Reputation Effects under Interdependent Values

Harry PEI*

October 7, 2019

Abstract: A patient player privately observes a persistent state, and interacts with a sequence of myopic uninformed players. The patient player is either a strategic type who maximizes his payoff, or one of several commitment types that mechanically plays the same action in every period. I focus on situations in which the uninformed player's best reply to a commitment action depends on the state, and the total probability of commitment types is sufficiently small. I show that the patient player's equilibrium payoff is bounded below his commitment payoff in some equilibria under some of his payoff functions. This is because he faces a tradeoff between building his reputation for commitment and signaling favorable information about the state. When players' stage-game payoff functions are *monotone-supermodular*, the patient player receives high payoffs in all states and in all equilibria. Under an additional condition on the state distribution, my reputation model yields a unique prediction on the patient player's equilibrium payoff and on-path behavior.

Keywords: reputation, interdependent values, commitment payoff, robust behavioral prediction

1 Introduction

Economists have long recognized that reputation lends credibility to agents' 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. These reputation results are *robust* as they apply across all equilibria,[1] 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 exclude settings in which reputation-building agents' private information directly affects their opponents' payoffs. For example, in markets for electronics and custom softwares, sellers benefit from reputations for providing good customer service, and they also have private information about the quality of their products. Their incentives to sustain reputations for providing good service interact dynamically with their motives to signal high product quality, which introduce new economic

*Northwestern University. I am indebted to Daron Acemoglu, Drew Fudenberg, Juuso Toikka, and Alex Wolitzky for guidance, and Jie Bai, Heski Bar-Isaac, Martin Cripps, Dan Clark, Joyee Deb, Miaomiao Dong, Esther Duflo, Mehmet Ekmekci, Yingni Guo, Yuhta Ishii, Qingmin Liu, Shuo Liu, Alessandro Pavan, Joel Sobel, Bruno Strulovici, Nicolas Vieille, and four referees for helpful comments.

[1] Throughout this paper, I say that a property is *robust* if and only if it is true for all equilibria of a given model.

forces that are omitted by existing reputation models. In addition, existing reputation results deliver robust predictions on players' equilibrium payoffs, but yield no robust prediction on their equilibrium *behaviors*.

This paper studies the effects of *interdependent values* on reputation-building players' payoffs and behaviors. In my model, a patient player 1 (e.g., seller) interacts with an infinite sequence of myopic player 2s (e.g., buyers), arriving one in each period and each plays the game only once. Different from existing reputation models, player 1 privately observes the realization of a payoff-relevant state (e.g., product quality) that is constant over time and affects both players' stage-game payoffs, in addition to knowing whether he is strategic or committed. If player 1 is strategic, then he maximizes his discounted average payoff. If player 1 is committed, then he mechanically plays the same action (i.e., a *commitment action*) in every period, which can depend on the state. Player 2s observe all the actions taken in the past, but not their predecessors' payoffs.

My results focus on properties that apply to *all* equilibria. I make three conceptual contributions. First, I identify conflicts between reputation building and signaling under interdependent values. Second, I provide sufficient conditions under which the patient player guarantees high returns from building reputations despite facing such conflicts. Third, I show that interdependent values can generate a *disciplinary effect*, which motivates the patient player to sustain his reputation, and leads to a unique prediction on his on-path behavior.

Theorem 1 shows that for every state, commitment action, and payoff function of player 2's, if player 2's best reply to this commitment action depends on the state, and the probability of commitment types is sufficiently small, then there *exists* a payoff function for player 1 and an equilibrium, in which his equilibrium payoff in the chosen state is strictly bounded below his complete information commitment payoff. This contrasts to Fudenberg and Levine (1989)'s result for private-value models, which says that player 1 receives at least his commitment payoff, regardless of his own payoff function and the probability of commitment types.

Intuitively, in order to motivate player 2s to play their best reply against the commitment action in the chosen state, player 1 needs to both convince them that this commitment action will be played in the future, and to signal them the correct information about the state. These two objectives are in conflict when player 2s believe that player 1 is more likely to play the commitment action in another state, under which her best reply against the commitment action is different. This captures for example, consumers becoming more skeptical about product quality after observing the seller provides good service.[2] When facing this conflict, player 1 either abandons his reputation, after which he loses the credibility for playing his commitment action in the future; or he signals negative information about the state, after which player 2s do not have incentives to choose player 1's desired best reply even when they are convinced that player 1 will play his commitment action.

Next, I restrict attention to games with *monotone-supermodular* (or *MS*) payoffs. My MS condition requires

[2]This effect is shown to be practically relevant by Miklos-Thal and Zhang (2013) via case studies and lab experiments. They find that sellers conduct demarketing activities, which includes providing limited services, in order to signal the quality of their products.

that states and actions be ranked such that (1) player 1's payoff is strictly decreasing in his own action and is strictly increasing in player 2's action; (2) action profiles and states are complements in both players' payoff functions. For example, the following product choice game between a firm and its customers satisfies MS when states and actions are ranked according to high state $\succ$ low state, good action $\succ$ bad action, trust $\succ$ not trust.

high product quality	trust	not trust
good action	$\mathbf{1}, 1$	$-1, 0$
bad action	$2, -1$	$0, 0$

low product quality	trust	not trust
good action	$\mathbf{2/3}, -1$	$-2, 0$
bad action	$2, -2$	$0, 0$

I establish robust predictions on player 1's payoff and behavior when he can build a reputation for playing his *highest action*. Theorems 2 and 3 consider two cases separately, depending on player 2's *prior belief about the state*. To yield clear comparisons with Theorem 1, these results allow commitment types to be arbitrarily rare and player 2's best reply against player 1's highest action to depend on the state.

When player 2's prior belief about the state is such that her highest action best replies against player 1's highest action,[3] Theorem 2 shows that in every state and in every equilibrium, a patient player 1 receives at least his payoff from the *highest action profile*, which is no less than his commitment payoff from playing his highest action. In the above example, it implies that when the probability of high state exceeds $1/2$, a high-quality firm secures payoff 1 and a low-quality firm secures payoff $2/3$ by playing the good action in every period.

The difference between Theorems 1 and 2 is driven by the MS condition on stage-game payoffs, which is assumed in the latter but not in the former. Intuitively, MS implies that player 1 has a stronger preference towards higher action profiles in higher states. When this stage-game is played *repeatedly*, despite MS *not ruling out* the aforementioned conflict between reputation-building and signaling,[4] it rules out situations in which player 1 *plays his highest action in every period* in a lower state, but not in a higher state. This leads to a *uniform lower bound* on player 2's posterior belief about the state, which applies to all histories such that player 1 has played his highest action in all previous periods. In the example, it implies that player 2's posterior attaches probability more than $1/2$ to the high-quality state as long as player 1 has played the good action in every period. This belief lower bound implies patient player 1's payoff lower bound, since player 2 has a strict incentive to choose player 1's desired best reply once she is convinced that player 1 will play his highest action.

Under the complementary condition on player 2's prior belief about the state, Theorem 3 shows that player 1 has a *unique equilibrium payoff* and a *unique on-path behavior*. Player 1's unique payoff is strictly lower than his payoff from the highest action profile, but is greater than his minmax payoff. Player 1's unique on-path

[3]This condition is equivalent to the *existence* of an equilibrium in a repeated game with the same state distribution but without commitment types, such that player 1 attains at least his payoff from the highest action profile in all states (Online Appendix E). The main text focuses on MS games in which player 2's action choice is binary. Generalizations can be found in Online Appendix G.

[4]This is because player 1's action not only signals the persistent state, but also affects the continuation equilibrium being played through repeated game effects. See Online Appendix H.1 for a $2 \times 2 \times 2$ example in which the high action signals the low state.

behavior is characterized by a *cutoff state* (in the example, low-quality state), such that he plays his highest action in every period when the state is above this cutoff, plays his lowest action in every period when the state is below this cutoff. At the cutoff state, he randomizes between playing his highest action in every period and playing his lowest action in every period. Player 1's mixing probability at the cutoff state is such that under player 2's posterior belief about the state after observing player 1's highest action, she is indifferent between her highest action and her lowest action (against player 1's highest action).

The unique behavioral prediction in Theorem 3 contrasts to the private-value reputation game in Fudenberg and Levine (1989) and the case studied by Theorem 2, in which there are multiple equilibria with different on-path behaviors. This result is driven by a novel *disciplinary effect*, namely, the strategic long-run player is guaranteed to receive a high payoff when pooling with commitment type, and is guaranteed to receive a low payoff after separating from commitment type.

I illustrate this effect using the product choice game. First, the disciplinary effect is absent in the private-value benchmark where product quality is known to be high. This is because after separating from the commitment type (that plays the good action), a patient strategic firm's continuation payoff can be anything between his minmax payoff 0 and his (pure) Stackelberg payoff 1. This multiplicity in continuation values leads to multiple on-path behaviors. For example, if at a given history, the strategic firm can still receive its Stackelberg payoff after separating from the commitment type, then it has a strict incentive to play the bad action at that history; if the firm can only receive its minmax payoff after separation, then it has a strict incentive to play the good action. A similar intuition applies when the probability of high state is above $1/2$. This is because after separating from the commitment type, there exist equilibria in which the firm receives high payoffs in all states, and equilibria in which it receives minmax payoffs in all states.

Next, consider the case studied by Theorem 3, which translates into the probability of high state being lower than $1/2$. Different from the private-value benchmark and games studied by Theorem 2, playing the bad action signals low quality *in all equilibria*, after which the strategic firm's continuation payoff equals his minmax payoff. To understand why, suppose towards a contradiction that playing the bad action signals high quality. Since belief is a martingale, playing the good action signals low quality, after which the buyer's posterior belief about quality is more pessimistic compared to her prior. As a result, there exists at least one state in the support of this posterior in which the strategic seller's continuation payoff equals his minmax payoff. Since playing the good action is costly, the seller has a strict incentive to deviate to the bad action in this state, which strictly increases his stage-game payoff while not lowering his continuation payoff. This leads to a contradiction.

In terms of contributions, my model unifies the commitment type approach in Fudenberg and Levine (1989) with the dynamic signaling approach in Bar-Isaac (2003), Kaya (2009), and Lee and Liu (2013). My results deliver novel insights on seller reputations once applied to the product choice game. For example, Theorem 1 implies that conflicts between building reputations for playing good actions and signaling high product quality can have persistent negative effects on sellers' profits. In addition, it also provides an explanation for why patient sellers may refrain from building reputations in markets with adverse selection on product quality. Theorem 2 implies that in markets where adverse selection on product quality is an important concern but stage-game payoffs are MS, firms can secure high profits in the long run by establishing reputations for taking consumer-friendly actions, despite it may occasionally trigger negative inferences about product quality.

Theorem 3 implies that when consumers are pessimistic about product quality, firms behave consummately in every period, despite receiving low payoffs for a long time. This provides an explanation to the findings in Bai, Gazze and Wang (2019), that *after* the 2008 scandal on Chinese dairy industry,[5] nearly all Chinese dairy firms experienced a significant revenue decline for at least five years, and surviving firms frequently conducted quality inspections in order to rectify the situation. Theorem 3 suggests a rationale for this combination of *persistent low profits* and *consistent reputation-building behaviors*. In particular, aside from firm's effort in conducting inspections, consumers' payoffs also depend on a persistent state, interpreted as attributes that affect consumers' health in the long run but are omitted by formal inspections. The scandal made consumers' beliefs about product safety pessimistic. This not only has persistent negative effects on firms' revenues, but also motivates firms to conduct inspections in order to avoid further negative inferences about product safety.

Theorems 2 and 3 can also be viewed as equilibrium refinements for repeated incomplete information games. Theorem 2 suggests that introducing reputational types can rule out low-payoff equilibria. Theorem 3 advances this line of research one-step further by delivering a unique prediction on a patient player's on-path behavior. This contrasts to dynamic signaling games without commitment types and private-value reputation games, both of which have multiple predictions on the informed player's on-path behavior. Since researchers are more likely to observe agents' behaviors rather than their payoffs, my behavioral uniqueness result brings us closer to empirically testing reputation models. It also provides a better explanation to certain behavioral patterns observed in practice, such as firms' reactions after scandals on product safety.

2 Baseline Model

Time is discrete, indexed by $t = 0, 1, 2....$ A long-lived player 1 (he) with discount factor $\delta \in (0, 1)$ interacts with an infinite sequence of short-lived player 2s (she), arriving one in each period and each plays the game only once. In period t, players simultaneously choose their actions $(a_{1,t}, a_{2,t}) \in A_1 \times A_2$.

Player 1 has perfectly persistent private information about a payoff-relevant state $\theta \in \Theta$, and whether he is

[5] The scandal broke out in July 2008 after 16 babies were diagnosed with kidney stones. These incidents were later related to the infant formula produced by Chinese firm Sanlu. The World Health Organization referred to it as *one of the largest food safety events it has to deal with in recent years*, and that *the crisis of confidence among Chinese consumers would be hard to overcome*.

strategic or *committed*. If player 1 is strategic, then he can flexibly choose his actions. If player 1 is committed, then he mechanically follows one of the several *commitment plans*. A typical commitment plan is denoted by $\gamma : \Theta \to A_1$, according to which the committed long-run player plays $\gamma(\theta)$ in every period when the state is θ. Let Γ be the set of possible commitment plans, and let γ^* stand for player 1 being strategic. Let

$$\mu \in \Delta\Big(\Theta \times \underbrace{(\{\gamma^*\} \cup \Gamma)}_{\text{player 1's characteristics}}\Big) \tag{2.1}$$

be player 2's prior belief about player 1's private information, which is a joint distribution of θ and player 1's *characteristics*, namely, whether he is strategic or committed, and if he is committed, which plan in Γ is he following. Let $\phi \in \Delta(\Theta)$ be the marginal distribution of μ on Θ, namely, the *prior distribution of states*. Let

$$A_1^* \equiv \{a_1 \in A_1 | \text{ there exist } \gamma \in \Gamma \text{ and } \theta \in \Theta \text{ such that } \gamma(\theta) = a_1\}, \tag{2.2}$$

be the set of *commitment actions*. Intuitively, an action a_1^* belongs to A_1^* if and only if it is played in some state under some commitment plan. For every $\theta \in \Theta$, player 1 is *strategic type* θ if he is strategic and knows that the state is θ. For every $a_1^* \in A_1^*$, player 1 is *commitment type* a_1^* if he is committed and plays a_1^* in every period.

Let $h^t \equiv (a_{1,s}, a_{2,s})_{s=0}^{t-1} \in \mathcal{H}^t$ be a public history. Let $\mathcal{H} \equiv \bigcup_{t=0}^{+\infty} \mathcal{H}^t$ be the set of public histories. Player 1's private history consists of the public history and his type. Player 2's private history coincides with the public history, which means that she cannot observe her predecessors' payoffs.[6] For every $\theta \in \Theta$, let $\sigma_\theta : \mathcal{H} \to \Delta(A_1)$ be strategic type θ's strategy. Let $\sigma_2 : \mathcal{H} \to \Delta(A_2)$ be player 2's strategy.

For $i \in \{1, 2\}$, player i's stage-game payoff in period t is $u_i(\theta, a_{1,t}, a_{2,t})$. This formulation allows for interdependent values since u_2 depends on θ. The solution concept is Bayesian Nash equilibrium (or *equilibrium* for short), which is a strategy profile $\sigma \equiv \big((\sigma_\theta)_{\theta\in\Theta}, \sigma_2\big)$ such that for every $\theta \in \Theta$, σ_θ maximizes the expected value of $\sum_{t=0}^{\infty}(1-\delta)\delta^t u_1(\theta, a_{1,t}, a_{2,t})$; and σ_2 maximizes player 2's stage-game payoff.

I assume that Θ, Γ, A_1, and A_2 are finite sets, $|A_1|, |A_2| \geq 2$, and μ has full support. This implies that an equilibrium exists and every type occurs with strictly positive probability. To simplify the exposition, I assume that the distribution of θ and the distribution of player 1's characteristics are *independent* according to μ.

2.1 Example: Product Choice Game

I introduce a variant of the product choice game in Mailath and Samuelson (2013) to highlight the key ingredients of my model. I use this example to discuss the implications of my results in section 3.3.

