Trust and Betrayals:
Reputational Payoffs and Behaviors without Commitment

Harry PEI*

February 12, 2020

I study a repeated game in which a patient player (e.g., a seller) wants to win the trust of some myopic opponents (e.g., buyers) but can strictly benefit from betraying them. Her benefit from betrayal is strictly positive and is her persistent private information. I provide a tractable formula for every type of patient player's highest equilibrium payoff. Her persistent private information affects this payoff only through the lowest benefit in the support of her opponents' prior belief. I also show that in every equilibrium which is optimal for the patient player, her behavior is nonstationary, and her long-run action frequencies are pinned down for all except for two types. Conceptually, my payoff-type approach incorporates a realistic concern that no type of reputation-building player is immune to reneging temptations. Compared to commitment-type models, the incentive constraints for all types of patient player lead to a sharp characterization of her highest attainable payoff and novel predictions on her behaviors.

Keywords: rational reputational types, lack-of-commitment problem, equilibrium behavior, reputation
JEL Codes: C73, D82, D83

1 Introduction

Trust is essential in many socioeconomic activities, yet it is also susceptible to opportunism and exploitation. To illustrate this, consider the example of a supplier who promises his clients about on-time deliveries. After the client agrees to purchase and makes a relationship-specific investment, the supplier has an incentive to delay in order to save cost. Similarly, sellers try to convince consumers of their high quality standards. But after receiving upfront payments, they may be tempted to undercut quality, especially on aspects that are hard to verify.

The common theme in these examples is a *lack-of-commitment problem* faced by suppliers and sellers. As a response, these agents build reputations for being trustworthy, from which they derive benefits in the future. In practice, a key challenge to reputation building is that all agents face temptations to renege, including those *role-models* that others wish to imitate. As a result, the heterogeneity across agents is more about how much temptation they are facing, rather than whether they are facing temptations or not. For example, all firms can save costs by undercutting quality, but their costs can be heterogenous due to differences in production technologies. This issue has not been addressed in the classic theories of Sobel (1985), Fudenberg and Levine (1989,1992),

*Department of Economics, Northwestern University. I thank Daron Acemoglu, Ricardo Alonso, Guillermo Caruana, Mehmet Ekmekci, Jeff Ely, Drew Fudenberg, Olivier Gossner, Bruno Jullien, Aditya Kuvalekar, Elliot Lipnowski, Debraj Ray, Bruno Strulovici, Saturo Takahashi, Jean Tirole, Juuso Toikka, Tristan Tomala, Alex Wolitzky, and Muhamet Yildiz for helpful comments.

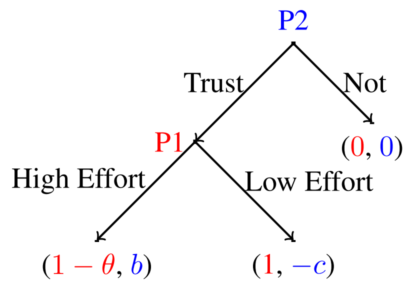

Figure 1: The stage game, where $\theta \in (0, 1)$, $b > 0$, $c > 0$

and Benabou and Laroque (1992), in which agents of some types are exogenously assumed to play prespecified strategies, and others can establish their reputations by imitating those *commitment types*.

To incorporate these realistic concerns, I introduce a reputation model *without* commitment types, with all types of the reputation-building player having strict incentives to betray others' trust. My results suggest that the *lack-of-commitment for all types* (including those that others want to imitate) introduces novel incentive constraints, which lead to a sharp characterization of patient reputation-building player's highest equilibrium payoff and novel predictions on her equilibrium behaviors. These results are not obtained in commitment-type reputation models due to the fact that commitment types face no incentive constraints.

To highlight the economic forces at work, my baseline model features a *sequential-move trust game* played repeatedly between a patient player 1 (e.g., a seller) and an infinite sequence of player 2s (e.g., buyers), arriving one in each period and each plays the game only once.[1] The stage game is depicted in Figure 1, in which player 1 wants to win her opponents' trust, but has a strict incentive to betray them (i.e., choosing low effort) once trust is granted. Her cost of exerting high effort is perfectly persistent and is her private information, which I call her *type*. Every player 2 observes the outcomes of all past interactions, and prefers to trust player 1 if and only if he expects high effort to be chosen with probability above some cutoff.

Theorem 1 characterizes every type of patient player's highest equilibrium payoff,[2] which is the product of her *complete information Stackelberg payoff* and an *incomplete-information multiplier*. The latter is a sufficient statistic for the effects of incomplete information. It depends only on the *lowest cost of effort* in the support of player 2s' prior belief, but *not* on the other types and the probability of each type. My formula for every type's highest equilibrium payoff originates from a constrained optimization problem, which unveils how the patient player's highest attainable payoff is limited by the lowest-cost type's incentive constraints (i.e., her payoff

[1] My results are robust when players move *simultaneously* in the stage game, and are robust under perturbations of player 2's discount factor (see section 5.1 and Appendix D). The assumption that buyers are myopic fits into applications such as durable good markets where each buyer has unit demand, and online platforms such as Uber where a buyer is unlikely to interact with the same seller twice.

[2] Since players move *sequentially* in the stage game, the patient player's lowest equilibrium payoff is 0. This is also the case in models with a small probability of commitment types, and is a general feature of sequential-move stage games where future player 2s can only observe the terminal node reached in each period, but cannot observe player 1's stage-game (extensive-form) strategy. This motivates my analysis of player 1's *highest equilibrium payoff* instead of her *lowest equilibrium payoff*.

cannot exceed her highest equilibrium payoff in the repeated complete information game) and the uninformed players' incentives and learning. This contrasts to commitment types in canonical reputation models who face no incentive constraints, and a tight characterization of the optimal equilibrium payoff is missing.[3]

Theorem 1 implies that for every type except for the lowest-cost one, her highest equilibrium payoff in the repeated incomplete information game is strictly greater than that in the repeated complete information game. This is somewhat puzzling since it requires every high-cost type to extract information rent (i.e., exerting low effort while receiving trust), which inevitably reveals information about her type and undermines her future informational advantage. The above tension grows as the informed player becomes more patient, since she needs to extract information rent in an increasing number of periods to obtain a given discounted average payoff strictly greater than her highest equilibrium payoff under complete information. I overcome these challenges by constructing equilibria which allow the patient player to extract information rent for a long time while respecting the incentive constraints of all types. In those equilibria, learning about the patient player's type is slow and ends in finite time. The lowest-cost type mixes between high and low effort in periods where active learning takes place, which enables the high-cost types to rebuild their reputations after milking them.

My next two results study properties of the patient player's *behavior* that apply to *all* equilibria in which she approximately attains her highest equilibrium payoff (which I call *high-payoff equilibria*). Theorem 2 shows that no type of patient player uses stationary strategies or has a completely mixed best reply in any high-payoff equilibrium. This conclusion extends to a type whose cost of exerting high effort is zero.

Theorem 2 implies that in every high-payoff equilibrium, every type of patient player has strict incentives at some on-path histories, and her equilibrium action depends nontrivially on past play. This conclusion contrasts to the Stackelberg commitment type in canonical reputation models that mechanically plays the same mixed action in every period. To strengthen my motivation for excluding Stackelberg commitment type, I allow the patient player to have arbitrary stage-game payoff functions in Proposition C.1, and show that the Stackelberg commitment behavior *cannot* arise unless the patient player is indifferent across all outcomes in the stage-game. In the buyer-seller application, it translates into a type of seller whose benefit from buyer's purchase and whose cost of exerting high effort are both zero.

Theorem 3 derives the following bounds on the patient player's action frequencies that apply to *all of her best replies* against her opponents' equilibrium strategies in *all high-payoff equilibria*. First, for every type except for the highest cost one, the relative frequency between high and low effort *cannot fall below* the ratio between their

[3] In the sequential-move stage-game, results from commitment-type reputation models conclude that a rational-type patient player's payoff cannot exceed her Stackelberg payoff and cannot fall below her minmax payoff. However, those results *do not imply* whether the Stackelberg payoff is attainable or not, what the patient player's highest equilibrium payoff is, and whether she can strictly benefit from incomplete information. In the simultaneous-move stage game, commitment-type models lead to a sharp characterization of the patient player's highest equilibrium payoff when there exists a commitment type that mechanically plays the Stackelberg action. However, the practical relevance of such mixed-strategy commitment type is somewhat questionable, which is substantiated in Proposition C.1.

probabilities in the Stackelberg action (i.e., the *critical ratio*). Second, for every type except for the lowest cost one, the relative frequency between high and low effort *cannot exceed* the aforementioned critical ratio. The two bounds together pin down the action frequencies for all types except for the highest and lowest cost types.

The significance of Theorem 3 is that the bounds apply to every *equilibrium best reply*, which lead to stronger testable implications compared to the ones in commitment-type models that only apply to the patient player's *equilibrium strategy*. This distinction is practically important when researchers can only observe a realized path of an informed player's actions, but not the entire distribution over her action paths, i.e., they can observe one of her equilibrium best replies, but cannot observe her equilibrium strategy. This is because when learning and signaling take place dynamically, an informed player's equilibrium strategy usually involves nontrivial mixing.

The conclusions of Theorems 2 and 3 rely on *all types of patient player's incentive constraints*, and therefore, are not obtained from commitment-type models. For a snapshot of the argument, suppose the lowest-cost type has a completely mixed equilibrium best reply, then exerting low effort at every history is also her equilibrium best reply. Due to the high-cost types' comparative advantages in exerting low effort, they exert low effort for sure at every on-path history. However, if they behave like this in equilibrium, then player 2s believe that low effort occurs with probability close to 1 in all future periods after observing low effort for a bounded number of periods. They will then stop trusting the patient player, leaving those high-cost types a low payoff. In general, this logic leads to an upper bound on the frequencies with which the low-cost types exert low effort under each of her equilibrium best replies.

Conversely, suppose it is optimal for a high-cost type to exert high effort at a given frequency. Then under every best reply of the lowest-cost type, her frequency of high effort must be weakly higher. In order for a high-cost type to hide behind the lowest-cost type and to extract information rent, the long-run frequencies of her actions cannot be too different from the equilibrium action frequencies of the lowest-cost type. Therefore, a lower bound on the high-cost type's equilibrium payoff leads to an upper bound on the frequency of high effort not only under each of her equilibrium strategies but also under each of her equilibrium best replies.

Related Literature: My paper contributes to the literature on reputations and repeated incomplete information games. My conceptual contribution lies both in the approach taken and in the research questions.

The canonical reputation models of Fudenberg and Levine (1989, 1992) fix the behavior of at least one type of patient player (call them *commitment types*), and derive payoff lower and upper bounds for a strategic-type patient player. When there is a type of patient player who mechanically plays her Stackelberg action, and the patient player's stage-game strategies are *statistically identified* (e.g., in simultaneous-move stage game with perfect monitoring), they show that the strategic player's equilibrium payoff approximately equals her Stackelberg payoff in all equilibria. Nevertheless, this approach is unsatisfactory from the following three perspectives:

1. The payoff upper and lower bounds are *not tight* when the patient player's stage-game strategies are *not statistically identified*, for example, when players move sequentially in the stage-game.

   In the sequential-move trust game, the canonical results imply that the patient player's payoff is between her minmax payoff and her Stackelberg payoff. However, it remains unclear whether the Stackelberg payoff is attainable or not, and if not, what the patient player's optimal equilibrium payoff is.

2. In simultaneous-move stage games, the tightness of these payoff bounds rely on the presence of Stackelberg commitment types whose behaviors require nontrivial mixing. The plausibility of such types is somewhat questionable, and so are the reputation results that rely on such types.

3. The commitment-type approach is not well-suited to answer questions related to players' reputation building *behaviors*. This is because first, canonical reputation results do not provide explicit characterization of the patient player's behavior, and there are many different behaviors that can lead to the same payoff. Second, the rational-type patient player's behavior is sensitive to the specifications of commitment types' strategies, with the latter being exogenously assumed.[4] Third, players' commitment strategies do not respond to changes in the payoff environment, such as the patient player's discount factor and her initial reputation.

