Games and Economic Behavior*, 73, 595-607.

A Proof of Theorem 1, Statements 1 & 3

Recall the definitions of strategic types and commitment types in section 2. Abusing notation, I use Θ to denote the set of strategic types and Ω to denote the set of commitment types. Let $\overline{\Omega} \equiv \Omega \bigcup \Theta$ be the entire set of types with ω a typical element of $\overline{\Omega}$. Let $\mu_t(\omega)$ be the probability of type ω under the posterior in period t.

Let $\alpha_1^* \in \Omega$ be the commitment action under consideration, in order to distinguish it from a generic action. If α_1^* is a Dirac measure on a_1^*, I will replace α_1^* with a_1^* for notation simplicity. Recall that $\text{BR}_2(a_1^*, \theta) \equiv \{a_2^*\}$ (or $\text{BR}_2(\alpha_1^*, \theta) \equiv \{a_2^*\}$). Since $\Lambda(a_1^*, \theta) = \{\varnothing\}$ (or $\underline{\Lambda}(\alpha_1^*, \theta) = \{\varnothing\}$) if $\text{BR}_2(a_1^*, \phi_{a_1^*}) \neq \{a_2^*\}$ (or $\text{BR}_2(\alpha_1^*, \phi_{\alpha_1^*}) \neq \{a_2^*\}$), 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 $\text{BR}_2(\alpha_1^*, \phi_{\alpha_1^*}) = \{a_2^*\}$).

A.1 Proof of Statement 1

When $\Omega = \{a_1^*\}$ and $\lambda \in \Lambda(a_1^*, \theta)$, for every $\tilde{\mu}$ with $\tilde{\mu}(\tilde{\theta}) \in [0, \mu(\tilde{\theta})]$ for all $\tilde{\theta} \in \Theta$, we have:

$$\{a_2^*\} = \arg\max_{a_2 \in A_2} \Big\{ \mu(a_1^*) u_2(\phi_{a_1^*}, a_1^*, a_2) + \sum_{\tilde{\theta} \in \Theta} \tilde{\mu}(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2) \Big\}.$$

Let h_*^t be the period t public history such that a_1^* is always played. For every $\omega \in \overline{\Omega}$, let $q_t(\omega)$ be the ex ante probability that the history is h_*^t and player 1 is type ω. By definition, $q_t(a_1^*) = \mu(a_1^*)$ for all t. Player 2's maximization problem at h_*^t is:

$$\max_{a_2 \in A_2} \Big\{ \mu(a_1^*) u_2(\phi_{a_1^*}, a_1^*, a_2) + \sum_{\tilde{\theta} \in \Theta} \Big[q_{t+1}(\tilde{\theta}) u_2(\tilde{\theta}, a_1^*, a_2) + (q_t(\tilde{\theta}) - q_{t+1}(\tilde{\theta})) u_2(\tilde{\theta}, \alpha_{1,t}(\tilde{\theta}), a_2) \Big] \Big\}$$

where $\alpha_{1,t}(\tilde{\theta}) \in \Delta(A_1 \backslash \{a_1^*\})$ 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 $\overline{T} \equiv \lceil 1/\rho \rceil$, which is independent of δ. There exist at most $\overline{T}$ periods in which a_2^* fails to be a strict best reply conditional on a_1^* has always been played. Therefore, type θ's payoff is bounded from below by:

$$(1 - \delta^{\overline{T}}) \min_{a \in A} u_1(\theta, a) + \delta^{\overline{T}} v_\theta(a_1^*),$$

which converges to $v_\theta(a_1^*)$ as $\delta \to 1$.

When there are other commitment types, let $\overline{p} \equiv \max_{\alpha_1 \in \Omega \backslash \{a_1^*\}} \alpha_1(a_1^*)$. which is strictly below 1. There exists $T \in \mathbb{N}$ independent of δ, such that for every $t \geq T$, a_2^* is player 2's strict best reply at h_*^t if: $\sum_{\tilde{\theta} \in \Theta} q_{t+1}(\tilde{\theta}) \geq \sum_{\tilde{\theta} \in \Theta} q_t(\tilde{\theta}) - \rho/2$. One can obtain the commitment payoff bound by considering the subgame starting from h_*^T.

A.2 Proof of Statement 3

Notation: For every $\alpha_1 \in \Omega \backslash \{\alpha_1^*\}$, $\theta \in \Theta$ and $\tilde{\mu} \in \Delta(\overline{\Omega})$ with $\tilde{\mu}(\alpha_1^*) \neq 0$, let

$$\tilde{\lambda}(\alpha_1) \equiv \tilde{\mu}(\alpha_1)/\tilde{\mu}(\alpha_1^*) \text{ and } \tilde{\lambda}(\theta) \equiv \tilde{\mu}(\theta)/\tilde{\mu}(\alpha_1^*)$$

Abusing notation, let $\tilde{\lambda} \equiv \Big(\big(\tilde{\lambda}(\alpha_1)\big)_{\alpha_1 \in \Omega \backslash \{\alpha_1^*\}}, \big(\tilde{\lambda}(\theta)\big)_{\theta \in \Theta} \Big)$ be the (expanded) likelihood ratio vector. Let $n \equiv |A_1|$ and $m \equiv |\overline{\Omega}| - 1$. For convenience, I write $\overline{\Omega} \backslash \{\alpha_1^*\} \equiv \{\omega_1, ..., \omega_m\}$ and $\tilde{\lambda} \equiv (\tilde{\lambda}_1, ..., \tilde{\lambda}_m)$. The proof consists of two parts.

A.2.1 Part I

Let Σ_2 be the set of player 2's strategies with σ_2 a typical element. Let

$$\text{NE}_2(\mu) \equiv \Big\{\sigma_2 \in \Sigma_2 \Big| \exists\, \delta \in (0,1) \text{ such that } (\sigma_1, \sigma_2) \in \text{NE}(\delta, \mu)\Big\}.$$

For every $\sigma_\omega : \mathcal{H} \to \Delta(A_1)$ and $\sigma_2 \in \Sigma_2$, 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 α_1^* to denote the strategy of playing α_1^* in every period.

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

$$\underline{\Lambda}(\psi, \chi) \equiv \Big\{\tilde{\lambda} \Big| \sum_{i=1}^{m} \tilde{\lambda}_i / \psi_i = \chi\Big\}.$$

Let λ be the likelihood ratio vector induced by player 2's prior belief μ. 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^∞, let h_t^∞ be its projection on $a_{1,t}$. Let $\alpha_1(\cdot|h^t) \in \Delta(A_1)$ 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 $\chi > 0$, $\lambda \in \underline{\Lambda}(\psi, \chi)$, $\sigma_2 \in NE_2(\mu)$ and $\epsilon > 0$, there exist $\overline{\delta} \in (0,1)$ and $T \in \mathbb{N}$ such that for every $\delta > \overline{\delta}$, there exists $\sigma_\omega : \mathcal{H} \to \Delta(A_1)$ that achieves the three objectives simultaneously:*

1. $$\lambda(h^t) \in \bigcup_{\widetilde{\chi} \in [0, \chi+\epsilon)} \underline{\Lambda}(\psi, \widetilde{\chi}) \quad \textit{for every} \quad h^t \in \mathcal{H}^{(\sigma_\omega, \sigma_2)}. \tag{A.1}$$

2. *For every $h^\infty \in \mathcal{H}^{(\sigma_\omega, \sigma_2)}$ and every $a_1 \in A_1$,*
$$\Big|\sum_{t=0}^{\infty} (1-\delta)\delta^t \mathbf{1}\{h_t^\infty = a_1\} - \alpha_1^*(a_1)\Big| < \frac{\epsilon}{2(2\chi + \epsilon)}. \tag{A.2}$$

3. $$\mathbb{E}^{(\sigma_\omega, \sigma_2)}\Big[\#\Big\{t \Big| d\big(\alpha_1^* \big\| \alpha_1(\cdot|h^t)\big) > \epsilon^2/2\Big\}\Big] < T. \tag{A.3}$$

In words, Proposition A.1 claims that for every equilibrium strategy of player 2's, there exists a mapping from the set of public histories to player 1's mixed actions under which the following three goals can be achieved simultaneously: (1) inducing favorable beliefs about the state; (2) the discounted average frequency of player 1's actions along every infinite action path is closely matched to α_1^*; (3) the expected number of periods in which player 2's believed action differs significantly from α_1^* is uniformly bounded from above by an integer independent of δ. My proof consists of three steps, which shows how to achieve each objective without compromising on the other two.

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 plays α_1^* in every period.

Let $\chi(h^t) \equiv \sum_{i=1}^m \lambda_i(h^t)/\psi_i$. Since $\lambda \in \underline{\Lambda}(\psi, \chi)$, 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 $\{\mathcal{F}^t\}_{t \in \mathbb{N}}$ is the filtration induced by the public history,[23] 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

[23]When α_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 ω_i plays action $a_1' \notin A_1^*$ with positive probability, then the expected value of $\lambda_i(h^t)$ can strictly decrease.

Upcrossing Inequality (see Chung 1974),

$$\mathcal{P}^{(\alpha_1^*,\sigma_2)}\Big\{U(\chi,\chi+\frac{\epsilon}{2})\geq 1\Big\}\leq\frac{2\chi}{2\chi+\epsilon}. \tag{A.4}$$

Let $\widetilde{\mathcal{H}}^\infty$ be the set of infinite histories that $\chi(h^t)$ is below $\chi+\frac{\epsilon}{2}$ for all periods. According to (A.4), it occurs with probability at least $\frac{\epsilon}{2\chi+\epsilon}$.

Step 2: In this step, I show that for large enough δ, 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 subset is $\epsilon-$close to α_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 $\mathcal{P}^{(\alpha_1^*,\sigma_2)}$, $X_t=1$ with probability $\alpha_1^*(a_1)$. Therefore, 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 show the following Lemma:

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

$$\limsup_{\delta\to 1}\mathcal{P}^{(\alpha_1^*,\sigma_2)}\Big(\Big|\sum_{t=0}^{+\infty}(1-\delta)\delta^t X_t-\alpha_1^*(a_1)\Big|\geq\varepsilon\Big)\leq\frac{\varepsilon}{n}. \tag{A.5}$$

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

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

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

$$\frac{\sum_{n=1}^k\widehat{X}_n}{1+\delta+...+\delta^{k-1}}=\sigma\sqrt{\frac{1-\delta^{2k}}{1-\delta^2}}\frac{1-\delta}{1-\delta^k}Z_k.$$

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

$$\sigma^2\frac{1-\delta^{2k}}{1-\delta^2}\frac{(1-\delta)^2}{(1-\delta^k)^2}.$$

The variance term converges to $\mathcal{O}\Big((1-\delta)\Big)$ as $k\to\infty$. According to Theorem 7.4.1 in Chung (1974), we have:

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

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

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

$$\mathcal{P}^{(\alpha_1^*,\sigma_2)}\Big(\Big|\frac{\sum_{i=1}^k\widehat{X}_n}{1+\delta+...+\delta^{k-1}}\Big|\geq\varepsilon\Big)<\frac{\varepsilon}{n}.$$

Taking the $k\to\infty$ limit, one obtains the conclusion of Lemma A.1. □

Step 3: According to Lemma A.1, for every $a_1 \in A_1$ and $\epsilon > 0$, there exists $\overline{\delta} \in (0,1)$, such that for all $\delta > \overline{\delta}$, there exists $\mathcal{H}^{\infty}_{\epsilon,a_1}(\delta) \subset \mathcal{H}^{\infty}$, such that

1. $$\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\mathcal{H}^{\infty}_{\varepsilon,a_1}(\delta)) \geq 1-\varepsilon/n, \tag{A.6}$$

2. For every $h^{\infty} \in \mathcal{H}^{\infty}_{\varepsilon,a_1}(\delta)$, the discounted average frequency of a_1 is ε-close to $\alpha_1^*(a_1)$.

Let $\mathcal{H}^{\infty}_{\varepsilon}(\delta) \equiv \bigcap_{a_1 \in A_1} \mathcal{H}^{\infty}_{\varepsilon,a_1}(\delta)$. According to (A.6), we have:

$$\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\mathcal{H}^{\infty}_{\varepsilon}(\delta)) \geq 1-\varepsilon. \tag{A.7}$$

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

$$\widehat{\mathcal{H}}^{\infty} \equiv \widetilde{\mathcal{H}}^{\infty} \bigcap \mathcal{H}^{\infty}_{\varepsilon}(\delta), \tag{A.8}$$

we have:

$$\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\widehat{\mathcal{H}}^{\infty}) \geq \frac{\epsilon}{2(2\chi+\epsilon)} \tag{A.9}$$

According to Gossner (2011), we have

$$\mathbb{E}^{(\alpha_1^*,\sigma_2)}\Big[\sum_{\tau=0}^{+\infty} d(\alpha^*||\alpha(\cdot|h^{\tau}))\Big] \leq -\log \mu(\alpha_1^*), \tag{A.10}$$

where $d(\cdot\|\cdot)$ is the Kullback-Leibler divergence between two action distributions. The Markov Inequality implies that:

$$\mathbb{E}^{(\alpha_1^*,\sigma_2)}\Big[\sum_{\tau=0}^{+\infty} d(\alpha^*||\alpha(\cdot|h^{\tau}))\Big|\widehat{\mathcal{H}}^{\infty}\Big] \leq -\frac{2(2\chi+\epsilon)\log \mu(\alpha_1^*)}{\epsilon}. \tag{A.11}$$

Let $\mathcal{P}^*$ be the probability measure over $\mathcal{H}^{\infty}$ such that for every $\mathcal{H}^{\infty}_0 \subset \mathcal{H}^{\infty}$,

$$\mathcal{P}^*(\mathcal{H}^{\infty}_0) \equiv \frac{\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\mathcal{H}^{\infty}_0 \bigcap \widehat{\mathcal{H}}^{\infty})}{\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\widehat{\mathcal{H}}^{\infty})}.$$

Let $\sigma_{\omega} : \mathcal{H} \to \Delta(A_1)$ be player 1's strategy that induces $\mathcal{P}^*$. The expected number of periods in which $d(\alpha_1^*||\alpha(\cdot|h^t)) > \epsilon^2/2$ is bounded from above by:

$$T \equiv \Big\lceil -\frac{4(2\chi+\epsilon)\log \mu(\alpha_1^*)}{\epsilon^3}\Big\rceil, \tag{A.12}$$

which is an integer independent of δ. The three steps together imply Proposition A.1.

A.2.2 Part II

Proposition A.1 and $\lambda \in \underline{\Lambda}(\alpha_1^*,\theta)$ do not imply that type θ can guarantee himself his commitment payoff. This is because due to the correlations between player 1's type and his action choice, player 2 may not have an incentive to play a_2^* even if $\lambda \in \underline{\Lambda}(\alpha_1^*,\theta)$ and the average action is close to α_1^*. I address this issue using two observations, which correspond to the two steps of my proof.

1. If $\lambda \in \underline{\Lambda}(\alpha_1^*,\theta)$, is small in all but at most one entry and player 1's average action is close to α_1^*, then player 2 has a strict incentive to play a_2^* regardless of the correlation. Let Λ^0 be the set of beliefs that has the above feature.

2. If $\lambda \in \underline{\Lambda}(\alpha_1^*, \theta)$ and player 1's average action is close to α_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 choice.

   I show that for every $\lambda \in \underline{\Lambda}(\alpha_1^*, \theta)$, there exists an integer K 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 K such periods, player 2's belief about his type will be in Λ^0, which concludes the proof.

Recall that $m \equiv |\overline{\Omega}| - 1$. 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 ψ_i equals to the intercept of $\Lambda(\alpha_1^*, \theta)$ on dimension ω_i.
- Otherwise, $\psi_i > 0$ is chosen to be large enough such that

$$\sum_{i=1}^{m} \lambda_i/\psi_i < 1. \tag{A.13}$$

Such ψ exists as $\lambda \in \underline{\Lambda}(\alpha_1^*, \theta)$. Let $\overline{\psi} \equiv \max\{\psi_j | j = 1, 2, ..., m\}$. Recall that Part I has established the existence of a strategy for player 1 under which:

1. Player 2's belief always satisfies (A.13), or more precisely, bounded from above by some $\chi < 1$.
2. The discounted average frequency of every action $a_1 \in A_1$ at every on-path history with infinite length is ϵ-close to the probability attached to a_1 in α_1^*.
3. In expectation, there exists at most T periods in which player 2's believed action differs significantly from α_1^*, where T is an integer independent of δ.

Step 1: For every $\xi > 0$, a likelihood ratio vector λ is of '**size $\boldsymbol{\xi}$**' if 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 moreover,

$$\lambda \in \Big\{\tilde{\lambda} \in \mathbb{R}_+^m \Big| \sum_{i=1}^{m} \tilde{\lambda}_i/\tilde{\psi}_i < 1\Big\} \subset \Big\{\tilde{\lambda} \in \mathbb{R}_+^m \Big| \#\{i | \tilde{\lambda}_i \leq \xi\} \geq m - 1\Big\}. \tag{A.14}$$

Intuitively, λ is of size ξ 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 ξ. By definition, for every $\xi' \in (0, \xi)$, if λ is of size ξ', then it is also of size ξ. Proposition A.2 establishes the commitment payoff bound when λ is of size ξ for ξ small enough.

Proposition A.2. *There exists $\xi > 0$, such that for every λ of size ξ, we have:*

$$\liminf_{\delta \to 1} \underline{V}_\theta(\mu, \delta) \geq u_1(\theta, \alpha_1^*, a_2^*).$$

In the proof, I show that using the strategy constructed in Proposition A.1, one 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 ω_i at history h^t. Let

$$B_{i,a_1}(h^t) \equiv \lambda_i(h^t)\Big(\alpha_1^*(a_1) - \alpha_1(a_1|h^t, \omega_i)\Big). \tag{A.15}$$

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 evert $\lambda \in \underline{\Lambda}(\alpha_1^*, \theta)$ and $\epsilon > 0$, there exists $\varepsilon > 0$ such that at every likelihood ratio vector $\widetilde{\lambda}$ satisfying:

$$\sum_{i=1}^{m} \widetilde{\lambda}_i/\psi_i < \frac{1}{2}\Big(1 + \sum_{i=1}^{m} \lambda_i/\psi_i\Big), \tag{A.16}$$

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

1. $|B_{i,a_1}(h^t)| < \varepsilon$ for all i and a_1.
2. $\big\|\alpha_1^* - \alpha_1(\cdot|h^t)\big\| \leq \epsilon$.

This is because when the prior belief satisfies (A.16), a_2^* is player 2's strict best reply when all types of player 1 are playing α_1^*. When ϵ and ε are both small enough, an ϵ-deviation of the average action together with an ε correlation between types and actions cannot overturn this strictness.