[6] This feature of my model makes reputation building challenging since player 2s can only learn the state from player 1's actions. This assumption is important for Theorem 1. It also highlights the novelty of Theorems 2 and 3.

Player 1 is a firm and player 2s are consumers. The firm privately observes $\theta \in \Theta \equiv \{\theta_h, \theta_l\}$, which is constant over time and is interpreted as the quality/safety/durability of its products. The firm chooses between a good action G and a bad action B. The good action is interpreted as good customer service, on-time deliveries, conducting inspections, or other efforts of the firm that can benefit consumers. Each consumer perfectly observes the firm's past actions, and chooses between a *trusting action* T (e.g., purchase a customized version) and a *non-trusting action* N (e.g., purchase a standardized version). Payoffs are given by:

$\theta = \theta_h$	T	N
G	$1, 1$	$-1, 0$
B	$2, -1$	$0, 0$

$\theta = \theta_l$	T	N
G	$1 - \eta, -1$	$-1 - \eta, 0$
B	$2, -2$	$0, 0$

in which $\eta \in (-1, 1)$ is a parameter. According to the payoff matrices, when the firm's quality is high, its cost of taking the good action is 1, and a consumer has an incentive to trust when the good action is played with probability more than $1/2$. When the firm's quality is low, its cost of taking the good action is $1 + \eta$, and a consumer has no incentive to trust regardless of the firm's action.

The firm is either *strategic* or *committed*. Suppose for simplicity that there is only one commitment plan γ, according to which the firm plays his pure Stackelberg action in every state:

$$\gamma(\theta) \equiv \begin{cases} G & \text{if } \theta = \theta_h \\ B & \text{if } \theta = \theta_l. \end{cases} \tag{2.3}$$

The set of commitment actions is $A_1^* = \{G, B\}$. The firm has four types: two strategic types, θ_h and θ_l; and two commitment types, G and B. The consumers' prior belief μ is a joint distribution of the state and whether the firm is committed or strategic, from which one can derive the prior state distribution $\phi \in \Delta\{\theta_h, \theta_l\}$.

3 Main Results

My main results focus on properties of a patient player 1's payoff and behavior that apply to *all* equilibria. To state the results, I introduce notation for player 2's best reply, player 1's commitment payoff, and his lowest equilibrium payoff. For every $a_1 \in A_1$, $\theta \in \Theta$, and u_2, let

$$\mathrm{BR}_2(\theta, a_1 | u_2) \equiv \arg\max_{a_2 \in A_2} u_2(\theta, a_1, a_2). \tag{3.1}$$

Given $\theta \in \Theta$ and commitment action $a_1^* \in A_1^*$, let

$$v_\theta(a_1^*, u_1, u_2) \equiv \min_{a_2 \in \mathrm{BR}_2(\theta, a_1^* | u_2)} u_1(\theta, a_1^*, a_2), \tag{3.2}$$

be type θ's *commitment payoff* from a_1^*. Let $\underline{v}_\theta(\delta, \mu, u_1, u_2)$ be type θ's *lowest equilibrium payoff* under parameter values (δ, μ, u_1, u_2). I make the following assumption that is satisfied for generic u_2:

Assumption 1. *For every $\theta \in \Theta$ and $a_1 \in A_1$, $BR_2(\theta, a_1|u_2)$ is a singleton.*

3.1 Reputation Failure in Games with Unrestricted Payoffs

Theorem 1 shows that if commitment types are rare and player 2's best reply against a commitment action depends on the state, then there *exists* u_1 under which patient player 1's lowest equilibrium payoff is strictly bounded below his complete information commitment payoff. Given u_2 and $a_1^* \in A_1^*$, I say that interdependent values are *nontrivial* under (u_2, a_1^*) if there exist θ' and $\theta'' \in \Theta$ such that:

$$\text{BR}_2(\theta', a_1^*|u_2) \bigcap \text{BR}_2(\theta'', a_1^*|u_2) = \{\varnothing\}. \tag{3.3}$$

Theorem 1. *For every u_2 that satisfies Assumption 1, $a_1^* \in A_1^*$, $\theta \in \Theta$, and full support $\phi \in \Delta(\Theta)$. If interdependent values are nontrivial under (u_2, a_1^*), then there exist u_1 and $\overline{\varepsilon} > 0$, such that for every prior belief μ which has state distribution ϕ and attaches probability less than $\overline{\varepsilon}$ to all commitment types, we have:*

$$\limsup_{\delta \to 1} \underline{v}_\theta(\delta, \mu, u_1, u_2) < v_\theta(a_1^*, u_1, u_2). \tag{3.4}$$

The proof is in Appendix B. Theorem 1 contrasts to the private-value reputation result in Fudenberg and Levine (1989), which implies that if $a_1^* \in A_1^*$ and player 2's best reply against a_1^* does not depend on the state, then $\liminf_{\delta \to 1} \underline{v}_\theta(\delta, \mu, u_1, u_2) \geq v_\theta(a_1^*, u_1, u_2)$ for every u_1, θ, and full support μ.

Intuitively, player 1's action not only shows his propensity to play a_1^* in the future, but also signals the persistent state. When interdependent values are nontrivial, player 2's belief about state affects her best reply against a_1^*. In order to motivate player 2s to play the action in $\text{BR}_2(\theta, a_1^*|u_2)$, player 1 needs to convince them that a_1^* will be played with high probability, and the state *does not* belong to the following subset:

$$\Big\{\theta' \in \Theta \Big| \text{BR}_2(\theta', a_1^*|u_2) \neq \text{BR}_2(\theta, a_1^*|u_2)\Big\}, \tag{3.5}$$

which is the set of states such that player 2's best reply against a_1^* differs from that under state θ. A conflict between these two objectives arises when player 2s believe that strategic types outside (3.5) separate from commitment type a_1^*, and strategic types in (3.5) play a_1^* in every period. Under this belief, player 1 cannot pool with commitment type a_1^* while separating away from strategic types in (3.5). Theorem 1 confirms that under some u_1, the above belief arises in equilibrium and negatively affects the patient player's payoff.

The above argument also explains why Theorem 1 applies to *every* full support state distribution ϕ, regardless of the probability it attaches to the states in (3.5). This is because under the aforementioned self-fulfilling belief, player 2 has no incentive to play the action in $\mathrm{BR}_2(\theta, a_1^*|u_2)$ as long as the probability of commitment type a_1^* is small relative to that of strategic types who knew that the state belongs to (3.5).

3.2 Reputation Results in Monotone-Supermodular Games

Motivated by Theorem 1, Theorems 2 and 3 derive robust predictions on player 1's payoff and behavior when u_1 and u_2 satisfy a *monotone-supermodularity* condition (or MS). To highlight the comparisons with Theorem 1, these results allow interdependent values to be nontrivial, and commitment types to be arbitrarily rare.

Assumption 2 (MS). *There exist an order on Θ, an order on A_1 and an order on A_2, under which:*

1. *$u_1(\theta, a_1, a_2)$ is strictly decreasing in a_1, and is strictly increasing in a_2.*
2. *$u_1(\theta, a_1, a_2)$ has strictly increasing differences in θ and (a_1, a_2).*
3. *$u_2(\theta, a_1, a_2)$ has strictly increasing differences in a_2 and (θ, a_1).*[7]

MS has three economic implications. First, player 1 faces a *lack-of-commitment* problem, namely, higher a_1 can induce higher a_2, but higher a_1 is costly. This is satisfied in the product choice game (section 2.1) when firm's actions are ranked according to $G \succ B$ and consumers' actions are ranked according to $T \succ N$. It means that providing good customer service is costly for the firm, the firm benefits from consumers' trust, and consumers have stronger incentives to trust when they expect the firm to provide good service. Second, player 1 wants to signal that the state is high regardless of the true state. This is the case in the product choice game when states are ranked according to $\theta_h \succ \theta_l$, which fits into the interpretation that θ stands for product quality, durability, and safety. Third, there are *complementarities* between state and action profile in players' payoff functions. In the product choice game with the above rankings over states and players' actions, MS is satisfied when $\eta > 0$, namely, providing good service is *less costly* when the firm's product quality is higher.

I examine a patient player 1's payoff and behavior when he can build a reputation for playing *his highest action*. Let $\overline{a}_i \equiv \max A_i$ and $\underline{a}_i \equiv \min A_i$ be player $i \in \{1, 2\}$'s highest action and lowest action, respectively. Under Assumption 2, strategic type θ's minmax payoff is $u_1(\theta, \underline{a}_1, \underline{a}_2)$. Let

$$\Theta^* \equiv \left\{\theta \in \Theta \middle| u_1(\theta, \overline{a}_1, \overline{a}_2) > u_1(\theta, \underline{a}_1, \underline{a}_2)\right\} \tag{3.6}$$

[7]Assumption 2 rules out zero-sum games, common interest games, and coordination games more generally. Reputation for commitment is not valuable in zero-sum games. In Online Appendix H.2, I provide an example of common interest game with nontrivial interdependent values, under which player 1's guaranteed payoff is arbitrarily low compared to his pure Stackelberg payoff. My MS condition is also different from the *monotone-submodular* condition in Liu (2011) and Liu and Skrzypacz (2014), as the key to my model is the payoff-relevant state, which is absent in theirs.

be the set of states under which $\overline{a}_1$ is *individually rational* for player 1. I focus on games in which Θ^* is nonempty (i.e., $\overline{a}_1$ is individually rational in some states), and player 2's action choice is binary:[8]

Assumption 3. *Θ^* is non-empty, and $|A_2| = 2$.*

A reputation for playing $\overline{a}_1$ is *potentially valuable* only when $\overline{a}_1$ is one of the commitment actions, and player 2 has an incentive to choose player 1's desired action $\overline{a}_2$ when she knew that player 1 is committed and plays $\overline{a}_1$ in every period. Formally, for every $a_1^* \in A_1^*$, let $\phi_{a_1^*} \in \Delta(\Theta)$ be the distribution of states *conditional on* player 1 being commitment type a_1^*, which can be derived from player 2's prior belief μ. Abusing notation, let

$$\mathrm{BR}_2(\phi_{a_1^*}, a_1^*|u_2) \equiv \arg\max_{a_2 \in A_2} \Big\{ \sum_{\theta \in \Theta} \phi_{a_1^*}(\theta) u_2(\theta, a_1^*, a_2) \Big\},$$

which is the set of player 2's pure best replies against commitment type a_1^*. The assumption is stated as follows:

Assumption 4. *$\overline{a}_1 \in A_1^*$ and $BR_2(\phi_{\overline{a}_1}, \overline{a}_1|u_2) = \{\overline{a}_2\}$.*

Assumptions 3 and 4 are satisfied in the product choice game under the aforementioned rankings over states and actions. This is because $\theta_h \in \Theta^*$, the set of commitment actions is $A_1^* = \{G, B\}$, and according to (2.3), ϕ_G is the Dirac measure on θ_h, which implies that $\mathrm{BR}_2(\phi_G, G|u_2) = \{T\}$.

Theorems 2 and 3 focus on opposite conditions on player 2's prior belief about state $\phi \in \Delta(\Theta)$. I say that ϕ is *optimistic* if:

$$\overline{a}_2 \in \arg\max_{a_2 \in A_2} \Big\{ \sum_{\theta \in \Theta^*} \phi(\theta) u_2(\theta, \overline{a}_1, a_2) \Big\}, \tag{3.7}$$

and ϕ is *pessimistic* otherwise. In the product choice game where $\eta \in (0,1)$, $\Theta^* = \{\theta_h, \theta_l\}$, and ϕ is optimistic if it attaches probability more than $1/2$ to θ_h. In general, ϕ is optimistic if it attaches high enough probability to high states, such that $\overline{a}_2$ best replies against $\overline{a}_1$ when $\overline{a}_1$ is played in all states under which it is individually rational. When the probability of commitment types is sufficiently small, ϕ being optimistic is equivalent to the *existence of equilibrium* under which every strategic type $\theta \in \Theta^*$ receives payoff at least $u_1(\theta, \overline{a}_1, \overline{a}_2)$. Theorem 2 shows that a patient player 1 receives at least this payoff in all states and in *all equilibria*:

Theorem 2. *If ϕ is optimistic, and the game satisfies Assumptions 1, 2, 3, and 4, then for every $\theta \in \Theta$:*

$$\liminf_{\delta \to 1} \underline{v}_\theta(\delta, \mu, u_1, u_2) \geq \max\{u_1(\theta, \overline{a}_1, \overline{a}_2), u_1(\theta, \underline{a}_1, \underline{a}_2)\}.^9 \tag{3.8}$$

[8]Binary action games have been a primary focus of the reputation literature, examples of which include Mailath and Samuelson (2001), Ekmekci (2011), and Liu (2011). Extensions to games with $|A_2| \geq 3$ with additional assumptions are in Online Appendix G.

[9]In Online Appendix F, I show that this lower bound is tight in the sense that no strategic type can guarantee a strictly higher payoff when (1) commitment types are rare, and (2) player 2's best reply against player 1's highest action depends on the state.

The proofs of Theorem 2 and the next theorem are in Appendix D. Since the RHS of (3.8) is weakly greater than type θ's commitment payoff from $\overline{a}_1$, player 1 is guaranteed to receive *at least* his commitment payoff from $\overline{a}_1$. This difference in conclusions between Theorem 2 and Theorem 1 is driven by the MS condition on stage-game payoffs, under which player 1 has a stronger preference towards higher action profiles in higher states. It implies that for every equilibrium strategy σ_2 of player 2's, suppose there exists θ such that playing $\overline{a}_1$ in every period best replies against σ_2 in state θ, then in every state higher than θ, $\overline{a}_1$ is chosen for sure in every period under every best reply against σ_2. For every equilibrium where such θ exists, player 2's posterior belief about the state cannot become more pessimistic upon observing $\overline{a}_1$.

However, MS *cannot* rule out conflicts between building reputation for playing $\overline{a}_1$ and signaling high θ in *all equilibria*. This is because for some equilibrium strategies of player 2's, playing $\overline{a}_1$ in every period is *not* a best reply in any state (call them *irregular equilibria*). Indeed, there exist irregular equilibria in which $\overline{a}_1$ signals low θ at some on-path histories. I circumvent this complication by establishing a *belief lower bound*, that in every irregular equilibrium, and at every history where $\overline{a}_1$ has always been played in the past, player 2's posterior belief about the state must be optimistic.[10] This belief lower bound implies that player 2 has a strict incentive to play $\overline{a}_2$ as long as she is convinced that $\overline{a}_1$ will be played. It further implies that a patient player 1 can secure payoff from the highest action profile by playing $\overline{a}_1$ in every period.