Motivated by these concerns, I propose a complementary approach in which all types of patient player face strict temptations to betray, and can flexibly choose their strategies in order to maximize their payoffs. Compared to the commitment-type approach, I obtain a sharp characterization of the patient player's highest equilibrium payoff in an interesting class of games. My characterization (1) applies both to sequential-move and simultaneous-move stage games, (2) relies only on variables that have clear economic interpretations, and (3) does not depend on the presence of any commitment type.

My approach also has the advantage of raising and answering novel questions related to the patient player's *equilibrium behavior*, such as (1) how the reputational types behave when they are also tempted to betray, and (2) how other types behave in order to exploit these rational reputational types. Compared to commitment-type models in which the reputational types' behaviors are exogenous, my approach treats the reputational type as a strategic player whose behavior responds to changes in the payoff environment.[5] By exploiting the incentive constraints of different types of patient player, I derive properties of her behaviors that are true for all high-payoff equilibria, which have not been obtained in commitment-type models. By studying the behavior of the lowest-cost type, my analysis complements the canonical reputation models by evaluating which of the many commitment strategies are more plausible in the sense that they can arise from maximizing reasonable payoff functions.

[4]Weinstein and Yildiz (2016) point out the sensitivity of the game's predictions to the specifications of the commitment types' strategies. Their findings call for a careful selection of types when setting up incomplete information game models, including types that occur with very low probability. My approach follows this spirit by requiring all types to have *reasonable* payoff functions.

[5]This differs from other approaches that endogenize commitment in dynamic games, which include introducing switching costs (Caruana and Einav 2008), and perturbing players' belief hierarchies and introducing interdependent values (Weinstein and Yildiz 2016).

My results are potentially useful for future applied work both in the computation of optimal equilibria and in identifying bounds for a seller's cost of supplying high quality.

The main drawback of my payoff-type approach is that it cannot rule out low-payoff equilibria, for example, equilibria in which all types of patient player receive their minmax payoffs. This is a disadvantage relative to commitment-type models when players move simultaneously in the stage-game. It is *not* a disadvantage in sequential-move stage-games since neither approach can rule out low-payoff equilibria, in which case the only role of reputation is to allow the patient player attaining higher payoffs relative to complete information.

My paper is related to the study of repeated incomplete information games pioneered by Aumann and Maschler (1995). Hart (1985) introduces bi-martingale techniques and characterizes the set of equilibrium payoffs in repeated games with one-sided private information and no discounting. In private value environments, Shalev (1994) simplifies Hart's characterization and describes the resulting payoff set by a linear program. Pęski (2014) characterizes the set of equilibrium payoffs when players' discount factors are close to but not equal to 1. Cripps and Thomas (2003) show that when the informed player is arbitrarily patient and the uninformed player's discount factor is bounded away from 1, Shalev's characterization remains to be a *necessary condition* for being an equilibrium payoff, but it is *not sufficient* in general, i.e., some payoffs in that set are not attainable when the uninformed player is impatient. In fact, conditions that are both *necessary and sufficient* for being equilibrium payoffs, and every type of informed player's optimal equilibrium payoff, remain open questions.[6]

My Theorem 1 contributes to this literature by identifying conditions that are both necessary and sufficient for being an equilibrium payoff. I characterize every type of patient informed player's highest equilibrium payoff when her opponent's discount factor is close to or equal zero. My characterization applies to an interesting class of games, which include but not limited to product choice games, entry deterrence games, with generalizations stated in Appendix D. The constraints I identified, together with all types of patient player's individual rationality conditions, lead to a full characterization of the patient informed player's limiting equilibrium payoff set.

2 The Baseline Model

I introduce a repeated *trust game* that highlights the lack-of-commitment problem in economic interactions. Different from reputation models with commitment types, all types of the reputation-building player are rational and have qualitatively similar payoff functions. This captures, for instance, all types of sellers are tempted to undercut quality but can strictly benefit from buyers' purchases.

[6] An exception is zero-sum games, in which Aumann and Maschler (1995)'s characterization applies under any discount factor of the uninformed player(s). However, the stage-game studied in this paper is not zero sum, and therefore, is not covered by the results of Aumann and Maschler (1995). In settings with short-lived players, the belief-free equilibrium approach in Hörner, Lovo, and Tomala (2011) is not applicable. This is because at every history in which some types of the patient player can extract information rent, the short-lived player's best reply depends on his belief about the patient player's type.

In section 5, I examine the robustness of my insights under variations of my baseline model, which include simultaneous-move stage games and the uninformed player being forward-looking, and then discuss alternative applications. Generalizations of my results beyond 2×2 trust games are stated in Appendix D.

Stage Game: Consider the trust game in Figure 1 (on page 2) between an informed seller (player 1, she) and an uninformed buyer (player 2, he). The buyer moves first, deciding whether to trust the seller (action T) or not (action N). If he chooses N, then both players' payoffs are normalized to 0. If he chooses T, then the seller chooses between high effort (action H) and low effort (action L). If the seller chooses L, then her payoff is 1 and the buyer's payoff is $-c$. If the seller chooses H, then her payoff is $1-\theta$ and the buyer's payoff is b, where:

- $b > 0$ is the buyer's benefit from the seller's high effort, $c > 0$ is the buyer's loss from the seller's low effort (or *betrayal*), both of which are common knowledge among players.
- $\theta \in \Theta \equiv \{\theta_1, ...\theta_m\} \subset (0,1)$ is the seller's cost of high effort and is her private information. Without loss of generality, I assume that $0 < \theta_1 < \theta_2 < ... < \theta_m < 1$.

The unique stage-game equilibrium outcome is N and the seller's payoff is 0. This is because every type of seller has a strict incentive to choose L after the buyer plays T. If the seller can optimally commit to an action $\alpha_1 \in \Delta\{H, L\}$ before the buyer moves, then every type's optimal commitment is to play H with probability $\gamma^* \equiv \frac{c}{b+c}$ and L with probability $1-\gamma^*$. Type θ_j's optimal commitment payoff (or *Stackelberg payoff*) is

$$v_j^{**} \equiv 1 - \gamma^* \theta_j \text{ for every } j \in \{1, 2, ..., m\}, \tag{2.1}$$

with $\gamma^* H + (1-\gamma^*)L$ her *Stackelberg action* and $\gamma^*(T, H) + (1-\gamma^*)(T, L)$ her *Stackelberg outcome*.

The comparison between the seller's Nash equilibrium payoff and her Stackelberg payoff highlights a *lack-of-commitment* problem, which is of first order importance not only in business transactions (Mailath and Samuelson 2001), but also in fiscal and monetary policies (Barro 1986, Phelan 2006) and political economy (Tirole 1996).

The rest of this article sets up a repeated version of this game in which *all types of sellers have strict incentives to betray*. I examine the extent to which a patient seller can overcome her lack-of-commitment problem, as well as different types of sellers' behaviors in seller-optimal equilibria.

Repeated Game: Time is discrete, indexed by $t = 0, 1, 2, ...$ A long-lived seller interacts with an infinite sequence of buyers, arriving one in each period and each plays the game only once.

Both b and c are common knowledge among players. The seller's cost θ is perfectly persistent and is her private information. The buyers have a full support prior belief $\pi \in \Delta(\Theta)$. Outcomes of all past interactions are perfectly observed. Let $y_t \in \{N, H, L\}$ be the stage-game outcome in period t. Let $h^t = \{y_s\}_{s=0}^{t-1} \in \mathcal{H}^t$

be a public history with $\mathcal{H} \equiv \bigcup_{t=0}^{+\infty} \mathcal{H}^t$ the set of public histories.[7] Let $A_1 \equiv \{H, L\}$ and $A_2 \equiv \{T, N\}$. Let $\sigma_2 : \mathcal{H} \to \Delta(A_2)$ be the buyers' strategy. Let $\sigma_1 \equiv (\sigma_\theta)_{\theta \in \Theta}$, with type θ seller's strategy being $\sigma_\theta : \mathcal{H} \to \Delta(A_1)$, which specifies this type of seller's action choices at every public history conditional on the buyer choosing T.

The seller's stage-game payoff is denoted by $u_1(\theta, y_t)$, which depends on her type θ and the stage-game outcome $y_t \in \{N, H, L\}$. The seller's discount factor is $\delta \in (0,1)$. Type θ seller chooses σ_θ in order to maximize:

$$\mathbb{E}^{(\sigma_\theta, \sigma_2)}\Big[\sum_{t=0}^{\infty} (1-\delta)\delta^t u_1(\theta, y_t)\Big], \tag{2.2}$$

with $\mathbb{E}^{(\sigma_\theta, \sigma_2)}[\cdot]$ the expectation over $\mathcal{H}$ under the probability measure induced by $(\sigma_\theta, \sigma_2)$.

3 Results

Theorem 1 characterizes every type of patient seller's highest equilibrium payoff. Theorem 2 shows that no type of seller uses stationary strategies or completely mixed strategies in any seller-optimal equilibrium. Theorem 3 derives bounds on the seller's action frequencies that apply to all of her equilibrium best replies. These novel predictions rely on *all types of seller's incentive constraints*, i.e., their strict incentives to betray and to receive trust. These are not obtained in commitment-type models since commitment types face no incentive constraints.

3.1 Patient Seller's Optimal Equilibrium Payoff

Since there are m types, the seller's *payoff* is an m-dimensional vector $v \equiv (v_1, ..., v_m) \in \mathbb{R}^m$, in which the jth entry represents the discounted average payoff of type θ_j. Let $v^* \equiv (v_1^*, ..., v_m^*)$, with

$$v_j^* \equiv \underbrace{(1-\gamma^*\theta_j)}_{\text{Type } \theta_j\text{'s Stackelberg payoff}} \underbrace{\frac{1-\theta_1}{1-\gamma^*\theta_1}}_{\text{incomplete-information multiplier}}, \text{ for every } j \in \{1, 2, ..., m\}. \tag{3.1}$$

Theorem 1. *If π has full support, then for every $\varepsilon > 0$, there exists $\underline{\delta} \in (0,1)$ such that for every $\delta \in (\underline{\delta}, 1)$,*

1. *There exists no Bayes Nash equilibrium (BNE) in which type θ_1's payoff is strictly more than v_1^*.*
   There exists no BNE in which type θ_j's payoff is more than $v_j^* + \varepsilon$ for some $j \in \{2, ..., m\}$.

2. *There exists a sequential equilibrium in which the seller's payoff is within ε of v^*.*[8]

[7]There is *no public randomization device* in my baseline model. My results apply as long as future buyers *cannot perfectly monitor the seller's mixed actions*, which is a standard assumption in the repeated game and reputation literature. This is satisfied in my baseline model, repeated simultaneous-move games with and without public randomizations, and repeated sequential-move games in which the public randomization is realized before player 2 moves or after player 1 moves. This assumption is violated when players move sequentially in the stage-game and the public randomization is realized in between player 2's and player 1's moves.

[8]I adopt the notion of sequential equilibrium in Pęski (2014, page 658). I use different solution concepts in the two statements of Theorem 1 to strengthen my result. In particular, the necessary conditions for being an equilibrium payoff applies under a weak solution concept (BNE), and the equilibrium that attains v^* can survive demanding refinements such as sequential equilibrium.

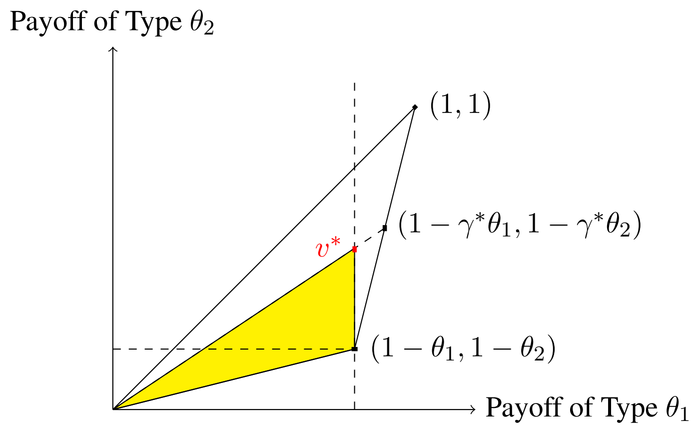

Figure 2: When $m = 2$, player 1's highest equilibrium payoff v^* in red and her equilibrium payoff set in yellow.