According to the Pinsker's Inequality, $\big\|\alpha_1^* - \alpha_1(\cdot|h^t)\big\| \leq \epsilon$ is implied by $d(\alpha_1^*||\alpha_1(\cdot|h^t)) \leq \epsilon^2/2$. Pick ϵ and ξ small enough such that:

$$\epsilon < \frac{\varepsilon}{2(1+\overline{\psi})} \tag{A.17}$$

and

$$\xi < \frac{\varepsilon}{(m-1)(1+\varepsilon)}. \tag{A.18}$$

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

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

The triangular inequality implies that:

$$\begin{aligned}\Big\|\lambda_1(\alpha_1^* - \alpha_1(a_1|h^t, \omega_1))\Big\| &\leq \sum_{i=2}^{m} \Big\|\lambda_i(\alpha_1^* - \alpha_1(a_1|h^t, \omega_i))\Big\| + \epsilon\Big(1 + \lambda_1 + \xi(m-1)\Big) \\ &\leq \xi(m-1) + \epsilon\Big(1 + \overline{\psi} + \xi(m-1)\Big) \leq \varepsilon. \end{aligned} \tag{A.19}$$

where the last inequality uses (A.17) and (A.18). Inequality (A.19) implies that $||B_{1,a_1}(h^t)|| \leq \varepsilon$. As a result, for every λ of size ξ, a_2^* is player 2's strict best reply at every history h^t satisfying $d(\alpha_1^*||\alpha_1(\cdot|h^t)) \leq \epsilon^2/2$. This in turn implies the validity of the commitment payoff bound. □

Step 2: In this step, I apply the conclusions of Propositions A.1 and A.2 to establish the mixed commitment payoff bound for every λ satisfying (A.13). Recall the definition of $B_{i,a_1}(h^t)$ in (A.15). According to Bayes Rule, if $a_1 \in A_1^*$ is observed at h^t, then

$$\lambda_i(h^t) - \lambda_i(h^t, a_1) = \frac{B_{i,a_1}(h^t)}{\alpha_1^*(a_1)} \quad \text{and} \quad \sum_{a_1 \in A_1^*} \alpha_1^*(a_1)\Big(\lambda_i(h^t) - \lambda_i(h^t, a_1)\Big) \geq 0.$$

Let

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

Suppose $B_{i,a_1}(h^t) \geq \varepsilon$ for some i and $a_1 \in A_1^*$, we have $||D(h^t, a_1)|| \geq \varepsilon$ where $||\cdot||$ denotes the $\mathcal{L}^2$-norm. Pick $\xi > 0$ small enough to meet the requirement in Proposition A.2. I define two sequences of subsets of $\underline{\Lambda}(\alpha_1^*, \theta)$, namely $\{\Lambda^k\}_{k=0}^{\infty}$ and $\{\widehat{\Lambda}^k\}_{k=1}^{\infty}$, recursively as follows:

– Let Λ^0 be the set of likelihood ratio vectors that are of size ξ,

– For every $k \geq 1$, let $\widehat{\Lambda}^k$ be the set of likelihood ratio vectors in $\underline{\Lambda}(\alpha_1^*, \theta)$ such that if $\lambda(h^t) \in \widehat{\Lambda}^k$, 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 $a_1 \in A_1^*$, there exists $a_1^* \in A_1^*$ such that $\lambda(h^t, a_1^*) \in \Lambda^{k-1}$.

– Let Λ^k be the set of likelihood ratio vectors in $\underline{\Lambda}(\alpha_1^*, \theta)$ 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

$$\lambda \in \Big\{\tilde{\lambda} \in \mathbb{R}_+^m \Big| \sum_{i=1}^m \tilde{\lambda}_i / \tilde{\psi}_i < 1\Big\} \subset \Big(\bigcup_{j=0}^{k-1} \Lambda^j\Big) \bigcup \widehat{\Lambda}^k. \tag{A.20}$$

By construction, we know that:

$$\Big\{\tilde{\lambda} \in \mathbb{R}_+^m \Big| \sum_{i=1}^m \tilde{\lambda}_i / \tilde{\psi}_i < 1\Big\} \subset \bigcup_{j=0}^{k} \Lambda^j = \Lambda^k. \tag{A.21}$$

Since $(0, ..., \psi_i - \upsilon, ..., 0) \in \Lambda^0$ for any $i \in \{1, 2, ..., m\}$ and $\upsilon > 0$, so $\text{co}(\Lambda^0) = \underline{\Lambda}(\alpha_1^*, \theta)$. By definition, $\{\Lambda^k\}_{k \in \mathbb{N}}$ is an increasing sequence with $\Lambda^k \subset \underline{\Lambda}(\alpha_1^*, \theta) = \text{co}(\Lambda^k)$ 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 $\text{clo}\Big(\underline{\Lambda}(\alpha_1^*, \theta)\Big)$. The next Lemma shows that $\text{clo}(\Lambda^\infty)$ coincides with $\text{clo}\Big(\underline{\Lambda}(\alpha_1^*, \theta)\Big)$.

Lemma A.2. *$clo(\Lambda^\infty) = clo\Big(\underline{\Lambda}(\alpha_1^*, \theta)\Big)$*

PROOF OF LEMMA A.2: Since $\Lambda^k \subset \underline{\Lambda}(\alpha_1^*, \theta)$ for every $k \in \mathbb{N}$, we know that $\text{clo}(\Lambda^\infty) \subset \text{clo}\Big(\underline{\Lambda}(\alpha_1^*, \theta)\Big)$. The rest of the proof establishes the other direction. Suppose towards a contradiction that

$$\text{clo}(\Lambda^\infty) \subsetneq \text{clo}\Big(\underline{\Lambda}(\alpha_1^*, \theta)\Big) \tag{A.22}$$

1. Let $\widehat{\Lambda} \subset \underline{\Lambda}(\alpha_1^*, \theta)$ be such that if $\lambda(h^t) \in \widehat{\Lambda}$, then either $\lambda(h^t) \in \Lambda^\infty$ or:

– For every $\{\alpha_1(\cdot|h^t, \omega_i)\}_{i=1}^m$ such that $||D(h^t, a_1)|| \geq \varepsilon$ for some $a_1 \in A_1^*$, there exists $a_1^* \in A_1^*$ such that $\lambda(h^t, a_1^*) \in \Lambda^\infty$.

2. Let $\breve{\Lambda}$ be the set of likelihood ratio vectors in $\underline{\Lambda}(\alpha_1^*, \theta)$ such that for every $\widetilde{\lambda} \in \breve{\Lambda}$, there exists $\widetilde{\psi} \equiv (\widetilde{\psi}_1, ..., \widetilde{\psi}_m) \in \mathbb{R}_+^m$ such that:

$$\widetilde{\psi}_i \in (0, \psi_i) \text{ for all } i \text{ and } \lambda \in \Big\{\widetilde{\lambda} \in \mathbb{R}_+^m \Big| \sum_{i=1}^m \widetilde{\lambda}_i / \widetilde{\psi}_i < 1\Big\} \subset \Big(\Lambda^\infty \bigcup \widehat{\Lambda}\Big). \tag{A.23}$$

Since Λ^∞ is defined as the limit of the above operator, so in order for (A.22) to be true, it has to be the case that $\breve{\Lambda} = \Lambda^\infty$, or $\Xi \bigcap \breve{\Lambda} = \{\varnothing\}$ where

$$\Xi \equiv \text{clo}\Big(\underline{\Lambda}(\alpha_1^*, \theta)\Big) \Big\backslash \text{clo}(\Lambda^\infty). \tag{A.24}$$

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

1. $\sum_{i=1}^m x_i / \psi_i = \chi$.

2. $\partial B(x,r)\bigcap H(\chi)\bigcap \underline{\Lambda}(\alpha_1^*,\theta)\subset \Lambda^\infty$ for every $r\geq \varrho$.

3. For every $r\geq\rho$ and $y\in\partial B(x,r)\bigcap\underline{\Lambda}(\alpha_1^*,\theta)$, 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(Λ^∞). Suppose $\lambda(h^t)=x$ for some h^t and there exists $a_1\in A_1^*$ such that $||D(h^t,a_1)||\geq\varepsilon$,

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

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

$$\sum_{a_1'\in A_1^*}\alpha_1^*(a_1')\lambda_i(h^t,a_1')\leq\lambda_i(h^t) \tag{A.25}$$

  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:

$$\sum_{a_1'\in A_1^*}\alpha_1^*(a_1')\sum_{i=1}^m\lambda_i(h^t,a_1')/\psi_i\leq\chi. \tag{A.26}$$

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

$$\rho\equiv\frac{1}{2}\min_{a_1\in A_1^*}\{\alpha_1^*(a_1)\}\varepsilon\kappa,$$

  (A.25) implies the existence of $a_1^*\in A_1^*\backslash\{a_1\}$ such that $\lambda(h^t,a_1^*)\in H(\chi)\bigcap B(x,\rho)$. Requirement 2 then implies that $x=\lambda(h^t)\in\breve{\Lambda}$. Since $x\in\Xi$, this leads to a contradiction.

Therefore, (A.22) cannot be true, which validates the conclusion of Lemma A.2. □

Lemma A.2 implies that for every $\lambda\in\underline{\Lambda}(\alpha_1^*,\theta)$, there exists an integer $K\in\mathbb{N}$ independent of δ such that $\lambda\in\Lambda^K$. Statement 3 of Theorem 1 can then be shown by induction on K. According to Proposition A.2, the statement holds for $K=0$. Suppose it applies to every $K\leq K^*-1$, let us consider the case when $K=K^*$. According to the construction of Λ^{K^*}, there exists a strategy for player 1 such that whenever a_2^* is not player 2's best reply despite $d(\alpha_1^*\|\alpha_1(\cdot|h^t))<\epsilon^2/2$, then the posterior belief after observing $a_{1,t}$ is in Λ^{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 show statement 2 and explain several related issues. The proof of statement 4 involves some additional technical steps, which will be relegated to Online Appendix B. The key intuition behind the distinctions between pure and mixed commitment actions will be summarized in Proposition B.3 and Proposition B.6 in Online Appendix B. I use a_1^* to denote the pure commitment action. Let $\overline{\Pi}(a_1^*,\theta)$, $\Pi(a_1^*,\theta)$ and $\underline{\Pi}(a_1^*,\theta)$ be the exteriors of $\overline{\Lambda}(a_1^*,\theta)$, $\Lambda(a_1^*,\theta)$ and $\underline{\Lambda}(a_1^*,\theta)$, respectively. I start with the following Lemma, which clarifies the role of the assumption that $\mathrm{BR}_2(a_1^*,\phi_{a_1^*})$ being a singleton.

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

$$\sum_{\tilde\theta\in\Theta^b_{(a_1^*,\theta)}}\lambda'(\tilde\theta)\Big(u_2(\tilde\theta,a_1^*,a_2')-u_2(\tilde\theta,a_1^*,a_2^*)\Big)>0 \tag{B.1}$$

as long as one the following conditions hold:

1. $\Lambda(a_1^*,\theta)\neq\{\varnothing\}$.

2. $\Lambda(a_1^*,\theta)=\{\varnothing\}$ *and* $BR_2(a_1^*,\phi_{a_1^*})$ *is a singleton.*

3. $\Lambda(a_1^*,\theta)=\{\varnothing\}$ *and* $a_2^*\notin BR_2(a_1^*,\phi_{a_1^*})$.

PROOF OF LEMMA B.1: When $\Lambda(a_1^*,\theta)\neq\{\varnothing\}$, by definition of $\Pi(a_1^*,\theta)$, there exists $0\ll\lambda'\ll\lambda$ and $a_2'\neq a_2^*$ such that:

$$\Big(u_2(\phi_{a_1^*},a_1^*,a_2')-u_2(\phi_{a_1^*},a_1^*,a_2^*)\Big)+\sum_{\tilde{\theta}\in\Theta^b_{(a_1^*,\theta)}}\lambda'(\tilde{\theta})\Big(u_2(\tilde{\theta},a_1^*,a_2')-u_2(\tilde{\theta},a_1^*,a_2^*)\Big)>0. \tag{B.2}$$

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, $\theta'\in\Theta^b_{(a_1^*,\theta)}$. Let

$$\lambda'(\tilde{\theta})\equiv\begin{cases}\lambda(\tilde{\theta}) & \text{if } \tilde{\theta}=\theta'\\ 0 & \text{otherwise}\end{cases}$$

λ' satisfies (B.1) due to the full support condition, i.e. $\mu(\omega)>0$ for every $\omega\in\overline{\Omega}$. □

Remark: As long as one of the three conditions in Lemma B.1 applies, one can dispense the assumption that $\text{BR}_2(a_1^*,\phi_{a_1^*})$ is a singleton.

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 $\Theta=\{\theta,\theta'\}$, $A_1=\{a_1^*,a_1'\}$, $A_2=\{a_2^*,a_2'\}$ and $\Omega=\{a_1^*,(1-\epsilon)a_1^*+\epsilon a_1'\}$ with $\phi_{a_1^*}(\theta')=1$ and $\phi_{(1-\epsilon)a_1^*+\epsilon a_1'}(\theta)=1$. Suppose $\{a_2^*\}=\text{BR}_2(a_1^*,\theta)=\text{BR}_2(a_1',\theta)$ and $\{a_2^*,a_2'\}=\text{BR}_2(a_1^*,\theta')=\text{BR}_2(a_1',\theta')$. Then type θ can guarantee himself payoff $u_1(\theta,a_1^*,a_2^*)$ by playing a_1^* in every period despite $\lambda\in\Pi(a_1^*,\theta)$ since a_1' is always player 2's strictly best reply given the presence of commitment type playing $(1-\epsilon)a_1^*+\epsilon a_1'$.

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

$$u_1(\tilde{\theta},a_1,a_2)\equiv\mathbf{1}\{\tilde{\theta}=\theta,a_1=a_1^*,a_2=a_2^*\}. \tag{B.3}$$

By definition, $v_\theta(a_1^*)=1$. I construct a sequential equilibrium that consists of a *normal phase* and an *abnormal phase*. Type θ's equilibrium action is pure at every history occurring with positive probability under (σ_θ,σ_2). Play starts from the normal phase and remains in it as long as the history of play is consistent with type θ's equilibrium strategy. Otherwise, play switches to the abnormal phase, which is absorbing. Let $A_1\equiv\{a_1^0,...,a_1^{n-1}\}$. I show there exists a constant $q\in(0,1)$ (independent of δ) such that:

1. After a bounded number of periods (uniform for all large enough δ), type θ obtains expected payoff $1-q$ in every period of the normal phase, i.e. his payoff is approximately $1-q$ when $\delta\to1$.

2. Type θ's continuation payoff is bounded below $1-2q$ in the beginning of the abnormal phase.

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

– **'Mechanical' Strategic Types:** Every strategic type in $\Theta_{(a_1^*,\theta)} \backslash \{\theta\}$ plays $\alpha_1 \in \Omega \backslash \{a_1^*\}$ in every period, with α_1 being arbitrarily chosen. If $\Omega = \{a_1^*\}$, then all types in $\Theta_{(a_1^*,\theta)} \backslash \{\theta\}$ play some arbitrarily chosen $\alpha_1 \in \Delta(A_1)$ such that α_1 is not the Dirac measure on a_1^*.

  Recall the construction of λ' in in Lemma B.1 and (B.2). For every strategic type $\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}$, he plays α_1 in every period with probability $x(\tilde{\theta}) \in [0,1]$ such that:

  > The likelihood ratio equals to λ' conditional on the union of the following set of events (1) player 1 is a bad strategic type and is *not* playing α_1 in every period; (2) player 1 is the commitment type that is playing a_1^* in every period.

  In what follows, I treat the strategic types that are playing α_1 in every period as the commitment type playing α_1. Formally, let

$$\widetilde{\Omega} \equiv \begin{cases} \{\alpha_1\} & \text{if } |\Omega| = 1 \\ \Omega \backslash \{a_1^*\} & \text{otherwise }. \end{cases}$$

  Let $l \equiv |\widetilde{\Omega}|$, which is at least 1. Let $\widetilde{\phi}_{\alpha_1} \in \Delta(\Theta)$ be the distribution of state conditional on the union of the following set of events: (1) player 1 is commitment type α_1; (2) player 1 is strategic type $\widetilde{\theta} \in \Theta_{(a_1^*,\theta)} \backslash \{\theta\}$ and is playing α_1 in every period.

– **Other Bad Strategic Types:** Conditional on not playing α_1 in every period, for every $\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}$, type $\tilde{\theta}$ plays a_1^* in every period with probability $p \in [0,1)$. The probability p is chosen such that there exists $a_2' \neq a_2^*$ with

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

  for every $\widetilde{p} \in [p, 1]$.

  This p exists according to (B.2). According to the construction of λ', Lemma B.1 also implies that

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

  For every $\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}$, type $\tilde{\theta}$ plays a_1^* at every history with probability p. For every $\alpha_1 \in \widetilde{\Omega}$, type $\tilde{\theta}$ plays α_1 at *every history of the normal phase* with probability $\frac{1-p}{l}$ and plays some other actions at histories of the abnormal phase, which will be specified later on.

  Call the bad strategic type(s) who play $\alpha_1 \in \widetilde{\Omega} \bigcup \{a_1^*\}$ in every period of the normal phase *type* $\theta(\alpha_1)$. Let $\mu_t(\theta(\alpha_1))$ be the total probability of such type in period t.

  My construction exhibits the following two properties. First, suppose $\mu_t(\alpha_1) = 0$, then $\mu_t(\theta(\alpha_1)) = 0$ throughout the normal phase. Second, suppose $\mu_t(\alpha_1) \neq 0$, then $\mu_t(\theta(\alpha_1))/\mu_t(\alpha_1) = \mu_0(\theta(\alpha_1))/\mu_0(\alpha_1)$ throughout the normal phase. That is to say during the normal phase, the likelihood ratio between the commitment type and the bad strategic type imitating him remains constant.

Next, I describe type θ's normal phase strategy:

1. **Preparation Sub-Phase:** This phase lasts from period 0 to $n-1$. Type θ plays a_1^i in period i for every $i \in \{0, 1, ..., n-1\}$. This is to separate from all the pure commitment types.

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

I claim that type θ's expected payoff is close to $1-q$ if he plays type θ's equilibrium strategy when δ 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 $\varrho > 1$ such that:

$$\mu_{t+1}(\theta)\Big/\Big(\mu_{t+1}(\alpha_1)+\mu_{t+1}(\theta(\alpha_1))\Big) \geq \varrho\mu_t(\theta)\Big/\Big(\mu_t(\alpha_1)+\mu_t(\theta(\alpha_1))\Big). \tag{B.6}$$

for every $\alpha_1 \in \widetilde{\Omega}$ such that $\mu_t(\alpha_1) \neq 0$.

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

$$\Big[(1-\delta^T)0+\delta^T(1-q),(1-\delta^T)+\delta^T(1-q)\Big].$$

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

Strategies in the Abnormal Phase: In the abnormal phase, player 2 has ruled out the possibility that player 1 is type θ. For every $\alpha_1 \in \widetilde{\Omega}$, type $\theta(\alpha_1)$ plays:

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

at every history of the abnormal phase where:

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

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

$$\sum_{\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}} \lambda'(\tilde{\theta})u_2(\tilde{\theta},\alpha_1',a_2') > \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*,\theta)}} \lambda'(\tilde{\theta})u_2(\tilde{\theta},\alpha_1',a_2^*).$$

Such η exists because of inequality (B.5).

Next, I verify that type θ 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. Let $\beta(\alpha_1) \equiv \mu_t(\theta(\alpha_1))/\mu_t(\alpha_1)$. Since $\max_{\alpha_1 \in \widetilde{\Omega}} \alpha_1(a_1^*) < 1-\eta$, whenever a_1^* is observed in period t, then

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

for every $\alpha_1 \in \widetilde{\Omega}$. Let $\gamma \equiv 1-\min_{\alpha_1 \in \widetilde{\Omega}} \alpha_1(a_1^*)$. 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 $\overline{k} \equiv \left\lceil \log \frac{2\gamma}{\eta} \Big/ \log \frac{1-\eta/2}{1-\eta} \right\rceil$. For every $\alpha_1 \in \widetilde{\Omega}$, let $\overline{\beta}(\alpha_1)$ be the smallest $\beta \in \mathbb{R}_+$ such that:

$$u_2(\tilde{\phi}_{\alpha_1}, \alpha_1, a_2') + \beta \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*, \theta)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, \hat{\alpha}_1(\alpha_1), a_2') \geq u_2(\tilde{\phi}_{\alpha_1}, \alpha_1, a_2^*) + \beta \sum_{\tilde{\theta} \in \Theta^b_{(a_1^*, \theta)}} \lambda'(\tilde{\theta}) u_2(\tilde{\theta}, \hat{\alpha}_1(\alpha_1), a_2^*)$$

The choice of η and (B.5) ensure the existence of such $\overline{\beta}(\alpha_1)$. Let $\overline{\beta} \equiv 2\max_{\alpha_1 \in \widetilde{\Omega}} \overline{\beta}(\alpha_1)$ and $\underline{\beta} \equiv \min_{\alpha_1 \in \widetilde{\Omega}} \frac{\mu(\theta(\alpha_1))}{\mu(\alpha_1)}$. Let $T_1 \equiv \left\lceil \log \frac{\overline{\beta}}{\underline{\beta}} \Big/ \log \frac{1-\eta/2}{1-\eta} \right\rceil$. In the beginning of the abnormal phase (regardless of when it is triggered), $\beta_t(\alpha_1) \geq \underline{\beta}$ for all $\alpha_1 \in \widetilde{\Omega}$. 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 $\overline{k}$ consecutive periods before a_2^* is being strictly dominated by a_2' again. Therefore, type θ's payoff in the abnormal phase is at most:

$$(1-\delta^{T_1}) + \delta^{T_1}\Big\{(1-\delta^{\overline{k}-1}) + \delta^{\overline{k}}(1-\delta^{\overline{k}-1}) + \delta^{2\overline{k}}(1-\delta^{\overline{k}-1}) + ...\Big\}$$

The term in the curly bracket converges to $\frac{\overline{k}}{1+\overline{k}}$ as $\delta \to 1$. Let

$$q \equiv \frac{\overline{k}}{2(\overline{k}+1)+1}.$$

By construction, type θ's payoff in beginning of the abnormal phase cannot exceed $1-2q$.