When ϕ is *pessimistic* and commitment types are rare, Theorem 3 uniquely pins down player 1's equilibrium payoff and on-path behavior. As a byproduct, it also pins down player 2's posterior beliefs upon observing $\overline{a}_1$. To characterize equilibria, let

$$A_1^g \equiv \Big\{a_1^* \in A_1^* \Big| \mathrm{BR}_2(\phi_{a_1^*}, a_1^* | u_2) = \{\overline{a}_2\}\Big\}, \tag{3.9}$$

which is the set of commitment types against which player 2 has a strict incentive to play $\overline{a}_2$. Assumption 4 implies that $\overline{a}_1 \in A_1^g$. For every pessimistic ϕ, let $\theta^*(\phi)$ be the *largest* $\theta \in \Theta^*$ such that:

$$\{\underline{a}_2\} = \arg\max_{a_2 \in A_2} \Big\{ \sum_{\theta' \succeq \theta} \phi(\theta') u_2(\theta', \overline{a}_1, a_2) \Big\}. \tag{3.10}$$

Let

$$r(\phi) \equiv \frac{u_1(\theta^*(\phi), \underline{a}_1, \underline{a}_2) - u_1(\theta^*(\phi), \overline{a}_1, \underline{a}_2)}{u_1(\theta^*(\phi), \overline{a}_1, \overline{a}_2) - u_1(\theta^*(\phi), \overline{a}_1, \underline{a}_2)}, \tag{3.11}$$

[10] Suppose towards a contradiction that in an irregular equilibrium, player 2's posterior belief about the state is pessimistic at some history where player 1 has played $\overline{a}_1$ in all previous periods. On one hand, the definition of irregular equilibrium implies that all strategic types eventually separate from commitment type $\overline{a}_1$, after which player 2's belief about state is optimistic given Assumption 4. On the other hand, some strategic types are supposed to separate from commitment type $\overline{a}_1$ at the *last history* where posterior belief is pessimistic, after which at least one of these types receives his minmax payoff. However, if this type deviates at this last history by pooling with commitment type $\overline{a}_1$, then his continuation payoff is no less than $u_1(\theta, \overline{a}_1, \overline{a}_2)$, which is strictly greater than his minmax payoff. This contradicts his incentive to separate from commitment type $\overline{a}_1$ at that last history.

which is strictly between 0 and 1 given that $\theta^*(\phi) \in \Theta^*$ and $u_1(\theta, \overline{a}_1, \overline{a}_2) > u_1(\theta, \underline{a}_1, \underline{a}_2) > u_1(\theta, \overline{a}_1, \underline{a}_2)$ for $\theta \in \Theta^*$. Let

$$w_\theta(\phi) \equiv \begin{cases} u_1(\theta, \underline{a}_1, \underline{a}_2) & \text{if } \theta \preceq \theta^*(\phi) \\ r(\phi) u_1(\theta, \overline{a}_1, \overline{a}_2) + \big(1 - r(\phi)\big) u_1(\theta, \overline{a}_1, \underline{a}_2) & \text{if } \theta \succ \theta^*(\phi). \end{cases} \tag{3.12}$$

Theorem 3. *Under Assumptions 1, 2, and 3. For every pessimistic ϕ, there exist $\overline{\varepsilon} \in (0,1)$ and $\underline{\delta} \in (0,1)$, such that if $\delta > \underline{\delta}$, μ has state distribution ϕ, attaches probability less than $\overline{\varepsilon}$ to all commitment types, and satisfies Assumption 4, then in every equilibrium, strategic type θ's payoff is $w_\theta(\phi)$ for every $\theta \in \Theta$, and*

1. *For every $\theta \succ \theta^*(\phi)$, strategic type θ plays $\overline{a}_1$ at each of his on-path history.*[11]
2. *For every $\theta \prec \theta^*(\phi)$, strategic type θ plays $\underline{a}_1$ at each of his on-path history.*
3. *In period 0, type $\theta^*(\phi)$ plays a mixed action supported in $A_1^g \cup \{\underline{a}_1\}$. His mixing probabilities are chosen such that for every $a_1 \in A_1^g \backslash \{\underline{a}_1\}$, after observing a_1 in period 0, player 2 is indifferent between $\overline{a}_2$ and $\underline{a}_2$ against a_1 under her posterior belief about the state. Starting from period 1, type $\theta^*(\phi)$ repeats the same action that he has played in period 0 at each of his on-path history.*

To better understand Theorem 3, I start from explaining the intuitions behind $\theta^*(\phi)$, $r(\phi)$, and $w_\theta(\phi)$. First, when payoffs are MS and commitment types occur with *small but positive probability*, Assumption 4 and the definition of $\theta^*(\phi)$ imply the existence of $q \in (0,1)$ such that:[12]

- if all strategic types above $\theta^*(\phi)$ play $\overline{a}_1$ with probability 1, all strategic types below $\theta^*(\phi)$ play $\underline{a}_1$ with probability 1, and strategic type $\theta^*(\phi)$ plays $\overline{a}_1$ with probability q, then after observing $\overline{a}_1$, player 2 is indifferent between $\overline{a}_2$ and $\underline{a}_2$ against $\overline{a}_1$ under her posterior belief about the state.

Second, since u_2 has strictly increasing differences in θ and a_2, we have $\{\underline{a}_2\} = \text{BR}_2(\theta^*(\phi), \overline{a}_1 | u_2)$. Given that $\underline{a}_1$ is played only when $\theta \preceq \theta^*(\phi)$, we know that player 2 plays $\underline{a}_2$ in every period if player 1 plays $\underline{a}_1$ in every period. Therefore, for every commitment action $a_1^* \in A_1^g \backslash \{\underline{a}_1\}$, type $\theta^*(\phi)$'s indifference condition in period 0 uniquely pins down the discounted average probability of $\overline{a}_2$ when a_1^* is played in every period. For commitment action $\overline{a}_1$, this discounted average probability is given by (3.11).

In terms of player 1's payoffs from the above strategy, every type below $\theta^*(\phi)$ receives his minmax payoff, and every type $\theta \succeq \theta^*(\phi)$ receives payoff $r(\phi) u_1(\theta, \overline{a}_1, \overline{a}_2) + \big(1 - r(\phi)\big) u_1(\theta, \overline{a}_1, \underline{a}_2)$. According to (3.12), every strategic type's unique equilibrium payoff is strictly lower than his guaranteed equilibrium payoff under an optimistic ϕ.[13] Intuitively, this is because when ϕ is pessimistic and commitment types are rare, player 2

[11] For any given equilibrium $\big((\sigma_\theta)_{\theta \in \Theta}, \sigma_2\big)$ and state $\theta \in \Theta$, a history h^t is an *on-path history* for strategic type θ if h^t occurs with positive probability under the probability measure induced by $(\sigma_\theta, \sigma_2)$.

[12] If Θ is an interval, ϕ has no atom, and u_2 is continuous in θ, then q is not needed to describe player 1's unique on-path behavior.

[13] In Online Appendix E, I show that for every $\theta \in \Theta$, $w_\theta(\phi)$ is a patient player 1's *highest equilibrium payoff* in a repeated incomplete information game with state distribution ϕ but without commitment types.

has no incentive to play $\overline{a}_2$ until some strategic types in Θ^* separate from commitment type $\overline{a}_1$, after which at least one of these types receives his minmax payoff. The definition of Θ^* implies that this type strictly prefers $(\overline{a}_1, \overline{a}_2)$ to $(\underline{a}_1, \underline{a}_2)$. In order to prevent this type from imitating other strategic types, the equilibrium payoff of every type in Θ^* must be strictly lower than his payoff from $(\overline{a}_1, \overline{a}_2)$.

The most novel part of Theorem 3 is that player 1's on-path behavior being unique. It has the interesting feature that player 1 repeats the same action over time and sustains his reputation.[14] This unique on-path behavior also pins down player 2's posterior beliefs upon observing player 1 playing $\overline{a}_1$. These sharp predictions on behavior and learning contrast to private-value reputation games in Fudenberg and Levine (1989), the optimistic prior case studied by Theorem 2, and repeated signaling games without commitment types, in which the patient player has multiple on-path behaviors, switching actions over time is his strict best reply in many equilibria, and player 2s' posterior beliefs upon observing player 1's commitment behavior differ across equilibria.

Intuitively, my behavioral uniqueness result is driven by a novel *disciplinary effect*, implied by the joint forces of interdependent values and commitment types. In particular, player 1 can guarantee payoff strictly greater than his minmax payoff by imitating commitment type $\overline{a}_1$ (i.e., *a guaranteed reward*), and is guaranteed to receive his minmax payoff after separating from commitment types (i.e., *a guaranteed punishment*).

The guaranteed reward part is driven by Assumption 4, which says that building a reputation for playing $\overline{a}_1$ is feasible and player 2 has a strict incentive to play $\overline{a}_2$ once she is convinced that player 1 is commitment type $\overline{a}_1$. This effect also occurs in private-value reputation games, and interdependent-value reputation games studied by Theorem 2, but is missing in repeated signaling games without commitment types.

The guaranteed punishment part is caused by interdependent values and the high likelihood of low states, which is absent in existing reputation models. The key observation is that in all equilibria, separating from commitment type $\overline{a}_1$ triggers negative inference about the state, after which player 1's continuation payoff equals his minmax payoff. To see this, take a simplified setting where $A_1 \equiv \{\overline{a}_1, \underline{a}_1\}$ and suppose towards a contradiction that playing $\underline{a}_1$ makes player 2's belief about the state more optimistic. Since player 2's prior belief is pessimistic and belief is a martingale, observing $\overline{a}_1$ leads to a more pessimistic belief about the state. This implies that there exists at least one strategic type in the support of player 2's posterior belief after observing $\overline{a}_1$, who receives his minmax payoff in the continuation game. Since playing $\overline{a}_1$ is strictly costly for player 1, the above strategic type has a strict incentive to deviate by playing $\underline{a}_1$. This contradicts his incentive to play $\overline{a}_1$.

In contrast, this guaranteed punishment is missing in private-value reputation games and the optimistic-prior case studied by Theorem 2. In case of private values, the result in Fudenberg, Kreps and Maskin (1990) implies that strategic player 1's continuation payoff after separating from commitment type $\overline{a}_1$ can be anything

[14] Player 2's on-path behavior is not unique. This is because the cutoff type's indifference condition only pins down the discounted average frequency of $\overline{a}_2$, but does not pin down how the play of $\overline{a}_2$ is allocated over time.

between his minmax payoff and his payoff from $(\overline{a}_1, \overline{a}_2)$. This multiplicity in continuation values leads to multiple on-path behaviors. This is because at a given history, whether player 1 has an incentive to pool with or separate from commitment type $\overline{a}_1$ depends on his continuation value after separation. In particular, he strictly prefers to pool with commitment type $\overline{a}_1$ if he can only receive his minmax payoff after separation, and strictly prefers to separate from commitment type $\overline{a}_1$ if he can still receive $u_1(\theta, \overline{a}_1, \overline{a}_2)$ after separation.

A similar intuition applies to the optimistic prior case studied by Theorem 2, in which every strategic type's continuation payoff (after separating from commitment type $\overline{a}_1$) can be anything between his minmax payoff and his payoff from $(\overline{a}_1, \overline{a}_2)$. This leads to multiple on-path behaviors for patient player 1, as well as multiple possible posterior beliefs for player 2s. In particular, playing $\overline{a}_1$ is interpreted as a positive signal about θ in some equilibria, and is interpreted as a negative signal about θ in other equilibria.

Player 1 also has multiple on-path behaviors in private-value reputation games with a persistent state, for example, when he has persistent private information about his discount factor or his cost of taking a high action. This is because a strategic type with high cost or low discount factor *either* separates from the commitment type, in which case the disciplinary effect disappears after separation; *or* he pools with the commitment type, in which case he is equivalent to the commitment type from player 2's perspective due to the private-value assumption.

3.3 Implications: Product Choice Game

I illustrate the economic insights of my results by applying them to the product choice game. My model and results are applicable to markets with two characteristics. First, buyers' willingness to pay depends on persistent factors such as product quality, safety, and durability, which the seller knows more about. Second, informative signals about the state, other than the seller's actions, are rarely available to consumers or are unlikely to arrive for a long time. This is the case when the state is the potential adaptability issues of custom softwares, the durability of equipments, the long-run health impact of food and drugs, and so on.

Tradeoff Between Reputation and Signaling: I apply Theorem 1 to commitment action G and state θ_h. First, consumers' payoff function satisfies Assumption 1. Second, interdependent values are nontrivial under (u_2, G) since $\text{BR}_2(\theta_h, G|u_2) = \{T\}$ and $\text{BR}_2(\theta_l, G|u_2) = \{N\}$. Take a full support state distribution $\phi(\theta_h) = 0.99$ and $\phi(\theta_l) = 0.01$, and let the firm's payoff function be the one in the matrices with $-1 < \eta \leq 0$. When the probability of commitment types is less than 0.01 and δ is large, the following strategy profile is an equilibrium:

1. Strategic type θ_l plays G if B has never been played before, and plays B otherwise.

2. In period 0, strategic type θ_h plays G with probability β, and plays B with probability $1-\beta$, in which $\beta \in (0,1)$ is such that when G is observed in period 0, player 2's posterior attaches probability $1/2$ to

state θ_h. Such β exists when the probability of commitment type G is less than that of strategic type θ_l. In period $t \geq 1$, strategic type θ_h plays G if B has never been played before, and plays B otherwise.

3. Consumer plays N in period 0. Starting from period 1, she plays T with probability $1/(2\delta)$ if B has never been played before, and plays N otherwise.

In the above equilibrium, the strategic high-quality firm's payoff is 0, which is strictly lower than his complete information commitment payoff from G, equals to 1. This is because consumers believe that in period 0, strategic low-quality firm is more likely to choose G compared to strategic high-quality firm. Despite the prior probability of state θ_h is close to 1, and θ_h occurs with probability 1 conditional on player 1 being commitment type G, observing G drives consumers' posterior belief about θ_h down to $1/2$, after which they do not have a strict incentive to trust the firm even when they are convinced that G is likely to be played.

The above equilibrium highlights the following economic mechanism that is missing in private-value reputation models. When a seller tries to build a reputation for providing good customer service, he is concerned that consumers will become more suspicious about product quality after observing his reliable behavior (e.g., buyers believe that low-quality sellers are more eager to sell, and hence are more likely to provide comprehensive service). The practical relevance of this concern is shown by Miklos-Thal and Zhang (2013), who present evidence showing that savvy consumers interpret seller's efforts (e.g., providing comprehensive service) as negative signals about product quality, and sellers respond by providing limited services in order to avoid such negative inferences. Theorem 1 unveils the persistent negative effects of this adverse belief from the perspective of a patient seller who tries to build a reputation. When consumers entertain the above belief, the seller either loses his reputation for providing good service, after which he cannot convince future consumers about his service standards, or he signals negative information about product quality, after which consumers do not have incentives to trust him despite knowing that he will provide good service.

My result suggests an explanation for empirical observations of reputation failures. For example, Bai (2018) studies watermelon markets and finds that sellers rarely exert effort to sort out sweet watermelons, and buyers do not trust sellers despite observing their efforts. By running a randomized control trial and comparing sellers' behaviors across different treatments, her empirical results suggest the significance of adverse selection in this market, and that sellers' lack of effort in the control group is not caused by their impatience or consumers' inability to observe sellers' efforts. These findings suggest the existence of bad equilibria in which the canonical reputation mechanism in Fudenberg and Levine (1989) fails. The results in Ely and Välimäki (2003) cannot explain her empirical observations either, since in bad reputation models, the long-run player has a strong incentive to build his reputation whenever his opponents participate. This is at odds with Bai's finding that sellers do not exert effort when having opportunities to do so.

MS Product Choice Game: To apply Theorems 2 and 3, I provide two applications which fit into my modeling assumptions. First, consider a software firm interacting with its clients. In every period, a client chooses between a custom software (action T) or a commercial off-the-shelf software (in short, COTS, action N), and the firm chooses whether to deliver the product on time (action G) or not (action B). One can also interpret action G as exerting high effort to reduce cost overruns, and action B as exerting low effort.[15]

A client has more incentive to choose the custom software when she expects an on-time delivery and the software's expected *quality* is high. I interpret quality as the software's versatility and adaptiveness. These factors are harder for buyers to observe (either by her direct observation or from previous buyers' experiences) compared to the history of the firm's delivery times and cost overruns. This is a reasonable assumption since information about delayed deliveries and cost-overruns can be received by future buyers soon after the firm delivers a product, while adaptability issues become salient only after major revolutions of operation systems or other complementary softwares, which typically take years to happen. Payoffs are MS in this market when the firm has private information about its *competence*. This is because the quality of the software is positively correlated with the firm's competence, and the cost of making a timely delivery and reducing cost overruns is negatively correlated with its competence.

Next, consider the interactions between a seller and his buyers in the food and drug industry. I provide an overview below, with a more detailed microfoundation in Appendix E. In every period, the seller chooses whether to conduct costly inspection, which can detect defective products being produced. After inspecting the products, the seller can pay additional cost per unit to improve defective products before selling them on the market. Motivated by situations in developing countries, I assume that formal inspections are informative only about a subset of factors related to product safety, but omit other relevant factors (call them *omitted factors*). The seller has persistent private information about the state, which is either high or low, interpreted as the quality of his production technology, the reliability of his upstream suppliers, the quality of his inputs, and so on. This state affects not only the fraction of products that can pass formal inspections, but also affects the omitted factors that are related to product safety.