The proof is in Appendix A and Online Appendix A. According to Theorem 1, v_j^* is type θ_j patient seller's highest equilibrium payoff, and the highest equilibrium payoffs for all types can be (approximately) attained in the same equilibrium. This sharp characterization has not been obtained in commitment-type reputation models. This is because players move *sequentially* in the stage game and buyers' observations *cannot* statistically identify the seller's stage-game strategy. In this setting, the canonical payoff lower and upper bound results can only conclude that a patient seller's equilibrium payoff is between 0 and $1-\gamma^*\theta$, but it remains unclear whether this payoff upper bound is attainable, and if it is not, what is every type of seller's optimal equilibrium payoff.

Equation (3.1) provides a tractable formula for type θ_j seller's highest equilibrium payoff, which is the product of her *complete information Stackelberg payoff* and an *incomplete information multiplier*. This multiplier is strictly below 1 and common for all types. Interestingly, it depends only on the *lowest cost type* in the support of buyers' prior belief, but not on the other types in the support and the probability of each type. To understand its intuition, I state a lemma that relates v_j^* to a constrained optimization problem, with proof in Appendix A.3:

Lemma 3.1. *For every $j \in \{1, 2, ..., m\}$, the value of the following constrained optimization problem is v_j^*:*

$$\max_{\alpha \in \Delta\{N,H,L\}} \Big\{ (1-\theta_j) \underbrace{\alpha(H)}_{\textit{probability of outcome } H} + \underbrace{\alpha(L)}_{\textit{probability of outcome } L} \Big\}, \tag{3.2}$$

subject to:

$$(1-\theta_1)\alpha(H) + \alpha(L) \leq 1-\theta_1, \tag{3.3}$$

$$\textit{and} \quad \alpha(H) \geq \frac{\gamma^*}{1-\gamma^*}\alpha(L). \tag{3.4}$$

I map α into the repeated incomplete information game. For given BNE $\sigma \equiv \big((\sigma_\theta)_{\theta\in\Theta}, \sigma_2\big)$, type θ_j's equilibrium payoff in the repeated game equals her expected stage-game payoff under $\alpha^j \in \Delta\{N, H, L\}$, where

$$\alpha^j(y) \equiv \mathbb{E}^{(\sigma_{\theta_j}, \sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = y\}\Big] \text{ for every } y \in \{N, H, L\}.$$

In what follows, I explain intuitively why constraints (3.3) and (3.4) are necessary for α^j. According to Lemma 3.1, the necessity of (3.3) and (3.4) implies that type θ_j's equilibrium payoff *cannot exceed* v_j^*.

Constraint (3.3) requires that by adopting the equilibrium strategy of type θ_j, type θ_1 cannot receive payoff strictly higher than $1-\theta_1$, the latter is type θ_1's highest equilibrium payoff when her type is common knowledge (Fudenberg, Kreps and Maskin 1990). Intuitively, type θ_1 is the most efficient type, she has no good candidate to imitate, so she cannot strictly benefit from incomplete information. The distribution α^j needs to satisfy constraint (3.3) since type θ_1 cannot receive strictly higher payoff by deviating to the equilibrium strategy of type θ_j. Such a constraint is missing in models with commitment types since commitment types face no incentive constraints.

Constraint (3.4) arises from buyers' incentive constraints and their learning. This is because first, in all except for a bounded number of periods, buyers' predictions about the seller's actions are arbitrarily close to the latter's equilibrium actions (according to the seller's true type). Second, a buyer has no incentive to play T at h^t unless he expects H to be played at h^t with probability at least γ^*. The two parts together imply that for every $\varepsilon > 0$, there exists $\underline{\delta} \in (0,1)$, such that in every equilibrium where $\delta > \underline{\delta}$,

$$\frac{\alpha^j(H)}{\alpha^j(L)} = \frac{\mathbb{E}^{(\sigma_{\theta_j}, \sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = H\}\Big]}{\mathbb{E}^{(\sigma_{\theta_j}, \sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = L\}\Big]} \geq \frac{\gamma^* - \varepsilon}{1-\gamma^*+\varepsilon}, \text{ for every } \theta_j \in \Theta.$$

The necessity of constraint (3.4) is obtained by sending ε to 0.

After understanding the *necessity* of these constraints, the challenging step is to establish their *sufficiency*, i.e., there exist equilibria that approximately attain v^*. This is somewhat puzzling since for every $j \geq 2$, type θ_j needs to extract information rent (i.e., playing L for sure in periods where the buyer plays T) for unbounded number of periods to receive discounted average payoff strictly greater than $1-\theta_j$. For this to be incentive compatible from the buyers' perspectives, other types of seller need to play H with high enough probability. As a result, a type that extracts information rent inevitably *reveals information about her type*. This undermines her ability to extract information rent in the future since her informational advantage deteriorates every time she receives high stage-game payoff. In section 4.1, I explain how to construct equilibrium in which high-cost types can extract information rent in unbounded number of periods while preserving their informational advantages.

Theorem 1 has four implications. First, aside from the lowest-cost type θ_1, every type can *strictly* benefit from incomplete information, i.e., for every $j \geq 2$, type θ_j's highest equilibrium payoff in the repeated incomplete information game v_j^* is strictly greater than her highest equilibrium payoff in the repeated complete information game $1-\theta_j$. Second, the multiplier term converges to 1 as θ_1 vanishes to 0, and every type's highest equilibrium payoff converges to her Stackelberg payoff. This suggests that although no type of seller is immune to reneging temptations, every type of patient seller can approximately attain her optimal commitment payoff.

Corollary 1. *For every $\varepsilon > 0$, there exist $\underline{\delta} \in (0,1)$ and $\overline{\theta}_1 > 0$ such that when $\delta > \underline{\delta}$ and $\theta_1 < \overline{\theta}_1$, there exists a sequential equilibrium in which type θ_j's equilibrium payoff is no less than $v_j^{**} - \varepsilon$ for all $j \in \{1, ..., m\}$.*

Third, in terms of buyers' learning, the lowest-cost type seller *fully reveals her private information* for unboundedly many times in every equilibrium where the high-cost types can attain high payoffs. Formally, let $\mathcal{H}^{(\sigma_\theta,\sigma_2)}$ be the set of histories that occur with positive probability under $(\sigma_\theta, \sigma_2)$. A subset of histories $\mathcal{H}'$ is called an *independent set* if no pair of elements in $\mathcal{H}'$ can be ranked via the predecessor-successor relationship. The result is stated as Corollary 2, with proof in Online Appendix B.

Corollary 2. *For every $N \in \mathbb{N}$ and $v = (v_1, ..., v_m)$ such that $v_j > 1 - \theta_j$ for every $j \in \{2, ..., m\}$, there exists $\underline{\delta} \in (0,1)$ such that in every BNE that attains v when $\delta > \underline{\delta}$, there exists an independent set $\mathcal{H}'$ with $|\mathcal{H}'| > N$ and $\mathcal{H}' \subset \mathcal{H}^{(\sigma_{\theta_1},\sigma_2)}$, such that player 2's belief attaches probability 1 to θ_1 for every $h^t \in \mathcal{H}'$.*

Fourth, in terms of social welfare, every payoff on the Pareto frontier is approximately attainable when the seller is patient and θ_1 is small. Formally, let $\overline{v} \equiv (v_0, v_1, ..., v_m) \in \mathbb{R}^{m+1}$ where v_0 is the buyers' (discounted average) payoff and v_j is type θ_j seller's payoff. I say that $\overline{v}$ is *incentive compatible* for buyers if there exists $(\alpha_1, a_2) \in \Delta(A_1) \times A_2$ such that: $a_2 \in \arg\max_{a_2' \in A_2} u_2(\alpha_1, a_2')$, $u_1(\theta_j, \alpha_1, a_2) = v_j$ for every $j \in \{1, 2, ..., m\}$, and $u_2(\alpha_1, a_2) = v_0$. Let $\overline{V}^* \subset \mathbb{R}^{m+1}$ be the convex hull of the set of incentive compatible payoffs. The Pareto frontier of $\overline{V}^*$ is a line connecting $(b, 1-\theta_1, ..., 1-\theta_m)$ and $(0, v_1^*, ..., v_m^*)$. According to statement 2 of Theorem 1, every payoff on this line is approximately attainable when θ_1 is small and δ is large.

Corollary 3. *For every $\varepsilon > 0$, there exist $\underline{\delta} \in (0,1)$ and $\overline{\theta}_1 > 0$, such that for every $\delta \in (\underline{\delta}, 1)$, $\theta_1 \in (0, \overline{\theta}_1)$, and $\overline{v}$ on the Pareto frontier of $\overline{V}^*$, there exists sequential equilibrium that attains payoff within ε of $\overline{v}$.*

3.2 Seller's Behavior in Optimal Equilibria

I focus on settings in which the seller has *at least two types*. I derive properties of the patient seller's *on-path behaviors* that apply to *all* BNEs in which the patient seller's payoff is approximately v^*.

First, I show that no matter how low the cost of playing H is, no type of patient seller has completely mixed equilibrium best replies, and moreover, every type's equilibrium strategy exhibits nontrivial history dependence. For every strategy profile $\sigma \equiv (\{\sigma_\theta\}_{\theta\in\Theta}, \sigma_2)$, I say σ_θ is *stationary* if it prescribes the same action at every history that occurs with positive probability under $(\sigma_\theta, \sigma_2)$; I say σ_θ is *completely mixed* if it prescribes a nontrivially mixed action at every history that occurs with positive probability under $(\sigma_\theta, \sigma_2)$.

Theorem 2. *When $m \geq 2$, for every small enough $\varepsilon > 0$, there exists $\underline{\delta} \in (0, 1)$, such that when $\delta > \underline{\delta}$, no type of player 1 uses stationary strategies or completely mixed strategies in any BNE that attains payoff within ε of v^*. Moreover, no type has a completely mixed best reply in any such BNE.*[9]

The conclusion of Theorem 2 contrasts to the *Stackelberg commitment type* in the reputation models of Fudenberg and Levine (1992) and Gossner (2011) that mechanically plays the same mixed action in every period.

The proof is in Online Appendix C, which highlights how type θ_1 seller's incentive constraints drive these behavioral predictions, which are present in my payoff-type model but are absent in commitment-type models. For an informal illustration, suppose toward a contradiction that type θ_j's best reply is to mix at every on-path history. Then both playing L at every on-path history and playing H at every on-path history are her best replies against σ_2. Since low-cost types enjoy comparative advantages in playing H and vice versa, every type with cost higher than θ_j plays L with probability 1 at every on-path history, and every type with cost lower than θ_j plays H with probability 1 at every on-path history.[10] In what follows, I argue that none of these pure stationary strategies are compatible with the requirement that type θ_j's payoff is strictly above $1 - \theta_j$ for every $j \geq 2$.

First, suppose there exists a type θ_k that plays L with probability 1 at every on-path history. According to Fudenberg and Levine (1989), if the seller is of type θ_k, then the buyers will eventually believe that L will be played with probability greater than $1 - \gamma^*$ in all future periods, after which they will have a strict incentive to play N. This implies that type θ_k seller's discounted average payoff is close to 0 when δ is close to 1.