Remark: My construction of 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 their analogical-based equilibrium, the strategic long-run player alternates between his actions in order to exploit the flaws in the short-run players' reasoning process and to manipulate their beliefs.

This leads to similar behavioral dynamics compared to the abnormal phase of my construction. This is because after reaching the abnormal phase, player 2's belief only attaches positive probability to types that are playing *stationary strategies* in the continuation game, i.e. types that are playing α_1 in every period and types that are playing $\widehat{\alpha}_1(\alpha_1)$ in every period. Let the long-run player's *reputation* be the likelihood ratio between the commitment type α_1 and the bad strategic types that are playing $\sigma(\alpha_1)$. At every history of the abnormal phase, type θ will be facing a trade-off between reaping high stage-game payoff (by playing a_1^*) and building-up his reputation (by playing actions other than a_1^*). 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 low, player 2 has a strict incentive to play a_2', which punishes player 1 for at least one period, making his payoff bounded away from 1 even in the $\delta \to 1$ limit.

C Proof of Theorem 2

I prove Theorem 2 for all games satisfying Assumptions 1-3. Compared to the proof sketch in the main text, the key difficulty arises from the fact that player 2s can observe their predecessors' actions and the past realizations of public randomization devices, i.e. $a_{2,t}$ can depend on $h^t \equiv \{a_{1,s}, a_{2,s}, \xi_s\}_{s \leq t-1}$. I show that when $\delta \to 1$,

1. Every strategic type in $\Theta_g \cup \Theta_p$ can secure himself payoff approximately $u_1(\theta, \overline{a}_1, \overline{a}_2)$ by playing $\overline{a}_1$ in every period when the prior is optimistic

2. For every $\theta \succsim \theta_p^*(\mu)$, type θ can secure himself payoff approximately $v_\theta^*(\mu)$ by playing $\overline{a}_1$ in every period when the prior is pessimistic.

To avoid cumbersome notation, I write v_θ^* instead of $v_\theta^*(\mu)$. Furthermore, I focus on the case where all commitment actions are pure. This is without loss of generality as the probability of other commitment types (pure and mixed) becomes arbitrarily small relative to commitment type $\overline{a}_1$ after player 2 observes $\overline{a}_1$ for a finite number of periods. Those finite number of periods have negligible payoff consequences as $\delta \to 1$.

The proof consists of seven parts. In subsections C.1-C.3, I introduce some useful concepts and define some constants that will be referred to later on in the proof. In subsection C.4, I present four useful observations, stated as Lemma C.1-C.4. Subsections C.5 and C.6 establish the first statement of Theorem 2, starting from the case where Θ_n is empty and then generalizing it to the case where Θ_n is non-empty. I establish the second statement of the theorem in subsection C.7 using conclusions from the previous subsections.

C.1 Several Useful Constants

I start from defining several useful constants which only depend on μ, u_1 and u_2, but are independent of σ and δ. Let $M \equiv \max_{\theta,a_1,a_2} |u_1(\theta,a_1,a_2)|$ and

$$K \equiv \max_{\theta\in\Theta}\Big\{u_1(\theta,\overline{a}_1,\overline{a}_2)-u_1(\theta,\overline{a}_1,\underline{a}_2)\Big\}\Big/\min_{\theta\in\Theta}\Big\{u_1(\theta,\overline{a}_1,\overline{a}_2)-u_1(\theta,\overline{a}_1,\underline{a}_2)\Big\}.$$

Since $\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)>0$, expression (4.2) implies the existence of $\kappa\in(0,1)$ such that:

$$\kappa\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)+\sum_{\theta\in\Theta}\mu(\theta)\mathcal{D}(\theta,\overline{a}_1)>0.$$

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

$$\rho_0(\kappa)\equiv\frac{(1-\kappa)\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)}{2\max_{(\theta,a_1)\in\Theta\times A_1}|\mathcal{D}(\theta,a_1)|}>0 \tag{C.1}$$

and

$$\overline{T}_0(\kappa)\equiv\lceil 1/\rho_0(\kappa)\rceil. \tag{C.2}$$

Let

$$\rho_1(\kappa)\equiv\frac{\kappa\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1},\overline{a}_1)}{\max_{(\theta,a_1)}|\mathcal{D}(\theta,a_1)|}. \tag{C.3}$$

and

$$\overline{T}_1(\kappa)\equiv\lceil 1/\rho_1(\kappa)\rceil. \tag{C.4}$$

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

$$(1-\delta^{\overline{T}_0(0)})u_1(\theta_p,\overline{a}_1,\underline{a}_2)+\delta^{\overline{T}_0(0)}u_1(\theta_p,\overline{a}_1,\overline{a}_2)>\frac{1}{2}\Big(u_1(\theta_p,\overline{a}_1,\overline{a}_2)+u_1(\theta_p,\underline{a}_1,\underline{a}_2)\Big). \tag{C.5}$$

C.2 Random History & Random Path

Let $\overline{\Omega}\equiv\Omega\bigcup\Theta$ be the entire set of types with ω a typical element of $\overline{\Omega}$. 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 $a_*^t\equiv(\overline{a}_1,...,\overline{a}_1)$. I call h^t a *public history*, r^t a *random history* and r^∞ a *random path*. Let $\mathcal{H}$ and $\mathcal{R}$ be the set of public histories and random histories, respectively, with $\succ$, $\succsim$, $\prec$ and $\precsim$ naturally defined. Recall that a strategy profile σ consists of $(\sigma_\omega)_{\omega\in\overline{\Omega}}$ with $\sigma_\omega:\mathcal{H}\to\Delta(A_1)$ and $\sigma_2:\mathcal{H}\to\Delta(A_2)$. Let $\mathcal{P}^\sigma(\omega)$ be the probability measure over public histories induced by (σ_ω,σ_2). Let $\mathcal{P}^\sigma\equiv\sum_{\omega\in\overline{\Omega}}\mu(\omega)\mathcal{P}^\sigma(\omega)$. Let $V^\sigma(h^t)\equiv(V_\theta^\sigma(h^t))_{\theta\in\Theta}\in\mathbb{R}^{|\Theta|}$ be the continuation payoff vector for strategic types at h^t.

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

$$\mathcal{R}_*^\sigma\equiv\Big\{r^\infty\Big|(a_*^t,r^t)\in\mathcal{H}^\sigma\text{ for all }t\text{ and }r^t\prec r^\infty\Big\}$$

be the set of random paths consistent with player 1 playing $\overline{a}_1$ in every period. For every $h^t = (a^t, r^t)$, let $\overline{\sigma}_1[h^t] : \mathcal{H} \to A_1$ be a continuation strategy at h^t satisfying $\overline{\sigma}_1[h^t](h^s) = \overline{a}_1$ for all $h^s \succsim h^t$ with $h^s = (a^t, \overline{a}_1, ..., \overline{a}_1, r^s) \in \mathcal{H}^\sigma$. Let $\underline{\sigma}_1[h^t] : \mathcal{H} \to A_1$ be a continuation strategy that satisfies $\underline{\sigma}_1[h^t](h^s) = \underline{a}_1$ for all $h^s \succsim h^t$ with $h^s = (a^t, \underline{a}_1, ..., \underline{a}_1, r^s) \in \mathcal{H}^\sigma$. For every $\theta \in \Theta$, let

$$\overline{\mathcal{R}}^\sigma(\theta) \equiv \Big\{ r^t \Big| \overline{\sigma}_1[a_*^t, r^t] \text{ is type } \theta\text{'s best reply to } \sigma_2 \Big\} \text{ and } \underline{\mathcal{R}}^\sigma(\theta) \equiv \Big\{ r^t \Big| \underline{\sigma}_1[a_*^t, r^t] \text{ is type } \theta\text{'s best reply to } \sigma_2 \Big\}.$$

C.3 Beliefs & Best Response Sets

Let $\mu(a^t, r^t) \in \Delta(\overline{\Omega})$ be player 2's posterior belief at (a^t, r^t) and specifically, let $\mu^*(r^t) \equiv \mu(a_*^t, r^t)$. Let

$$\mathcal{B}_\kappa \equiv \Big\{ \tilde{\mu} \in \Delta(\overline{\Omega}) \Big| \kappa \tilde{\mu}(\overline{a}_1) \mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta} \tilde{\mu}(\theta) \mathcal{D}(\theta, \overline{a}_1) \geq 0 \Big\}. \tag{C.6}$$

By definition, $\mathcal{B}_{\kappa'} \subsetneq \mathcal{B}_\kappa$ for every $\kappa, \kappa' \in [0, 1]$ with $\kappa' < \kappa$.

For every $r^t \in \mathcal{R}^t$ and $\omega \in \overline{\Omega}$, let $q^*(r^t)(\omega)$ be the (ex ante) probability that (1) player 1 is type ω; (2) player 1 has played $\overline{a}_1$ from period 0 to $t-1$, conditional on the realization of random history being r^t. Let $q^*(r^t) \in \mathbb{R}_+^{|\overline{\Omega}|}$ be the corresponding vector of probabilities. For every δ and $\sigma \in \text{NE}(\delta, \mu)$,

1. For every a^t and $r^t, \hat{r}^t \succ r^{t-1}$ satisfying $(a^t, r^t), (a^t, \hat{r}^t) \in \mathcal{H}^\sigma$, we have $\mu(a^t, r^t) = \mu(a^t, \hat{r}^t)$.
2. For every $r^t, \hat{r}^t \succ r^{t-1}$ with $(a_*^t, r^t), (a_*^t, \hat{r}^t) \in \mathcal{H}^\sigma$, we have $q^*(r^t) = q^*(\hat{r}^t)$.

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 plays $\overline{a}_1$ in every period, we have $q^*(r^t)(\overline{a}_1) = \mu_0(\overline{a}_1)$.

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

$$\mathcal{R}_g^\sigma \equiv \Big\{ r^t \Big| (a_*^t, r^t) \in \mathcal{H}^\sigma \text{ and } \mu^*(r^t)\big(\Theta_p \cup \Theta_n\big) = 0 \Big\}, \tag{C.7}$$

and let

$$\widehat{\mathcal{R}}_g^\sigma \equiv \Big\{ r^t \Big| \exists r^T \succsim r^t \text{ such that } r^T \in \mathcal{R}_g^\sigma \Big\}. \tag{C.8}$$

Intuitively, $\widehat{\mathcal{R}}_g^\sigma$ is the set of on-path random histories under which all the strategic types in $\Theta_p \cup \Theta_n$ will be separated from commitment type $\overline{a}_1$ at some random histories in the future.

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 θ's strategy. The first one shows the implications of MSM on player 1's equilibrium strategy:

Lemma C.1. *Suppose $\sigma \in NE(\delta, \mu)$, $\theta \succ \tilde{\theta}$ and $h_*^t = (a_*^t, r^t) \in \mathcal{H}^\sigma(\theta) \cap \mathcal{H}^\sigma(\tilde{\theta})$,*

1. *If $r^t \in \overline{\mathcal{R}}^\sigma(\tilde{\theta})$, then $\sigma_\theta(a_*^s, r^s)(\overline{a}_1) = 1$ for every $(a_*^s, r^s) \in \mathcal{H}^{(\overline{\sigma}_1(h_*^t), \sigma_2)}(\theta)$ with $r^s \succsim r^t$.*
2. *If $r^t \in \underline{\mathcal{R}}^\sigma(\theta)$, then $\sigma_{\tilde{\theta}}(a^s, r^s)(\underline{a}_1) = 1$ for every $(a^s, r^s) \in \mathcal{H}^{(\underline{\sigma}_1(h_*^t), \sigma_2)}(\tilde{\theta})$ with $(a^s, r^s) \succsim (a_*^t, r^t)$.*

PROOF OF LEMMA C.1: I only need to show the first part, as the second part is symmetric after switching signs. Without loss of generality, I focus on history h^0. For notation simplicity, let $\overline{\sigma}_1[h^0] = \overline{\sigma}_1$. For every σ_ω and σ_2, let $P^{(\sigma_\omega, \sigma_2)} : A_1 \times A_2 \to [0, 1]$ be defined as:

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

where $p_t^{(\sigma_\omega,\sigma_2)}(a_1,a_2)$ is the probability of (a_1,a_2) occurring in period t under (σ_ω,σ_2). Let $P_i^{(\sigma_1,\sigma_2)} \in \Delta(A_2)$ be $P^{(\sigma_1,\sigma_2)}$'s marginal distribution on A_i, for $i \in \{1,2\}$.

Suppose towards a contradiction that $\overline{\sigma}_1$ is type $\tilde{\theta}$'s best reply and there exists σ_θ with $P_1^{(\sigma_\theta,\sigma_2)}(\overline{a}_1) < 1$ such that σ_θ is type θ's best reply, then type $\tilde{\theta}$ and θ's incentive constraints require that:

$$\sum_{a_2\in A_2}\Big(P_2^{(\overline{\sigma}_1,\sigma_2)}(a_2)-P_2^{(\sigma_\theta,\sigma_2)}(a_2)\Big)u_1(\tilde{\theta},\overline{a}_1,a_2) \geq \sum_{a_2\in A_2, a_1\neq \overline{a}_1} P^{(\sigma_\theta,\sigma_2)}(a_1,a_2)\Big(u_1(\tilde{\theta},a_1,a_2)-u_1(\tilde{\theta},\overline{a}_1,a_2)\Big),$$

and

$$\sum_{a_2\in A_2}\Big(P_2^{(\overline{\sigma}_1,\sigma_2)}(a_2)-P_2^{(\sigma_\theta,\sigma_2)}(a_2)\Big)u_1(\theta,\overline{a}_1,a_2) \leq \sum_{a_2\in A_2, a_1\neq \overline{a}_1} P^{(\sigma_\theta,\sigma_2)}(a_1,a_2)\Big(u_1(\theta,a_1,a_2)-u_1(\theta,\overline{a}_1,a_2)\Big).$$

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

$$\sum_{a_2\in A_2, a_1\neq \overline{a}_1} P^{(\sigma_\theta,\sigma_2)}(a_1,a_2)\Big(u_1(\tilde{\theta},a_1,a_2)-u_1(\tilde{\theta},\overline{a}_1,a_2)\Big)$$

$$> \sum_{a_2\in A_2, a_1\neq \overline{a}_1} P^{(\sigma_\theta,\sigma_2)}(a_1,a_2)\Big(u_1(\theta,a_1,a_2)-u_1(\theta,\overline{a}_1,a_2)\Big)$$

which implies that:

$$\sum_{a_2\in A_2}\Big(P_2^{(\sigma_\theta,\sigma_2)}(a_2)-P_2^{(\overline{\sigma}_1,\sigma_2)}(a_2)\Big)\Big(u_1(\theta,\overline{a}_1,a_2)-u_1(\tilde{\theta},\overline{a}_1,a_2)\Big) > 0. \tag{C.9}$$

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

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

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

$$\sum_{a_2\in A_2}\Big(P_2^{(\overline{\sigma}_1,\sigma_2)}(a_2)-P_2^{(\sigma_\theta,\sigma_2)}(a_2)\Big)u_1(\tilde{\theta},\overline{a}_1,a_2) > 0. \tag{C.10}$$

Since both $P_2^{(\sigma_\theta,\sigma_2)}$ and $P_2^{(\overline{\sigma}_1,\sigma_2)}$ are probability distributions, we have

$$\sum_{a_2\in A_2}\Big(P_2^{(\sigma_\theta,\sigma_2)}(a_2)-P_2^{(\overline{\sigma}_1,\sigma_2)}(a_2)\Big) = 0.$$

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

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

Lemma C.2. *If $\mu^*(r^t) \in \mathcal{B}_\kappa$ and $\overline{a}_2$ is not a strict best reply at (a_*^t, r^t), then for every $r^{t+1} \succ r^t$ with $(a_*^{t+1}, r^{t+1}) \in \mathcal{H}^\sigma$, we have:*

$$\sum_{\theta\in\Theta}\Big(q^*(r^t)(\theta)-q^*(r^{t+1})(\theta)\Big) \geq \rho_0(\kappa). \tag{C.11}$$

PROOF OF LEMMA C.2: If $\mu^*(r^t) \in \mathcal{B}_\kappa$, then:[24]

$$\kappa\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta\in\Theta} q^*(r^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0.$$

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

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta\in\Theta} q^*(r^{t+1})(\theta)\mathcal{D}(\theta, \overline{a}_1) + \sum_{\theta\in\Theta}\Big(q^*(r^t)(\theta) - q^*(r^{t+1})(\theta)\Big)\mathcal{D}(\theta, \underline{a}_1) \leq 0,$$

for every $r^{t+1} \succ r^t$ with $(a_*^{t+1}, r^{t+1}) \in \mathcal{H}^\sigma$, or equivalently

$$\underbrace{\kappa\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta\in\Theta} q^*(r^t)(\theta)\mathcal{D}(\theta, \overline{a}_1)}_{\geq 0} + \underbrace{(1-\kappa)\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1)}_{>0}$$

$$+\sum_{\theta\in\Theta}\Big(q^*(r^{t+1})(\theta) - q^*(r^t)(\theta)\Big)\mathcal{D}(\theta, \overline{a}_1) + \sum_{\theta\in\Theta}\Big(q^*(r^t)(\theta) - q^*(r^{t+1})(\theta)\Big)\mathcal{D}(\theta, \underline{a}_1) \leq 0,$$

According to (C.1), we have:

$$\sum_{\theta\in\Theta}\Big(q^*(r^t)(\theta) - q^*(r^{t+1})(\theta)\Big) \geq \frac{(1-\kappa)\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1)}{2\max_{(\theta,a_1)\in\Theta\times A_1}|\mathcal{D}(\theta, a_1)|} = \rho_0(\kappa).$$

□

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

Lemma C.3. *For every $\sigma \in NE(\delta, \mu)$ and $h^t \in \mathcal{H}^\sigma$ with*

$$\mu(h^t)(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta} \mu(h^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0. \tag{C.12}$$

Let $\underline{\theta} \equiv \min\Big\{supp\big(\mu(h^t)\big)\Big\}$, then:

$$V_{\underline{\theta}}(h^t) = u_1(\underline{\theta}, \underline{a}_1, \underline{a}_2).$$

PROOF OF LEMMA C.3: Let

$$\Theta^* \equiv \Big\{\tilde{\theta} \in \Theta_p \cup \Theta_n \Big| \mu(h^t)(\tilde{\theta}) > 0\Big\}.$$

Since $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$, (C.12) implies that $\Theta^* \neq \{\varnothing\}$. The rest of the proof is done via induction on $|\Theta^*|$. When $|\Theta^*| = 1$, there exists a pure strategy $\sigma_{\underline{\theta}}^* : \mathcal{H} \to A_1$ in the support of $\sigma_{\underline{\theta}}$ such that (C.12) holds for all h^s satisfying $h^s \in \mathcal{H}^{(\sigma_{\underline{\theta}}^*, \sigma_2)}$ and $h^s \succsim h^t$. At every such h^s, $\underline{a}_2$ is player 2's strict best reply. When playing $\sigma_{\underline{\theta}}^*$, type $\underline{\theta}$'s stage game payoff is no more than $u_1(\underline{\theta}, \underline{a}_1, \underline{a}_2)$ in every period.

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

1. $\mu(h^\tau) \notin \mathcal{B}_\kappa$ for all $h^s \succsim h^\tau \succsim h^t$.
2. $V_{\underline{\theta}}(h^s) > u_1(\underline{\theta}, \underline{a}_1, \underline{a}_2)$.