Mapping this into the product choice game, the seller's two actions are: not conducting inspections (action B), and conducting inspections and improve defective products (action G). I assume that future consumers can observe the seller's past actions (or equivalently, whether the seller has sold products that cannot pass formal inspections), but cannot directly observe the omitted factors. This is reasonable because a firm's effort in conducting inspections can be revealed via factory visits, its annual reports (which contain information

[15]The software firm needs to exert effort no matter which version of the product the buyer chooses to buy. In case of custom software, the firm needs to develop functions required by its client. In case of COTS, the firm needs to create packages to accommodate the buyer's specific needs. The necessity of effort creates problems such as delayed delivery and cost overruns, which are major concerns for buyers in this market in addition to the software's quality. See Banerjee and Duflo (2000) for detailed descriptions of this industry.

about the number of quality-control personnel being hired), or through governmental inspections which can tell whether the firm has sold defective products or not. Consumers' lack of information about these omitted factors can be attributed to imperfect screening technologies and regulatory loopholes. In developing countries, consumers' learning can also be slowed down *misdiagnoses*, making it hard for them to learn about the health impact of a particular food or drug through previous consumers' symptoms after consumption.[16]

Payoffs are MS in this market since a seller who has high-quality production technology faces lower cost to take action G compared to a low-quality seller. This is because when the seller's quality is high, a higher fraction of his products can pass quality control, so his net cost of quality control (which not only includes the direct cost of inspections but also the cost the seller incurs to improve defective products) is strictly lower.

Implication of Theorem 2: Recall from section 3.2 that Assumptions 1 to 4 are satisfied in the product choice game when $\eta > 0$. In the interesting case where $\eta \in (0, 1)$, the state distribution ϕ is optimistic if and only if it attaches probability more than $1/2$ to the high-quality state. Under these conditions, Theorem 2 implies that in every equilibrium, a patient strategic firm receives payoff at least 1 in state θ_h and receives payoff at least $1 - \eta$ in state θ_l. These are payoffs from taking the good action while receiving consumers' trust. In particular, a firm can secure these high profits despite it is facing tradeoffs between establishing reputation for playing the good action and signaling that its product quality is high.

Economically, Theorem 2 identifies market conditions under which a firm can secure high profits in the long run by sustaining reputations for providing good customer service (e.g., on-time deliveries, reduce cost overruns, etc.), despite doing so may possibly trigger negative inferences about product quality. This has implications for designing business strategies in markets where adverse selection on product quality is an important concern, but players' payoffs are MS. For example, Theorem 2 provides an explanation to the empirical finding in Banerjee and Duflo (2000) on the Indian custom software industry, in which sellers who enjoy reputations for making on-time deliveries and reducing cost overruns receive high profits. This finding is nontrivial given the amount of heterogeneity in software quality, the importance of quality for buyers, and the lack of informative signals about key aspects of quality, such as versatility and adaptiveness.

Implication of Theorem 3: Recall that when $\eta \in (0, 1)$, the product choice game satisfies Assumptions 1 to 4, and ϕ is pessimistic if and only if it attaches probability less than $1/2$ to state θ_h. Under these conditions, Theorem 3 says that when the total probability of commitment types is small enough, a patient strategic firm's payoff is η in state θ_h, and is 0 in state θ_l. In addition, the firm's on-path behavior is the same for all equilibria,

[16] Adhvarya (2014) finds that more than half of the patients receiving treatment for malaria in Tanzania were not infected. He also provides evidence showing that misdiagnosis slows down consumers' social learning. In the case of dairy products, consumers are unaware of the safety hazards of Sanlu's infant formula until 2008, many years after the product was introduced to the market.

in which strategic type θ_h plays G in every period, and strategic type θ_l mixes between playing G in every period and playing B in every period. His mixing probability in period 0 is such that after observing G, buyers' posterior belief attaches probability $1/2$ to state θ_h.

Theorem 3 implies that when consumers are initially pessimistic about product quality or safety, firms receive low profits for a long time, and conditional on them not exiting the market (e.g., never receiving consumers' trust in the future), firms take the good action in every period and sustain their reputations. This combination of persistent low payoff and consistent reputation-building behavior is a novel prediction in the reputation literature. My result provides an explanation to some recent empirical findings on the food and drug industries in developing countries, where product quality and safety are important concerns, but consumers receive limited information about these variables based on previous consumers' experiences.

For example, Bai, Gazze and Wang (2019) find that after the 2008 scandal of the Chinese dairy industry, nearly all Chinese dairy firms suffered a 68% revenue decline for at least five years, including firms that were deemed innocent by formal inspections. More interestingly, firms that survived the scandal took costly measures, such as frequently conducting inspections, to rectify the situation. Theorem 3 provides an explanation for this combination of persistent low payoffs and consistent reputation-building behaviors. As shown empirically by Luong et al. (2019), the scandal led to an unexpected negative shock on consumers' beliefs about product safety, after which the strategic firm's unique payoff under a pessimistic prior is strictly lower than its guaranteed payoff under an optimistic one. When consumers are pessimistic about product safety and a reputation for conducting inspections is feasible, surviving firms establish this reputation because failing to do so in any period makes consumers more pessimistic about product safety (in particular, factors that are omitted by formal inspections), after which firms are not trusted by future consumers even if they conduct inspections.

4 Concluding Remarks

My analysis unveils some challenges to reputation building when uninformed players' learning is confounded. This is related to a contemporary paper of Deb and Ishii (2019) that studies reputation building when uninformed players do not know the monitoring structure. Their model assumes that there exists a public signal which can statistically identify the state.[17] They construct a commitment type that plays a *nonstationary* strategy under which the patient informed player can secure his complete information commitment payoff.

In contrast, I study a model in which the state affects the uninformed players' best replies, but can only

[17] Their Assumption 2.3 requires that for every $\theta, \theta' \in \Theta$, there exists $\alpha_1 \in \Delta(A_1)$ such that the signal distribution under (θ, α_1) *cannot* be induced by *any action distributions* in state θ'. Their identification assumption is violated in my model, as well as repeated incomplete information game models of Aumann and Maschler (1995), Hörner and Lovo (2009), and Pęski (2014), and repeated signaling games models in Kaya (2009).

be learnt via the informed player's actions. The lack of exogenous signals that can statistically identify the state introduces new challenges for uninformed players to learn the correct best reply against the commitment action. In terms of results, I derive a unique prediction on the informed player's on-path behavior, in addition to deriving lower bounds on his equilibrium payoff.

In terms explaining reputation failures, Theorem 1 is related to Ely and Välimäki (2003) and Ely, Fudenberg and Levine (2008), that study a class of *private-value* reputation games called *participation games*. They show that a patient informed player's equilibrium payoff is low when *bad commitment types*, namely ones that discourage the uninformed players from participating, occur with high enough probability compared to the Stackelberg commitment type. Their bad reputation result relies on uninformed players' ability to choose a *non-participating action*, under which the public signal becomes uninformative about the informed player's action. The informed player receives low payoff because his opponents have no incentive to participate. However, once they participate, the informed player has a strong incentive to build his reputation.

In my model, the uninformed players *cannot* stop the informed player from signaling his type, but the informational content of the informed player's actions is sensitive to equilibrium selection. In particular, there exist equilibria in which uninformed players believe that the strategic informed player is more likely to choose the commitment action in some alternative state, under which they have no incentive to play the informed player's desired best reply. As a result, strategic informed players refrain from building reputations despite having opportunities to do so, which leads to a different prediction compared to models of bad reputations.

My Theorem 3 suggests that interdependent values can contribute to reputation sustainability. Following Cripps, Mailath and Samuelson (2004), I say that a reputation for playing $\overline{a}_1$ is sustained in an equilibrium if *conditional on* player 1 being strategic, there exists a positive probability event under which player 1's reputation (i.e., probability with which player 2's posterior assigns to commitment type $\overline{a}_1$) does not vanish as $t \to \infty$. According to this definition, reputation for playing $\overline{a}_1$ is sustained in all equilibria in games studied by Theorem 3. This contrasts to Fudenberg and Levine (1989) in which player 1 loses his reputation in some equilibria, and Cripps, et al. (2004) in which player 1 loses his reputation in all equilibria.

Moreover, player 1's unique on-path behavior in Theorem 3 is close to one of his equilibrium behaviors in a benchmark repeated game with the same state distribution but *without* commitment types, which is also the case in Cripps, et al.(2004). The difference is, playing a costly action (such as $\overline{a}_1$) in every period is *suboptimal* in any equilibrium of a repeated game with private values, full support monitoring, but without commitment types. This explains why reputation vanishes in all equilibria of Cripps, et al.(2004). In contrast, playing $\overline{a}_1$ in every period is optimal in *some equilibria* of the repeated game with interdependent values, perfect monitoring, but without commitment types. Theorem 3 shows that introducing commitment type selects equilibria with this particular on-path behavior, which leads to the sustainability of reputation in my model.

A Characterization Theorem

I generalize the baseline model by allowing for (1) arbitrary correlation between the distribution of θ the distribution of player 1's characteristics, and (2) the committed long-run player to play *mixed actions*.

Let $m \equiv |\Theta|$ be the number of states. Let $\gamma : \Theta \to \Delta(A_1)$ be a typical commitment plan, and let Γ be a finite set of commitment plans. The set of commitment actions is redefined as:

$$\mathcal{A}_1^* \equiv \{\alpha_1 \in \Delta(A_1) | \text{ there exist } \gamma \in \Gamma \text{ and } \theta \in \Theta \text{ such that } \gamma(\theta) = \alpha_1\}, \tag{A.1}$$

which is a finite subset of $\Delta(A_1)$. Let α_1^* be a typical element of $\mathcal{A}_1^*$. Let μ be player 2's prior belief, which is a joint distribution defined in (2.1). For every $\theta \in \Theta$, let $\mu(\theta)$ be the probability of *strategic type* θ. For every $\alpha_1^* \in \mathcal{A}_1^*$, let $\mu(\alpha_1^*)$ be the probability of *commitment type* α_1^*. Let $\phi_{\alpha_1^*} \in \Delta(\Theta)$ be the state distribution conditional on player 1 being commitment type α_1^*. Let $\lambda_\theta(\mu, \alpha_1^*) \equiv \mu(\theta)/\mu(\alpha_1^*)$. Let $\lambda(\mu, \alpha_1^*) \equiv \big\{\lambda_\theta(\mu, \alpha_1^*)\big\}_{\theta \in \Theta} \in \mathbb{R}_+^m$ be the *likelihood ratio vector* with respect to α_1^*.

For every $\phi \in \Delta(\Theta)$, $\alpha_1 \in \Delta(A_1)$, and $u_2 : \Theta \times A_1 \times A_2 \to \mathbb{R}$, let:

$$\mathrm{BR}_2(\phi, \alpha_1 | u_2) \equiv \arg\max_{a_2 \in A_2} \Big\{ \sum_{\theta \in \Theta} \sum_{a_1 \in A_1} \phi(\theta) \alpha_1(a_1) u_2(\theta, a_1, a_2) \Big\}.$$

For every $\alpha_1^* \in \mathcal{A}_1^*$, let $v_\theta(\alpha_1^*, u_1, u_2)$ be type θ's *commitment payoff* from α_1^*, which extends the definition in (3.3). Let $\underline{v}_\theta(\delta, \mu, u_1, u_2)$ be type θ's *lowest equilibrium payoff* under parameter (δ, μ, u_1, u_2). The following assumption is analogous to Assumption 1 and is satisfied for generic u_2:

Assumption 1'. *For every $\alpha_1^* \in \mathcal{A}_1^*$ and $\theta \in \Theta$, $BR_2(\theta, \alpha_1^* | u_2)$ is a singleton.*

For given $\alpha_1^* \in \mathcal{A}_1^*$, $\theta^* \in \Theta$, and u_2, Theorem 1' provides *sufficient and (almost) necessary condition* on player 2's prior belief μ such that:

$$\liminf_{\delta \to 1} \underline{v}_\theta(\delta, \mu, u_1, u_2) \geq v_\theta(\alpha_1^*, u_1, u_2), \text{ for all } u_1.$$

Let $a_2^*(\theta^*, \alpha_1^* | u_2)$ be the unique element in $\mathrm{BR}_2(\theta^*, \alpha_1^* | u_2)$. Let $\Lambda(\theta^*, \alpha_1^*, u_2)$ be the subset of $\mathbb{R}_+^m$ such that $\lambda \in \Lambda(\theta^*, \alpha_1^*, u_2)$ *if and only if*:

$$\big\{a_2^*(\theta^*, \alpha_1^* | u_2)\big\} = \arg\max_{a_2 \in A_2} \big\{ u_2(\phi_{\alpha_1^*}, \alpha_1^*, a_2) + \sum_{\theta \in \Theta} \lambda'_\theta u_2(\theta, \alpha_1^*, a_2) \big\}. \tag{A.2}$$

for all $\lambda' \equiv \{\lambda'_\theta\}_{\theta\in\Theta}$ with $0 \leq \lambda' \leq \lambda$. Let $co(\cdot)$ denote the convex hull of a set, and let

$$\underline{\Lambda}(\theta^*, \alpha_1^*, u_2) \equiv \mathbb{R}_+^m \Big\backslash \mathrm{co}\Big(\mathbb{R}_+^m \backslash \Lambda(\theta^*, \alpha_1^*, u_2)\Big). \tag{A.3}$$

According to (A.2) and (A.3), whether $\lambda(\mu, \alpha_1^*)$ belongs to $\Lambda(\theta^*, \alpha_1^*, u_2)$ and $\underline{\Lambda}(\theta^*, \alpha_1^*, u_2)$ does not depend on the probability of other commitment types and the probability of good strategic types (i.e., ones under which player 2's best reply against α_1^* is the same as that under state θ^*). This is because first, when player 1's action frequency matches α_1^*, the probability of other commitment types vanishes in the long run. Second, in the worst equilibrium, the good strategic types separate from commitment type α_1^* and the bad ones pool with commitment type α_1^*. Let $cl(\cdot)$ denote the closure.

Theorem 1'. *Under Assumption 1': For every $(\theta^*, \alpha_1^*) \in \Theta \times \mathcal{A}_1^*$ with α_1^* being a pure action,*

1. *If $\lambda(\mu, \alpha_1^*) \in \Lambda(\theta^*, \alpha_1^*, u_2)$, then $\liminf_{\delta\to 1} \underline{v}_{\theta^*}(\delta, \mu, u_1, u_2) \geq v_{\theta^*}(\alpha_1^*, u_1, u_2)$ for every u_1;*
2. *If $\lambda(\mu, \alpha_1^*) \notin cl\Big(\Lambda(\theta^*, \alpha_1^*, u_2)\Big)$ and $BR_2(\phi_{\alpha_1^*}, \alpha_1^*|u_2)$ is a singleton, then there exists u_1 such that $\limsup_{\delta\to 1} \underline{v}_{\theta^*}(\delta, \mu, u_1, u_2) < v_{\theta^*}(\alpha_1^*, u_1, u_2)$;*

for every $(\theta^*, \alpha_1^*) \in \Theta \times \mathcal{A}_1^*$ with α_1^* being a nontrivially mixed action,

3. *If $\lambda(\mu, \alpha_1^*) \in \underline{\Lambda}(\theta^*, \alpha_1^*, u_2)$, then $\liminf_{\delta\to 1} \underline{v}_{\theta^*}(\delta, \mu, u_1, u_2) \geq v_{\theta^*}(\alpha_1^*, u_1, u_2)$ for every u_1;*
4. *If $\lambda(\mu, \alpha_1^*) \notin cl\Big(\underline{\Lambda}(\theta^*, \alpha_1^*, u_2)\Big)$, $BR_2(\phi_{\alpha_1^*}, \alpha_1^*|u_2)$ is a singleton and $\alpha_1^* \notin co\Big(\mathcal{A}_1^* \Big\backslash \{\alpha_1^*\}\Big)$, then there exists u_1 such that $\limsup_{\delta\to 1} \underline{v}_{\theta^*}(\delta, \mu, u_1, u_2) < v_{\theta^*}(\alpha_1^*, u_1, u_2)$.*

The proof is in Appendix B (statement 2 and Theorem 1), Appendix C (statements 1 and 3), and Online Appendices A and B (statement 4 and remaining steps for statement 3). In what follows, I replace α_1^* with a_1^* when it is pure, and suppress the dependence of a_2^*, λ, Λ, and $\underline{\Lambda}$ on α_1^*, θ^*, and u_2. Note that Theorem 1' generalizes Theorem 1 since all entries of λ go to infinity when the probability of commitment types vanishes to 0. If player 2's best reply against α_1^* depends on the state, then there exists a cutoff real number such that when all entries of λ exceed this cutoff, λ belongs neither to Λ nor $\underline{\Lambda}$. Later, I show that $\mathrm{BR}_2(\phi_{\alpha_1^*}, \alpha_1^*|u_2)$ being a singleton, is not needed for the conclusion when all commitment actions are pure (Appendix B).