Next, suppose $j \geq 2$ and types θ_1 to θ_{j-1} play H with probability 1 at every on-path history. Then after type θ_j plays L for one period, she becomes the lowest-cost type in the support of buyers' posterior belief. This implies that type θ_j cannot extract information rent in the continuation game (Proposition A.1), in which case her discounted average payoff cannot exceed $(1 - \delta) + \delta(1 - \theta_j)$. The latter converges to $1 - \theta_j$ as $\delta \to 1$. This contradicts type θ_j's equilibrium payoff being close to v_j^* given that $v_j^* > 1 - \theta_j$ for every $j \geq 2$.

Remark: The conclusion of Theorem 2 applies more broadly. It extends to simultaneous-move stage games, as well as to a type whose cost of high effort equals 0. In Appendix C, I allow player 1 to have *arbitrary stage-game payoff functions*. I show that if a type plays a nontrivially mixed action at every history, then she is *indifferent across all outcomes* in the stage game (Proposition C.1). In the buyer-seller application, it translates into a type of seller who faces *zero cost* of supplying high quality, and receives *zero benefit* from buyers' purchases. This strengthens the motivation for excluding Stackelberg commitment type and other mixed commitment types, since their behaviors can only arise under knife-edge payoff functions, so their plausibility is questionable.

[9] Neither $m \geq 2$ nor the seller attains payoff approximately v^* is redundant. In Online Appendix D, I construct counterexamples in which the conclusion of Theorem 2 fails when either one of these conditions fails.

[10] If we order the states and actions according to $T \succ N$, $H \succ L$ and $\theta_1 \succ \theta_2 \succ ... \succ \theta_m$, then the stage-game payoff satisfies a monotone-supermodularity condition in Liu and Pei (2018). This is sufficient to guarantee the monotonicity of the sender's strategy with respect to her type in one-shot signalling games. I use an implication of their result on repeated signalling games (Pei 2018).

3.3 Action Frequencies in Optimal Equilibria

Focusing on seller-optimal equilibria, my next result derives lower and upper bounds on the seller's action frequencies that apply not only to all of her equilibrium strategies, but also to all of her *equilibrium best replies*.

Theorem 3. *For every small enough $\varepsilon > 0$, there exists $\underline{\delta} \in (0,1)$, such that for every $\delta > \underline{\delta}$, in every BNE $\sigma \equiv \big((\sigma_\theta)_{\theta\in\Theta}, \sigma_2\big)$ that attains payoff within ε of v^*,*

1. *For every $\theta \neq \theta_m$ and for every best reply $\widehat{\sigma}_\theta$ of type θ's against σ_2:*

$$\frac{\mathbb{E}^{(\widehat{\sigma}_\theta,\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t = H\}\Big]}{\mathbb{E}^{(\widehat{\sigma}_\theta,\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t = L\}\Big]} \geq \frac{\gamma^*-\varepsilon}{1-(\gamma^*-\varepsilon)}. \tag{3.5}$$

2. *For every $\theta \neq \theta_1$ and for every best reply $\widehat{\sigma}_\theta$ of type θ's against σ_2:*

$$\frac{\mathbb{E}^{(\widehat{\sigma}_\theta,\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t = H\}\Big]}{\mathbb{E}^{(\widehat{\sigma}_\theta,\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t = L\}\Big]} \leq \frac{\gamma^*+\varepsilon}{1-(\gamma^*+\varepsilon)}. \tag{3.6}$$

The proof is in Appendix B. According to Theorem 3, when the patient seller's cost is not the highest one in the support, the relative frequency between outcomes (T,H) and (T,L) cannot be lower than $\gamma^*/(1-\gamma^*)$ under each of her equilibrium best replies. When her cost is not the lowest one in the support, this relative frequency cannot be greater than $\gamma^*/(1-\gamma^*)$ under each of her equilibrium best replies. The two bounds together pin down the action frequencies for all types of sellers except for the lowest-cost and the highest-cost types.

Since type θ_j seller's payoff is approximately v_j^* in every seller-optimal equilibrium, Theorem 3 also pins down the discounted average frequency of every stage-game outcome under every equilibrium best reply:

Corollary 4. *For every small enough $\varepsilon > 0$, there exists $\underline{\delta} \in (0,1)$ such that for every $\delta > \underline{\delta}$, in every BNE $\sigma \equiv \big((\sigma_\theta)_{\theta\in\Theta}, \sigma_2\big)$ that attains payoff within ε of v^*, for every $\widehat{\sigma}_\theta$ that is type $\theta \in \{\theta_2, ..., \theta_{m-1}\}$'s best reply against σ_2, we have:*

$$\mathbb{E}^{(\widehat{\sigma}_\theta,\sigma_2)}\big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t = H\}\big] \in \Big(\gamma^*\frac{1-\theta_1}{1-\gamma^*\theta_1}-\varepsilon,\ \gamma^*\frac{1-\theta_1}{1-\gamma^*\theta_1}+\varepsilon\Big), \tag{3.7}$$

$$\mathbb{E}^{(\widehat{\sigma}_\theta,\sigma_2)}\big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t = L\}\big] \in \Big((1-\gamma^*)\frac{1-\theta_1}{1-\gamma^*\theta_1}-\varepsilon,\ (1-\gamma^*)\frac{1-\theta_1}{1-\gamma^*\theta_1}+\varepsilon\Big), \tag{3.8}$$

and

$$\mathbb{E}^{(\widehat{\sigma}_\theta,\sigma_2)}\big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t = N\}\big] \in \Big(\frac{(1-\gamma^*)\theta_1}{1-\gamma^*\theta_1}-\varepsilon,\ \frac{(1-\gamma^*)\theta_1}{1-\gamma^*\theta_1}+\varepsilon\Big). \tag{3.9}$$

The bounds in Theorem 3 and the predictions of Corollary 4 differ from Fudenberg and Levine (1992)'s result, which implies that under every type of seller's *equilibrium strategy*, the ratio between the discounted average frequency of (T, H) and the discounted average frequency of (T, L) is no less than $\gamma^*/(1-\gamma^*)$.

This is because Fudenberg and Levine (1992)'s result only applies to the seller's *equilibrium strategies*, while the bounds in Theorem 3 apply to all of her *equilibrium best replies*, which include but not limited to her equilibrium strategies. As an illustrative example, the strategy of playing the Stackelberg action at every history satisfies the requirement in Fudenberg and Levine (1992), but has been ruled out by Theorem 3. This is because for some pure strategies in the support of the seller's equilibrium strategy (such as playing L at every history), the frequency of L exceeds (3.5). For other pure strategies in its support (such as playing H at every history), the frequency of L falls below (3.6).

Conceptually, my bounds that apply to *all equilibrium best replies* have stronger testable implications compared to the ones that only apply to *equilibrium strategies*. This distinction is economically important given that in many markets, researchers can only observe one or a few *realized paths* of the game's outcomes instead of an entire distribution over outcome paths. The predictions of Theorem 3 apply to every pure-strategy best reply, and therefore, can be tested by observing a realized path of equilibrium play. For example, Corollary 4 implies that the frequencies of H and L are strictly decreasing in θ_1, and the frequency of N is strictly increasing in θ_1. By observing the frequency of each outcome along a realized path of play, a researcher can identify whether there are more efficient types under the buyers' belief, and if yes, the researcher can infer the value of θ_1 based on his estimation of γ^* and the observed frequencies of outcomes.

4 Constructing Optimal Equilibria

I explain how to construct equilibria in which a patient seller's payoff is approximately v^*. Let $v^N \equiv (0, 0, ..., 0)$, $v^H \equiv (1-\theta_1, 1-\theta_2, ..., 1-\theta_m)$, and $v^L \equiv (1, 1, ..., 1)$, which are the seller's payoffs under stage-game outcomes N, (T, H), and (T, L), respectively. For every $\gamma \in [\gamma^*, 1]$, let

$$v(\gamma) \equiv \frac{\theta_1(1-\gamma)}{1-\gamma\theta_1} v^N + \frac{(1-\theta_1)\gamma}{1-\gamma\theta_1} v^H + \frac{(1-\theta_1)(1-\gamma)}{1-\gamma\theta_1} v^L \in \mathbb{R}^m. \tag{4.1}$$

One can verify that $v(\cdot)$ is a continuous function of γ with $v(\gamma^*) = v^*$. Statement 2 of Theorem 1 is implied by:

Proposition 4.1. *For every $\eta \in (0, 1)$ and $\gamma \in (\gamma^*, 1)$, there exists $\underline{\delta} \in (0, 1)$, such that when $\delta > \underline{\delta}$ and π attaches probability more than η to type θ_1, there exists sequential equilibrium that attains payoff $v(\gamma)$.*

A constructive proof is presented in Online Appendix A. In section 4.1, I provide an overview of the equilibrium construction in an environment with *two types*. In section 4.2, I summarize the main ideas and key

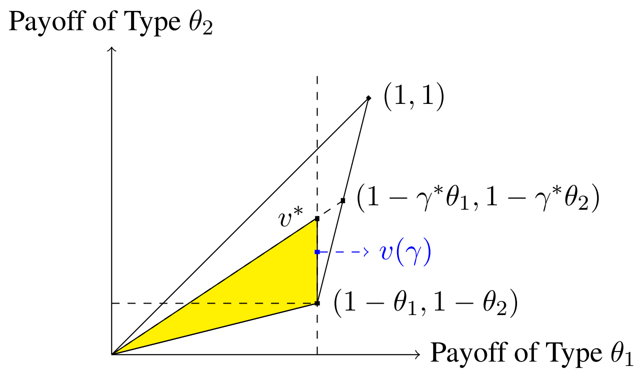

Figure 3: The set of attainable payoffs in yellow and $v(\gamma)$ in blue for some $\gamma \in (\gamma^*, 1)$.

features of my construction. In section 4.3, I compare the equilibrium dynamics with other reputation models.

4.1 Equilibrium Construction with Two Types: Strategies and Continuation Values

I specify *players' actions* and *the evolution of player 1's continuation value* at on-path histories. At every off-path history, player 2 plays N and all types of player 1 play L. I keep track of two state variables:

1. $\eta(h^t) \in [0,1]$, which is the probability of type θ_1 under player 2's belief at h^t, and is called player 1's *reputation*, with $\eta(h^0)$ equals the prior probability of type θ_1.

2. $v(h^t) \in \mathbb{R}^2$, which is P1's continuation value at h^t, with $v(h^0) \equiv v(\gamma)$. I verify in Online Appendix A that for every h^t, $v(h^t)$ is a convex combination of v^N, v^H and v^L, i.e., $v(h^t) = p^N(h^t)v^N + p^H(h^t)v^H + p^L(h^t)v^L$. Hence, keeping track of $v(h^t)$ is equivalent to keeping track of $p^N(h^t)$, $p^H(h^t)$, and $p^L(h^t)$.

I partition the set of histories into *three classes*, depending on the value of $p^L(h^t)$:

- **Class 1 Histories:** $p^L(h^t) \geq 1-\delta$.
- **Class 2 Histories:** $p^L(h^t) \in (0, 1-\delta)$.
- **Class 3 Histories:** $p^L(h^t) = 0$.

Play starts from a Class 1 history, and reaches a Class 3 history in finite time, after which it stays there forever. Active learning about player 1's type happens at Class 1 and Class 2 histories, but stops at Class 3 histories.

Class 1 Histories: Play starts from a Class 1 history. At every Class 1 history, i.e., one in which $p^L(h^t) \geq 1-\delta$:

- Player 2 plays T for sure.

- Player 2's posterior beliefs, $\eta(h^t, H)$ and $\eta(h^t, L)$, are functions of $\eta(h^t)$:

$$\eta(h^t, H) = \eta^* + \min\Big\{1 - \eta^*, \big(1 + \lambda(1 - \gamma^*)\big)\big(\eta(h^t) - \eta^*\big)\Big\}, \tag{4.2}$$

$$\text{and} \quad \eta(h^t, L) = \eta^* + (1 - \lambda\gamma^*)(\eta(h^t) - \eta^*). \tag{4.3}$$

  - $\eta^* \in (\gamma^*\eta(h^0), \eta(h^0))$ is a constant, which satisfies (A.5) in Online Appendix A. Given that $\eta(h^0) > \eta^*$, one can verify by induction that $\eta(h^t) > \eta^*$ for every Class 1 (and Class 2) history h^t.
  - $\lambda > 0$ is a constant that measures the speed of player 2's learning. I require

$$\big(1 - \lambda\gamma^*\big)^{1-\gamma}\big(1 + \lambda(1 - \gamma^*)\big)^{\gamma} > 1. \tag{4.4}$$

    Intuitively, (4.4) is satisfied when λ is sufficiently small given that $\gamma > \gamma^*$.