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

3. For all a_1 such that $\mu(h^s, a_1) \notin \mathcal{B}_\kappa$, $\sigma_{\underline{\theta}}(h^s)(a_1) = 0$.[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 $\mu(h^s, a_1)(\underline{\theta}) = 0$, there exists $\tilde{\theta} \in \Theta^* \backslash \{\underline{\theta}\}$ such that $(h^s, a_1) \in \mathcal{H}^\sigma(\tilde{\theta})$. Our induction hypothesis suggests that:

$$V_{\tilde{\theta}}(h^s) = u_1(\tilde{\theta}, \underline{a}_1, \underline{a}_2).$$

The incentive constraints of type $\underline{\theta}$ and type $\tilde{\theta}$ 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:

$$\mathbb{E}\Big[\sum_{\tau=0}^{\infty}(1-\delta)\delta^\tau\Big(u_1(\underline{\theta}, \alpha_{1,\tau}, \alpha_{2,\tau}) - u_1(\underline{\theta}, \underline{a}_1, \underline{a}_2)\Big)\Big] > 0 \geq \mathbb{E}\Big[\sum_{\tau=0}^{\infty}(1-\delta)\delta^\tau\Big(u_1(\tilde{\theta}, \alpha_{1,\tau}, \alpha_{2,\tau}) - u_1(\tilde{\theta}, \underline{a}_1, \underline{a}_2)\Big)\Big],$$

where $\mathbb{E}[\cdot]$ is taken over probability measure $\mathcal{P}^\sigma$. However, the supermodularity condition implies that,

$$u_1(\underline{\theta}, \alpha_{1,\tau}, \alpha_{2,\tau}) - u_1(\underline{\theta}, \underline{a}_1, \underline{a}_2) \leq u_1(\tilde{\theta}, \alpha_{1,\tau}, \alpha_{2,\tau}) - u_1(\tilde{\theta}, \underline{a}_1, \underline{a}_2),$$

leading to a contradiction. □

The next Lemma outlines an important implication of $r^t \notin \widehat{\mathcal{R}}_g^\sigma$.

Lemma C.4. *If $r^t \notin \widehat{\mathcal{R}}_g^\sigma$ and $(a_*^t, r^t) \in \mathcal{H}^\sigma$, then there exists $\theta \in \big(\Theta_p \cup \Theta_n\big) \bigcap \text{supp}\big(\mu^*(r^t)\big)$ such that $r^t \in \overline{R}^\sigma(\theta)$.*

PROOF OF LEMMA C.4: Suppose towards a contradiction that $r^t \notin \widehat{\mathcal{R}}_g^\sigma$ but no such θ exists. Let

$$\theta_1 \equiv \max\Big\{\big(\Theta_p \cup \Theta_n\big) \bigcap \text{supp}\big(\mu^*(r^t)\big)\Big\}.$$

The set on the RHS is non-empty according to the definition of $\widehat{\mathcal{R}}_g^\sigma$ and $\mathcal{R}_g^\sigma$

Let $(a_*^{t_1}, r^{t_1}) \succsim (a_*^t, r^t)$ be the history at which type θ_1 has a strict incentive not to play $\overline{a}_1$ with $(a_*^{t_1}, r^{t_1}) \in \mathcal{H}^\sigma$. For any $(a_*^{t_1+1}, r^{t_1+1}) \succ (a_*^{t_1}, r^{t_1})$ with $(a_*^{t_1+1}, r^{t_1+1}) \in \mathcal{H}^\sigma$, on one hand, we have $\mu^*(r^{t_1+1})(\theta_1) = 0$. On the other hand, the fact that $r^t \notin \widehat{\mathcal{R}}_g^\sigma$ implies that $\mu^*(r^{t_1+1})(\Theta_n \cup \Theta_p) > 0$.

Let

$$\theta_2 \equiv \max\Big\{\big(\Theta_p \cup \Theta_n\big) \bigcap \text{supp}\big(\mu^*(r^{t_1+1})\big)\Big\},$$

and let us examine type θ_1 and θ_2's incentive constraints at $(a_*^{t_1}, r^{t_1})$. According to Lemma C.1, there exists $r^{t_2} \succ r^{t_1}$ such that type θ_2 has a strict incentive not to play $\overline{a}_1$ at $(a_*^{t_2}, r^{t_2}) \in \mathcal{H}^\sigma$.

Therefore, we can iterate this process and obtain $r^{t_3} \succ r^{t_4}$... Since

$$\Big|\text{supp}\Big(\mu^*(r^{t_{k+1}})\Big)\Big| \leq \Big|\text{supp}\Big(\mu^*(r^{t_k})\Big)\Big| - 1,$$

for any $k \in \mathbb{N}$, there exists $m \leq |\Theta_p \cup \Theta_n|$ such that $(a_*^{t_m}, r^{t_m}) \in \mathcal{H}^\sigma$, $r^{t_m} \succsim r^t$ and $\mu^*(r^{t_m})(\Theta_n \cup \Theta_p) = 0$, which contradicts $r^t \notin \widehat{\mathcal{R}}_g^\sigma$. □

C.5 Proof of Statement 1 Theorem 2: $\Theta_n = \{\varnothing\}$

This subsection examines the case in which $\Theta_n = \{\varnothing\}$. I will incorporate states in Θ_n in subsection C.6. The main result in this part is the following Proposition:

Proposition C.1. *If $\Theta_n = \{\varnothing\}$ and $\mu \in \mathcal{B}_\kappa$, then for every $\theta \in \Theta$, we have:*

$$V_\theta(a_*^0, r^0) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - 2M(1 - \delta^{\overline{T}_0(\kappa)}).$$

[25] I omit $(a_{2,s}, \xi_s)$ in the expression for histories since they play no role in the posterior belief on $\overline{\Omega}$ at every on-path history.

Despite Proposition C.1 is stated in terms of player 1's guaranteed payoff at h^0, the conclusion applies to all r^t and $\theta \in \Theta_g \cup \Theta_p$ as long as $\mu^*(r^t) \in \mathcal{B}_\kappa$ and $(a_*^t, r^t) \in \mathcal{H}^\sigma(\theta) \backslash \bigcup_{\theta_n \in \Theta_n} \mathcal{H}^\sigma(\theta_n)$. I show Lemma C.5 and Lemma C.6, which together imply Proposition C.1.

Lemma C.5. *For every $\sigma \in NE(\delta, \mu)$, if $\mu^*(r^t) \in \mathcal{B}_\kappa$ for all $r^t \in \widehat{\mathcal{R}}_g^\sigma$, then for every $r^\infty \in \mathcal{R}_*^\sigma$,*

$$\left|\left\{t \in \mathbb{N} \middle| r^\infty \succ r^t \text{ and } \overline{a}_2 \text{ is not a strict best reply at } (a_*^t, r^t)\right\}\right| \leq \overline{T}_0(\kappa). \tag{C.13}$$

PROOF OF LEMMA C.5: Pick any $r^\infty \in \mathcal{R}_*^\sigma$, if $r^0 \notin \widehat{\mathcal{R}}_g^\sigma$, then let $t^* = -1$. Otherwise, let

$$t^* \equiv \max\left\{t \in \mathbb{N} \cup \{+\infty\} \middle| r^t \in \widehat{\mathcal{R}}_g^\sigma \text{ and } r^\infty \succ r^t\right\}.$$

According to Lemma C.2, for every $t \leq t^*$, if $\overline{a}_2$ is not a strict best reply at (a_*^t, r^t), then we have inequality (C.11).

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

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

$$\theta \in \left(\Theta_p \cup \Theta_n\right) \bigcap \text{supp}\left(\mu^*(r^{t^*+1})\right)$$

such that $r^{t^*+1} \in \overline{R}^\sigma(\theta)$. Since $\theta_g \succ \theta$ for all $\theta_g \in \Theta_g$, Lemma C.1 implies that for every θ_g and $r^\infty \succ r^t \succsim r^{t^*+1}$, we have $\sigma_{\theta_g}(a_*^t, r^t) = 1$, and therefore, $q^*(r^t)(\theta_g) = q^*(r^{t+1})(\theta_g)$. This implies that $\mu^*(r^t) \in \mathcal{B}_\kappa$ for every $r^\infty \succ r^t \succsim r^{t^*+1}$. If $\overline{a}_2$ is not a strict best reply at (a_*^t, r^t) for any $t > t^*$, inequality (C.11) again applies.

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

$$\sum_{\theta \in \Theta}\left(q^*(r^t)(\theta) - q^*(r^{t+1})(\theta)\right) \geq \rho_0(\kappa),$$

from which we obtain (C.13). $\square$

The next result shows that the condition required in Lemma C.5 holds in every equilibrium when δ is large enough. Moreover, it applies regardless of the short-run players' prior belief, which will be useful in the proof of the second statement in subsection C.7.

Lemma C.6. *For every $\sigma \in NE(\delta, \mu)$ with $\delta > \overline{\delta}$, $\mu^*(r^t) \in \mathcal{B}_0$ for every $r^t \in \widehat{\mathcal{R}}_g^\sigma$ with $\mu^*(r^t)(\Theta_n) = 0$.*

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

$$(1 - \delta^{\overline{T}_0(\kappa^*)})u_1(\theta_p, \overline{a}_1, \underline{a}_2) + \delta^{\overline{T}_0(\kappa^*)}u_1(\theta_p, \overline{a}_1, \overline{a}_2) > \frac{1}{2}\Big(u_1(\theta_p, \overline{a}_1, \overline{a}_2) + u_1(\theta_p, \underline{a}_1, \underline{a}_2)\Big). \tag{C.14}$$

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

▷ $r^{T_1} \succ r^{t_1}$, $r^{T_1} \in \mathcal{R}_g^\sigma$ 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 $r^{T_1} \succ r^t \succsim r^{t_1}$. Then there exists $a_1 \neq \overline{a}_1$ and $r^{t_1^*+1} \succ r^{t_1^*}$ such that $\mu\big((a_*^{t_1^*}, a_1), r^{t_1^*+1}\big) \notin \mathcal{B}_0$ and $\big((a_*^{t_1^*}, a_1), r^{t_1^*+1}\big) \in \mathcal{H}^\sigma$. This also implies the existence of $\theta_p \in \Theta_p \bigcap \text{supp}\Big(\mu\big((a_*^{t_1^*}, a_1), r^{t_1^*+1}\big)\Big)$.

According to Lemma C.3, type θ_p's continuation payoff at $(a_*^{t_1^*}, r^{t_1^*})$ by playing a_1 is at most

$$(1 - \delta)u_1(\theta_p, \underline{a}_1, \overline{a}_2) + \delta u_1(\theta_p, \underline{a}_1, \underline{a}_2). \tag{C.15}$$

His incentive constraint at $(a_*^{t_1^*}, r^{t_1^*})$ requires that his expected payoff from $\overline{\sigma}_1$ is weakly lower than (C.15), i.e. there exists $r^{t_1^*+1} \succ r^{t_1^*}$ satisfying $(a_*^{t_1^*+1}, r^{t_1^*+1}) \in \mathcal{H}^\sigma$ and type θ_p's continuation payoff at $(a_*^{t_1^*+1}, r^{t_1^*+1})$ is no more than:

$$\frac{1}{2}\Big(u_1(\theta_p, \overline{a}_1, \overline{a}_2) + u_1(\theta_p, \underline{a}_1, \underline{a}_2)\Big). \tag{C.16}$$

If $\mu^*(r^t) \in \mathcal{B}_{\kappa^*}$ for every $r^t \in \widehat{\mathcal{R}}_g^\sigma \cap \{r^t \succsim r^{t_1^*}\}$, then according to Lemma C.5, his continuation payoff at $(a_*^{t_1^*}, r^{t_1^*})$ by playing $\overline{\sigma}_1$ is at least:

$$(1-\delta^{\overline{T}_0(\kappa^*)})u_1(\theta_p, \overline{a}_1, \underline{a}_2) + \delta^{\overline{T}_0(\kappa^*)}u_1(\theta_p, \overline{a}_1, \overline{a}_2),$$

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

Suppose on the other hand, there exists $r^{t_2} \succ r^{t_1^*}$ such that:

▷ $r^{t_2} \in \hat{\mathcal{R}}_g^\sigma$ while $\mu^*(r^{t_2}) \notin \mathcal{B}_{\kappa^*}$.

There exists $r^{T_2} \succ r^{t_2}$ such that $r^{T_2} \in \mathcal{R}_g^\sigma$ 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 $\mu^*(r^t) \notin \mathcal{B}_0$ for $r^{T_2} \succ r^t \succsim r^{t_2}$. Then there exists $a_1 \neq \overline{a}_1$ and $r^{t_2^*+1} \succ r^{t_2^*}$ such that $\mu\big((a_*^{t_2^*}, a_1), r^{t_2^*+1}\big) \notin \mathcal{B}_0$ and $\big((a_*^{t_2^*}, a_1), r^{t_2^*+1}\big) \in \mathcal{H}^\sigma$.

Iterating the above process and repeatedly apply the aforementioned argument, we know that for every $k \geq 1$, in order to satisfy player 1's incentive constraint to play $a_1 \neq \overline{a}_1$ at $(a_*^{t_k^*}, r^{t_k^*})$, 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 $\mu^*(r^{t_k}) \notin \mathcal{B}_{\kappa^*}$ but $\mu^*(r^{t_k^*+1}) \in \mathcal{B}_0$ as well as $r^{t_{k+1}} \succ r^{t_k^*+1}$, we have:

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

for every $k \geq 2$. (C.17) and (C.4) together suggest that this iteration process cannot last for more than $\overline{T}_1(\kappa^*)$ rounds, which is an integer independent of δ, 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 every $\delta \geq \overline{\delta}$ and $\sigma \in NE(\delta, \mu)$. If r^t satisfies $(a_*^t, r^t) \in \mathcal{H}^\sigma$, $\mu^*(r^t)(\Theta_n) = 0$, $r^t \notin \widehat{\mathcal{R}}_g^\sigma$ and*

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

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

PROOF OF LEMMA C.7: Since $\mu^*(r^t)(\Theta_n) = 0$ and $r^t \notin \widehat{\mathcal{R}}_g^\sigma$, Lemma C.4 implies the existence of $\theta_p \in \Theta_p \cap \text{supp}(\mu^*(r^t))$ such that $r^t \in \overline{R}^\sigma(\theta_p)$. According to Lemma C.1, $\sigma_\theta(a_*^s, r^s)(\overline{a}_1) = 1$ for every $(a_*^s, r^s) \in \mathcal{H}^\sigma(\theta)$ with $r^s \succsim r^t$. 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 $\overline{\kappa} \in (0,1)$ such that:[26]

$$\overline{\kappa}\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta\in\Theta} q^*(r^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) > 0.$$

According to (C.11), we have:

$$\sum_{\theta\in\Theta_p} \Big(q^*(r^s)(\theta) - q^*(r^{s+1})(\theta)\Big) \geq \rho_0(\overline{\kappa})$$

whenever $\overline{a}_2$ is not a strict best reply at $(a_*^s, r^s) \succsim (a_*^t, r^t)$. Therefore, there can be at most $\overline{T}_0(\overline{\kappa})$ such periods. Hence, there exists r^N with $(a_*^N, r^N) \in \mathcal{H}^\sigma$ such that:

[26]There are two reasons for why one cannot directly apply the conclusion in Lemma C.2. First, a stronger conclusion is required for Lemma C.7. Second, $\overline{\kappa}$ can be arbitrarily close to 1, while κ is uniformly bounded below 1 for any given μ.

1. $\overline{a}_2$ is not a strict best reply at (a_*^N, r^N).

2. $\overline{a}_2$ is a strict best reply for all $(a_*^s, r^s) \succ (a_*^N, r^N)$ with $(a_*^s, r^s) \in \mathcal{H}^\sigma$.

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

C.6 Proof of Statement 1 Theorem 2: Incorporating Types in Θ_n

Next, we extend the proof in subsection C.5 by allowing for types in Θ_n. Lemmas C.5 and C.6 imply the following result in this general environment:

Proposition C.2. *For every $\delta > \overline{\delta}$ and $\sigma \in NE(\delta, \mu)$, there exists no $\theta_p \in \Theta_p$, random histories r^{t+1} and r^t with $r^{t+1} \succ r^t$ and $a_1 \neq \overline{a}_1$ that simultaneously satisfy the following three requirements:*

1. $r^{t+1} \in \widehat{\mathcal{R}}_g^\sigma$.

2. $\big((a_*^t, a_1), r^{t+1}\big) \in \mathcal{H}^\sigma(\theta_p)$.

3. $V_{\theta_p}\Big(\big((a_*^t, a_1), \hat{r}^{t+1}\big)\Big) = u_1(\theta_p, \underline{a}_1, \underline{a}_2)$ *for all* $\hat{r}^{t+1} \succ r^t$.

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 $r^t \in \underline{\mathcal{R}}^\sigma(\theta_p)$. According to Lemma 4.1, $\theta_n \prec \theta_p$ for all $\theta_n \in \Theta_n$. The second part of Lemma C.1 then implies that $\mu^*(\hat{r}^{t+1})(\Theta_n) = 0$ for all $\hat{r}^{t+1} \succ r^t$ with $(a_*^{t+1}, \hat{r}^{t+1}) \in \mathcal{H}^\sigma$.

If $\mu^*(r^{t+1}) \in \mathcal{B}_\kappa$, then requirement 2 and Proposition C.1 result in a contradiction when examining type θ_p's incentive at (a_*^t, r^t) to play a_1 as opposed to $\overline{a}_1$. If $\mu^*(r^{t+1}) \notin \mathcal{B}_\kappa$, since $\delta > \overline{\delta}$ and $r^{t+1} \in \widehat{\mathcal{R}}_g^\sigma$, we obtain a contradiction from Lemma C.6. □

The rest of the proof consists of several steps by considering a given $\sigma \in \text{NE}(\delta, \mu)$ when δ is large enough. First,

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta} q^*(r^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0 \tag{C.19}$$

for all $t \geq 1$ and r^t satisfying $(a_*^t, r^t) \in \mathcal{H}^\sigma$. This is because otherwise, according to Lemma C.3, there exists $\theta \in \text{supp}(\mu^*(r^t))$ such that $V_\theta(a_*^t, r^t) = u_1(\theta, \underline{a}_1, \underline{a}_2)$. But then, at (a_*^{t-1}, r^{t-1}) with $r^{t-1} \prec r^t$, he could obtain strictly higher payoff by playing $\underline{a}_1$ instead of $\overline{a}_1$, leading to a contradiction.

Next comes the following Lemma:

Lemma C.8. *If μ is optimistic, then $V_\theta(a_*^t, r^t) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - 2M(K+1)(1-\delta)$ for every θ and $r^t \notin \widehat{\mathcal{R}}_g^\sigma$ satisfying the following two requirements:*

1. $(a_*^t, r^t) \in \mathcal{H}^\sigma$.

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

PROOF OF LEMMA C.8: If $\mu^*(r^t) \in \mathcal{B}_\kappa$ and $r^t \notin \widehat{\mathcal{R}}_g^\sigma$, then Lemmas C.1 and C.4 suggest that $\mu^*(r^s) \in \mathcal{B}_\kappa$ for all $r^s \succsim r^t$ and the conclusion is straightforward from Lemma C.2.

Therefore, for the rest of the proof, I consider the adverse circumstance in which $\mu^*(r^t) \notin \mathcal{B}_\kappa$. I consider two cases. First, when $\mu^*(r^t)(\Theta_n) > 0$, then according to (C.19),[27]

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta_p \cup \Theta_g} q^*(r^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) > 0.$$

[27] To see this, consider three cases. If $\Theta_p = \{\varnothing\}$, then this inequality is obvious. If $\Theta_p \neq \{\varnothing\}$, then $\mathcal{D}(\theta_n, \overline{a}_1) \leq 0$ for all $\theta_n \in \Theta_n$ according to Lemma 4.1. When $\mathcal{D}(\theta_n, \overline{a}_1) < 0$ for all θ_n, then the inequality follows from (C.19). When $\mathcal{D}(\theta_n, \overline{a}_1) = 0$ for some $\theta_n \in \Theta_n$, then $\mathcal{D}(\theta_p, \overline{a}_1) = 0$ for all $\theta_p \in \Theta_p$. The inequality then follows from $\mathcal{D}(\theta_g, \overline{a}_1) > 0$ for all $\theta_g \in \Theta_g$ as well as $\hat{\theta} \in \Theta_g$.