To better understand Theorem 1', I explain the intuition behind Λ and $\underline{\Lambda}$. First, for type θ^* to secure payoff $v_{\theta^*}(\alpha_1^*, u_1, u_2)$ under every u_1, it is necessary that a_2^* is player 2's best reply against α_1^* under her prior belief about the state. However, this is *not sufficient* since player 2's belief is updated over time. As a result, player 1 needs to find a strategy under which he can pool with commitment type α_1^*, while making sure that player 2 has an incentive to play a_2^* under her posterior belief about the state. Whether α_1^* is pure or mixed matters since it affects the set of posterior beliefs that can possibly arise under a given prior belief.

When α_1^* is pure, as long as player 1 plays α_1^*, each entry of λ is nonincreasing. As a result, Λ requires a_2^* to be a best reply against α_1^* after some strategic types separate from commitment type α_1^* with positive probability and the posterior likelihood ratio vector is λ', which is less than the prior likelihood ratio λ.

When α_1^* is nontrivially mixed, playing some actions in the support of α_1^* may *increase* some (or all) entries of the likelihood ratio vector. Having $\lambda \in \Lambda$ is no long sufficient, which is explained via the following example:

- Suppose $\alpha_1^* = \frac{1}{2}a_1' + \frac{1}{2}a_1''$, there are two states θ_1 and θ_2 under which player 2's best reply against α_1^* differs from that under state θ^*, and the prior likelihood ratio vector λ belongs to Λ but does not belong to $\underline{\Lambda}$. Suppose player 2 believes that strategic type θ_1 plays a_1' with probability 1 and strategic type θ_2 plays a_1'' with probability 1. No matter which action player 1 plays in the support of α_1^*, the posterior likelihood ratio vector is bounded away from Λ.

The above problem disappears when $\lambda \in \underline{\Lambda}$. This is because the likelihood ratio vector is a supermartingale (conditional on α_1^*), and $\mathbb{R}^m \backslash \underline{\Lambda}$ is convex. As a result, for every belief updating process that respects Bayes Rule and at every history in which the probability of commitment type α_1^* is positive, there exists at least one pure action in the support of α_1^*, such that the posterior likelihood ratio vector also belongs to $\underline{\Lambda}$.

B Proof of Theorem 1 and Theorem 1' Statement 2

Step 1: I show that when $\lambda \notin \mathrm{cl}(\Lambda)$, there exist $a_2 \neq a_2^*$ and $\lambda' \equiv \{\lambda'_\theta\}_{\theta \in \Theta}$, such that first, $0 \leq \lambda' \leq \lambda$ and $\lambda'_{\theta^*} = 0$, second,

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

and third,

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

According to the definition of Λ, there exists $\lambda'' \equiv \{\lambda''_\theta\}_{\theta \in \Theta}$ such that $0 \leq \lambda'' \leq \lambda$, and

$$a_2^* \notin \arg\max_{a_2 \in A_2} \Big\{ u_2(\phi_{a_1^*}, a_1^*, a_2') + \sum_{\theta \in \Theta} \lambda''_\theta u_2(\theta, a_1^*, a_2') \Big\}.$$

Let $\lambda' \in \mathbb{R}_+^m$ be such that $\lambda'_{\theta^*} \equiv 0$, and $\lambda'_\theta \equiv \lambda''_\theta$ for all $\theta \neq \theta^*$. Since $\{a_2^*\} = \mathrm{BR}_2(\theta^*, a_1^* | u_2)$, there exists $a_2' \neq a_2^*$:

$$u_2(\phi_{a_1^*}, a_1^*, a_2') + \sum_{\theta \in \Theta} \lambda'_\theta u_2(\theta, a_1^*, a_2') > u_2(\phi_{a_1^*}, a_1^*, a_2^*) + \sum_{\theta \in \Theta} \lambda'_\theta u_2(\theta, a_1^*, a_2^*).$$

If the unique element in $\mathrm{BR}_2(\phi_{a_1^*}, a_1^* | u_2)$ is a_2^*, then (B.1) and (B.2) hold for $a_2 = a_2'$. If the unique element in $\mathrm{BR}_2(\phi_{a_1^*}, a_1^* | u_2)$ is $a_2'' \neq a_2^*$, then there exists $\theta' \in \Theta$ such that $u_2(\theta', a_1^*, a_2'') > u_2(\theta', a_1^*, a_2^*)$. Let $\lambda' \in \mathbb{R}_+^m$ be defined as: $\lambda'_{\theta'} \equiv \lambda_{\theta'}$, and $\lambda'_\theta \equiv 0$ for all $\theta \neq \theta'$, then (B.1) and (B.2) hold for λ' and $a_2 = a''_2$.

For the conditions in Theorem 1, let $\theta' \neq \theta^*$ be such that $\mathrm{BR}_2(\theta', a_1^*|u_2) = \{a'_2\}$ with $a'_2 \neq a_2^*$. Let $\lambda' \in \mathbb{R}^m_+$ be such that $\lambda'_{\theta'} = \lambda_{\theta'}$ and $\lambda'_\theta = 0$ for all $\theta \neq \theta'$. Then (B.1) and (B.2) hold for $a_2 = a'_2$.

Step 2: Let

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

By definition, $v_{\theta^*}(a_1^*, u_1, u_2) = 1$. I describe players' equilibrium strategies. On the equilibrium path, strategic type θ^* plays a different pure action in each period from period 0 to $|A_1| - 1$. Starting from period $|A_1|$, he plays a_1^* for $k^* \in \mathbb{N}$ periods and then some prespecified $a_1 \neq a_1^*$ in the next period, and his behavior rotates every $k^* + 1$ periods. I will specify the value of integer k^* by the end of step 3.

I construct $\lambda' \in \mathbb{R}^m_+$ and $a'_2 \neq a_2^*$ according to Step 1. Inequality (B.2) implies the existence of $\epsilon > 0$ such that:

$$u_2(\phi_{a_1^*}, a_1^*, a'_2) - u_2(\phi_{a_1^*}, a_1^*, a_2^*) + (1-\epsilon) \sum_{\theta \in \Theta} \lambda'_\theta \Big(u_2(\theta, a_1^*, a'_2) - u_2(\theta, a_1^*, a_2^*) \Big) > 0. \tag{B.4}$$

For every $\widetilde{\theta} \neq \theta^*$, with probability $\Big(\lambda_{\widetilde{\theta}} - \lambda'_{\widetilde{\theta}}\Big)\Big/\lambda_{\widetilde{\theta}}$, strategic type $\widetilde{\theta}$ plays $a'_1 \neq a_1^*$ in every period; with probability $(1-\epsilon)\lambda'_{\widetilde{\theta}}/\lambda_{\widetilde{\theta}}$, strategic type $\widetilde{\theta}$ plays a_1^* in every period. For every $\alpha_1 \in \mathcal{A}_1^*$ that is nontrivially mixed, strategic type $\widetilde{\theta}$ plays strategy σ_{α_1} with probability $\frac{\epsilon}{k}\lambda'_{\widetilde{\theta}}\Big/\lambda_{\widetilde{\theta}}$, with $k \in \mathbb{N}$ being the number of nontrivially mixed commitment actions in $\mathcal{A}_1^*$ and σ_{α_1} will be specified in the next paragraph. If $k = 0$, then one can set $\epsilon = 0$.

Next, I describe strategy σ_{α_1}. If h^t occurs with positive probability under strategic type θ^*'s equilibrium strategy, then $\sigma_{\alpha_1}(h^t) = \alpha_1$. If h^t occurs with zero probability under strategic type θ^*'s equilibrium strategy, then $\sigma_{\alpha_1}(h^t) = \widehat{\alpha}_1$, in which:

$$\widehat{\alpha}_1(\alpha_1) \equiv (1 - \frac{\eta}{2})a_1^* + \frac{\eta}{2}\widetilde{\alpha}_1(\alpha_1) \tag{B.5}$$

and

$$\widetilde{\alpha}_1(\alpha_1)[a_1] \equiv \begin{cases} 0 & \text{when } a_1 = a_1^* \\ \alpha_1(a_1)/(1-\alpha_1(a_1^*)) & \text{otherwise .} \end{cases} \tag{B.6}$$

Since $\mathcal{A}_1^*$ is finite, there exists $\eta > 0$ such that $\max_{\alpha_1 \in \mathcal{A}_1^* \setminus \{a_1^*\}} \alpha_1(a_1^*) < 1 - \eta$. According to (B.1), for every $\alpha'_1 \in \Delta(A_1)$ with $\alpha'_1(a_1^*) \geq 1 - \eta$, we have:

$$\sum_{\theta \in \Theta} \lambda'_\theta u_2(\theta, \alpha'_1, a'_2) > \sum_{\theta \in \Theta} \lambda'_\theta u_2(\theta, \alpha'_1, a_2^*). \tag{B.7}$$

Step 3: I verify type θ^*'s incentive constraints. Instead of explicitly constructing type θ^*'s strategy at histories after he has deviated, I derive a *uniform upper bound* on his continuation payoff *after his first deviation*. For every $\alpha_1 \in \mathcal{A}_1^*$, let $\mu_t(\theta(\alpha_1))$ be the probability that player 1 is strategic and follows strategy σ_{α_1}. Let $\beta_t(\alpha_1) \equiv \mu_t(\theta(\alpha_1))/\mu_t(\alpha_1)$. The value of $\beta_t(\alpha_1)$ equals $\beta_0(\alpha_1)$ at histories that occur with positive probability under type θ^*'s equilibrium strategy.

Next, consider histories that occur with zero probability under type θ^*'s equilibrium strategy. Since

$$\max_{\alpha_1 \in \mathcal{A}_1^* \backslash \{a_1^*\}} \alpha_1(a_1^*) < 1 - \eta,$$

then for every $\alpha_1 \in \mathcal{A}_1^* \backslash \{a_1^*\}$,

$$\beta_{t+1}(\alpha_1) \geq \frac{1-\eta/2}{1-\eta} \beta_t(\alpha_1). \tag{B.8}$$

when a_1^* is observed in period t. Let $\kappa \equiv 1 - \min_{\alpha_1 \in \mathcal{A}_1^* \backslash \{a_1^*\}} \alpha_1(a_1^*)$. If $a_1 \neq a_1^*$ is observed in period t, then by definition of $\widetilde{\alpha}_1(\alpha_1)$, we have:

$$\beta_{t+1}(\alpha_1) \geq \frac{\eta}{2\kappa} \beta_t(\alpha_1). \tag{B.9}$$

Let $\overline{k} \equiv \left\lceil \log \frac{2\kappa}{\eta} \Big/ \log \frac{1-\eta/2}{1-\eta} \right\rceil$. For every $\alpha_1 \in \mathcal{A}_1^*$, let $\overline{\beta}(\alpha_1)$ be the smallest $\beta \in \mathbb{R}_+$ such that:

$$u_2(\phi_{\alpha_1}, \alpha_1, a_2') + \beta \sum_{\theta \in \Theta} \lambda_\theta' u_2(\theta, \widehat{\alpha}_1(\alpha_1), a_2') \geq u_2(\phi_{\alpha_1}, \alpha_1, a_2^*) + \beta \sum_{\theta \in \Theta} \lambda_\theta' u_2(\theta, \widehat{\alpha_1}(\alpha_1), a_2^*) \tag{B.10}$$

Let $\overline{\beta} \equiv 2 \max_{\alpha_1 \in \mathcal{A}_1^* \backslash \{a^*\}} \overline{\beta}(\alpha_1)$ and $\underline{\beta} \equiv \min_{\alpha_1 \in \mathcal{A}_1^* \backslash \{a^*\}} \frac{\mu(\theta(\alpha_1))}{\mu(\alpha_1)}$. Let $T_1 \equiv \left\lceil \log \frac{\overline{\beta}}{\underline{\beta}} \Big/ \log \frac{1-\eta/2}{1-\eta} \right\rceil$.

At any history right after type θ^*'s first deviation, $\beta_t(\alpha_1) \geq \underline{\beta}$ for all $\alpha_1 \in \mathcal{A}_1^*$. After player 2 observes a_1^* for T_1 consecutive periods, a_2^* is strictly dominated by a_2' until some $a_1' \neq a_1^*$ is observed. Moreover, every time player 1 plays some $a_1' \neq a_1^*$, he can induce outcome (a_1^*, a_2^*) for at most $\overline{k}$ consecutive periods before a_2^* is strictly dominated by a_2' again. Therefore, type θ^*'s continuation payoff after his first deviation is at most:

$$(1-\delta^{T_1}) + \delta^{T_1} \Big\{ (1-\delta^{\overline{k}-1}) + \delta^{\overline{k}}(1-\delta^{\overline{k}-1}) + \delta^{2\overline{k}}(1-\delta^{\overline{k}-1}) + ... \Big\}, \tag{B.11}$$

which converges to $\frac{\overline{k}}{1+\overline{k}}$ as $\delta \to 1$. Let $k^* \equiv 2\overline{k}$. When $\delta \to 1$, type θ^*'s payoff at any on-path history converges to $\frac{2\overline{k}}{2\overline{k}+1}$, which is strictly greater than (B.11).

C Proof of Theorem 1': Statements 1 and 3

Proof of Statement 1: If $a_1^* \in \mathcal{A}_1^*$ is a pure action and $\lambda \in \Lambda$, then for every $\lambda' \equiv \{\lambda_\theta'\}_{\theta \in \Theta}$ with $0 \leq \lambda' \leq \lambda$, we have:

$$\{a_2^*\} = \arg\max_{a_2 \in A_2} \Big\{ u_2(\phi_{a_1^*}, a_1^*, a_2) + \sum_{\theta \in \Theta} \lambda_\theta' u_2(\theta, a_1^*, a_2) \Big\}. \tag{C.1}$$

Let $\overline{h}^t$ be a public history in which a_1^* has been played in every period. For every $\theta \in \Theta$, let $q_t(\theta)$ be the ex ante probability of the event that $h^t = \overline{h}^t$ and player 1 is strategic type θ. Player 2's maximization problem at $\overline{h}^t$ is:

$$\max_{a_2 \in A_2} \Big\{ \mu(a_1^*) u_2(\phi_{a_1^*}, a_1^*, a_2) + \sum_{\theta \in \Theta} \Big[q_{t+1}(\theta) u_2(\theta, a_1^*, a_2) + (q_t(\theta) - q_{t+1}(\theta)) u_2(\theta, \alpha_{1,t}(\theta), a_2) \Big] \Big\} \tag{C.2}$$

in which $\alpha_{1,t}(\theta) \in \Delta(A_1 \backslash \{a_1^*\})$ can be arbitrary if type θ plays a_1^* at $\overline{h}^t$, and is the distribution of type θ's action at $\overline{h}^t$ conditional on $a_{1,t} \neq a_1^*$ if type θ does not play a_1^* at $\overline{h}^t$.

According to (C.1) and (C.2), there exists $\rho > 0$, such that player 2 has a strict incentive to play a_2^* at $\overline{h}^t$ if:

$$\sum_{\theta \in \Theta} q_{t+1}(\theta) > \sum_{\theta \in \Theta} q_t(\theta) - \rho. \tag{C.3}$$

Therefore, if player 1 plays a_1^* in every period, then there exist at most $\overline{T} \equiv \lceil 1/\rho \rceil$ periods in which player 2 does not have a strict incentive to play a_2^*. Therefore, type θ^*'s equilibrium payoff is at least:

$$(1 - \delta^{\overline{T}}) \min_{(a_1, a_2) \in A_1 \times A_2} u_1(\theta^*, a_1, a_2) + \delta^{\overline{T}} v_{\theta^*}(a_1^*, u_1, u_2). \tag{C.4}$$

Since $\overline{T}$ is independent of δ, the value of (C.4) converges to $v_{\theta^*}(a_1^*, u_1, u_2)$ as $\delta \to 1$.

Proof of Statement 3: I construct a *nonstationary strategy* $\sigma_1 : \mathcal{H} \to \Delta\Big(\text{supp}(\alpha_1^*)\Big)$, under which first, the posterior likelihood ratio vector belongs to $\underline{\Lambda}$ in every period; second, the discounted average frequency of every $a_1 \in A_1$ is approximately $\alpha_1^*(a_1)$; third, player 2 believes that actions close to α_1^* will be played in all except for a bounded number of periods.