Equations (4.2) and (4.3) pin down both types of player 1's actions at h^t. This is because according to Bayes Rule, the probability with which type θ_1 plays H at h^t is:

$$\frac{\eta(h^t) - \eta(h^t, L)}{\eta(h^t, H) - \eta(h^t, L)} \cdot \frac{\eta(h^t, H)}{\eta(h^t)}, \tag{4.5}$$

and the probability with which type θ_2 plays H at h^t is:

$$\frac{\eta(h^t) - \eta(h^t, L)}{\eta(h^t, H) - \eta(h^t, L)} \cdot \frac{1 - \eta(h^t, H)}{1 - \eta(h^t)}. \tag{4.6}$$

Plugging (4.2) and (4.3) into (4.5) and (4.6), every type's action at h^t can be written as a function of $\eta(h^t)$.

Let $p_H(h^t)$ be the probability of H being played at h^t according to player 2's belief at h^t. Since player 2's belief is a martingale, we have $p_H(h^t)\eta(h^t, H) + (1 - p_H(h^t))\eta(h^t, L) = \eta(h^t)$. Equation (4.3) and $\eta(h^t) > \eta^*$ imply that $\eta(h^t, L) \neq \eta(h^t)$, so the martingale condition can be rewritten as:

$$\frac{1 - p_H(h^t)}{p_H(h^t)} = \frac{\eta(h^t, H) - \eta(h^t)}{\eta(h^t) - \eta(h^t, L)}.$$

Plugging (4.2) and (4.3) into the above equation, we have:

$$\frac{1 - p_H(h^t)}{p_H(h^t)} = \frac{\eta(h^t, H) - \eta(h^t)}{\eta(h^t) - \eta(h^t, L)} \leq \frac{1 - \gamma^*}{\gamma^*}. \tag{4.7}$$

This implies that $p_H(h^t) \geq \gamma^*$, i.e., player 2 has an incentive to play T at h^t. Player 1's continuation value after playing L at h^t is:

$$v(h^t, L) = \frac{p^N(h^t)}{\delta} v^N + \frac{p^L(h^t) - (1-\delta)}{\delta} v^L + \frac{p^H(h^t)}{\delta} v^H. \tag{4.8}$$

If h^t is such that $\eta(h^t, H) < 1$, then player 1's continuation value after playing H at h^t is given by:

$$v(h^t, H) = \frac{p^N(h^t)}{\delta} v^N + \frac{p^L(h^t)}{\delta} v^L + \frac{p^H(h^t) - (1-\delta)}{\delta} v^H. \tag{4.9}$$

If h^t is such that $\eta(h^t, H) = 1$, then player 1's continuation value after playing H at h^t is given by:

$$v(h^t, H) = \frac{v_1(h^t, H)}{1-\theta_1} v^H + \Big(1 - \frac{v_1(h^t, H)}{1-\theta_1}\Big) v^N, \tag{4.10}$$

$$\text{with } v_1(h^t, H) \equiv \frac{v_1(h^t) - (1-\delta)(1-\theta_1)}{\delta} \text{ and } v_1(h^t) \text{ is the first entry of } v(h^t). \tag{4.11}$$

If h^t is such that $\eta(h^t, H) < 1$, then (4.8) and (4.9) imply that both types of player 1 are indifferent between H and L. If h^t is such that $\eta(h^t, H) = 1$, then (4.8), (4.10) and (4.11) imply that type θ_1 is indifferent while type θ_2 strictly prefers L. Player 1's incentive constraints at h^t are satisfied since when $\eta(h^t, H) < 1$, both types are required to mix; and when $\eta(h^t, H) = 1$, type θ_1 is required to mix while type θ_2 plays L for sure.

Class 2 Histories: At every Class 2 history h^t, i.e., one in which $p^L(h^t) \in (0, 1-\delta)$,

- Player 2 plays T for sure.
- Type θ_1 plays H for sure. Type θ_2 plays L with probability $\min\{1, \frac{1-\gamma^*}{1-\eta(h^t)}\}$. Therefore, the probability with which L being played at h^t is $(1-\eta(h^t)) \min\{1, \frac{1-\gamma^*}{1-\eta(h^t)}\}$, which is no more than $1-\gamma^*$. This implies that player 2 has an incentive to play T.

After player 1 plays L at h^t, $\eta(h^t, L) = 0$ given that type θ_1 plays H for sure, and player 1's continuation value is given by:

$$v\big(h^t, L\big) \equiv \frac{Q(h^t)}{\delta} v^H + \frac{\delta - Q(h^t)}{\delta} v^N, \tag{4.12}$$

where

$$Q(h^t) \equiv p^H(h^t) - \frac{1 - \delta - p^L(h^t)}{1-\theta_2} \tag{4.13}$$

Player 1's continuation value after playing H depends on whether $\eta(h^t, H)$ equals 1 or not, the latter can be computed according to Bayes Rule via $\eta(h^t)$ and different types' mixing probabilities specified above.

1. If $\eta(h^t, H) < 1$, then player 1's continuation value after playing H at h^t is given by (4.9).

2. If $\eta(h^t, H) = 1$, then player 1's continuation value after playing H at h^t is given by (4.10).

If h^t is such that $\eta(h^t, H) < 1$, then (4.12) and (4.9) imply that type θ_2 is indifferent and type θ_1 strictly prefers to play H. If h^t is such that $\eta(h^t, H) = 1$, then (4.12) and (4.10) imply that type θ_2 strictly prefers to play L and type θ_1 strictly prefers to play H. Player 1's incentive constraints at h^t are satisfied since type θ_1 is required to play H, while type θ_2 is required to mix only if $\eta(h^t, H) < 1$, and is required to play L if $\eta(h^t, H) = 1$.

Class 3 Histories: If h^t is such that $p^L(h^t) = 0$, then all types of player 1 play the same action at h^t so learning about her type stops. Moreover, her continuation value at every subsequent history is also a convex combination of v^H and v^N, i.e., all subsequent histories belong to Class 3. The construction of equilibrium play after reaching any Class 3 history uses Lemma 3.7.2 in Mailath and Samuelson (2006, page 99):

Lemma 3.7.2 in MS (2006). *For all $\varepsilon > 0$, there exists $\underline{\delta} \in (0,1)$ such that for every $\delta \in (\underline{\delta}, 1)$ and every $v \in \mathbb{R}^m$ that is a convex combination of $v(1)$, $v(2)$,...,$v(k)$, there exists $\{v^s\}_{s=0}^{\infty}$ with $v^s \in \{v(1), ..., v(k)\}$ such that (1) $v = \sum_{s=0}^{\infty}(1-\delta)\delta^s v^s$, and (2) for every $l \in \mathbb{N}$, $\sum_{s=l}^{\infty}(1-\delta)\delta^{s-l}v^s$ is within ε of v.*

Since $v(h^t)$ is a convex combination of v^H and v^N, the above lemma implies that when δ is above some cutoff, there exists an infinite sequence $\{v^s\}_{t=0}^{\infty}$ with $v^s \in \{v^N, v^H\}$ such that:

1. $v(h^t) = (1-\delta)\sum_{s=0}^{\infty}\delta^s v^s$,
2. For every $l \in \mathbb{N}$, $(1-\delta)\sum_{s=l}^{\infty}\delta^{s-l}v^s$ is within ε of $v(h^t)$.

The continuation play following h^t is:

- For every $s \in \mathbb{N}$ such that $v^s = v^H$, P2 plays T and all types of P1 play H in period $t+s$.
- For every $s \in \mathbb{N}$ such that $v^s = v^N$, P2 plays N and all types of P1 play L in period $t+s$.

Player 2's incentive constraints at Class 3 histories are trivially satisfied. To verify player 1's incentive constraints, I show in Lemma A.5 of Online Appendix A that $p^H(h^t)$ is bounded from below by some strictly positive number for every h^t belonging to Class 1 or Class 2. This implies that if h^t belongs to Class 3 but none of its predecessors belong to Class 3, then $p^H(h^t)$ is also bounded from below by a positive number. Pick ε in the above lemma to be small enough, one can ensure that player 1's continuation value at every on-path history succeeding h^t is strictly bounded away from 0. This implies that at every on-path history where player 1 is asked to play H, all types of patient player 1 have strict incentives to comply. This is because every type's continuation payoff equals 0 if she does not comply, and her continuation payoff is strictly bounded away from 0 if she complies.

Promise Keeping Constraints: I have verified player 2s' incentive constraints, and have constructed continuation values under which player 1's incentive constraints are satisfied. What remains to be verified is the promise keeping constraint, that the continuation play at every on-path history delivers all types of player 1 their respective continuation values. This is established by showing that under the above strategy profile, play reaches a Class 3 history in finite time with probability 1 (implied by Lemmas A.1 and A.4 in Online Appendix A), and player 1's continuation value at Class 3 histories can be delivered via a sequence of payoffs consisting of v^H and v^N.

4.2 Summary of Ideas Behind the Construction

To begin with, learning about player 1's type is indispensable for her to attain payoff v^*, or anything above v^H. According to Corollary 2, player 1 needs to be able to *rebuild* her reputation after playing L.[11] This is because otherwise, she cannot extract information rent in the long run. To make reputation rebuilding feasible, both types of player 1 mix between H and L, except for histories where $\eta(h^t)$ is close to 1, in which case the low-cost type θ_1 mixes between H and L while the high-cost type θ_2 plays L for sure. Intuitively, this arrangement reduces the high-cost type's reputation loss when she shirks for sure and extracts information rent.[12]

The presence of the absorbing phase (i.e., Class 3 histories) is to provide all types of player 1 the incentives to mix in the active learning phase (Class 1 & 2 histories). Despite player 1 can flexibly choose her actions when active learning takes place, her action choices affect her continuation payoff after reaching the absorbing phase, as well as the calendar time at which play enters the absorbing phase. For example, if player 1 plays L too frequently in the beginning, then $p^L(h^t)$ decreases more quickly and play reaches a Class 3 history as soon as $p^L(h^t) = 0$, after which player 1's continuation value is low and can no longer extract information rent.

A more subtle situation occurs when player 1 plays H too frequently, which decreases $p^H(h^t)$ and increases $p^L(h^t)$. It raises a concern that $p^H(h^t)/p^L(h^t)$ may fall below $\frac{\gamma^*}{1-\gamma^*}$, after which player 1's continuation payoff cannot be delivered in any equilibrium. To address this, play enters the absorbing phase when $\eta(h^t)$ reaches 1, after which the continuation play consists only of outcomes N and (T, H). Type θ_1 is indifferent between entering the absorbing phase (by playing H) and remaining in the active learning phase (by playing L), while type θ_2 strictly prefers the latter option. This is because type θ_1 has a comparative advantage in playing H. Moreover, requirement (4.4) on λ implies that if $p^H(h^0)/p^L(h^0)$ is strictly greater than $\frac{\gamma^*}{1-\gamma^*}$, and play remains in the active learning phase at h^t, then $p^H(h^t)/p^L(h^t)$ is also strictly greater than $\frac{\gamma^*}{1-\gamma^*}$. This is formally stated as Lemma A.1 in Online Appendix A, which ensures that the equilibrium play eventually reaches a Class 3 history.

What needs to be considered next is the speed of learning λ. In order to maximize player 1's equilibrium

[11] This is true except for histories such that $p^L(h^t) \leq 1 - \delta$, which happens only if L has been played too frequently before.