Since $r^t \notin \widehat{\mathcal{R}}_g^\sigma$, according to Lemma C.4, there exists $\theta \in \Theta_p \cup \Theta_n$ with $(a_*^t, r^t) \in \mathcal{H}^\sigma(\theta)$ such that $r^t \in \overline{\mathcal{R}}^\sigma(\theta)$. According to Lemma C.1, for all $\theta_g \in \Theta_g$ with $(a_*^t, r^t) \in \mathcal{H}^\sigma(\theta_g)$ and every $(a_*^s, r^s) \in \mathcal{H}^\sigma(\theta)$ with $r^s \succsim r^t$, we have $\sigma_{\theta_g}(a_*^s, r^s)(\overline{a}_1) = 1$.

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

$$V_\theta(h^s) = u_1(\theta, \underline{a}_1, \underline{a}_2). \tag{C.20}$$

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

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

Let

$$\overline{\theta}_n \equiv \max\Big\{\text{supp}(\mu^*(r^t)) \bigcap \Theta_n\Big\}.$$

The second part of Lemma C.1 and (C.20) together imply that $\mu^*(r^\tau)(\overline{\theta}_n) > 0$. Let us examine type $\overline{\theta}_n$'s incentive at (a_*^t, r^t) to play his equilibrium strategy as opposed to play $\underline{a}_1$ in every period. This requires that:

$$\mathbb{E}\Big[\sum_{s=t}^{\tau-1}(1-\delta)\delta^{s-t}u_1(\overline{\theta}_n, \overline{a}_1, \alpha_{2,s}) + (\delta^{\tau-t}-\delta^{\tau+1-t})u_1(\overline{\theta}_n, a_{1,\tau}, \alpha_{2,\tau}) + \delta^{\tau+1-t}u_1(\overline{\theta}_n, \underline{a}_1, \underline{a}_2)\Big] \geq u_1(\overline{\theta}_n, \underline{a}_1, \underline{a}_2).$$

where $\mathbb{E}[\cdot]$ 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 $u_1(\overline{\theta}_n, \underline{a}_1, \underline{a}_2) \geq u_1(\overline{\theta}_n, \overline{a}_1, \overline{a}_2)$, the above inequality implies that:

$$\mathbb{E}\Big[\sum_{s=t}^{\tau-1}(1-\delta)\delta^{s-t}\Big(u_1(\overline{\theta}_n, \overline{a}_1, \alpha_{2,s}) - u_1(\overline{\theta}_n, \overline{a}_1, \overline{a}_2)\Big) + (\delta^{\tau-t}-\delta^{\tau+1-t})\Big(u_1(\overline{\theta}_n, \underline{a}_1, \alpha_{2,\tau}) - u_1(\overline{\theta}_n, \underline{a}_1, \underline{a}_2)\Big)\Big] \leq 0.$$

According to the definitions of K and M, we know that for all θ,

$$\mathbb{E}\Big[\sum_{s=t}^{\tau}(1-\delta)\delta^{s-t}\Big(u_1(\theta_n, \overline{a}_1, \alpha_{2,s}) - u_1(\theta_n, \overline{a}_1, \overline{a}_2)\Big)\Big] \leq 2M(K+1)(1-\delta). \tag{C.21}$$

This bounds the loss (relative to the payoff from the highest action profile) from above in periods before all types in Θ_n separate from the commitment type. For every $r^\infty \in \mathcal{R}_*^\sigma$, since $r^t \notin \widehat{\mathcal{R}}_g^\sigma$, we have:

$$\begin{aligned}
\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta\in\Theta} q^*(r^{\tau(r^\infty)+1})(\theta)\mathcal{D}(\theta, \overline{a}_1) \quad &\geq \quad \mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta\in\Theta_p\cup\Theta_g} q^*(r^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) \\
&> \quad \mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta\in\Theta} q^*(r^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0
\end{aligned}$$

According to Lemma C.7, we know that $V_\theta(a_*^{\tau(r^\infty)+1}, r^{\tau(r^\infty)+1}) = u_1(\theta, \overline{a}_1, \overline{a}_2)$ for all $\theta \in \Theta_g \cup \Theta_p$ and $r^\infty \in \mathcal{R}_*^\sigma$. This together with (C.21) gives the conclusion.

Second, when $\mu^*(r^t)(\Theta_n) = 0$. If $t = 0$, the conclusion directly follows from Proposition C.1. If $t \geq 1$ and there exists $\hat{r}^t$ such that $r^t, \hat{r}^t \succ r^{t-1}$, $(a_*^t, \hat{r}^t) \in \mathcal{H}^\sigma$ and $\hat{r}^t \in \widehat{\mathcal{R}}_g^\sigma$. Then, since

$$\mu^*(r^t) = \mu^*(\hat{r}^t),$$

we have $\mu^*(\hat{r}^t)(\Theta_n) = 0$. Since $\hat{r}^t \in \widehat{\mathcal{R}}_g^\sigma$, according to Lemma C.6, $\mu^*(\hat{r}^t) = \mu^*(r^t) \in \mathcal{B}_\kappa$. 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_*^t, r^t) with $r^t \notin \widehat{\mathcal{R}}_g^\sigma$.

Lemma C.9. *For every $\theta_n \in \Theta_n$ such that $\overline{a}_2 \notin BR_2(\overline{a}_1, \theta_n)$ and $r^t \notin \widehat{\mathcal{R}}^\sigma_g$ with $(a^t_*, r^t) \in \mathcal{H}^\sigma_{\theta_n}$ and $\mu^*(r^t) \notin \mathcal{B}_\kappa$, we have:*

$$V_{\theta_n}(a^t_*, r^t) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M. \tag{C.22}$$

This is implied by Lemma C.8 (Part I). Let

$$A(\delta) \equiv 2M(K+1)(1-\delta), \quad B(\delta) \equiv 2M(1-\delta^{\overline{T}_0(\kappa)})$$

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 $r^t \in \widehat{\mathcal{R}}^\sigma_g$.

Lemma C.10. *When $\delta > \overline{\delta}$ and $\sigma \in NE(\delta, \mu)$, for every $r^t \in \widehat{\mathcal{R}}^\sigma_g$,*

$$V_\theta(a^t_*, r^t) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - \Big(A(\delta) + B(\delta)\Big) - 2\overline{T}_1(\kappa)\Big(A(\delta) + B(\delta) + C(\delta)\Big).^{28} \tag{C.23}$$

for all θ such that $(a^t_*, r^t) \in \mathcal{H}^\sigma(\theta)$.

PROOF OF LEMMA C.10: The non-trivial part of the proof deals with situations where $\mu^*(r^t) \notin \mathcal{B}_\kappa$. Since $r^t \in \widehat{\mathcal{R}}^\sigma_g$, Lemma C.6 implies that $\mu^*(r^t)(\Theta_n) \neq 0$. Without loss of generality, assume $\Theta_n \subset \text{supp}\Big(\mu^*(r^t)\Big)$. Let me introduce $|\Theta_n| + 1$ integer valued random variables on the space $\mathcal{R}^\sigma_*$.

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

1. $\mu^*(r^{s+1}) \in \mathcal{B}_{\kappa/2}$ for $r^{s+1} \succ r^s$ with $(a^{s+1}_*, r^{s+1}) \in \mathcal{H}^\sigma$.
2. $r^s \notin \widehat{\mathcal{R}}^\sigma_g$.

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

- $\big((a^\tau_*, a_1), r^{\tau+1}\big) \in \mathcal{H}^\sigma(\tilde{\theta})$ for some $\tilde{\theta} \in \Theta_p \cup \Theta_n$.
- $\mu\big((a^\tau_*, a_1), r^{\tau+1}\big) \notin \mathcal{B}_0$.

Lemma C.3 implies the existence of $\theta \in \Theta_p \cup \Theta_n$ with $\big((a^\tau_*, a_1), r^{\tau+1}\big) \in \mathcal{H}^\sigma(\theta)$ such that

$$V_\theta\big((a^\tau_*, a_1), r^{\tau+1}\big) = u_1(\theta, \underline{a}_1, \underline{a}_2).$$

Suppose towards a contradiction that $\theta \in \Theta_p$, then Lemma C.1 implies that $\mu^*(r^{\tau+1})(\Theta_n) = 0$. Since $\mu^*(r^{\tau+1}) \in \mathcal{B}_{\kappa/2}$, Proposition C.1 implies that type θ's continuation payoff by always playing $\overline{a}_1$ is at least

$$(1-\delta^{\overline{T}_0(\kappa/2)})u_1(\theta, \overline{a}_1, \underline{a}_2) + \delta^{\overline{T}_0(\kappa/2)}u_1(\theta, \overline{a}_1, \overline{a}_2),$$

which is strictly larger than his payoff from playing a_1, which is at most $2M(1-\delta) + u_1(\theta, \underline{a}_1, \underline{a}_2)$, leading to a contradiction.

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

In the second case, Lemma C.9 implies that $V_{\theta_n}\big(a^\tau_*, r^\tau\big) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M$ for all $\theta_n \in \Theta_n$ with $r^\tau \in \mathcal{H}^\sigma(\theta_n)$.

[28] One can further tighten this bound. However, (C.23) is sufficient for the purpose of proving Theorem 2.

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

1. $\mu^*(r^{s+1}) \in \mathcal{B}_{\kappa/2}$ for $r^{s+1} \succ r^s$ with $(a_*^{s+1}, r^{s+1}) \in \mathcal{H}^\sigma$.
2. $r^s \notin \widehat{\mathcal{R}}_g^\sigma$.
3. $\mu^*(r^{s+1})(\theta_n) = 0$ for $r^{s+1} \succ r^s$ with $(a_*^{s+1}, r^{s+1}) \in \mathcal{H}^\sigma$, .

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

$$\tau = \max_{\theta_n \in \Theta_n}\{\tau_{\theta_n}\}. \tag{C.24}$$

Suppose on the contrary that $\tau > \max_{\theta_n \in \Theta_n}\{\tau_{\theta_n}\}$ for some $r^\infty \in \mathcal{R}_*^\sigma$. Then there exists $(a_*^s, r^s) \succsim (a_*^t, r^t)$ such that $r^s \in \widehat{\mathcal{R}}_g^\sigma$, $\mu^*(r^s) \notin \mathcal{B}_\kappa$ and $\mu^*(r^s)(\Theta_n) = 0$, which contradicts Lemma C.6 when $\delta > \overline{\delta}$.

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

$$\mathbb{E}\Big[\sum_{s=t}^{\tau}(1-\delta)\delta^{\tau-t}\Big(u_1(\theta, \overline{a}_1, \overline{a}_2) - u_1(\theta, \overline{a}_1, \hat{\alpha}_{2,s})\Big)\Big] \leq 2MK|\Theta_n|(1-\delta), \tag{C.25}$$

for all $\theta \in \Theta$ and

$$V_{\tilde{\theta}_n}(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}}) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M, \tag{C.26}$$

for

$$\tilde{\theta} \equiv \min\Big\{\Theta_n \bigcap \text{supp}\Big(\mu^*(r^{\tau_{\theta_n}+1})\Big)\Big\}$$

with $\theta_n, \tilde{\theta}_n \in \Theta_n$, where $\mathbb{E}[\cdot]$ is taken over $\mathcal{P}^\sigma$ and $\hat{\alpha}_{2,s} \in \Delta(A_2)$ is player 2's (mixed) action at (a_*^s, r^s).

When $|\Theta_n| = 1$, let θ_n be its unique element. Consider player 1's pure strategy of playing $\overline{a}_1$ until r^τ and then play $\underline{a}_1$ forever. This is one of type θ_n's best responses according to (C.24), which results in payoff at most:

$$\mathbb{E}\Big[\sum_{s=t}^{\tau-1}(1-\delta)\delta^{s-t}u_1(\theta_n, \overline{a}_1, \hat{\alpha}_{2,s}) + \delta^{\tau-t}\Big(u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M\Big)\Big].$$

The above expression cannot be smaller than $u_1(\theta_n, \underline{a}_1, \underline{a}_2)$, which is the payoff he can guarantee by playing $\underline{a}_1$ in every period. Since $u_1(\theta_n, \underline{a}_1, \underline{a}_2) \geq u_1(\theta_n, \overline{a}_1, \overline{a}_2)$, and from the definition of K, we get for all θ,

$$\mathbb{E}\Big[\sum_{s=t}^{\tau-1}(1-\delta)\delta^{s-t}\Big(u_1(\theta, \overline{a}_1, \overline{a}_2) - u_1(\theta, \overline{a}_1, \hat{\alpha}_{2,s})\Big)\Big] \leq 2MK(1-\delta).$$

We can then obtain (C.26) for free since $\tau = \tau_{\theta_n}$ and type θ_n's continuation value at (a_*^τ, r^τ) is at most $u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M$ 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 $(a_*^\tau, r^\tau) \notin \mathcal{H}^\sigma(\theta_n)$, then there exists $(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}}) \prec (a_*^\tau, r^\tau)$ with $(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}}) \in \mathcal{H}^\sigma(\theta_n)$ at which type θ_n plays $\overline{a}_1$ with probability 0. I put an upper bound on type θ_n's continuation payoff at $(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}})$ by examining type $\tilde{\theta}_n \in \Theta_n \backslash \{\theta_n\}$'s incentive to play $\overline{a}_1$ at $(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}})$, where

$$\tilde{\theta} \equiv \min\Big\{\Theta_n \bigcap \text{supp}\Big(\mu^*(r^{\tau_{\theta_n}+1})\Big)\Big\}$$

This requires that:

$$\mathbb{E}\Big[\sum_{s=0}^{\infty}(1-\delta)\delta^s u_1(\tilde{\theta}_n, \alpha_{1,s}, \alpha_{2,s})\Big] \leq \underbrace{u_1(\tilde{\theta}_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M}_{\text{by induction hypothesis}}.$$

where $\{(\alpha_{1,s}, \alpha_{2,s})\}_{s\in\mathbb{N}}$ is the equilibrium continuation play following $(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}})$. By definition, $\tilde{\theta}_n \succ \theta_n$, so the supermodularity condition implies that:

$$u_1(\theta_n, \underline{a}_1, \underline{a}_2) - u_1(\tilde{\theta}_n, \underline{a}_1, \underline{a}_2) \geq u_1(\theta_n, \alpha_{1,s}, \alpha_{2,s}) - u_1(\tilde{\theta}_n, \alpha_{1,s}, \alpha_{2,s}).$$

Therefore, we have:

$$\begin{aligned} V_{\theta_n}(a_*^{\tau_{\theta_n}}, r^{\tau_{\theta_n}}) &= \mathbb{E}\Big[\sum_{s=0}^{\infty}(1-\delta)\delta^s u_1(\theta_n, \alpha_{1,s}, \alpha_{2,s})\Big] \\ &\leq \mathbb{E}\Big[\sum_{s=0}^{\infty}(1-\delta)\delta^s \Big(u_1(\tilde{\theta}_n, \alpha_{1,s}, \alpha_{2,s}) + u_1(\theta_n, \underline{a}_1, \underline{a}_2) - u_1(\tilde{\theta}_n, \underline{a}_1, \underline{a}_2)\Big)\Big] \\ &\leq u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M. \end{aligned}$$

Back to type θ_n's incentive constraint. Since it is optimal for him to play $\overline{a}_1$ until $r^{\tau_{\theta_n}}$ and then play $\underline{a}_1$ forever, doing so must give him a higher payoff than playing $\underline{a}_1$ forever starting from r^t, which gives:

$$\mathbb{E}\Big[\sum_{s=t}^{\tau_{\theta_n}-1}(1-\delta)\delta^{s-t} u_1(\theta_n, \overline{a}_1, \hat{\alpha}_{2,s}) + \delta^{\tau_{\theta_n}}\Big(u_1(\theta_n, \underline{a}_1, \underline{a}_2) + 2(1-\delta)M\Big)\Big] \geq u_1(\theta_n, \underline{a}_1, \underline{a}_2).$$

This implies that:

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

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

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

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

$$\mathbb{E}\Big[\sum_{s=\tau_{\theta_n}}^{\tau_\theta-1}(1-\delta)\delta^{s-\tau_{\theta_n}}\Big(u_1(\theta, \overline{a}_1, \overline{a}_2) - u_1(\theta, \overline{a}_1, \alpha_{2,s})\Big)\Big] \leq 2MK(k-1)(1-\delta). \tag{C.28}$$

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

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

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 $(a_*^{\tau+1}, r^{\tau+1}) \in \mathcal{H}^\sigma$. I consider three cases:

1. If $r^{\tau+1} \notin \widehat{\mathcal{R}}_g^\sigma$, by Lemma C.8, then for every θ,

$$V_\theta(a_*^{\tau+1}, r^{\tau+1}) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - A(\delta).$$

2. If $r^{\tau+1} \in \widehat{\mathcal{R}}_g^\sigma$ and $\mu^*(r^s) \in \mathcal{B}_\kappa$ for all r^s satisfying $r^s \succsim r^{\tau+1}$ and $r^s \in \widehat{\mathcal{R}}_g^\sigma$, then for every θ,

$$V_\theta(a_*^{\tau+1}, r^{\tau+1}) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - B(\delta).$$

3. If there exists r^s such that $\mu^*(r^s) \notin \mathcal{B}_\kappa$ with $r^s \succsim r^{\tau+1}$ and $r^s \in \widehat{\mathcal{R}}_g^\sigma$, then repeat the procedure in the beginning of this proof by defining random variables

   ▷ $\tau' : \mathcal{R}_*^\sigma \to \{n \in \mathbb{N} \cup \{+\infty\} | n \geq s\}$

   ▷ $\tau'_{\theta_n} : \mathcal{R}_*^\sigma \to \{n \in \mathbb{N} \cup \{+\infty\} | n \geq s\}$

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

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

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

   Therefore, such iterations can last for at most $2\overline{T}_1(\kappa)$ rounds.

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

$$\mathcal{R}_b^\sigma \equiv \Big\{ r^t \Big| \mu^*(r^t) \notin \mathcal{B}_\kappa \text{ and } r^t \in \hat{\mathcal{R}}_g^\sigma \Big\}$$

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

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

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

1. $\mu^*(r^{s+1}) \in \mathcal{B}_{\kappa/2}$ for $r^{s+1} \succ r^s$ with $(a_*^{s+1}, r^{s+1}) \in \mathcal{H}^\sigma$.
2. $r^s \notin \widehat{\mathcal{R}}_g^\sigma$.

I call

$$\mathcal{T}(r^t) \equiv \Big\{ r^s \Big| r^{\tau[r^{t_1}]} \succsim r^s \succsim r^t \Big\}$$

a '*tree*' with root r^t. For any $k \geq 2$, let

$$\mathcal{R}^\sigma(k) \equiv \Big\{ r^t \Big| r^t \in \mathcal{R}_b^\sigma,\ r^t \succ r^{\tau[r^s]} \text{ for some } r^s \in \mathcal{R}^\sigma(k-1) \text{ and there exists no } r^s \prec r^t \text{ that satisfy these two conditions} \Big\}.$$

Let T be the largest integer such that $\mathcal{R}^\sigma(T) \neq \{\varnothing\}$. According to (C.29), we know that $T \leq 2\overline{T}_1(\kappa)$. Similarly, we can define trees with roots in $\mathcal{R}(k)$ for every $k \leq T$.

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

$$V_\theta(a_*^t, r^t) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - (T + 1 - k)\Big(A(\delta) + B(\delta) + C(\delta) \Big). \tag{C.30}$$

The proof is done by inducting on k from backwards. When $k = T$, player 1's continuation value at $(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$ is at least $u_1(\theta, \overline{a}_1, \overline{a}_2) - A(\delta) - B(\delta)$ according to Lemma C.2 and Lemma C.8. His continuation value at r^t is at least:

$$u_1(\theta, \overline{a}_1, \overline{a}_2) - A(\delta) - B(\delta) - C(\delta).$$

Suppose the conclusion holds for all $k \geq n+1$, then when $k = n$, type θ's continuation payoff at (a_*^t, r^t) is at least:

$$\mathbb{E}\Big[(1 - \delta^{\tau[r^t]-t}) u_1(\theta, \overline{a}_1, \overline{a}_2) + \delta^{\tau[r^t]-t} V_\theta(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1}) \Big] - C(\delta)$$

Pick any $(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$, consider the set of random paths r^∞ that it is consistent with, let this set be $\mathcal{R}^\infty(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$. Partition it into two subsets:

1. $\mathcal{R}_{+}^{\infty}(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$ consists of r^∞ such that for all $s \geq \tau[r^t]+1$ and $r^s \prec r^\infty$, we have $r^s \notin \mathcal{R}_b^\sigma$.