Let $\overline{A_1} \equiv \text{supp}(\alpha_1^*)$. For every $\sigma_1 : \mathcal{H} \to \Delta(A_1)$ and $\sigma_2 : \mathcal{H} \to \Delta(A_2)$, let $\mathcal{P}^{(\sigma_1, \sigma_2)}$ be the probability measure over $\mathcal{H}$ induced by (σ_1, σ_2), let $\mathcal{H}^{(\sigma_1, \sigma_2)}$ be the set of histories that occur with positive probability under $\mathcal{P}^{(\sigma_1, \sigma_2)}$ and let $\mathbb{E}^{(\sigma_1, \sigma_2)}$ be its expectation operator. Let

$$\Theta^b_{(\alpha_1^*, \theta^*)} \equiv \big\{ \theta \in \Theta \big| a_2^* \notin \text{BR}_2(\theta, \alpha_1^* | u_2) \big\}. \tag{C.5}$$

be the set of states under which player 2's best reply against α_1^* is different from that under state θ^*. For every $\theta \in \Theta^b_{(\alpha_1^*, \theta^*)}$, let ψ_θ^* be the largest $\psi \in \mathbb{R}_+$ such that:

$$a_2^* \in \arg\max_{a_2 \in A_2} \Big\{ u_2(\phi_{\alpha_1^*}, \alpha_1^*, a_2) + \psi u_2(\theta, \alpha_1^*, a_2) \Big\}. \tag{C.6}$$

One can show that ψ_θ^* is the intercept of Λ on the axis for λ_θ. Assumption 1' requires $\text{BR}_2(\theta^*, \alpha_1^* | u_2)$ to be a singleton, which implies:

$$\underline{\Lambda} = \Big\{\widetilde{\lambda} \in \mathbb{R}^m_+ \Big| \sum_{\theta \in \Theta^b_{(\alpha_1^*, \theta^*)}} \frac{\widetilde{\lambda}_\theta}{\psi^*_\theta} < 1\Big\}. \tag{C.7}$$

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

$$\underline{\Lambda}(\psi, \chi) \equiv \Big\{\widetilde{\lambda} \in \mathbb{R}^m_+ \Big| \sum_{i=1}^m \frac{\widetilde{\lambda}_i}{\psi_i} < \chi\Big\}. \tag{C.8}$$

Let $\mu(h^t)$ be player 2's posterior belief at h^t. I write $\lambda(h^t)$ in short for $\lambda(\mu(h^t), \alpha_1^*)$, which is the likelihood ratio vector at h^t. Let h^∞ be an infinite history and let h^∞_t be player 1's action in period t under h^∞. Let $d(\cdot\|\cdot)$ be the KL divergence. The key step is stated as Proposition C.1:

Proposition C.1. *For every $\psi \in \mathbb{R}^m_+$ and $\chi > 0$, if $\lambda \in \underline{\Lambda}(\psi, \chi)$, then for every strategy profile σ and every $\epsilon > 0$, there exist $\overline{\delta} \in (0, 1)$ and $T \in \mathbb{N}$ such that for every $\delta > \overline{\delta}$, there exists $\widehat{\sigma}_1 : \mathcal{H} \to \Delta(\overline{A}_1)$ such that if player 2s update their beliefs according to σ and player 1 plays $\widehat{\sigma}_1$ then:*

1.

$$\lambda(h^t) \in \underline{\Lambda}(\psi, \chi + \epsilon) \quad \textit{for every} \quad h^t \in \mathcal{H}^{(\widehat{\sigma}_1, \sigma_2)}. \tag{C.9}$$

2.

$$\Big|\sum_{t=0}^\infty (1-\delta)\delta^t \mathbf{1}\{h^\infty_t = a_1\} - \alpha_1^*(a_1)\Big| < \frac{\epsilon}{2(2\chi + \epsilon)} \textit{ for every } h^\infty \in \mathcal{H}^{(\widehat{\sigma}_1, \sigma_2)} \textit{ and } a_1 \in A_1. \tag{C.10}$$

3.

$$\mathbb{E}^{(\widehat{\sigma}_\theta, \sigma_2)}\Big[\#\Big\{t\Big|d\big(\alpha_1^*\big\|\alpha_1(\cdot|h^t)\big) > \epsilon^2/2\Big\}\Big] < T. \tag{C.11}$$

This is crucial for proving statement 3 of Theorem 1' as it implies the following corollary:

Corollary C.1. *If $\lambda \in \underline{\Lambda}$ and δ is large enough, then for every strategy profile σ, there exists $\widehat{\sigma}_1 : \mathcal{H} \to \Delta(\overline{A}_1)$, such that if player 2's beliefs are updated according to σ and player 1 deviates to $\widehat{\sigma}_1$, then player 1 can achieve the following three objectives simultaneously:*

1. *Player 2's posterior likelihood ratio vector belongs to $\underline{\Lambda}$ in every period.*

2. *The discounted average frequency of every $a_1 \in A_1$ is approximately $\alpha_1^*(a_1)$.*

3. *Player 2's prediction about player 1's action is close to α_1^* in all but a bounded number of periods.*

To see this, let $\psi_\theta = \psi_\theta^*$ for all $\theta \in \Theta^b_{(\alpha_1^*,\theta^*)}$. When $\lambda \in \underline{\Lambda}$, there exists $\psi_\theta \in \mathbb{R}_+$ for every $\theta \notin \Theta^b_{(\alpha_1^*,\theta^*)}$ such that $\lambda \in \underline{\Lambda}(\psi, 1)$. In what follows, I show Proposition C.1 in three steps. The remaining proof of statement 3 after establishing Proposition C.1 is in Online Appendix B, the ideas of which are summarized in section **??**

Proof of Proposition C.1: Step 1 Let $\mathcal{P}^{(\alpha_1^*,\sigma_2)}$ be the probability measure over $\mathcal{H}$ when player 1 plays α_1^* in every period and player 2 plays according to σ_2. Let $\chi(h^t) \equiv \sum_{i=1}^m \lambda_i(h^t)/\psi_i$. When $\lambda \in \underline{\Lambda}(\psi, \chi)$, we have $\chi(h^0) < \chi$. Let $\{\mathcal{F}^t\}_{t\in\mathbb{N}}$ be the filtration induced by the public history. Since $\{\lambda_i(h^t), \mathcal{P}^{(\alpha_1^*,\sigma_2)}, \mathcal{F}^t\}_{t\in\mathbb{N}}$ is a non-negative supermartingale for every $i \in \{1, 2, ..., m\}$, 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. The Doob's Upcrossing Inequality implies:

$$\mathcal{P}^{(\alpha_1^*,\sigma_2)}\Big\{U(\chi, \chi + \frac{\epsilon}{2}) \geq 1\Big\} \leq \frac{2\chi}{2\chi + \epsilon}. \tag{C.12}$$

Let $\widetilde{\mathcal{H}}^\infty$ be the set of infinite histories that $\chi_t \leq \chi + \frac{\epsilon}{2}$ for all periods. According to (C.12), it occurs with probability at least $\frac{\epsilon}{2\chi+\epsilon}$.

Proof of Proposition C.1: Step 2 I show that when δ is close enough to 1, 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 close to α_1^*. For fixed $a_1 \in \overline{A}_1$, 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 C.1. *For any $\varepsilon > 0$, there exists $\overline{\delta} \in (0, 1)$, such that for all $\delta \in (\overline{\delta}, 1)$,*

$$\limsup_{\delta\to 1} \mathcal{P}^{(\alpha_1^*,\sigma_2)}\Big(\Big|\sum_{t=0}^{+\infty}(1-\delta)\delta^t X_t - \alpha_1^*(a_1)\Big| \geq \varepsilon\Big) \leq \frac{\varepsilon}{n}. \tag{C.13}$$

PROOF OF LEMMA C.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:

$$\sup_{x\in\mathbb{R}} |F_k(x)-\Phi(x)| \leq C_0\sum_{n=1}^{k}|X_{k,n}|^3 \sim C_1(1-\delta)^{\frac{3}{2}},$$

where C_0 and C_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 $\delta\to 1$.

Therefore, 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 $k>K$,

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

Lemma C.1 is obtained by taking $k\to\infty$. □

Proof of Proposition C.1: Step 3 According to Lemma C.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. $$\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\mathcal{H}^{\infty}_{\varepsilon,a_1}(\delta))\geq 1-\varepsilon/n, \tag{C.14}$$

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 (C.14):

$$\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\mathcal{H}^{\infty}_{\varepsilon}(\delta))\geq 1-\varepsilon. \tag{C.15}$$

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

$$\widehat{\mathcal{H}}^{\infty}\equiv\widetilde{\mathcal{H}}^{\infty}\bigcap\mathcal{H}^{\infty}_{\varepsilon}(\delta), \tag{C.16}$$

we have:

$$\mathcal{P}^{(\alpha_1^*,\sigma_2)}(\widehat{\mathcal{H}}^{\infty}) \geq \frac{\epsilon}{2(2\chi+\epsilon)} \tag{C.17}$$

According to Gossner (2011),

$$\mathbb{E}^{(\alpha_1^*,\sigma_2)}\Big[\sum_{\tau=0}^{+\infty} d(\alpha_1^*||\alpha_1(\cdot|h^\tau))\Big] \leq -\log\mu(\alpha_1^*). \tag{C.18}$$

Since KL divergence is non-negative, the Markov Inequality implies:

$$\mathbb{E}^{(\alpha_1^*,\sigma_2)}\Big[\sum_{\tau=0}^{+\infty} d(\alpha^*||\alpha(\cdot|h^\tau))\Big|\widehat{\mathcal{H}}^{\infty}\Big] \leq -\frac{2(2\chi+\epsilon)\log\mu(\alpha_1^*)}{\epsilon}. \tag{C.19}$$

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

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

Let $\widehat{\sigma}_1 : \mathcal{H} \to \Delta(\overline{A}_1)$ 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 at most:

$$T \equiv \Big\lceil -\frac{4(2\chi+\epsilon)\log\mu(\alpha_1^*)}{\epsilon^3}\Big\rceil. \tag{C.20}$$

Since T is independent of δ, the three steps together imply Proposition C.1.

Summary of Remaining Steps: Proposition C.1 and Corollary C.1 do not directly imply that type θ^* can guarantee $v_{\theta^*}(\alpha_1^*, u_1, u_2)$ for all u_1. This is because due to the correlations between player 1's action and the state, player 2s may not have incentives to play a_2^* despite $\lambda \in \underline{\Lambda}$ and player 1's average action is close to α_1^*. I address this issue in Online Appendix B, with the main ideas described below. Suppose $\lambda \in \underline{\Lambda}$,

1. If all entries of λ except for at most one is sufficiently small, then player 2 has a strict incentive to play a_2^* when player 1's average action is close to α_1^*. Let Λ^0 be the set of beliefs with this feature. By construction, one can directly apply Proposition C.1 to establish the commitment payoff theorem.

2. If 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 at that history must be significantly different. This implies that player 1's action at that history must be very informative about his type, in which case he can pick a particular action that induces player 2 to learn. I show that for every $\lambda \in \underline{\Lambda}$, there exists a finite integer $K(\lambda)$ and a strategy for type θ^* player 1 such that if type θ^* player 1 follows this strategy, then after at most $K(\lambda)$ such periods, player 2's belief about his type is in Λ^0, which concludes the proof.

D Proof of Theorems 2 and 3

I show Theorems 2 and 3 under a simplifying assumption that player 2 can only observe player 1's past actions, which conveys the key ideas and intuition. The full proof is relegated to Online Appendices C and D.

D.1 Partition of Θ

For every $\phi \in \Delta(\Theta)$ and $\alpha_1 \in \Delta(A_1)$, let $\mathcal{D}(\phi, \alpha_1) \equiv u_2(\phi, \alpha_1, \overline{a}_2) - u_2(\phi, \alpha_1, \underline{a}_2)$. Let

$$\Theta_g \equiv \Big\{\theta \Big| \mathcal{D}(\theta, \overline{a}_1) \geq 0 \text{ and } \theta \in \Theta^*\Big\}, \quad \Theta_p \equiv \Big\{\theta \Big| \mathcal{D}(\theta, \overline{a}_1) < 0 \text{ and } \theta \in \Theta^*\Big\} \tag{D.1}$$

and $\Theta_n \equiv \Theta \backslash \Theta^*$. One can verify that first, $\{\Theta_g, \Theta_p, \Theta_n\}$ is a partition of Θ, and second, $\Theta^* = \Theta_g \cup \Theta_p$. I focus on the nontrivial case in which neither Θ_g nor $\Theta_p \cup \Theta_n$ is empty. I show the following lemma:

Lemma D.1. *When u_1 and u_2 satisfy Assumption 2:*

1. *If $\theta_g \in \Theta_g$, $\theta_p \in \Theta_p$, and $\theta_n \in \Theta_n$, then $\theta_g \succ \theta_p$, $\theta_p \succ \theta_n$ and $\theta_g \succ \theta_n$.*
2. *If both Θ_p and Θ_n are nonempty, then $\mathcal{D}(\theta_n, \overline{a}_1) < 0$ for every $\theta_n \in \Theta_n$.*

Proof. To show statement 1, first consider the case in which $\Theta_p \neq \{\varnothing\}$. Since $\mathcal{D}(\theta_g, \overline{a}_1) \geq 0$, $\mathcal{D}(\theta_p, \overline{a}_1) < 0$, and u_2 has SID in θ and a_2, we have $\theta_g \succ \theta_p$. Since $u_1(\theta_p, \overline{a}_1, \overline{a}_2) > u_1(\theta_p, \underline{a}_1, \underline{a}_2)$, $u_1(\theta_n, \overline{a}_1, \overline{a}_2) \leq u_1(\theta_n, \underline{a}_1, \underline{a}_2)$, and u_1 has SID in θ and (a_1, a_2), we have $\theta_p \succ \theta_n$. Next, consider the case in which $\Theta_p = \{\varnothing\}$. 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 $\theta_g \succ \theta_n$. To show statement 2, given $\Theta_p, \Theta_n \neq \{\varnothing\}$, u_2 has SID in θ and a_2 implies that $\mathcal{D}(\theta_n, \overline{a}_1) < \mathcal{D}(\theta_p, \overline{a}_1) < 0$. $\square$

D.2 Implication of Stage-Game MS & Two Classes of Equilibria

I derive an implication of *MS stage-game payoffs* on a *repeated MS game*. For any strategy profile $\sigma \equiv \big((\sigma_\theta)_{\theta \in \Theta}, \sigma_2\big)$ and $\theta \in \Theta$, let $\mathcal{P}^{(\sigma_\theta, \sigma_2)}$ be the probability measure over public histories induced by $(\sigma_\theta, \sigma_2)$.

Lemma D.2. *Under Assumptions 2 and 3, for every $\widehat{\theta} \succ \widetilde{\theta}$ and in every equilibrium $\sigma \equiv \big((\sigma_\theta)_{\theta \in \Theta}, \sigma_2\big)$,*

1. *if playing $\overline{a}_1$ in every period is type $\widetilde{\theta}$'s best reply against σ_2, then according to $\sigma_{\widehat{\theta}}$, strategic type $\widehat{\theta}$ plays $\overline{a}_1$ with probability 1 at every history that occurs with positive probability under $\mathcal{P}^{(\sigma_{\widehat{\theta}}, \sigma_2)}$.*
2. *if playing $\underline{a}_1$ in every period is type $\widehat{\theta}$'s best reply against σ_2, then according to $\sigma_{\widetilde{\theta}}$, strategic type $\widetilde{\theta}$ plays $\underline{a}_1$ with probability 1 at every history that occurs with positive probability under $\mathcal{P}^{(\sigma_{\widetilde{\theta}}, \sigma_2)}$.*

This lemma is implied by Theorem 1 in Liu and Pei (2018), which says that in a *1-shot signaling game* where the sender's payoff u_1 and the receiver's payoff u_2 satisfy Assumption 2, and the receiver's action choice is binary, then in every Nash equilibrium, the sender's action is nondecreasing in θ.