[12] Histories at which type θ_2 has a strict incentive to play L is indispensable for her to attain payoff v_2^*. This is because otherwise, playing H at every history where T is played with positive probability is one of type θ_2's equilibrium best reply, and under this strategy of hers, her payoff cannot exceed $1 - \theta_2$, which is strictly lower than v_2^*.

payoff, one needs to maximize the frequency of outcome (T, L) while simultaneously providing incentives for player 2s to trust. This leads to the role of *slow learning*, i.e., λ being low. To understand why, first, player 2's incentive to trust translates into an upper bound on the *relative rate of learning*, given by (4.7). In a nutshell, it requires that the magnitude of reputation improvement after playing H to be small enough relative to the magnitude of reputation deterioration after playing L. Second, fixing the relative rate of learning and the long-run frequencies of H and L, the amount of reputation loss per period vanishes as the *absolute speed of learning* goes to zero. As a result, lowering the absolute speed of learning improves player 1's long-term reputation without compromising on player 2s' willingness to trust. This allows for an increase in the long-run frequency of outcome L without sacrificing player 1's reputation, which helps to improve her payoff.

4.3 Comparing Equilibrium Dynamics

I compare the equilibrium dynamics in my model to those in reputation models with behavioral biases (Jehiel and Samuelson 2012), models of reputation cycles (Sobel 1985, Phelan 2006, Liu 2011, Liu and Skryzpacz 2014), models of gradual learning (Benabou and Laroque 1992, Ekmekci 2011, Wiseman 2005), and the capital-theoretic models of reputations (Board and Meyer-ter-Vehn 2013, Bohren 2018, Dilmé 2018).

Analogical-Based Reasoning Equilibria: The patient player alternates between her actions to manipulate her opponents' belief is reminiscent of the *analogy-based reasoning equilibria* in Jehiel and Samuelson (2012). In their model, there are multiple commitment types who are playing stationary mixed strategies, and one strategic type who can flexibly choose her actions. The short-run players mistakenly believe that the strategic type is playing a stationary strategy. In the trust game, their results imply that the strategic long-run player's behavior experiences a *reputation building phase* in which she plays H for a bounded number of periods, followed by a *reputation manipulation phase* that resembles the active learning phase in my model where she alternates between H and L according to the Stackelberg frequencies. The short-run players' posterior belief fluctuates within a small neighborhood of the cutoff belief, implying that the long-run player's type is never fully revealed.

Comparing my model to theirs, there are two qualitative differences in the reputation dynamics that highlight the distinctions between rational and analogical-based short-run players. First, learning stops in finite time in my model while it lasts forever in theirs. This is driven by the constraint that type θ_1's equilibrium payoff cannot exceed $1 - \theta_1$, which comes from the rational short-run players' ability to correctly predict the long-run player's average action in *every period*. This constraint is absent when short-run players use analogy-based reasoning since they can only correctly predict *the long-run player's average action across all periods*. Second, the short-run players learn the true state with positive probability in every high-payoff equilibrium of my model, while in Jehiel and Samuelson (2012), the probability with which they learn the true state is zero. This is because

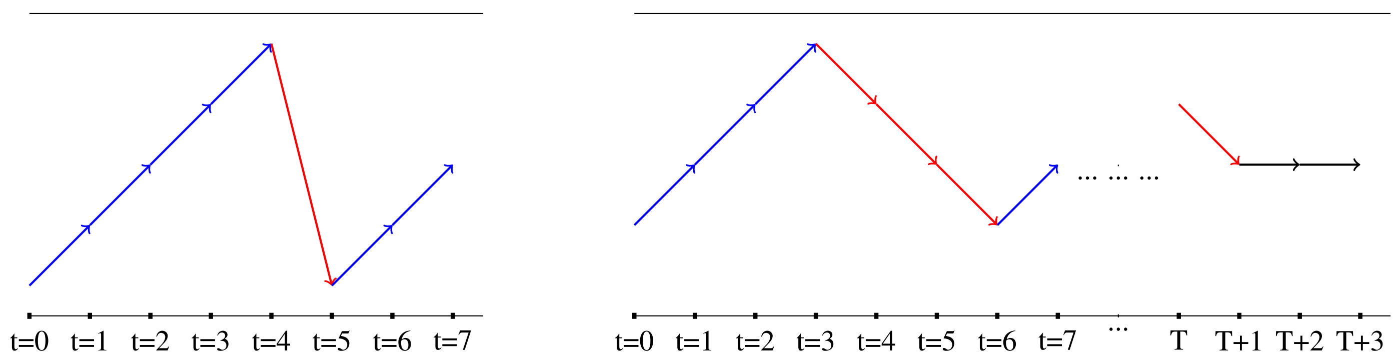

Figure 4: The horizontal axis represents the timeline and the vertical axis measures the informed player's reputation, that is, the probability of the commitment type or lowest-cost type. Left: Reputation cycles in Phelan (2006). Right: A sample path of the reputation cycle in my model.

analogy-based short-run players' posterior beliefs depend only on the empirical frequencies of the observed actions. That is to say, their beliefs are not responsive enough to each individual observation.

Reputation Building-Milking Cycles: The behavioral pattern that a patient player builds her reputation in order to milk it in the future has been identified in commitment-type models with either changing types (Phelan 2006), or limited memories (Liu 2011, Liu and Skrzypacz 2014). In terms of differences, first, the reputation cycles in Phelan (2006), Liu (2011) and Liu and Skrzypacz (2014) can last forever while learning stops in finite time in mine This is driven by the constraint that the lowest-cost type's equilibrium payoff cannot exceed $1-\theta_1$, which arises because this type is rational and has strict incentives to betray.

Second, reputations are built and milked *gradually* in my model while in theirs, the agent's reputation falls to its lower bound every time she milks it. This is because the commitment types in their models never betray. As a result, one misbehavior reveals the long-run player's rationality. In my model, the behaviors of good and bad types are close since all types share the same ordinal preferences over stage-game outcomes and face strict temptations to betray. In the long-run player's optimal equilibrium, the lowest-cost type betrays with positive probability for unbounded number of periods, which does not reduce her own payoff while at the same time, covering up the other types when they milk reputations. These differences are depicted in Figure 4.

This feature of gradual learning is supported empirically by several studies of online markets. As documented in Dellarocas (2006) and Bar-Isaac and Tadelis (2008), consumers judge the quality of sellers based on their reputation scores, which are usually obtained via averaging the ratings they have obtained in the past. A number of empirical works document that one recent negative rating neither significantly affects the amount of sales nor the prices of a reputable seller who has obtained many positive ratings in the past. This empirical observation is better explained by the reputation dynamics in my model.

Reputation Models with Gradual Learning: Benabou and Laroque (1992) and Ekmekci (2011) study reputation games with commitment types and the long-run player's actions are imperfectly monitored. In their equilibria, learning also happens gradually since the short-run players cannot tell the difference between intended cheating and exogenous noise. In contrast, my model has perfect monitoring of stage-game outcomes but has no commitment type. Gradual learning occurs since the reputational type cheats with positive probability. The different driving forces behind gradual learning also lead to different long-run outcomes. In my model, reputation building-milking cycles stop in finite time while in theirs, reputation cycles last forever. As mentioned before, this is driven by the rational reputational type's strict incentive to betray, which implies an upper bound on her equilibrium payoff. Another difference is that the short-run players never fully learn the long-run player's type in Benabou and Laroque (1992) and Ekmekci (2011), while in mine, the lowest-cost type fully reveals her private information for unboundedly many times in every high-payoff equilibrium.

The class of equilibria I construct start from a phase where active learning takes place followed by an absorbing phase where learning stops. Different from the learning phases in Wiseman (2005), during which players experiment and learn about their payoffs, the learning phase in my model is constructed so that the patient player can extract information rent and can attain strictly higher payoff compared to the complete information benchmark. As a result, the length of the learning phase in my model increases with the discount factor, while it does not vary with the discount factor in Wiseman (2005).

Capital-Theoretic Reputation Models: Reputation cycles also occur in the Poisson good news models of Board and Meyer-ter-Vehn (2013) and Dilmé (2018). They characterize Markov equilibria in which the long-run player exerts effort when her reputation is above some cutoff. Different from my model, reputation jumps up immediately after the arrival of good news. Moreover, the long-run player's reputation depends only on the most recent time of news arrival in their models while it depends on the frequencies of her past actions in mine. These distinctions are caused by the difference sources of learning. In my model, learning arises from the differences in different types' behaviors, while in their models, all types adopt the same behavior but face different news arrival rates. In terms of the applications, my model fits into online platforms where feedback arrives frequently while their models fit into markets with infrequent news arrival.[13]

5 Concluding Remarks

I conclude by discussing the robustness of my results to simultaneous-move stage games, forward-looking buyers, and imperfect monitoring of the seller's actions (section 5.1). Then I list some alternative applications of my

[13]In the bad news model of Board and Meyer-ter-Vehn (2013) and the Brownian model of Bohren (2018), the informed player's effort increases in her reputation. This differs from my model in which the high-cost types cheat for sure when her reputation is close to 1.

model and results (section 5.2). Generalizations beyond 2×2 trust games are stated in Appendix D.

5.1 Robustness of Results

Simultaneous-Move Stage Game: Consider a simultaneous-move trust game with stage-game payoffs:

-	T	N
H	$1-\theta, b$	$-d(\theta), 0$
L	$1, -c$	$0, 0$

where $b, c > 0$, $\theta \in \Theta \equiv \{\theta_1, \theta_2, ..., \theta_m\} \subset (0,1)$ is player 1's persistent private information, and $d(\theta) \geq 0$ measures player 1's loss when she exerts high effort while player 2 does not trust. In the repeated version of this game, players' past action choices are perfectly monitored and the public history $h^t \equiv \{a_{1,s}, a_{2,s}\}_{s=0}^{t-1}$ consists of both players' past action choices. Other features of the game remain the same as in the baseline model.

For the results on equilibrium payoffs, recall the definition of v^* in (3.1). A construction similar to the one in section 4 implies that v^* is approximately attainable when δ is close to 1. Under a supermodularity condition on the stage-game payoffs:

$$0 \leq d(\theta_j) - d(\theta_i) \leq \theta_j - \theta_i \quad \text{for every} \quad j < i, \tag{5.1}$$

one can show that for every $j \in \{1, 2, ..., m\}$, v_j^* is type θ_j patient long-run player's highest equilibrium payoff.

Under the supermodularity condition in (5.1), the conclusions in Theorem 2 and Theorem 3 extend to the simultaneous-move stage game. In particular, no type of the long-run player uses a stationary strategy or has a completely best reply in any equilibrium that approximately attains v^*. For the bounds on the long-run player's action frequencies that apply to all of her pure-strategy best replies, one needs to replace $y_t = H$ and $y_t = L$ in (3.5) and (3.6) with $a_{1,t} = H$ and $a_{1,t} = L$, respectively.

Stage Game with Imperfect Monitoring: Player 1 is an agent, for example a worker, a supplier or a private contractor. In every period, a principal (player 2, for example an employer or a final good producer) is randomly matched with the agent. The principal then decides whether to incur a fixed cost and interact with the agent or to skip the interaction.[14] The agent chooses her effort from a closed interval unbeknownst to the principal. The probability with which the service quality being high increases with her effort. In line with the literature on incomplete contracts, the service quality is not contractible but is observable to the agent and all the subsequent principals. The cost of effort is linear and the marginal cost of effort is the agent's persistent private information.[15]

[14] Interpretations of this fixed cost includes, an upfront payment made by the final good producer to his supplier and a relationship specific investment the principal needs to make in order to collaborate with the agent.

[15] Chassang (2010) studies a game in which players face similar incentives. The main difference is that the agent's cost of effort is common knowledge but the set of actions that are available in each period is the agent's private information. Tirole (1996) uses a similar model to study the collective reputation of commercial firms and that of bureaucrats.