2. $\mathcal{R}_{-}^{\infty}(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$ consists of r^∞ such that there exists $s \geq \tau[r^t]+1$ and $r^s \prec r^\infty$ at which $r^s \in \mathcal{R}^\sigma(n+1)$.

Conditional on $r^\infty \in \mathcal{R}_{+}^{\infty}(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$, we have:

$$V_\theta(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1}) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - A(\delta) - B(\delta).$$

Conditional on $r^\infty \in \mathcal{R}_{-}^{\infty}(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1})$, type θ's continuation payoff is no less than

$$V_\theta(a_*^s, r^s) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - (T-n)\Big(A(\delta) + B(\delta) + C(\delta)\Big)$$

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

$$V_\theta(a_*^{\tau[r^t]+1}, r^{\tau[r^t]+1}) \geq u_1(\theta, \overline{a}_1, \overline{a}_2) - (T+1-n)\Big(A(\delta) + B(\delta) + C(\delta)\Big).$$

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 every $r^t \in \mathcal{R}^\sigma(0)$. □

The first statement of Theorem 2 is implied by Lemmas C.8, C.9 and C.10.

C.7 Proof of Statement 2 Theorem 2

Let $\kappa \in (0,1)$. Given $\delta > \overline{\delta}$ and $\sigma \in \text{NE}(\delta, \mu)$, let us examine r^1 such that $(a_*^1, r^1) \in \mathcal{H}^\sigma$.[29] If $\mu^*(r^1) \in \mathcal{B}_\kappa$, then for every $\hat{r}^1$ with $(a_*^1, \hat{r}^1) \in \mathcal{H}^\sigma$, we have $\mu^*(\hat{r}^1) \in \mathcal{B}_\kappa$. The conclusion is then implied by statement 1 of Theorem 2. If $\mu^*(r^1) \notin \mathcal{B}_\kappa$, then we still have:

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta} q^*(r^1)(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0. \tag{C.31}$$

This is because otherwise, there exists $\theta \in \text{supp}\mu^*(r^1)$ such that $V_\theta(a_*^1, r^1) = u_1(\theta, \underline{a}_1, \underline{a}_2)$ according to Lemma C.3, contradicting type θ's incentive to play $\overline{a}_1$ in period 0.

In what follows, I consider two cases separately.

1. If $\Theta_n \bigcap \text{supp}\mu^*(r^1) = \{\varnothing\}$, then Lemma C.6 implies that $r^1 \notin \widehat{\mathcal{R}}_g^\sigma$. According to Lemma C.4, there exists $\theta \in (\Theta_p \cup \Theta_n) \bigcap \text{supp}\mu^*(r^1)$ such that $r^1 \in \overline{\mathcal{R}}^\theta$. According to Lemma C.1, for every $\theta_g \in \Theta_g$, type θ_g will play $\overline{a}_1$ at every $(a_*^t, r^t) \succsim (a_*^1, r^1)$ with $(a_*^t, r^t) \in \mathcal{H}^\sigma(\theta_g)$.

   According to the definition of v_θ^*, for every $\theta \in \Theta$, type θ can secure payoff v_θ^* at r^1. Since $\mu^*(r^1) \notin \mathcal{B}_\kappa$, $\mu^*(\hat{r}^1) \notin \mathcal{B}_\kappa$ for every $\hat{r}^1$ with $(a_*^1, \hat{r}^1) \in \mathcal{H}^\sigma$. The argument in the previous paragraph applies symmetrically, which implies that type θ's discounted average payoff at h^0 is at least

$$(1-\delta)u_1(\theta, \overline{a}_1, \underline{a}_2) + \delta v_\theta^*.$$

2. If $\Theta_n \bigcap \text{supp}\mu^*(r^1) \neq \{\varnothing\}$, then according to Lemma C.10, type θ can guarantee payoff at least the RHS of (C.23), which leads to the same conclusion.

[29] I consider random histories in period 1 as other pure strategy commitment types will be separated from type $\overline{a}_1$ in period 1. If there are commitment types playing mixed strategies, then one needs to examine random histories in period T such that after T periods, the probability of the mixed strategy commitment type becomes negligible compared to that of commitment type $\overline{a}_1$.

D Proof of Theorem 3 & Extensions

I show Theorem 3 and present an extension by allowing for mixed strategy commitment types. I will inherit the notation from Appendix B. In particular, recall the definitions of $\overline{\Omega}$, $\mathcal{H}^\sigma$, $\mathcal{H}^\sigma(\omega)$, $q(h^t)$ and $\widehat{\mathcal{R}}^\sigma_g$.

D.1 Proof of Theorem 3

Step 1: Let

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

and

$$Y(h^t) \equiv \mu(\Omega)\mathcal{D}(\overline{\theta}, \overline{a}_1) + \sum_{\theta \in \Theta_g \cup \Theta_p} q(h^t)(\theta)\mathcal{D}(\theta, \overline{a}_1). \tag{D.2}$$

According to (4.7), $X(h^0) < 0$ and $Y(h^0) < 0$. Moreover, at every $h^t \in \mathcal{H}^\sigma$ with $Y(h^t) < 0$, player 2 has a strict incentive to play $\underline{a}_2$. According to Lemma C.3, there exists $\theta_p \in \Theta_p$ with $h^t \in \mathcal{H}(\theta_p)$ such that type θ_p's continuation value at h^t is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, which further implies that playing $\underline{a}_1$ in every period is one of his best replies. According to Lemma C.1 and using the implication that $Y(h^0) < 0$, every $\theta_n \in \Theta_n$ plays $\underline{a}_1$ with probability 1 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 $h^1 = (\overline{a}_1, r^1) \in \mathcal{H}^\sigma$, we have:

$$\sum_{\theta \in \Theta_g \cup \Theta_p} q(h^1)(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0. \tag{D.3}$$

Suppose towards a contradiction that $\sum_{\theta \in \Theta_g \cup \Theta_p} q(h^1)(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$, then $X(h^1) \geq \mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1)$. According to Proposition C.1, there exists $K \in \mathbb{R}_+$ independent of δ such that type θ's continuation payoff is at least $u_1(\theta, \overline{a}_1, \overline{a}_2) - (1-\delta)K$ at every $h^1_* \in \mathcal{H}^\sigma$. When δ is large enough, this contradicts the conclusion in the previous step that there exists $\theta_p \in \Theta_p$ such that type θ_p's continuation value at h^0 is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, as he can profitably deviate by playing $\overline{a}_1$ in period 0.

Step 3: According to (D.3), we have $\mu^*(r^1) \notin \mathcal{B}_0$. Step 1 also implies that $\mu^*(r^1)(\Theta_n) = 0$. According to Lemma C.6, we have $r^1 \notin \widehat{\mathcal{R}}^\sigma_g$. According to Lemma C.1, type θ_g plays $\overline{a}_1$ at every $h^t \in \mathcal{H}(\theta_g)$ with $t \geq 1$ for every $\theta_g \in \Theta_g$.

Next, I show that $r^0 \notin \widehat{\mathcal{R}}^\sigma_g$. Suppose towards a contradiction that $r^0 \in \widehat{\mathcal{R}}^\sigma_g$, then there exists $h^T = (a^T_*, r^T) \in \mathcal{H}^\sigma$ such that $\mu(h^T)(\Theta_p \cup \Theta_n) = 0$. If $T \geq 2$, it contradicts our previous conclusion that $r^1 \notin \widehat{\mathcal{R}}^\sigma_g$. If $T = 1$, then it contradicts (D.3). Therefore, we have $r^0 \notin \widehat{\mathcal{R}}^\sigma_g$. This implies that type θ_g plays $\overline{a}_1$ 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 θ_p by showing that $X(h^t) = 0$ for every $h^t = (a^t_*, r^t) \in \mathcal{H}^\sigma$ 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 θ_p's continuation payoff at (a^{t+1}_*, r^{t+1}) is $u_1(\theta_p, \overline{a}_1, \overline{a}_2)$ by playing $\overline{a}_1$ in every period, while his continuation payoff at (a^t_*, a_1, r^{t+1}) is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, leading to a contradiction. Suppose next that $X(h^1) < 0$, similar to the previous argument, 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 $t \geq 1$, $h^t = (a^t_*, r^t) \in \mathcal{H}^\sigma$. This establishes the uniqueness of player 1's equilibrium behavior.

D.2 Generalizations to Mixed Strategy Commitment Types

Next, I generalize Theorem 3 by accommodating mixed strategy commitment types. For every $\theta \in \Theta$, let $\lambda(\theta)$ be the prior likelihood ratio between strategic type θ and the lowest strategic type $\underline{\theta} \equiv \min \Theta$ 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. Recall that

$$\Omega^g \equiv \{\alpha_1 \in \Omega | \mathcal{D}(\alpha_1, \phi_{\alpha_1}) > 0\}.$$

Let $\mathcal{H}_1^t$ be the set of action paths with length t and let $\overline{\mathcal{H}}_1^t \equiv \{(\overline{a}_1, ...\overline{a}_1), (\underline{a}_1, ...\underline{a}_1)\}$, which is a subset of $\mathcal{H}_1^t$. For every strategy profile $\sigma \equiv \Big((\sigma_\theta)_{\theta \in \Theta}, \sigma_2\Big)$ and state θ, let $\mathcal{P}_{1,t}^{\sigma}(\theta)$ be the probability measure over $\mathcal{H}_1^t$ induced by $(\sigma_\theta, \sigma_2)$ and let $\mathcal{H}^\sigma(\theta)$ be the set of histories that occur with positive probability under $(\sigma_\theta, \sigma_2)$. For every $\gamma \geq 0$ and two equilibria σ and σ', strategic type θ's on-path behaviors in these equilibria are *γ-close* if for every $t \geq 1$,

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

where $D_B(p, q)$ denotes the Bhattacharyya distance between distributions p and q.[30] If $\gamma = 0$, then type θ's on-path behavior in these two equilibria are the same. Intuitively, the above distance measures the difference between the *ex ante* distributions over player 1's action paths. The generalization of Theorem 3 that allows for mixed strategy commitment types is stated below.

Theorem 3'. *Suppose $\overline{a}_1 \in \Omega$ and $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$, then for every $\lambda \in [0, +\infty)^{|\Theta|}$ satisfying:*

$$\sum_{\theta \in \Theta_p \cup \Theta_g} \lambda(\theta) \mathcal{D}(\theta, \overline{a}_1) < 0, \tag{D.4}$$

there exist $\overline{\epsilon} > 0$ and $\gamma : (0, \overline{\epsilon}) \to \mathbb{R}_+$ satisfying $\lim_{\epsilon \downarrow 0} \gamma(\epsilon) = 0$, such that for every μ with $\{\mu(\theta)/\mu(\underline{\theta})\}_{\theta \in \Theta} = \lambda$ and $\mu(\Omega) < \overline{\epsilon}$, there exist $\overline{\delta} \in (0, 1)$ and $\theta_p^* \in \Theta_p$ such that for every $\delta > \overline{\delta}$ and $\sigma, \sigma' \in NE(\delta, \mu)$:

- *For every $\theta \succ \theta_p^*$ and $h^t \in \mathcal{H}^\sigma(\theta)$, type θ plays $\overline{a}_1$ at h^t.*
- *For every $\theta \prec \theta_p^*$ and $h^t \in \mathcal{H}^\sigma(\theta)$, type θ plays $\underline{a}_1$ at h^t.*
- *Type θ_p^*'s on-path behavior is $\gamma(\epsilon)$-close between σ and σ'.*
- *$\mathcal{P}_{1,t}^{\sigma}(\theta_p^*)(\overline{\mathcal{H}_1^t}) > 1 - \gamma(\overline{\epsilon})$ for every $t \geq 1$.*

If all the actions in Ω^g are pure, then type θ_p^*'s on-path behavior is the same across all equilibria under generic parameter values, according to which he behaves consistently over time with probability 1.

To comment on the conditions in Theorem 3', first of all, (D.4) is implied by (4.3) given that $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$. Second, when Ω^g contains elements other than $\overline{a}_1$, obtaining sharp predictions on player 1's on-path behavior requires the total probability of commitment types to be small enough. This is because the presence of multiple good commitment types gives the strategic types many good reputations to choose from. In particular, if a good commitment type other than $\overline{a}_1$ occurs with sufficiently high probability, then the bad strategic types can imitate

[30] One can replace the Bhattacharyya distance with the Rényi divergence or Kullback-Leibler divergence in the following way: strategic type θ's on-path behavior is γ-close between σ and σ' if there exists a probability measure $\mathcal{P}$ on $\mathcal{H}$ such that for every $t \geq 1$,

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

where $D(\cdot||\cdot)$ is either the Rényi divergence of order greater than 1 or the Kullback-Leibler divergence.

this type with high probability which weakens the punishment for behaving inconsistently. An example of this is provided in Appendix G.5.

Next, I provide a sufficient condition on $\overline{\epsilon}$, namely the upper bound on the total probability of commitment types. Let

$$Y(h^t) \equiv \mu(h^t)(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\alpha_1 \in \Omega^g} \mu(h^t)(\alpha_1)\mathcal{D}(\phi_{\alpha_1}, \alpha_1) + \sum_{\theta \in \Theta_p \cup \Theta_g} \mu(h^t)(\theta)\mathcal{D}(\theta, \overline{a}_1), \tag{D.5}$$

which is an upper bound on player 2's incentive to play $\overline{a}_2$ at h^t. I require $\overline{\epsilon}$ to be small enough such that

$$\overline{\epsilon}\mathcal{D}(\overline{\theta}, \overline{a}_1) + (1-\overline{\epsilon}) \sum_{\theta \in \Theta_p \cup \Theta_g} \frac{\lambda(\theta)}{\sum_{\tilde{\theta} \in \Theta} \lambda(\tilde{\theta})} \mathcal{D}(\theta, \overline{a}_1) < 0. \tag{D.6}$$

Such $\overline{\epsilon}$ exists since $\sum_{\theta \in \Theta_p \cup \Theta_g} \lambda(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0$. Inequality (D.6) implies that $Y(h^0) < 0$, which is also equivalent to (4.3) when $\Omega^g = \{\varnothing\}$.

Third, when there are mixed strategy commitment types, the probabilities with which type θ_p^* mixes may not be the same across all equilibria for two reasons.

1. Suppose player 2 has no incentive to play $\overline{a}_2$ against any mixed commitment type, then given that all strategic types either plays $\overline{a}_1$ in every period or plays $\underline{a}_1$ in every period, player 2's incentive to play $\overline{a}_2$ is increasing over time as long as $\overline{a}_1$ has been observed in every period of the past. As a result, there will be $T(\delta)$ periods in which player 2 has a strict incentive to play $\underline{a}_2$, followed by at most one period in which she is indifferent between $\overline{a}_2$ and $\underline{a}_2$, followed by periods in which she has a strict incentive to play $\overline{a}_2$, with $T(\delta)$ and the probabilities with which she mix between $\overline{a}_2$ and $\underline{a}_2$ in period $T(\delta)$ pinned down by type θ_p^*'s indifference condition in period 0. Under degenerate parameter values in which there exists an integer T such that type θ_p^* is just indifferent between playing $\underline{a}_1$ in every period and playing $\overline{a}_1$ in every period when $\underline{a}_2$ will be played in the first T periods, his mixing probability between playing $\overline{a}_1$ in every period and playing $\underline{a}_1$ in every period is not unique. Nevertheless, when the ex ante probability of Ω is smaller than ϵ, his probability of mixing cannot vary by more than $\gamma(\epsilon)$ even in this degenerate case, with $\gamma(\cdot)$ diminishes as $\epsilon \downarrow 0$.

2. When there are good mixed strategy commitment types, the probability with which type θ_p^* 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 θ_p^* imitates the mixed commitment type with sufficiently high probability relative to the probability of that good commitment type, then player 2 will have a strict incentive to play $\underline{a}_2$. This implies that as the probability of commitment type vanishes, the upper bound on the probability with which type θ_p^* builds a mixed reputation also vanishes.

D.3 Proof of Theorem 3'

Unique Equilibrium Behavior for Strategic Types in Θ_n and Θ_g: 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 $\underline{a}_1$ at every $h^t \in \mathcal{H}^\sigma(\theta_n)$ in every equilibrium σ. This is similar to Step 1 in the proof of Theorem 3. Since $Y(h^0) < 0$ and at every $h^t \in \mathcal{H}^\sigma$ with $Y(h^t) < 0$, player 2 has a strict incentive to play $\underline{a}_2$. Applying Lemma C.3, there exists $\theta_p \in \Theta_p$ with $h^t \in \mathcal{H}^\sigma(\theta_p)$ such that type θ_p's continuation value at h^t is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$. Therefore, playing $\underline{a}_1$ in every period is his best reply. Type θ_n's on-path behavior is pinned down by Lemma C.1.

Next, I establish (D.3). Suppose towards a contradiction that $\sum_{\theta \in \Theta_g \cup \Theta_p} q(h^1)(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$, then $Y(h^1) \geq \mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1)$. According to Theorem 2, there exists $K \in \mathbb{R}_+$ independent of δ such that type θ's continuation payoff is at least $u_1(\theta, \overline{a}_1, \overline{a}_2) - (1-\delta)K$ at every $h^1_* \in \mathcal{H}^\sigma$. When δ is large enough, this contradicts the conclusion in the previous step that there exists $\theta_p \in \Theta_p$ such that type θ_p's continuation value at h^0 is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, as he can profitably deviate by playing $\overline{a}_1$ in period 0. According to (D.3), we have $\mu^*(r^1) \notin \mathcal{B}_0$. Following the same procedure, one can show that $r^1 \notin \widehat{\mathcal{R}}^\sigma_g$ and $r^0 \notin \widehat{\mathcal{R}}^\sigma_g$ for every r^1 satisfying $(a^1_*, r^1) \in \mathcal{H}^\sigma$. This implies that for every equilibrium σ and every $\theta_g \in \Theta_g$, type θ_g plays $\overline{a}_1$ at every $h^t \in \mathcal{H}^\sigma(\theta_g)$.

Consistency of Equilibrium Behavior and Generic Uniqueness of Equilibrium Payoff when $\theta \in \Theta_p$: Let

$$\Omega^{gm} \equiv \Big\{\alpha_1 \in \Omega \Big| \alpha_1 \text{ is non-trivially mixed}\Big\}, \tag{D.7}$$

be the set of good mixed commitment actions. I start from showing that when $\Omega^{gm} = \{\varnothing\}$, type θ_p will behave consistently over time for every $\theta_p \in \Theta_p$ in every equilibrium. For every $t \geq 1$, let

$$Z(h^t) \equiv \mu(h^t)(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\alpha_1 \in \widehat{\Omega}^b} q(h^t)(\alpha_1)\mathcal{D}(\phi_{\alpha_1}, \alpha_1) + \sum_{\theta \in \Theta_p \cup \Theta_g} q(h^t)(\theta)\mathcal{D}(\theta, \overline{a}_1) \tag{D.8}$$

where

$$\widehat{\Omega}^b \equiv \Big\{\alpha_1 \in \Omega \backslash \{\overline{a}_1\} \Big| \mathcal{D}(\alpha_1, \phi_{\alpha_1}) < 0\Big\}. \tag{D.9}$$

Intuitively, $Z(h^t)$ provides a lower bound on player 2's incentives to play $\overline{a}_2$. If $\Omega^{gm} = \{\varnothing\}$, then $\mu(h^t)(\Omega^g \backslash \{\overline{a}_1\}) = 0$ for every $h^t = (a^t_*, r^t) \in \mathcal{H}^\sigma$ with $t \geq 1$. Therefore, player 2 has a strict incentive to play $\underline{a}_2$ if $Z(h^t) < 0$. Moreover, according to the conclusion in the previous step that type $\theta_g \in \Theta_g$ plays $\overline{a}_1$ for sure at every $h^t = (a^t_*, r^t)$, we know that for every $h^t \succ h^{t-1}$, we have $Z(h^t) \geq Z(h^{t-1})$.