In the *repeated MS game* of this paper, each type of player 1 chooses $\sigma_\theta : \mathcal{H} \to \Delta(A_1)$ and induces a (discounted average) distribution over player 2s' actions. Liu and Pei (2018)'s theorem implies that if a lower type player 1 finds it optimal to *play $\overline{a}_1$ in every period* of the repeated game, then playing actions other than $\overline{a}_1$ on the equilibrium path must be suboptimal for any higher type. A similar argument applies to $\underline{a}_1$. However, it *does not* imply that at any *given history*, a higher strategic type is more likely to play $\overline{a}_1$ than a lower strategic type. This is because player 1's action affects the equilibrium being played in the continuation game.

Motivated by the above discussions, I categorize the set of equilibria into two classes. An equilibrium $\sigma \equiv \big((\sigma_\theta)_{\theta \in \Theta}, \sigma_2\big)$ is *regular* if there exists $\theta \in \Theta_p \cup \Theta_n$ such that playing $\overline{a}_1$ in every period is type θ's best reply against σ_2. Otherwise, σ is *irregular*. For future reference, let $\overline{q}$ be the probability player 2's prior belief μ attaches to commitment type $\overline{a}_1$. Let $\overline{h}^t \equiv (\overline{a}_1, \overline{a}_1, ..., \overline{a}_1)$. For given strategy profile σ, let $q_t^\sigma(\theta)$ be the (ex ante) probability of the event that the state is θ, player 1 is strategic, and the history is $\overline{h}^t$.

D.3 Analysis of Regular Equilibria

I show that first, if ϕ is optimistic, then there exists a constant $C \in \mathbb{R}_+$ such that for every $\theta \in \Theta^*$, type θ's payoff in any regular equilibrium is at least $u_1(\theta, \overline{a}_1, \overline{a}_2) - (1-\delta)C$. Second, if ϕ is pessimistic and the probability of commitment types is small enough, then player 1's payoff and on-path behavior are the same in all regular equilibria, and are given by the ones characterized in Theorem 3.

For any given regular equilibrium σ, Lemmas D.1 and D.2 imply that for every $\theta_g \in \Theta_g$ and $t \in \mathbb{N}$, type θ_g plays $\overline{a}_1$ with probability 1 at $\overline{h}^t$. Let θ^* be the *lowest* $\theta \in \Theta_p \cup \Theta_n$ such that playing $\overline{a}_1$ in every period is type θ's best reply against σ_2. In what follows, I show that $\theta^* \in \Theta_p$.

Suppose towards a contradiction that $\theta^* \in \Theta_n$. In order for type θ^*'s equilibrium payoff to be no less than his minmax payoff $u_1(\theta, \underline{a}_1, \underline{a}_2)$, player 2 needs to play $\overline{a}_2$ with probability 1 at $\overline{h}^t$ for every $t \in \mathbb{N}$. But then, type θ^*'s payoff from playing $\underline{a}_1$ in every period is at least $(1-\delta)u_1(\theta^*, \underline{a}_1, \overline{a}_2) + \delta u_1(\theta^*, \underline{a}_1, \underline{a}_2)$, which is strictly greater than $u_1(\theta^*, \overline{a}_1, \overline{a}_2)$. The latter is type θ^*'s highest possible payoff from playing $\overline{a}_1$ in every period. This leads to a contradiction which implies that $\theta^* \in \Theta_p$.

Let t^* be the smallest $t \in \mathbb{N}$ such that $q_t(\theta) = 0$ for all $\theta \in \Theta_n$. If $t^* \geq 1$, then there exists $\theta_n \in \Theta_n$ such that one of type θ_n's best reply against σ_2 is to play $\overline{a}_1$ from period 0 to t^*-1, under which his payoff is at most:

$$\sum_{t=0}^{t^*-1} (1-\delta)\delta^t u_1(\theta_n, \overline{a}_1, \alpha_{2,t}) + (1-\delta)\delta^{t^*} u_1(\theta_n, \underline{a}_1, \alpha_{2,t^*}) + \delta^{t^*+1} u_1(\theta_n, \underline{a}_1, \underline{a}_2), \tag{D.2}$$

in which $\alpha_{2,t} \in \Delta(A_2)$ is player 2's action at $\overline{h}^t$. The above payoff must be higher than type θ_n's minmax payoff $u_1(\theta_n, \underline{a}_1, \underline{a}_2)$. Since $u_1(\theta_n, \underline{a}_1, \underline{a}_2) \geq u_1(\theta_n, \overline{a}_1, \overline{a}_2)$, we have:

$$\sum_{t=0}^{t^*-1} (1-\delta)\delta^t \Big(u_1(\theta_n, \overline{a}_1, \overline{a}_2) - u_1(\theta_n, \overline{a}_1, \alpha_{2,t}) \Big) \leq (1-\delta)\delta^{t^*} \Big(u_1(\theta_n, \underline{a}_1, \alpha_{2,t^*}) - u_1(\theta_n, \overline{a}_1, \overline{a}_2) \Big). \quad \text{(D.3)}$$

Since Θ is finite, and u_1 is strictly increasing in a_2, the value of the following expression is positive and finite:

$$\max_{\theta', \theta'' \in \Theta} \Big\{ \frac{u_1(\theta', \overline{a}_1, \overline{a}_2) - u_1(\theta', \overline{a}_1, \underline{a}_2)}{u_1(\theta'', \overline{a}_1, \overline{a}_2) - u_1(\theta'', \overline{a}_1, \underline{a}_2)} \Big\}.$$

Therefore, (D.3) implies the existence of a constant $C_0 > 0$ such that for every $\theta \in \Theta^*$,

$$\sum_{t=0}^{t^*-1} (1-\delta)\delta^t \Big(u_1(\theta, \overline{a}_1, \overline{a}_2) - u_1(\theta, \overline{a}_1, \alpha_{2,t}) \Big) \leq (1-\delta) C_0. \quad \text{(D.4)}$$

For periods after t^*, I examine optimistic and pessimistic ϕ separately.

Case 1: ϕ is Optimistic For every $t \geq t^*$, player 2 does not have a strict incentive to play $\overline{a}_2$ at $\overline{h}^t$ *only if*:

$$\overline{q}\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta^*} q_{t+1}^\sigma(\theta) \mathcal{D}(\theta, \overline{a}_1) + \sum_{\theta \in \Theta^*} (q_t^\sigma(\theta) - q_{t+1}^\sigma(\theta)) \mathcal{D}(\theta, \underline{a}_1) \leq 0. \quad \text{(D.5)}$$

The LHS of (D.5) is a lower bound on the difference between player 2's expected payoff from playing $\overline{a}_2$ and $\underline{a}_2$ at $\overline{h}^t$. Since ϕ is optimistic and every type in Θ_g plays $\overline{a}_1$ with probability 1 at every $\overline{h}^t$, we know that for every $t \geq t^*$, $\sum_{\theta \in \Theta} q_t^\sigma(\theta)\mathcal{D}(\theta, \overline{a}_1) = \sum_{\theta \in \Theta^*} q_t^\sigma(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$. Therefore, (D.5) implies that:

$$\overline{q}\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) \leq \sum_{\theta \in \Theta^*} \big(q_t^\sigma(\theta) - q_{t+1}^\sigma(\theta) \big) \big(\mathcal{D}(\theta, \overline{a}_1) - \mathcal{D}(\theta, \underline{a}_1) \big),$$

or equivalently,

$$\sum_{\theta \in \Theta^*} \big(q_t^\sigma(\theta) - q_{t+1}^\sigma(\theta) \big) \geq C_1 \equiv \frac{\overline{q}\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1)}{\max_{\theta \in \Theta^*} \{ \mathcal{D}(\theta, \overline{a}_1) - \mathcal{D}(\theta, \underline{a}_1) \}}. \quad \text{(D.6)}$$

The MS condition implies that $\max_{\theta \in \Theta^*} \{ \mathcal{D}(\theta, \overline{a}_1) - \mathcal{D}(\theta, \underline{a}_1) \} > 0$. Therefore, C_1 is a strictly positive constant, which is independent of δ.

Since $\sum_{\theta \in \Theta_p} q_t^\sigma(\theta) \leq 1$ for all $t \in \mathbb{N}$, the number of periods such that $\overline{a}_2$ is not a strict best reply at $\overline{h}^t$ is no more than $\lceil 1/C_1 \rceil$. Therefore, for every $\theta \in \Theta^*$, if type θ plays $\overline{a}_1$ in every period, his loss relative to $u_1(\theta, \overline{a}_1, \overline{a}_2)$ is no more than $(1-\delta)C_0$ from period 0 to $t^* - 1$, and is no more than $1 - \delta^{\lceil 1/C_1 \rceil}$ after period t^*. As a result, his payoff in any regular equilibrium is no less than $u_1(\theta, \overline{a}_1, \overline{a}_2)$ as $\delta \to 1$.

Case 2: ϕ is Pessimistic For every regular equilibrium σ, let

$$\mathcal{I}_t^\sigma \equiv \overline{q}\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta} q_t^\sigma(\theta)\mathcal{D}(\theta, \overline{a}_1). \tag{D.7}$$

Given that $\theta^* \in \Theta_p$, which implies that $\Theta_p \neq \{\varnothing\}$, Lemmas D.1 and D.2 imply that first, $\mathcal{D}(\theta, \overline{a}_1) > 0$ only if $\theta \in \Theta_g$, and second, every type in Θ_g plays $\overline{a}_1$ with probability 1 at every $\overline{h}^t$. Therefore, $\mathcal{I}_t^\sigma$ is *nondecreasing* in t for every regular equilibrium σ. I show the following lemma:

Lemma D.3. *If ϕ is pessimistic and the probability of commitment types is small enough, then in every regular equilibrium σ, $t^* = 1$ and $\mathcal{I}_t^\sigma = 0$ for every $t \geq 1$.*

Proof of Lemma D.3: I proceed in two steps. In step 1, I show that $\mathcal{I}_t^\sigma = 0$ for every $t \geq \max\{1, t^*\}$. In step 2, I show that $t^* = 1$. The two steps together lead to the conclusion of Lemma D.3.

Step 1: First, suppose that $\mathcal{I}_t^\sigma < 0$ for some $t \geq \max\{1, t^*\}$. Since player 2's belief is a martingale, there exists $\theta_p \in \Theta_p$ such that (1) $q_t^\sigma(\theta_p) > 0$, and (2) under one of type θ_p's pure strategy best replies to σ_2, $\underline{a}_2$ is player 2's strict best reply in all subsequent periods. If type θ_p plays according to this pure strategy best reply against σ_2, then his continuation payoff at $\overline{h}^t$ is $u_1(\theta, \underline{a}_1, \underline{a}_2)$, which is his minmax payoff. He can profitably deviate at $\overline{h}^{t-1}$ by playing $\underline{a}_1$ in every period, which leads to a strictly higher stage-game payoff in period $t-1$. This leads to a contradiction.

Next, suppose that $\mathcal{I}_t^\sigma > 0$ for some $t \geq \max\{1, t^*\}$. I start from showing that player 2 has a strict incentive to play $\overline{a}_2$ at $\overline{h}^s$ for every $s \geq t$. Suppose towards a contradiction that $\overline{a}_2$ is not a strict best reply at $\overline{h}^{s_0}$ for some $s_0 \geq t$, then there exists $\theta_p \in \Theta_p$ such that type θ_p plays $a_1 \neq \overline{a}_1$ with positive probability at $\overline{h}^{s_0}$. Similar to (D.6), we have:

$$\sum_{\theta \in \Theta_p} \left(q_{s_0}^\sigma(\theta) - q_{s_0+1}^\sigma(\theta)\right) \geq \frac{\mathcal{I}_{s_0}^\sigma}{\max_{\theta \in \Theta_p}\{\mathcal{D}(\theta, \overline{a}_1) - \mathcal{D}(\theta, \underline{a}_1)\}}. \tag{D.8}$$

Since all types in Θ_g play $\overline{a}_1$ with probability 1, type θ_p's payoff at $\overline{h}^{s_0}$ is at most $(1-\delta)u_1(\theta_p, \underline{a}_1, \overline{a}_2) + \delta u_1(\theta_p, \underline{a}_1, \underline{a}_2)$. Since $u_1(\theta_p, \underline{a}_1, \underline{a}_2) < u_1(\theta_p, \overline{a}_1, \overline{a}_2)$, type θ_p has an incentive to play $\underline{a}_1$ instead of $\overline{a}_1$ at $\overline{h}^{s_0}$ only if there exists $s_1 > s_0$ such that $\overline{a}_2$ is not a strict best reply at $\overline{h}^{s_1}$. Iterate this process, one can obtain an *infinite sequence* $\{s_0, s_1, s_2, ...\}$ such that for every $i \in \mathbb{N}$,

$$\sum_{\theta \in \Theta_p} \left(q_{s_i}^\sigma(\theta) - q_{s_i+1}^\sigma(\theta)\right) \geq \frac{\mathcal{I}_{s_i}^\sigma}{\max_{\theta \in \Theta_p}\{\mathcal{D}(\theta, \overline{a}_1) - \mathcal{D}(\theta, \underline{a}_1)\}}.$$

Since $\mathcal{I}_s^\sigma$ is non-decreasing in s, the RHS is bounded away from 0. Since $\sum_{\theta \in \Theta_p} q_s^\sigma(\theta) \leq 1$ for every $s \in \mathbb{N}$, this sequence *ends in finite time*. This contradiction implies that $\overline{a}_2$ is a strict best reply at $\overline{h}^s$ for every $s \geq t$.

Therefore, for every $\theta \in \Theta^*$ and $t \in \mathbb{N}$, by playing $\overline{a}_1$ in every period, type θ's payoff at $\overline{h}^t$ is no less than $u_1(\theta, \overline{a}_1, \overline{a}_2) - (1-\delta)C_0$, where C_0 is the constant defined in (D.4). Since ϕ is pessimistic, $\mathcal{I}_0^\sigma < 0$ when the total probability of commitment types is small enough. Suppose $\mathcal{I}_t^\sigma > 0$ for some $t \geq \max\{1, t^*\}$, then there exists $\theta_p \in \Theta_p$ that plays $a_1 \neq \overline{a}_1$ with positive probability at $\overline{h}^s$ for some $s < t$, after which his continuation payoff is $u_1(\theta_p, \underline{a}_1, \underline{a}_2)$. This payoff is strictly less compared to his continuation payoff from playing $\overline{a}_1$ in every period, which leads to a contradiction. This implies that $\mathcal{I}_t^\sigma = 0$ for every $t \geq \max\{1, t^*\}$.

Step 2: I show that $t^* = 1$. Suppose towards a contradiction that $t^* > 1$. Then there exists $\theta_n \in \Theta_n$ whose best response to σ_2 is to play $\overline{a}_1$ until period $t^* - 1$. Since $\mathcal{I}_t^\sigma = 0$ for all $t \geq t^*$ and $\mathcal{I}_t^\sigma$ is non-decreasing in t, then $\mathcal{I}_t^\sigma < 0$ for all $t < t^*$. This implies that type θ_n's payoff following his equilibrium strategy is at most $(1-\delta^{t^*-1})u_1(\theta_n, \overline{a}_1, \underline{a}_2) + \delta^{t^*-1}u_1(\theta_n, \underline{a}_1, \underline{a}_2)$, which is strictly less than his minmax payoff $u_1(\theta_n, \underline{a}_1, \underline{a}_2)$. This leads to a contradiction. □

Lemma D.3 implies that in every regular equilibrium under a pessimistic ϕ, first, all types in Θ_n play $\underline{a}_1$ in every period. Second, player 2's posterior after observing $\overline{a}_1$ is such that she is indifferent between $\overline{a}_2$ and $\underline{a}_2$ against $\overline{a}_1$. Third, $\mathcal{I}_t^\sigma = 0$ for all $t \geq 1$ implies that if player 1 plays $\overline{a}_1$ in period 0 on the equilibrium path, then he plays $\overline{a}_1$ at $\overline{h}^t$ for every $t \in \mathbb{N}$.