Players move sequentially in the stage game. Different from the baseline model, after player 2 chooses to trust, player 1 chooses among a continuum of effort levels $e \in [0,1]$. The quality of the output being produced is denoted by $z \in \{G, B\}$, which is *good* (or $z = G$) with probability e and is *bad* (or $z = B$) with complementary probability. The cost of effort for type θ_i is $\theta_i e$. Player 1's benefit from her opponent's trust is normalized to 1. Therefore, her stage-game payoff under outcome N is 0 and that under outcome (T, e) is $1 - \theta_i e$. Player 2's payoff is 0 if he chooses N. His benefit from good output is b while his loss from bad output is c, with $b, c > 0$. Therefore, player 2 is willing to trust only when player 1's expected effort exceeds $\gamma^* \equiv \frac{c}{b+c}$.

Consider the repeated version of this game in which the public history consists of player 2's actions and the realized output quality, i.e., player 1's effort choice is her private information. In period t, let $a_{1,t}$ be player 1's action, let $a_{2,t}$ be player 2's action, and let z_t be the realized output quality. Let $h^t = \{a_{2,s}, z_s\}_{s=0}^{t-1} \in \mathcal{H}^t$ be a public history with $\mathcal{H} \equiv \bigcup_{t=0}^{+\infty} \mathcal{H}^t$ the set of public histories. Let $h_1^t = \{a_{1,s}, a_{2,s}, z_s\}_{s=0}^{t-1} \in \mathcal{H}_1^t$ be player 1's private history with $\mathcal{H}_1 \equiv \bigcup_{t=0}^{+\infty} \mathcal{H}_1^t$ the set of private histories. Let $\sigma_2 : \mathcal{H} \to \Delta(A_2)$ be player 2's strategy and let $\sigma_\theta : \mathcal{H}_1 \to \Delta(A_1)$ be type θ player 1's strategy, with $\sigma_1 \equiv (\sigma_\theta)_{\theta \in \Theta}$.

The above game with a continuum of effort, linear effort cost, and imperfect monitoring is equivalent to the baseline model with binary effort and perfect monitoring. To see this, choosing effort level e under imperfect monitoring is equivalent to choosing a mixed action $eH + (1-e)L$ under perfect monitoring. In terms of the results on payoffs, one can show that v_j^* is type θ_j's highest equilibrium payoff when she is patient, and payoff vector v^* is approximately attainable when δ is close to 1. In terms of behaviors, the bounds on the relative frequencies can be applied to realized paths of public signals, namely, one needs to replace $y_t = H$ and $y_t = L$ in (3.5) and (3.6) with $(a_{2,t}, z_t) = (T, G)$ and $(a_{2,t}, z_t) = (T, B)$, respectively.

Forward-Looking Buyer: My results are robust when the seller faces a single buyer whose discount factor δ_2 is strictly positive but close to 0. To begin with, the constructed equilibrium that approximately attains v^* remains to be an equilibrium under any δ_2. This is because at every off-path history, the buyer plays N and all types of player 1 play L, in which case the buyer receives his minmax payoff. Hence, the buyer's strategy in the constructed equilibrium maximizes his stage-game payoff while it cannot lower his continuation payoff.

The necessity of constraint (3.3) relies on the observation that at every on-path history, the buyer has no incentive to play T unless he expects H to be played with positive probability. This remains valid when $\delta_2 < \gamma^*$. Suppose toward a contradiction that at some on-path history h^t, all types of seller play L for sure, but the buyer plays T with strictly positive probability. The buyer's discounted average payoff by playing T at h^t is at most:

$$\underbrace{(1-\delta_2)(-c)}_{\text{P2's stage-game payoff if he plays } T \text{ while P1 plays } L \text{ for sure}} + \underbrace{\delta_2 b}_{\text{P2's maximal continuation payoff after playing } T}.$$

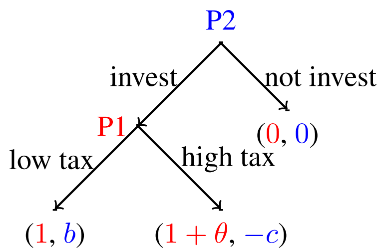

Figure 5: Capital Taxation Game Between Government and Investors

Since $\delta_2 < \gamma^* \equiv \frac{c}{b+c}$, the above expression is strictly less than 0. This contradicts the buyer's incentive to play T at h^t since he can secure payoff 0 by playing N in every subsequent period.

In addition, when δ_2 is close to 0, the buyer has no incentive to play T at h^t unless he expects H to be played with probability more than $\gamma^* - \varepsilon$, with ε vanishes to 0 as $\delta_2 \to 0$. This implies an *approximate version* of constraint (3.4) when the seller's discount factor δ_1 is close enough to 1:

$$\alpha^j(H) \geq \frac{\gamma^* - \varepsilon}{1 - \gamma^* + \varepsilon} \alpha^j(L), \tag{5.2}$$

with

$$\alpha^j(y) \equiv \mathbb{E}^{(\sigma_{\theta_j}, \sigma_2)} \Big[\sum_{t=0}^{\infty} (1-\delta_1)\delta_1^t \mathbf{1}\{y_t = y\} \Big] \text{ for every } y \in \{N, H, L\}.$$

Replacing (3.4) with constraint (5.2), the value of the constrained optimization problem is close to v_j^*, which converges to v_j^* as $\delta_2 \to 0$. This implies the robustness of Theorem 1 to perturbations of δ_2. Given that the proofs of Theorems 2 and 3 do not use buyers' incentive constraints aside from the conclusion that v^* is a patient seller's highest equilibrium payoff, those results are also robust to small perturbations of δ_2.

5.2 Alternative Applications

Capital Taxation: Player 1 is a government and player 2s are a sequence of foreign investors. The stage-game is depicted in Figure 5, where $\theta \in \{\theta_1, ..., \theta_m\}$ is the government's private information that measures its benefit from expropriating investors via high tax rates. I assume that $0 < \theta_1 < ..., < \theta_m$, namely, all types of government strictly benefit from expropriation. This game can be analyzed using similar techniques as my baseline model. Let $\gamma^* \equiv \frac{c}{b+c}$. One can show that type θ government's highest equilibrium payoff v_j^* is the value of:

$$\max_{\alpha \in \Delta\{\text{not invest,high tax,low tax}\}} \big\{ \alpha(\text{low tax}) + (1+\theta_j)\alpha(\text{high tax}) \big\}, \tag{5.3}$$

subject to:

$$\alpha(\text{low tax}) + (1+\theta_1)\alpha(\text{high tax}) \leq 1, \tag{5.4}$$

$$\text{and} \quad \alpha(\text{low tax}) \geq \frac{\gamma^*}{1-\gamma^*}\alpha(\text{high tax}). \tag{5.5}$$

Solving this problem, one can obtain $v_j^* = \frac{1+\theta_j-\gamma^*\theta_j}{1+\theta_1-\gamma^*\theta_1}$. Similar to the baseline model, v_j^* depends only on θ_j and the lowest possible temptation to expropriate θ_1, but not on the other aspects of incomplete information. In addition, the lowest benefit type θ_1 cannot strictly benefit from incomplete information, while types θ_2 to θ_m can receive payoff strictly higher than 1. The results on the government's on-path behaviors in equilibria that approximately attain $v^* \equiv (v_1^*, ..., v_m^*)$ extend as well.

Entry Deterrence/Limit Pricing Game Player 1 is an incumbent choosing between a low price (or *fight*) and a normal price (or *accommodate*). Player 2 is an entrant deciding whether to enter or not. Players' payoffs are:

-	Out	Enter
Low Price	$1-\theta, 0$	$-d(\theta), -b$
Normal Price	$1, 0$	$0, c$

where θ and $d(\theta)$ are the incumbent's costs from lowering prices, which is interpreted as limit pricing if the entrant stays out, and is interpreted as predation if the entrant enters. As argued in Milgrom and Roberts (1982), θ depends on the efficiency of the incumbent's production technology, which tends to be its persistent private information. This maps into the simultaneous-move version of the stage game once we replace Low Price with H, Normal Price with L, Out with T, and Enter with N.

Monetary Policy: Player 1 is a central bank, that interacts with a continuum of households (player 2s), each has negligible mass. In every period, the central bank chooses the inflation level while at the same time, households form their expectations about inflation. To simplify matters, I assume that both actual inflation and expected inflation are binary variables. In line with the classic work of Barro (1986), players' stage game payoffs are:

-	Low Expectation	High Expectation
Low Inflation	$1-\theta, x_1$	$-d(\theta), -y_1$
High Inflation	$1, -y_2$	$0, x_2$

where $x_1, x_2, y_1, y_2 > 0$ are parameters, $\theta \in \Theta \subset (0,1)$ is the central bank's private information. This game maps into the simultaneous-move version of my stage game.

To interpret these payoffs, households want to match their expectations with the actual inflation. The central bank's payoff decreases with the actual inflation and increases with the amount of surprised inflation (defined as actual inflation minus expected inflation). As argued in Barro (1986), the central bank can strictly benefit from surprised inflation as it can increase real economic activities, decrease unemployment rate and increase governmental revenue. How the central bank trades-off these benefits with the costs of inflation is captured by θ, which depends on the central banker's ideology and tends to be her persistent private information. The assumption that $\theta < 1$ implies that inflation is costly for the central bank if it is fully anticipated by households.

A Statement 1 of Theorem 1: Necessity of Constraints

In section A.1, I establish the necessity of constraint (3.3). In section A.2, I establish the necessity of constraint (3.4). In section A.3, I show Lemma 3.1, namely, v_j^* is the value of the constrained optimization problem.

A.1 Necessity of Constraint (3.3)

For every strategy profile σ, let $\mathcal{H}^\sigma$ be the set of histories that occur with positive probability under σ. For every $h^t \in \mathcal{H}^\sigma$, let $\Theta^\sigma(h^t) \subset \Theta$ be the support of player 2's belief at h^t. The necessity of (3.3) is implied by:

Proposition A.1. *For every prior belief π, including those that do not have full support, if type θ_i is the lowest-cost type in the support of this prior belief, then her equilibrium payoff is no more than $1-\theta_i$ in all BNEs.*

Proof. Rank player 1's actions according to $H \succ L$. Given strategy profile σ and history $h^t \in \mathcal{H}^\sigma$, let

$$\overline{a}_1^\sigma(h^t) \equiv \max\Big\{ \bigcup_{\theta \in \Theta^\sigma(h^t)} \text{supp}\Big(\sigma_\theta(h^t)\Big)\Big\} \tag{A.1}$$

be the highest action played by player 1 with positive probability at h^t. By definition, for every BNE σ and $h^t \in \mathcal{H}^\sigma$, if $\sigma_2(h^t)$ assigns positive probability to T, then $\overline{a}_1^\sigma(h^t) = H$. The rest of my proof is done by induction on the number of types in the support of player 2's prior belief.

1. I establish the conclusion when $|\Theta^\sigma(h^0)| = 1$.
2. Suppose the conclusion holds when $|\Theta^\sigma(h^0)| \leq n$, it also holds when $|\Theta^\sigma(h^0)| = n+1$.

Step 1: I show that when $|\Theta^\sigma(h^0)| = 1$, the only type in the support of player 2's prior belief, denoted by θ_i, receives payoff no more than $1-\theta_i$. This also implies that for every equilibrium σ and for every $h^t \in \mathcal{H}^\sigma$, if $\Theta^\sigma(h^t) = \{\theta_i\}$ for some $\theta_i \in \Theta$, then type θ_i's continuation payoff at h^t cannot exceed $1-\theta_i$.