Subcase 1: No Mixed Commitment Types Consider the case where there exists no $\alpha_1 \in \widehat{\Omega}^b$ such that $\alpha_1 \notin A_1$, i.e. there are no mixed strategy commitment types that affect player 2's best reply. By definition, $Z(h^t) = X(h^t)$ for every $t \geq 1$. As shown in Theorem 3, we know that $Z(h^t) = 0$ for every $h^t = (a^t_*, r^t) \in \mathcal{H}^\sigma$ and $t \geq 1$. Let $\Omega^g \equiv \{a^1_1, ..., a^{n-1}_1, a^n_1\}$ with $a^1_1 \prec a^2_1 \prec ... \prec a^{n-1}_1 \prec a^n_1 \equiv \overline{a}_1$. When $n \geq 2$, there exists $q : \Theta_p \to \Delta(\Omega^g \cup \{\underline{a}_1\})$ such that:

- **Monotonicity:** For every $\theta_p \succ \theta'_p$ and $a^i_1 \in \Omega^g \cup \{\underline{a}_1\}$. First, if $q(\theta_p)(a^i_1) > 0$, then $q(\theta'_p)(a^j_1) = 0$ for every $a^j_1 \succ a^i_1$. Second, if $q(\theta'_p)(a^i_1) > 0$, then $q(\theta_p)(a^j_1) = 0$ for every $a^j_1 \prec a^i_1$.
- **Indifference:** For every $a^i_1 \in \Omega^g \backslash \{\underline{a}_1\}$, we have:

$$\mu(a^i_1)\mathcal{D}(\phi_{a^i_1}, a^i_1) + \sum_{\theta_p \in \Theta_p} \mu(\theta_p)q(\theta_p)(a^i_1)\mathcal{D}(\theta_p, a^i_1) = 0. \tag{D.10}$$

These two conditions uniquely pin down function $q(\cdot)$, and therefore, the behavior of every type in Θ_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. Recall the definition of consistent action path. Since all strategic types in Θ_g are playing $\overline{a}_1$ in every period, so type θ's continuation value at every on-path inconsistent history must be $u_1(\theta, \underline{a}_1, \underline{a}_2)$ for every $\theta \in \Theta$. I show that in every equilibrium, type θ_p's behavior must be consistent for every $\theta_p \in \Theta_p$. Let

$$W(h^t) \equiv \mu(h^t)(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta_p \cup \Theta_g} q(h^t)(\theta)\mathcal{D}(\theta, \overline{a}_1). \tag{D.11}$$

For every consistent history h^t where a_1 is the consistent action, we know that $W(h^t) \leq Z(h^t)$ since

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

From the proof of Theorem 3, we have $W(h^t) \geq 0$. A similar argument can show that:

1. If there exists $\alpha_1 \in \widehat{\Omega}^b$ such that $\alpha_1(a_1) > 0$, then $W(h^t) > 0$.
2. If there exists no such α_1, then $W(h^t) = 0$.

The consistency of type θ_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 $\underline{\theta}_p$ 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 $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, i.e. playing $\underline{a}_1$ 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 $\underline{\theta}_p$ must be playing inconsistently at h^t with probability 1.

Suppose type $\underline{\theta}_p$ plays inconsistently with positive probability at h^t with $Z(h^t) \leq 0$, then his continuation value at h^t is $u_1(\underline{\theta}_p, \underline{a}_1, \underline{a}_2)$. He strictly prefers to deviate and play $\underline{a}_1$ forever at $h^{t-1} \prec h^t$ unless there exists $\hat{h}^T \succ h^{t-1}$ such that $Z(\hat{h}^T) \geq 0$ and type $\underline{\theta}_p$ strictly prefers to play consistently from h^{t-1} to $\hat{h}^T$. This implies that every θ_p plays consistently with probability 1 from h^{t-1} to $\hat{h}^T$, i.e. for every $h^t \succ h^{t-1}$ in which type θ_p plays inconsistently with positive probability and $h^T \succ h^t$, we have $Z(h^T) > Z(\hat{h}^T) \geq 0$. This implies that at h^t, type $\underline{\theta}_p$'s continuation payoff by playing consistently until $Z \geq 0$ is strictly higher than behaving inconsistently, leading to a contradiction.

Suppose type $\underline{\theta}_p$ 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 $u_1(\underline{\theta}_p, a_1, \overline{a}_2)$, which is no less than $u_1(\underline{\theta}_p, \overline{a}_1, \overline{a}_2)$, while his continuation value by playing inconsistently is at most $(1-\delta)u_1(\underline{\theta}_p, \underline{a}_1, \overline{a}_2) + \delta u_1(\underline{\theta}_p, \underline{a}_1, \underline{a}_2)$, which is strictly less when δ is large enough, leading to a contradiction.

Consider generic μ such that there exist $\theta_p^* \in \Theta_p$ and $q \in (0,1)$ such that:

$$\mu(\overline{a}_1)\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + q\mu(\theta_p^*)\mathcal{D}(\theta_p^*, \overline{a}_1) + \sum_{\theta \succ \theta_p^*} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) = 0; \tag{D.12}$$

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

$$(1-\delta^T)u_1(\theta_p^*, a_1, \underline{a}_2) + \delta^T u_1(\theta_p^*, a_1, \overline{a}_2) = u_1(\theta_p^*, \underline{a}_1, \underline{a}_2). \tag{D.13}$$

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

$$\sum_{\theta \succ \theta_p^*} \mu(\theta)\mathcal{D}(\theta, \overline{a}_1) + \sum_{\alpha_1 \in \widehat{\Omega}^b} \mu(\Omega)\mathcal{D}(\phi_{\alpha_1}, \alpha_1) > 0 \tag{D.14}$$

and

$$(1-q)\mu(\theta_p^*)\mathcal{D}(\theta_p^*, \overline{a}_1) + \mu(\Omega) \max_{\alpha_1 \in \Omega} \mathcal{D}(\phi_{\alpha_1}, \alpha_1) < 0, \tag{D.15}$$

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

$$(1-\delta^T)u_1(\theta_p^*, \overline{a}_1, \underline{a}_2) + \delta^T u_1(\theta_p^*, \overline{a}_1, \overline{a}_2) > u_1(\theta_p^*, \underline{a}_1, \underline{a}_2) > (1-\delta^{T+1})u_1(\theta_p^*, \overline{a}_1, \underline{a}_2) + \delta^{T+1}u_1(\theta_p^*, \overline{a}_1, \overline{a}_2).$$

The probability with which type θ_p^* plays $\overline{a}_1$ in every period, denoted by $q^*(\overline{a}_1) \in (0,1)$, is pinned down via:

$$q^*(\overline{a}_1)\mu(\theta_p^*)\mathcal{D}(\theta_p^*,\overline{a}_1)+\sum_{\theta\succ\theta_p^*}\mu(\theta)\mathcal{D}(\theta,\overline{a}_1)+\sum_{\alpha_1\in\Omega}\mu(\alpha_1)\underbrace{\alpha_1(\overline{a}_1)^T}_{\text{prob that type } \alpha_1 \text{ plays } \overline{a}_1 \text{ for } T \text{ consecutive periods}}\mathcal{D}(\phi_{\alpha_1},\alpha_1)=0. \tag{D.16}$$

Similarly, the probability with which type θ_p^* plays $a_1 \in \Omega^g$ in every period, denoted by $q^*(a_1)$, is pinned down via:

$$q^*(a_1)\mu(\theta_p^*)\mathcal{D}(\theta_p^*,\overline{a}_1)+\sum_{\alpha_1\in\Omega}\mu(\alpha_1)\alpha_1(\overline{a}_1)^{T(a_1)}\mathcal{D}(\phi_{\alpha_1},\alpha_1)=0.$$

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

$$\begin{aligned}(1-\delta^{T(a_1)})u_1(\theta_p^*,a_1,\underline{a}_2)+\delta^{T(a_1)}u_1(\theta_p^*,a_1,\overline{a}_2) &> u_1(\theta_p^*,\underline{a}_1,\underline{a}_2)\\ &> (1-\delta^{T(a_1)+1})u_1(\theta_p^*,a_1,\underline{a}_2)+\delta^{T(a_1)+1}u_1(\theta_p^*,a_1,\overline{a}_2).\end{aligned}$$

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

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

Consider the expressions of $Y(h^t)$ in (D.5) and $Z(h^t)$ in (D.8) which provide upper and lower bounds, respectively, on player 2's incentive to play $\overline{a}_2$ at h^t. When $\mu(\Omega)<\overline{\epsilon}$, previous arguments imply the existence of $\overline{\gamma}(\overline{\epsilon})$ with $\lim_{\overline{\epsilon}\downarrow 0}\overline{\gamma}(\overline{\epsilon})=0$, such that for every equilibrium,

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

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

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

$$s\mathcal{D}(\theta_p,\overline{a}_1)+\mathcal{D}(\phi_{\alpha_1},\alpha_1)<0. \tag{D.17}$$

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

$$s^*\overline{\epsilon}\Big\{\min_{\theta_p\in\Theta_p}(1-\overline{\epsilon})\frac{\lambda(\theta_p)}{\sum_{\theta\in\Theta}\lambda(\theta)}\Big\}^{-1}. \tag{D.18}$$

Expression (D.18) provides an upper bound for $\gamma(\overline{\epsilon})$, which vanishes as $\overline{\epsilon}\downarrow 0$. When $\mu(\Omega)$ is sufficiently small, Lemma C.1 implies the existence of a cutoff type θ_p^* such that all types strictly above θ_p^* always plays $\overline{a}_1$ and all types strictly below θ_p^* always plays $\underline{a}_1$, and type θ_p^* plays consistently with probability at least $1-\gamma(\overline{\epsilon})$, concluding the proof.

E Proof of Propositions 4.1 and 4.3

The proof consists of three parts. In Part I, I show the sufficiency part of Proposition 4.1. In Part II, I show the necessity part of Proposition 4.1 and establish the upper bound in Proposition 4.3 for types below the cutoff type. In Part III, I establish the upper bound in Proposition 4.3 for types above the cutoff type.

E.1 Part I

In this part, I show the *if* direction of Proposition 4.1. If π satisfies (4.9), then the following strategy profile is an equilibrium when δ is large enough.

1. Types in $\Theta_g \cup \Theta_p$ plays $\overline{a}_1$ in every period on the equilibrium path.

2. Types in Θ_n plays $\underline{a}_1$ in every period on the equilibrium path.

3. Player 2's action in period 0 depends on the probability of the states in Θ_n.[31]

4. Starting from period 1, player 2 plays $\overline{a}_2$ if and only if $\overline{a}_1$ has been played in every previous period and plays $\underline{a}_1$ vice versa.

One can verify that for every $\theta \in \Theta$, type θ's equilibrium payoff when $\delta \to 1$ is no less than $u_1(\theta, \overline{a}_1, \overline{a}_2)$.

E.2 Part II

In this part, I show the *only if* direction of Proposition 4.1. The conclusion in this step also establishes Proposition 4.3 for types no greater than $\theta_p^*(\pi)$. The result is stated as Lemma E.1:

Lemma E.1. *For every π satisfying (4.9) and every $\theta \precsim \theta_p^*(\pi)$, type θ's payoff is $u(\theta, \underline{a}_1, \underline{a}_2)$ in every equilibrium.*

Proof of Lemma E.2: Let $\Theta \equiv \{\theta_1, ..., \theta_m\}$ with $\theta_1 \prec \theta_2 \prec ... \prec \theta_m$. I introduce a new element θ_0 and let $\theta_0 \prec \theta_1$. Let $\overline{\Theta} \equiv \Theta \cup \{\theta_0\}$. For a given belief $\pi \in \Delta(\Theta)$, if π satisfies (4.9), then let $\theta_p^*(\pi) \equiv \theta_0$; if π satisfies (4.10), then $\theta_p^*(\pi) \succsim \theta_1$ and is pinned down via (4.15). Let $\pi(h^t) \in \Delta(\Theta)$ be player 2's belief at h^t and let $\pi(h^t, a_1) \in \Delta(\Theta)$ be her posterior after observing a_1 at h^t.

I directly apply the conclusion of Lemma C.1 in Appendix C which is also valid in the benchmark repeated game without commitment types. It is done by induction on $\Big|\text{supp}(\pi)\Big|$, namely the number of types in the support of player 2's prior belief. The case in which $\Big|\text{supp}(\pi)\Big| = 1$ is trivial. Suppose towards a contradiction that the conclusion holds for all π satisfying $\Big|\text{supp}(\pi)\Big| \leq k-1$ but fails for some π with $\Big|\text{supp}(\pi)\Big| = k$. Let type $\theta \precsim \theta_p^*(\pi)$ be a type that obtains payoff strictly greater than $u_1(\theta, \underline{a}_1, \underline{a}_2)$. Then there exists $h^t \in \mathcal{H}^\sigma(\theta)$ and $a_1, a_1' \in A_1$ such that:

1. Type θ plays a_1 with positive probability at h^t.

2. For every h^s with $h^t \succsim h^s \succsim h^0$, we have $\theta_p^*(\pi(h^s)) = \theta_p^*(\pi)$.

3. $$\theta_p^*\Big(\pi(h^t, a_1)\Big) \prec \theta_p^*(\pi) \precsim \theta_p^*\Big(\pi(h^t, a_1')\Big), \tag{E.1}$$

where (E.1) comes from the martingale property of beliefs.

In what follows, I consider two cases separately, depending on whether type θ plays a_1' with positive probability at h^t or not.

[31] Player 2's action choice in period 0 is irrelevant for the proof as it has negligible payoff consequences for player 1 as $\delta \to 1$.

1. If type θ plays a_1' with probability 0 at h^t, then $\left|\Theta(h^t, a_1')\right| \leq k-1$ and according to the induction hypothesis, for every $\widetilde{\theta}$ that is in the support of $\pi(h^t, a_1')$ but is no greater than $\theta_p^*(\pi(h^t, a_1'))$, type $\widetilde{\theta}$ will receive continuation payoff $u_1(\widetilde{\theta}, \underline{a}_1, \underline{a}_2)$ after playing a_1' at h^t. This implies that his continuation payoff is no more than $u_1(\widetilde{\theta}, \underline{a}_1, \underline{a}_2)$ at h^t. Therefore, playing $\underline{a}_1$ in every period is type $\theta_p^*(\pi(h^t, a_1'))$'s best reply at h^t. Since $\theta \precsim \theta_p^*(\pi) \precsim \theta_p^*\Big(\pi(h^t, a_1')\Big)$, Lemma C.1 implies that playing $\underline{a}_1$ in every period is also type θ's best reply at h^t, from which his payoff should be no more than $u_1(\theta, \underline{a}_1, \underline{a}_2)$. This leads to a contradiction.

2. The above argument suggests that for every $h^t \in \mathcal{H}^\sigma(\theta)$, there exists some $a_1'(h^t) \in A_1$ such that:

   (a) $\theta_p^*(\pi) \precsim \theta_p^*\Big(\pi(h^t, a_1'(h^t))\Big)$.

   (b) Type θ plays $a_1'(h^t)$ with positive probability at h^t.

   By construction, the strategy of playing $a_1'(h^t)$ at every h^t is type θ's best reply, from which his stage-game payoff in every period is no more than $u_1(\theta, \underline{a}_1, \underline{a}_2)$. This contradicts the hypothesis that type θ's equilibrium payoff is strictly greater than $u_1(\theta, \underline{a}_1, \underline{a}_2)$.

□

E.3 Part III

In this part, I establish Proposition 4.3 for types strictly greater than $\theta_p^*(\pi)$. Without loss of generality, I normalize $u_1(\theta, \underline{a}_1, \underline{a}_2)$ to 0 for every $\theta \in \Theta$. Let $x_\theta(a_1) \equiv -u_1(\theta, a_1, \underline{a}_2)$ and $y_\theta(a_1) \equiv u_1(\theta, a_1, \overline{a}_2)$. Assumptions 1 and 2 imply that:

1. $x_\theta(a_1) \geq 0$, with "$=$" holds only when $a_1 = \underline{a}_1$.

2. $y_\theta(a_1) > 0$ for every $\theta \in \Theta$ and $a_1 \in A_1$.

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

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

I start with defining '*pessimistic belief path*' for a given equilibrium $\sigma \in \text{NE}(\delta, \pi)$. 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

$$(a_{1,0}, ..., a_{1,t-1}, r^t) \in \mathcal{H}^\sigma \text{ and } \sum_{\theta \succsim \theta^*} \pi_t(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0,$$

where π_t is player 2's posterior belief after observing $(a_{1,0}, ..., a_{1,t-1})$. For every $\theta \in \Theta$, 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 playing $\underline{a}_1$ in every period is type θ's best reply.

For every $\theta \succ \theta_p^*(\pi)$ and every action path $a_1^\infty = (a_{1,0}, a_{1,1}, ...)$ that type θ plays with strictly positive probability under $\sigma \in \text{NE}(\delta, \pi)$, we have:

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

and

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

where $\mathcal{P}^{a_1^\infty}(a_1,a_2)$ is the occupation measure of (a_1,a_2) induced by a_1^∞ and player 2's equilibrium strategy σ_2. This implies that $V_\theta^\sigma(\delta)$ cannot exceed the value of the following linear program:

$$\max_{\{\beta(a_1),\gamma(a_1)\}_{a_1\in A_1}} \Big\{ \sum_{a_1\in A_1} \beta(a_1)y_\theta(a_1)-\gamma(a_1)x_\theta(a_1)\Big\}, \tag{E.2}$$

subject to

$$\sum_{a_1\in A_1} \gamma(a_1)+\beta(a_1)=1,$$

$$\gamma(a_1),\beta(a_1)\geq 0 \text{ for every } a_1\in A_1,$$

and

$$\sum_{a_1\in A_1} \beta(a_1)y_{\theta_p^*(\pi)}(a_1)-\gamma(a_1)x_{\theta_p^*(\pi)}(a_1)\leq 0. \tag{E.3}$$

Since the objective function and the constraints are both linear, it is without loss of generality to focus on solutions where there exist $a_1^*, a_1^{**}\in A_1$ such that

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

According to (E.3), we have:

$$\beta(a_1^*)y_{\theta_p^*(\pi)}(a_1^*)\leq \big(1-\beta(a_1^*)\big)x_{\theta_p^*(\pi)}(a_1^{**}). \tag{E.4}$$

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

$$\max_{a_1^*,a_1^{**}\in A_1}\Big\{\frac{y_\theta(a_1^*)x_{\theta_p^*(\pi)}(a_1^{**})-x_\theta(a_1^{**})y_{\theta_p^*(\pi)}(a_1^*)}{x_{\theta_p^*(\pi)}(a_1^{**})+y_{\theta_p^*(\pi)}(a_1^*)}\Big\}. \tag{E.5}$$

Expression (E.5) is maximized when $a_1^*=a_1^{**}=\overline{a}_1$, which gives the following upper bound for $V_\theta^\sigma(\delta)$:

$$V_\theta^\sigma(\delta)\leq r(\pi)u_1(\theta,\overline{a}_1,\overline{a}_2)+\big(1-r(\pi)\big)u_1(\theta,\overline{a}_1,\underline{a}_2), \tag{E.6}$$

with $r(\pi)\in(0,1)$ is pinned down via:

$$r(\pi)u_1\big(\theta_p^*(\pi),\overline{a}_1,\overline{a}_2\big)+\big(1-r(\pi)\big)u_1\big(\theta_p^*(\pi),\overline{a}_1,\underline{a}_2\big)=u_1\big(\theta_p^*(\pi),\underline{a}_1,\underline{a}_2\big). \tag{E.7}$$

The upper bound in the right-hand-side of (E.6) can be asymptotically achieved when $\delta\to 1$, as there exists an equilibrium such that:

- Type θ plays $\underline{a}_1$ in every period if $\theta\prec\theta_p^*(\pi)$, plays $\overline{a}_1$ in every period if $\theta\succ\theta_p^*(\pi)$.
- Type $\theta_p^*(\pi)$ randomizes between playing $\overline{a}_1$ in every period and playing $\underline{a}_1$ in every period with probabilities $q(\pi)$ and $1-q(\pi)$, respectively.