According to Lemma D.2, for every regular equilibrium σ, there exists a cutoff state $\theta^* \in \Theta^*$ such that strategic player 1 plays $\overline{a}_1$ in every period with probability 1 when $\theta \succ \theta^*$. When the total probability of commitment types is sufficiently small, and given that player 1's characteristics and the state are *independent*, the cutoff state equals $\theta^*(\phi)$, defined in (3.10), for all regular equilibria under ϕ. According to the definition of $\theta^*(\phi)$, we have

$$\sum_{\theta' \succeq \theta^*(\phi)} \phi(\theta')\mathcal{D}(\theta', \overline{a}_1) < 0 \quad \text{and} \quad \sum_{\theta' \succ \theta^*(\phi)} \phi(\theta')\mathcal{D}(\theta', \overline{a}_1) \geq 0. \tag{D.9}$$

This implies that $\theta^*(\phi) \in \Theta_p$. Since $\mathcal{I}_1^\sigma = 0$, $\sum_{\theta' \succ \theta^*(\phi)} \phi(\theta')\mathcal{D}(\theta', \overline{a}_1) \geq 0$, and $\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) > 0$, the cutoff type $\theta^*(\phi)$ plays $\overline{a}_1$ with strictly positive probability in period 0. This implies that the cutoff type's equilibrium payoff is bounded from above by $u_1(\theta^*(\phi), \overline{a}_1, \overline{a}_2)$.

For every $a_1^* \in A_1^g \backslash \{\overline{a}_1, \underline{a}_1\}$, after observing player 1 playing a_1^* in period 0, I show that player 2 must be indifferent between $\overline{a}_2$ and $\underline{a}_2$ against a_1^*. This is because if player 2 strictly prefers $\overline{a}_2$, then a similar argument to Step 1 in the proof of Lemma D.3 implies that type $\theta^*(\phi)$'s discounted average payoff by playing a_1^* in every period converges to $u_1(\theta^*(\phi), a_1^*, \overline{a}_2)$ as $\delta \to 1$, which is strictly greater than his payoff from playing $\overline{a}_1$ in every period. This leads to a contradiction. If player 2 strictly prefers $\underline{a}_2$, then according to the definition of A_1^g, there exists $\theta \prec \theta^*(\phi)$ such that strategic type θ plays a_1^* with positive probability in period 0, and his continuation payoff in period 1 is no more than $u_1(\theta, \underline{a}_1, \underline{a}_2)$. Therefore, this type strictly prefers to play $\underline{a}_1$ in period 0, which leads to a contradiction.

Player 2's indifference after observing $a_1^* \in A_1^g \backslash \{\overline{a}_1, \underline{a}_1\}$ implies that in any regular equilibrium, if a strategic type $\theta \preceq \theta^*(\phi)$ plays a_1^* with positive probability in period 0, then he plays a_1^* with probability 1 at every subsequent history such that he has played a_1^* in all previous periods.

When the total probability of commitment types is sufficiently small and given the MS condition on stage-game payoffs, type $\theta^*(\phi)$ plays every action in $A_1^g \backslash \{\overline{a}_1, \underline{a}_1\}$ with positive probability. Since player 2 is indifferent between $\overline{a}_2$ and $\underline{a}_2$ against a_1^* after observing $a_1^* \in A_1^g \backslash \{\overline{a}_1, \underline{a}_1\}$, the probability with which strategic type $\theta^*(\phi)$ plays actions in $A_1^g \backslash \{\overline{a}_1, \underline{a}_1\}$ converges to 0 as the probability of commitment type goes to 0. This together with (4.9) implies that strategic type $\theta^*(\phi)$ plays $\underline{a}_1$ with positive probability in period 0.

When the total probability of commitment types is small enough, player 2 has no incentive to play $\overline{a}_2$ after observing $\underline{a}_1$, which implies that type $\theta^*(\phi)$'s equilibrium payoff equals its minmax payoff. Therefore, playing $\underline{a}_1$ in every period is type $\theta^*(\phi)$'s best reply against σ_2. Lemma D.2 then implies that types lower than $\theta^*(\phi)$ plays $\underline{a}_1$ with probability 1 in every period. The cutoff type's equilibrium payoff pins down the discounted average frequency with which player 2 plays $\overline{a}_2$ conditional on player 1 plays $\overline{a}_1$ in every period, given by $r(\phi)$. This together with types $\theta \succ \theta^*(\phi)$'s on-path behavior pins down every type's payoff in all regular equilibria.

D.4 Analysis of Irregular Equilibria

I show that first, if ϕ is optimistic, then there exists $C \in \mathbb{R}_+$ such that type $\theta \in \Theta^*$'s payoff in any irregular equilibrium is at least $u_1(\theta, \overline{a}_1, \overline{a}_2) - (1-\delta)C$. This together with the conclusion on regular equilibria establishes Theorem 2. Second, if ϕ is pessimistic, then irregular equilibria do not exist. Therefore, the unique payoff and unique on-path behavior in regular equilibria apply to *all* equilibria.

Recall that $t^* \in \mathbb{N}$ is the smallest $t \in \mathbb{N}$ such that $q_t^\sigma(\theta) = 0$ for all $\theta \in \Theta_n$. Similar to the analysis of regular equilibria, there exists a constant $C_0 > 0$ such that in every equilibrium for every $\theta \in \Theta^*$, type θ's loss from period 0 to t^* is no more than $(1-\delta)C_0$ relative to $u_1(\theta, \overline{a}_1, \overline{a}_2)$. I show the following lemma:

Lemma D.4. *There exists $C_2 > 0$ such that for every $t \geq t^*$, if $\sum_{\theta \in \Theta^*} q_t^\sigma(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$ and $\overline{a}_2$ is not a strict best reply at $\overline{h}^t$, then*

$$\sum_{\theta \in \Theta^*} \left(q_t^\sigma(\theta) - q_{t+1}^\sigma(\theta)\right) \geq C_2. \tag{D.10}$$

Proof of Lemma D.4: By definition, when $t \geq t^*$, $q_t^\sigma(\theta) = 0$ for every $\theta \notin \Theta^*$. If $\overline{a}_2$ is not a strict best reply at $\overline{h}^t$, then

$$\overline{q}\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \sum_{\theta \in \Theta^*} q_{t+1}^\sigma(\theta)\mathcal{D}(\theta, \overline{a}_1) + \sum_{\theta \in \Theta^*} (q_t^\sigma(\theta) - q_{t+1}^\sigma(\theta))\mathcal{D}(\theta, \underline{a}_1) \leq 0, \tag{D.11}$$

where the LHS is a lower bound on player 2's relative payoff from playing $\overline{a}_2$ instead of $\underline{a}_2$ at history $\overline{h}^t$.

Inequality (D.11) can be rewritten as:

$$\overline{q}\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1) + \underbrace{\sum_{\theta \in \Theta^*} q_t^{\sigma}(\theta)\mathcal{D}(\theta, \overline{a}_1)}_{\geq 0} + \sum_{\theta \in \Theta^*} (q_t^{\sigma}(\theta) - q_{t+1}^{\sigma}(\theta))\Big(\mathcal{D}(\theta, \underline{a}_1) - \mathcal{D}(\theta, \overline{a}_1)\Big) \leq 0, \tag{D.12}$$

which implies:

$$\sum_{\theta \in \Theta^*} \big(q_t^{\sigma}(\theta) - q_{t+1}^{\sigma}(\theta)\big) \geq C_2 \equiv \frac{\overline{q}\mathcal{D}(\phi_{\overline{a}_1}, \overline{a}_1)}{\max_{\theta \in \Theta^*}\{\mathcal{D}(\theta, \overline{a}_1) - \mathcal{D}(\theta, \underline{a}_1)\}}. \tag{D.13}$$

□

Next, I establish a uniform lower bound on player 2's posterior belief about the state at $\overline{h}^t$.

Lemma D.5. *In every irregular equilibrium,* $\sum_{\theta \in \Theta^*} q_t^{\sigma}(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$ *for every* $t \in \mathbb{N}$.

Proof of Lemma D.5: The definition of irregular equilibrium implies the existence of $t \in \mathbb{N}$ such that $q_t^{\sigma}(\theta) = 0$ for all $\theta \in \Theta_p \cup \Theta_n$. Let $\bar{t}$ be the smallest of such t. By definition, $\sum_{\theta \in \Theta^*} q_t^{\sigma}(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$ for every $t \geq \bar{t}$.

Suppose towards a contradiction that there exists $t \leq \bar{t}$ such that $\sum_{\theta \in \Theta^*} q_t^{\sigma}(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0$. Let $\widehat{t}$ be the largest of such t. By definition, $\sum_{\theta \in \Theta^*} q_{\widehat{t}}^{\sigma}(\theta)\mathcal{D}(\theta, \overline{a}_1) < 0$, but $\sum_{\theta \in \Theta^*} q_t^{\sigma}(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$ for all $t > \widehat{t}$. Consider player 1's incentives at history $\overline{h}^{\widehat{t}}$.

1. By definition, there exists $a_1 \neq \overline{a}_1$ that is played with positive probability by some types in Θ^* at $\overline{h}^{\widehat{t}}$, such that player 2's belief at $(\overline{h}^{\widehat{t}}, a_1)$ is pessimistic.

2. For every $\theta \in \Theta^*$ such that $q_{\widehat{t}}^{\sigma}(\theta) > 0$, if type θ plays $\overline{a}_1$ in every period, then his continuation payoff at $\overline{h}^{\widehat{t}}$ is at least $u_1(\theta, \overline{a}_1, \overline{a}_2) - (1-\delta)C_0 - (1-\delta^{1/C_2})$, which converges to 1 as $\delta \to 1$.

The two parts together imply that when δ is close enough to 1, there exists $\theta \in \Theta^*$ with $q_{\widehat{t}}^{\sigma}(\theta) > 0$, such that type θ receives a strictly higher continuation payoff by playing $\overline{a}_1$ at $\overline{h}^{\widehat{t}}$, compared to playing a_1 at $\overline{a}_1$. This violates his incentive constraint, which leads to a contradiction. □

D.5 Summary

When ϕ is optimistic, Lemma D.5 implies that $\sum_{\theta \in \Theta^*} q_t^{\sigma}(\theta)\mathcal{D}(\theta, \overline{a}_1) \geq 0$ for every $t \in \mathbb{N}$ and in every irregular equilibrium. Lemma D.4 then implies that for every $\theta \in \Theta^*$, type θ's guaranteed payoff in any irregular equilibrium is no less than $u_1(\theta, \overline{a}_1, \overline{a}_2)$ as $\delta \to 1$. This together with the analysis of regular equilibria establishes Theorem 2. When ϕ is pessimistic, Lemma D.5 implies that irregular equilibria do not exist and all equilibria are regular. Therefore, player 1's equilibrium payoff and on-path behavior in regular equilibria (see section D.3) are player 1's unique payoff and unique on-path behavior for all equilibria.

D.6 Overview of Full Proof

When player 2's strategy depends on her predecessors' actions, extra complications arise in the analysis of irregular equilibria, such as the proof of Lemma D.5. This is because conditional on player 1 playing $\overline{a}_1$ in every period, there may not exist a *last history* at which player 2's posterior belief about the state is pessimistic.

To overcome this challenge, I show that every time a switching from a pessimistic to an optimistic belief happens, strategic types in Θ_p must be separating from commitment type $\overline{a}_1$ with an ex ante probability bounded from below. This implies that such switching happens at most in a finite number of times conditional on every realized path of play. On the other hand, strategic types in Θ_p only have incentives to separate at those switching histories when his continuation payoff from imitating type $\overline{a}_1$ is low. This implies that there exists at least another switching following that history, meaning that such switching happens infinitely many times if it happens once. This leads to a contradiction.

E Application: Food and Drug Industry

I explain how to map the application of the food and drug industry to my product choice game framework. A firm privately observes a perfectly persistent state $\theta \in \{\theta_l, \theta_h\}$, with $0 < \theta_l < \theta_h < 1$. This is interpreted as the quality of the firm's production technology or the quality of its upstream supplier, which is persistent over time and affects the quality of its products. In period $t \in \mathbb{N}$, the firm produces a continuum of goods indexed by $i \in [0, 1]$ at cost normalized to 0. An individual product's *quality* depends on two factors: $y_{i,t} \in \{-1, 1\}$ that can be revealed via conducting inspections, and $z_{i,t} \in \{-1, 1\}$ that is omitted by inspections. A consumer's willingness to pay for product i equals $y_{i,t} + z_{i,t}$. I assume $\{y_{i,t}, z_{i,t}\}_{i\in[0,1],t\in\mathbb{N}}$ are i.i.d. across product and across time, with $\Pr(y_{i,t} = 1) = \Pr(z_{i,t} = 1) = \theta$.

The firm chooses whether to conduct inspection in each period. Inspection costs the firm $\varepsilon > 0$, after which it observes $\{y_{i,t}\}_{i\in[0,1]}$. Based on these results, the firm can improve the quality of its defective products. This is modeled as changing $y_{i,t} = -1$ to $y_{i,t} = 1$ at cost $c > 0$ per unit mass. There is a unit mass of consumers, each has unit demand and chooses whether to buy or not at an exogenous price $p \in (0, 2)$. They can observe whether the firm has sold defective products in all previous periods, but cannot directly observe the firm's behavior in the current period.

Focusing on the case in which both ε and c are strictly positive but sufficiently small compared to p, one can argue that conditional on conducting inspection in period t, the firm strictly prefers to improve all of its defective products. This is because otherwise, the firm can strictly increase its stage-game payoff by not conducting inspection. Therefore, one can write the stage-game as:

$\theta = \theta_h$	Buy	Not Buy
Inspection	$p - c_h, 2\theta_h - p$	$-c_h, 0$
Not	$p, 4\theta_h - 2 - p$	$0, 0$

$\theta = \theta_l$	Buy	Not Buy
Inspection	$p - c_l, 2\theta_l - p$	$-c_l, 0$
Not	$p, 4\theta_l - 2 - p$	$0, 0$

where $c_h \equiv \varepsilon + c(1-\theta_h)$ and $c_l \equiv \varepsilon + c(1-\theta_l)$ are the costs of inspection in the high and the low state, respectively. The game satisfies Assumption 1 for generic θ_h and θ_l. It satisfies Assumption 2 since $c_h, c_l > 0$, the firm benefits from consumers' purchases, consumers are more likely to purchase when θ is high or the probability of quality control is high, and $c_h < c_l$ given that $\theta_h > \theta_l$. This game can be mapped into the example in section 2.1 when:

$$2\theta_h - p > 0 > \max\{2\theta_l - p, 4\theta_h - 2 - p\} \quad \Leftrightarrow \quad \frac{2+p}{4} > \theta_h > \frac{p}{2} > \theta_l,$$

namely, consumers prefer to buy only when $\theta = \theta_h$ and the firm conducts quality control. Under these conditions, Theorem 3 suggests that when consumers' prior attaches probability less than $\frac{p-2\theta_l}{2(\theta_h-\theta_l)}$ to state θ_h, then in every equilibrium, the strategic high-quality firm conducts inspections in every period and receives discounted average payoff $c(\theta_h - \theta_l)$; the strategic low-quality firm receives discounted average payoff 0 by mixing between conducting inspections in every period and not conducting inspections in every period.

References

[1] Adhvaryu, Achyuta (2014) "Learning, Misallocation, and Technology Adoption: Evidence from New Malaria Therapy in Tanzania," *Review of Economic Studies*, 81(4), 1331-1365.

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

[3] Bai, Jie (2018) "Melons as Lemons: Asymmetric Information, Consumer Learning and Quality Provision," Working Paper.

[4] Bai, Jie, Lodovica Gazze and Yukun Wang (2019) "Collective Reputation in Trade: Evidence from Chinese Dairy Industry," Working Paper.

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

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

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

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

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

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

[13] 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.

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

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

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

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

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

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

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

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

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

[23] Liu, Shuo and Harry Pei (2018) "Monotone Equilibria in Signalling Games," Working Paper.

[24] Luong, Tuan Anh, Ce Shi and Zheng Wang (2019) "The Impact of Media on Trade: Evidence from the 2008 China Milk Scandal," Working Paper.

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

[26] Mailath, George and Larry Samuelson (2013) "Reputations in Repeated Games," Handbook of Game Theory with Economic Applications, Elsevier.

[27] Miklos-Thal, Jeanine and Juanjuan Zhang (2013) "Demarketing to Manage Consumer Quality Inferences," *Journal of Marketing Research*, 50(1), 55-69.

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

[29] Pęski, Marcin (2014) "Repeated Games with Incomplete Information and Discounting," *Theoretical Economics*, 9, 651-694.