F Proof of Proposition 4.2

For every $\sigma\in\text{NE}(\delta,\pi)$, I define the set of *high histories* recursively. Let $\overline{\mathcal{H}}^0\equiv\{h^0\}$ and

$$\overline{a}_1(h^0)\equiv\max\Big\{\bigcup_{\theta\in\Theta}\text{supp}\Big(\sigma_\theta(h^0)\Big)\Big\}.$$

Let

$$\overline{\mathcal{H}}^1\equiv\{h^1|\text{ there exists } h^0\in\overline{\mathcal{H}}^0 \text{ such that } h^1\succ h^0 \text{ and } \overline{a}_1(h^0)\in h^1\}.$$

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

$$\overline{a}_1(h^t) \equiv \max\Big\{ \bigcup_{\theta \in \Theta(h^t)} \text{supp}\Big(\sigma_\theta(h^t)\Big)\Big\} \tag{F.1}$$

and

$$\overline{\mathcal{H}}^{t+1} \equiv \{h^{t+1} | \text{ there exists } h^t \in \overline{\mathcal{H}}^t \text{ such that } h^{t+1} \succ h^t \text{ and } \overline{a}_1(h^t) \in h^{t+1}\}. \tag{F.2}$$

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

1. For every $h^s \succsim h^t$ with $h^s \in \overline{\mathcal{H}}$, we have $\theta \in \Theta(h^s)$.
2. If $h^{t-1} \prec h^t$, then for every $\tilde{\theta} \in \Theta(h^{t-1})$, there exists $h^s \in \overline{\mathcal{H}}$ with $h^s \succ h^{t-1}$ such that $\tilde{\theta} \notin \Theta(h^s)$.

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

1. $\overline{\mathcal{H}}(\Theta) \subset \overline{\mathcal{H}}$.
2. For every $h^t, h^s \in \overline{\mathcal{H}}(\Theta)$, neither $h^t \succ h^s$ nor $h^t \prec h^s$.

In what follows, I show the following Lemma:

Lemma F.1. *For every $h^t \in \overline{\mathcal{H}}$, if $\theta = \max \Theta(h^t)$, then type θ's continuation payoff at h^t is no more than* $\max\{u_1(\theta, \overline{a}_1, \overline{a}_2), u_1(\theta, \underline{a}_1, \underline{a}_2)\}$.

Lemma F.1 implies the conclusion in Proposition 4.2 as $h^0 \in \overline{\mathcal{H}}$ and $\overline{\theta} = \max \Theta(h^0)$. A useful conclusion to show Lemma F.1 is the following observation:

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

PROOF OF LEMMA F.2: Since u_1 has SID in θ and (a_1, a_2), so for every $\theta \succ \widetilde{\theta}$,

$$(\overline{a}_1, \overline{a}_2) \in \arg \max_{(a_1, a_2)} \Big\{ u_1(\theta, a_1, a_2) - u_1(\widetilde{\theta}, a_1, a_2)\Big\} \tag{F.3}$$

which yields the upper bound on the difference between type θ and type $\widetilde{\theta}$'s continuation payoffs. $\square$

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

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

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

– $\theta \in \Theta(h^s)$ but $\theta \notin \Theta(h^{s+1})$ for every $h^{s+1} \succ h^s$ and $h^{s+1} \in \overline{\mathcal{H}}$.

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

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

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

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

1. For every $h^s \in \overline{\mathcal{H}}^B(h^t)$, at h^{s+1} with $h^{s+1} \succ h^s$ and $h^{s+1} \in \overline{\mathcal{H}}$, by definition,
$$|\Theta(h^{s+1})| \leq n.$$
Let $\widetilde{\theta} \equiv \max \Theta(h^{s+1})$. By induction hypothesis, type $\widetilde{\theta}$'s continuation payoff at h^{s+1} is at most $u_1(\widetilde{\theta}, \overline{a}_1, \overline{a}_2)$. This applies to every such h^{s+1}.

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

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

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

G Counterexamples

I present several counterexamples to complement the analysis in the main text. For future reference, I abuse notation and use θ to denote the Dirac measure on θ and a_i to denote the Dirac measure on a_i for $i \in \{1, 2\}$.

G.1 Counterexample when Supermodularity is Violated

Example 1: To begin with, I construct low-payoff equilibria in games where the supermodularity condition on u_1 is violated but all other assumptions are satisfied.[32] Let the stage game payoff be:

θ_1	O	I
F	$1, 0$	$-1, -1$
A	$2, 0$	$0, 1$

θ_0	O	I
F	$5/2, 0$	$1/2, 1/2$
A	$3, 0$	$1, 3/2$

This payoff matrix models the situation studied by Harrington (1986), in which an incumbent firm has persistent private information about the cost of operating in an industry and his cost is positively correlated with those of the potential entrants'.

To see how this game's payoff fails the supermodularity condition, let us rank the states and players' actions according to $\theta_1 \succ \theta_0$, $F \succ A$ and $O \succ I$. Player 1's cost of fighting is 1 in state θ_1 and is $1/2$ in state θ_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 higher, and the entrant's profit from entry decreases with the operating cost and increases with the incumbent's price.

When $\Omega = \{F, A\}$ and $\mu(F) \leq 2\mu(\theta_0)$, I construct low payoff equilibria in each of the following three cases, depending on the signs of:

$$X \equiv \frac{\mu(\theta_0)}{2} + \Big(\frac{\mu(F)\phi_F(\theta_0)}{2} - \mu(F)\phi_F(\theta_1)\Big) - \mu(\theta_1) \tag{G.1}$$

[32] The case in which u_2 has decreasing differences between a_2 and θ is similar once we reverse the order on the states.

and

$$Y \equiv \frac{\mu(F)\phi_F(\theta_0)}{2} - \mu(F)\phi_F(\theta_1). \tag{G.2}$$

1. If $X \leq 0$, then type θ_0 plays F in every period, type θ_1 mixes between playing F in every period and playing A in every period, with the probability of playing F equals to $1 + X/\mu(\theta_1)$. 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 $\frac{1}{2\delta}O + (1 - \frac{1}{2\delta})I$ otherwise. Despite the probability of type θ_1 is large relative to that of type θ_0, type θ_1's equilibrium payoff is 0 and type θ_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 θ_1 always plays A, type θ_0 mixes between playing F in every period and playing A in every period, 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 θ_1's equilibrium payoff is 0 and type θ_0's equilibrium payoff is 1.

3. If $X > 0$ and $Y > 0$, then both types play A in every period. Player 2 plays I no matter what. Type θ_1's equilibrium payoff is 0 and type θ_0's equilibrium payoff is 1.

Example 2: Next, I construct low-payoff equilibria in the entry deterrence game when the supermodularity condition on u_2 is violated. I focus on the case in which u_2 has *strictly decreasing differences* between a_2 and a_1.[33] Consider the following $2 \times 2 \times 2$ game with stage-game payoffs given by:

$\theta = \theta_1$	h	l
H	$1, -1$	$-1, 0$
L	$2, 1$	$0, 0$

$\theta = \theta_0$	h	l
H	$1 - \eta, -2$	$-1 - \eta, 0$
L	$2, -1$	$0, 0$

with $\eta \in (0, 1)$. The states and players' actions are ranked according to $H \succ L$, $h \succ l$ and $\theta_1 \succ \theta_0$. Let $\Omega = \{H, L\}$. Theorem 2 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 played in every period or (L, l) is played in every period, 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 θ_0 plays L and type θ_1 plays H. Player 2 plays h in period t (≥ 2) if and only if $t \geq T(\delta) + 1$ and player 1's past play coincides with type θ_1's equilibrium strategy. The integer $T(\delta)$ is chosen such that:

$$(1 - \delta^{T(\delta)})(-1) + 2\delta^{T(\delta)} > 0 > (1 - \delta^{T(\delta)})(-1 - \eta) + 2\delta^{T(\delta)}.$$

Such $T(\delta)$ exists when δ 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 in Theorems 2 and 3. For this purpose, I consider two counterexamples in which Assumption 1 is violated in different ways.

[33] The case in which u_2 has decreasing differences between a_2 and θ is similar to the previous example. One only needs to reverse the order between the states.

Example 1: Consider the following $2 \times 2 \times 2$ game:

$\theta = \theta_1$	h	l
H	$3/2, 2$	$0, 0$
L	$1, 1$	$0, 0$

$\theta = \theta_0$	h	l
H	$-1, -1/2$	$1, 0$
L	$0, -1$	$5/2, 1/4$

One can verify that this game satisfies the supermodularity assumption once we rank the states and actions according to $\theta_1 \succ \theta_0$, $H \succ L$ and $h \succ l$.[34] However, the monotonicity condition fails as player 1's ordinal preferences over a_1 and a_2 depend on the state.

Suppose $\Omega = \{H, L\}$. The total probability of commitment types is small enough such that $4\mu(H) < \mu(\theta_0)$ and $\frac{5}{4}\mu(L) < \mu(\theta_1)$. The correlations between the commitment action and the state are irrelevant for this example. Consider the following equilibrium in which player 2 plays a '*tit-for-tat*' like strategy. Type θ_1 plays L in every period on the equilibrium path and type θ_0 plays H in every period on the equilibrium path. 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 δ is high enough, notice that first, player 2's incentive constraints are always satisfied. As for player 1,

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. Therefore, he has no incentive to deviate as long as $\delta > 1/2$.

2. If $\theta = \theta_0$, 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 $t \in \mathbb{N}$,
$$(1-\delta)\frac{5}{2} - (\delta^t - \delta^{t+1}) \leq 1 - \delta^{t+1}. \tag{G.3}$$
which is equivalent to:
$$\frac{5}{2} \leq 1 + \delta + ... + \delta^{t-1} + 2\delta^t.$$
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 where type θ_1 plays H in every period, type θ_0 plays L in every period. 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: The conclusions in Theorems 2 and 3 will also fail 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 repeated version of the following stage-game:

$\theta = \theta_1$	h	l
H	$2, 2$	$0, 0$
L	$1, 1$	$-1/2, 0$

$\theta = \theta_0$	h	l
H	$1/4, -1/2$	$1/8, 0$
L	$0, -1$	$-1/16, 1/4$

[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 (θ, a_1, a_2). The definition of supermodular function can be found in Topkis (1998).

Both players' payoffs are strict supermodular functions of (θ, 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 $\Omega = \{H, L\}$, $4\mu(H) < \mu(\theta_0)$ and $\frac{5}{4}\mu(L) < \mu(\theta_1)$. The following strategy profile is an equilibrium. Type θ_1 plays L in every period and type θ_0 plays H in every period. 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 θ_1 and type θ_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. Verifying players' incentive constraints follows the same steps as in the previous example, which is omitted.

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.

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

This is because aside from the equilibrium constructed above, there also exists an equilibrium in which type θ_1 plays H in every period, type θ_0 mixes between playing H in every period and playing L in every period. The mixture probabilities are chosen such that player 2 becomes indifferent between h and l starting from period 1 conditional on H having 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:

$\theta = \theta_1$	l	m	r
H	$0, 0$	$5/2, 2$	$6, 3$
L	$\epsilon, 0$	$5/2 + \epsilon, -1$	$6 + \epsilon, -2$

$\theta = \theta_0$	l	m	r
H	$0, 0$	$2, -1$	$3, -2$
L	$2\epsilon, 0$	$2 + 2\epsilon, -2$	$3 + 2\epsilon, -3$

where $\epsilon > 0$ is small enough. Let the rankings on actions and states be $H \succ L$, $r \succ m \succ l$ and $\theta_1 \succ \theta_0$. One can check that the stage game payoffs are MSM.

Suppose $\Omega = \{H, L\}$ with $\mu(\theta_0) = 2\eta$, $\mu(H) = \eta$ and $\phi_H(\theta_1) = 1$, with $\eta \in (0, 1/3)$. Type θ_1's commitment payoff from playing H is 6. However, consider the following equilibrium:

– Type θ_0 plays H in every period. Type θ_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 θ_1's equilibrium strategy. She plays l at every off-path history.

Type θ_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 the strategic 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 Inconsistent Equilibrium Play in Private Value Reputation Games

I construct an equilibrium in the private value product choice game (Mailath and Samuelson 2001,2006) such that despite there exists a commitment type that exerts high effort in every period, the strategic long-run player abandons his reputation early on in the relationship and the frequency with which L is played does not vanish as $\delta \to 1$. The construction can be generalized to other private value reputation games such as the entry deterrence game (Kreps and Wilson 1982, Milgrom and Roberts 1982), capital taxation game (Phelan 2006), etc. Players' stage-game payoffs are given by:

$-$	C	S
H	$1,3$	$-1,2$
L	$2,0$	$0,1$

Suppose $H \in \Omega$ and $\mu(H)$ is small enough. Consider the following strategy profile, which is an 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 $t-1$.

Intuitively, starting from period 1, every time player 1 shirks, he will be punished in the next period as player 2 will play C with smaller 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. It is straightforward to verify that

1. The strategic long-run player's equilibrium payoff is δ, which is arbitrarily close to his pure Stackelberg commitment payoff 1 as $\delta \to 1$.
2. The strategic long-run player's equilibrium play is very different from that of the commitment type's. In particular, (i) imitating the commitment type is a strictly dominated strategy, which yields payoff $\delta - (1-\delta)$; (ii) the occupation measure of L equals to $1/2$ as $\delta \to 1$.

G.5 Low Probability of Commitment Type for Behavioral Uniqueness

The following example illustrates why $\mu(\Omega)$ being small is not redundant to obtain sharp predictions on player 1's on-path behavior (Theorems 3 and 3'). Consider the following $2 \times 3 \times 2$ stage game:

$\theta = \theta_1$	C	S
H	$1,2$	$-2,0$
M	$2,1$	$-1,0$
L	$3,-1$	$0,0$

$\theta = \theta_0$	C	S
H	$1/2,-1$	$-5/2,0$
M	$3/2,-2$	$-3/2,0$
L	$3,-3$	$0,0$

Let $\Omega \equiv \{H, M, L\}$ with $\mu(H) = \mu(\theta_1) = 1/98$, $\mu(\theta_0) = 3/49$, $\mu(M) = 6/49$ and $\mu(L) = 39/49$. Let $\phi_H = \phi_M$ be the Dirac measure on θ_1 and let ϕ_L be the Dirac measure on θ_0. One can check that $M \in \Omega^g$ and μ satisfies (4.3). However, for every $\delta > 5/6$, one can construct the following class of equilibria indexed by $T \in \{1, 2, ...\}$:

- **Equilibrium σ^T:** Type θ_0 plays M in every period. Type θ_1 plays M from period 0 to period T, and plays H starting from period $T+1$. Player 2 plays S in period 0. From period 1 to $T+1$, she plays C with probability 1 if player 1 has played H in every period or player 1 has played M in every period. From period $T+2$ and onwards, she plays C with probability 1 if player 1 has played H in every period, and plays a mixed action $\frac{3\delta-1}{3\delta}C + \frac{1}{3\delta}S$ if player 1 has played M in every period. At all other histories, player 2 plays S with probability 1.

One can verify players' incentive constraints. In particular in period $T+1$ conditional on player 1 has always played M in the past, type θ_1 is indifferent between playing H and M while type θ_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 θ_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

In this subsection, I show that the patient long-run player has multiple possible on-path behaviors when the stage-game payoffs are monotone-supermodular and the prior belief about the state is optimistic. In particular, I construct an equilibrium in the repeated product choice game with MSM payoffs and an optimistic prior belief such that at some on-path histories, player 2's belief about the state deteriorates after observing player 1 playing his highest action. One can also verify that the constructed strategy profile is also part of a sequential equilibrium under its induced belief system.

In this example, players' stage-game payoffs are given by:

$\theta=\theta_1$	h	l
H	$1,3$	$-1,2$
L	$2,0$	$0,1$

$\theta=\theta_0$	h	l
H	$1-\eta,0$	$-1-\eta,1$
L	$2,-2$	$0,0$

with $\eta \in (0,1)$. Let $\Omega \equiv \{H,L\}$ with $\mu(H)=0.06(1-\epsilon)$, $\mu(\theta_0)=0.04(1-\epsilon)$, $\mu(\theta_1)=0.9(1-\epsilon)$, $\mu(L)=\epsilon$ and ϕ_H is the Dirac measure on θ_1. I assume that $\epsilon>0$ is small enough. Consider the following strategy profile:

- In period 0, type θ_1 plays H with probability $2/45$ and type θ_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.
- 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) in every future period on the equilibrium path.
  - With probability $1-\frac{1-\delta}{\delta^2}-\frac{1-\delta}{\delta}$, play (H,h) in every future period 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 θ_0 plays L for sure and type θ_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.

[35]Notice that under a generic prior belief, type θ_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.

In period 0, player 2's belief about θ deteriorates after observing H. This is true no matter whether we only count the strategic types (as strategic type θ_0 plays H with strictly higher probability) or also count the commitment type (probability of θ_1 decreases from $24/25$ to $10/11$).

The existence of the above equilibrium also shows that the long-run player's on-path behavior is not unique, as there exists another equilibrium where both strategic types play H in every period and player 2 plays h if and only if L has never been played before and plays l otherwise. One can also construct an infinite sequence of equilibria indexed by $t \in \mathbb{N}$ such that both strategic types play H in every period aside from period t. Player 2 plays l in period t. She plays h in period $s \neq t$ if and only if L has not been played in any periods other than t, and plays l otherwise.

G.7 Multiple Equilibrium Behaviors when Player 1 is Impatient

I present an example in which the game's payoff satisfies Assumptions 1, 2 and 3, player 2's prior belief is pessimistic but player 1 has multiple equilibrium behaviors since δ is not high enough. Consider the following product choice game:

$\theta = \theta_1$	C	S
H	$1, 3$	$-1, 2$
L	$2, 0$	$0, 1$

$\theta = \theta_0$	C	S
H	$1-\eta, 0$	$-1-\eta, 1$
L	$2, -2$	$0, 0$

with $\eta \in (0, 1)$, $\Omega \equiv \{H, L\}$, ϕ_H be the Dirac measure on θ_1 and ϕ_L is irrelevant for the gist of this example. Player 2's prior is given by:

$$\mu(\theta_0) = 0.7, \mu(\theta_1) + \mu(H) = 0.3 \text{ with } \mu(H) \in (0, 0.1).$$

One can verify the condition on pessimistic prior belief and the total probability of commitment type being small, namely (4.7), is satisfied. 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.

- Type θ_0 plays L in every period.
- Type θ_1 plays H in every period besides period $t \in \{1, 2, ...\}$, 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 θ_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.

G.8 Why $\lambda \in \Lambda(\alpha_1^*, \theta)$ is not sufficient when α_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 α_1^* is mixed. Players' payoffs are given by:

θ_1	l	m	r
H	$1, 3$	$0, 0$	$0, 0$
L	$2, -1$	$0, 0$	$0, 0$
D	$3, -1$	$1/2, 0$	$1/2, 0$

θ_2	l	m	r
H	$0, 1/2$	$0, 3/2$	$0, 0$
L	$0, 1/2$	$0, 3/2$	$0, 0$
D	$0, 0$	$0, 0$	$0, 0$

θ_3	l	m	r
H	$0, 1/2$	$0, 0$	$0, 3/2$
L	$0, 1/2$	$0, 0$	$0, 3/2$
D	$0, 0$	$0, 0$	$0, 0$

Suppose $\Omega = \{\alpha_1^*, D\}$ with $\alpha_1^* \equiv \frac{1}{2}H + \frac{1}{2}L$ and $\phi_{\alpha_1^*}$ is the Dirac measure on θ_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 $\lambda = (1/2, 1/2) \in \Lambda(\alpha_1^*, \theta_1)$. When δ is large enough, the following strategy profile constitutes an equilibrium in which type θ_1's payoff is $1/2$ in the $\delta \to 1$ limit.

– Type θ_1 plays D in every period.

– In period 0, type θ_2 plays H and type θ_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.

– Starting from period 1, if she observes H or D in period 0, then she plays m in every subsequent period. If she observes L in period 0, then she plays r in every subsequent period.

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 θ_1 an incentive to play D in every subsequent period, making player 2's belief self-fulfilling. This situation only arises when α_1^* is mixed and $k(\alpha_1^*, \theta) \geq 2$.