This is because $\Theta^\sigma(h^0) = \{\theta_i\}$ implies that $\Theta^\sigma(h^t) = \{\theta_i\}$ for every $h^t \in \mathcal{H}^\sigma$. Therefore, $\overline{a}_1^\sigma(h^t)$ is played by type θ_i with positive probability at every $h^t \in \mathcal{H}^\sigma$. Given type θ_i's equilibrium strategy σ_{θ_i}, the following strategy $\widetilde{\sigma}_{\theta_i} : \mathcal{H} \rightarrow \Delta(A_1)$, defined as:

$$\widetilde{\sigma}_{\theta_i}(h^t) \equiv \begin{cases} \overline{a}_1^\sigma(h^t) & \text{if } h^t \in \mathcal{H}^\sigma \\ \sigma_{\theta_i}(h^t) & \text{otherwise.} \end{cases} \tag{A.2}$$

also best replies against player 2's equilibrium strategy σ_2, from which type θ_i receives his equilibrium payoff. If type θ_i plays according to $\widetilde{\sigma}_{\theta_i}$ against σ_2, then according to Step 1, the outcome at every history in $\mathcal{H}^\sigma$ is either (T, H) or N. Therefore, type θ_i's stage-game payoff at every history in $\mathcal{H}^\sigma$ cannot exceed $1-\theta_i$, so his discounted average payoff cannot exceed $1-\theta_i$.

Step 2: I show that if the conclusion holds when $|\Theta^\sigma(h^0)| \leq n$, then it also holds when $|\Theta^\sigma(h^0)| = n+1$. I define $\overline{\mathcal{H}}_t^\sigma$ for every $t \in \mathbb{N}$ recursively. Let $\overline{\mathcal{H}}_0^\sigma \equiv \{h^0\}$. Given the definition of $\overline{\mathcal{H}}_t^\sigma$, let

$$\overline{\mathcal{H}}_{t+1}^\sigma \equiv \Big\{h^{t+1} \in \mathcal{H}^\sigma \Big| \exists h^t \in \overline{\mathcal{H}}_t^\sigma \text{ s.t. } h^{t+1} \succ h^t \text{ and either } h^{t+1} = (h^t, N) \text{ or } h^{t+1} = \big(h^t, (T, \overline{a}_1^\sigma(h^t))\big)\Big\}.$$

Intuitively, $\overline{\mathcal{H}}_{t+1}^\sigma$ is the set of period $t+1$ on-path histories such that player 1 has played his *highest on-path action* from period 0 to t. Let $\overline{\mathcal{H}}^\sigma \equiv \cup_{t=0}^\infty \overline{\mathcal{H}}_t^\sigma$.

Recall that θ_i is the notation for the lowest-cost type in the support of player 2's prior belief. Given type θ_i's equilibrium strategy σ_{θ_i}, let $\widehat{\sigma}_{\theta_i} : \mathcal{H} \to \Delta(A_1)$ be defined as:

$$\widehat{\sigma}_{\theta_i}(h^t) \equiv \begin{cases} \overline{a}_1^\sigma(h^t) & \text{if } h^t \in \mathcal{H}^\sigma \text{ and } \overline{a}_1^\sigma(h^t) \in \text{supp}\big(\sigma_{\theta_i}(h^t)\big) \\ \sigma_{\theta_i}(h^t) & \text{otherwise.} \end{cases} \tag{A.3}$$

By construction, $\widehat{\sigma}_{\theta_i}$ is type θ_i's best reply against σ_2. Let $\mathcal{H}^{(\widehat{\sigma}_{\theta_i}, \sigma_2)}$ be the set of histories that occur with positive probability under $(\widehat{\sigma}_{\theta_i}, \sigma_2)$. Let

$$\overline{\mathcal{H}}^{\sigma,\theta_i} \equiv \Big\{h^t \in \overline{\mathcal{H}}^\sigma \Big| \theta_i \in \Theta^\sigma(h^t) \text{ and } \overline{a}_1^\sigma(h^t) \notin \text{supp}\big(\sigma_{\theta_i}(h^t)\big)\Big\}. \tag{A.4}$$

Intuitively, $h^t \in \overline{\mathcal{H}}^{\sigma,\theta_i}$ if and only if

1. At every $h^s \prec h^t$, type θ_i's equilibrium strategy σ_{θ_i} plays $\overline{a}_1^\sigma(h^s)$ with positive probability. This comes from $h^t \in \overline{\mathcal{H}}^\sigma$ and $\theta_i \in \Theta^\sigma(h^t)$.
2. At h^t, type θ_i plays $\overline{a}_1^\sigma(h^t)$ with zero probability. This comes from $\overline{a}_1^\sigma(h^t) \notin \text{supp}\big(\sigma_{\theta_i}(h^t)\big)$.

Consider type θ_i's payoff if he plays $\widehat{\sigma}_{\theta_i}$ and player 2 plays σ_2. For any given $h^t \in \mathcal{H}^{(\widehat{\sigma}_{\theta_i}, \sigma_2)}$,

1. If there *does not exist* $h^s \preceq h^t$ such that $h^s \in \overline{\mathcal{H}}^{\sigma,\theta_i}$, then type θ_i's stage-game payoff at h^t and at all histories preceding h^t is no more than $1-\theta_i$.
2. If there *exists* $h^s \preceq h^t$ such that $h^s \in \overline{\mathcal{H}}^{\sigma,\theta_i}$, then I show below that type θ_i's continuation payoff at h^s is no more than $1-\theta_i$.

   First, since $h^s \in \overline{\mathcal{H}}^{\sigma,\theta_i}$, after player 2 observes $\overline{a}_1^\sigma(h^s)$ at h^s, θ_i is no longer in the support of player 2's posterior belief. Therefore, for every $h^{s+1} \succ h^s$ with $\overline{a}_1^\sigma(h^s)$ being played at h^s, there exist at most n types in the support of player 2's posterior belief at h^{s+1}.

   Let θ_j be the lowest-cost type in the support of P2's belief at h^{s+1}. According to the induction hypothesis, type θ_j's continuation payoff *after* playing $\overline{a}_1^\sigma(h^s)$ at h^s is no more than $1-\theta_j$. Type θ_j's stage-game payoff by playing $\overline{a}_1^\sigma(h^s)$ at h^s is also no more than $1-\theta_j$. This implies that his continuation payoff at h^s is at most $1-\theta_j$.

   Therefore, type θ_j's continuation payoff by deviating to $\widehat{\sigma}_{\theta_i}$ starting from h^s is no more than $1-\theta_j$. Since $\theta_i < \theta_j$, and the maximal difference between type θ_i and θ_j's stage-game payoff is $\theta_j - \theta_i$, we know that type θ_i's continuation payoff at h^s by playing $\widehat{\sigma}_{\theta_i}$ is no more than $1-\theta_i$.

The two parts together imply that when $|\Theta^\sigma(h^0)| = n+1$.

1. type θ_i's stage-game payoff before reaching any history that belong to $\overline{\mathcal{H}}^{\sigma,\theta_i}$ is no more than $1-\theta_i$,
2. type θ_i's continuation payoff at any history belonging to $\overline{\mathcal{H}}^{\sigma,\theta_i}$ is no more than $1-\theta_i$.

Therefore, θ_i's discounted average payoff in period 0 is no more than $1-\theta_i$. $\square$

A.2 Necessity of Constraint (3.4)

Suppose toward a contradiction that $(v_1, ..., v_m)$ is an equilibrium payoff and there exists $j \in \{1, 2, ..., m\}$ such that $v_j > v_j(\gamma^*)$. Then given the constraint established in the first part that $v_1 \leq 1 - \theta_1$, we know that $j > 1$. Under the probability measure over $\mathcal{H}$ induced by $(\sigma_{\theta_j}, \sigma_2)$, let $X^{(\sigma_{\theta_j}, \sigma_2)}$ be the occupation measure of outcome (T, H) and let $Y^{(\sigma_{\theta_j}, \sigma_2)}$ be the occupation measure of outcome (T, L). Since $v_j > v_j(\gamma^*)$, we have:

$$\frac{X^{(\sigma_{\theta_j}, \sigma_2)}}{Y^{(\sigma_{\theta_j}, \sigma_2)}} < \frac{\gamma^*}{1 - \gamma^*}. \tag{A.5}$$

Let the value of the left-hand-side be $\frac{\gamma}{1-\gamma}$ for some $\gamma \in [0, \gamma^*)$.

For every $h^\tau \in \mathcal{H}$, let $\sigma_{\theta_j}(h^\tau) \in \Delta(A_1)$ be the (mixed) action prescribed by σ_{θ_j} at h^τ and let $\alpha_1(\cdot|h^\tau)$ be player 2's expected action of player 1's at h^τ. Let $d(\cdot\|\cdot)$ be the Kullback-Leibler divergence between two action distributions. Suppose player 1 plays according to σ_{θ_j}, the result in Gossner (2011) implies that:

$$\mathbb{E}^{(\sigma_{\theta_j}, \sigma_2)}\Big[\sum_{\tau=0}^{+\infty} d(\sigma_{\theta_j}(h^\tau)||\alpha_1(\cdot|h^\tau))\Big] \leq -\log \pi_0(\theta_j). \tag{A.6}$$

This implies that for every $\epsilon > 0$, the expected number of periods such that $d(\sigma_{\theta_j}(h^\tau)||\alpha_1(\cdot|h^\tau)) > \epsilon$ is no more than

$$T(\epsilon) \equiv \Big\lceil \frac{-\log \pi_0(\theta_j)}{\epsilon} \Big\rceil. \tag{A.7}$$

Let

$$\epsilon \equiv d\Big(\frac{\gamma + 2\gamma^*}{3} H + (1 - \frac{\gamma + 2\gamma^*}{3}) L \Big\| \gamma^* H + (1 - \gamma^*) L\Big), \tag{A.8}$$

and let δ be large enough such that:

$$\frac{X^{(\sigma_{\theta_j}, \sigma_2)}}{Y^{(\sigma_{\theta_j}, \sigma_2)} - (1 - \delta^{T(\epsilon)})} < \frac{2\gamma + \gamma^*}{3 - 2\gamma - \gamma^*}. \tag{A.9}$$

According to (A.6) and (A.7), if type θ_j plays according to her equilibrium strategy, then there are at most $T(\epsilon)$ periods in which player 2's expectation over player 1's action differs from σ_{θ_j} by more than ϵ. According to (A.8), aside from $T(\epsilon)$ periods, player 2 will trust player 1 at h^t only when $\sigma_{\theta_j}(h^t)$ assigns probability at least $\frac{\gamma+2\gamma^*}{3}$ to H. Therefore, under the probability measure induced by $(\sigma_{\theta_j}, \sigma_2)$, the occupation measure with which player 2 trusts player 1 is at most:

$$\underbrace{(1 - \delta^{T(\epsilon)})}_{\text{periods in which player 2's prediction is wrong}} + \underbrace{\Big(X^{(\sigma_{\theta_j}, \sigma_2)} + Y^{(\sigma_{\theta_j}, \sigma_2)} - (1 - \delta^{T(\epsilon)})\Big)\frac{2\gamma + \gamma^*}{\gamma + 2\gamma^*}}_{\text{maximal frequency with which player 2 trusts after he learns}}, \tag{A.10}$$

which is strictly less than $X^{(\sigma_{\theta_j}, \sigma_2)} + Y^{(\sigma_{\theta_j}, \sigma_2)}$ when δ is close enough to 1. This leads to a contradiction.

A.3 Proof of Lemma 3.1

Constraint (3.4) implies that:

$$(1 - \theta_1)\alpha(H) + \alpha(L) \leq (1 - \theta_1)\alpha(H) + \frac{1 - \gamma^*}{\gamma^*}\alpha(H) = (\frac{1}{\gamma^*} - \theta_1)\alpha(H),$$

or equivalently,

$$\alpha(H) \geq \frac{\gamma^*}{1 - \gamma^* \theta_1}\Big((1 - \theta_1)\alpha(H) + \alpha(L)\Big).$$

Recall that θ_1 is the lowest-cost type. The objective function (3.2) can be rewritten as:

$$(1-\theta_j)\alpha(H)+\alpha(L) = (1-\theta_1)\alpha(H)+\alpha(L) - \underbrace{(\theta_j-\theta_1)}_{\geq 0} \underbrace{\alpha(H)}_{\geq \frac{\gamma^*}{1-\gamma^*\theta_1}\Big((1-\theta_1)\alpha(H)+\alpha(L)\Big)}$$

$$\leq \Big(1-(\theta_j-\theta_1)\frac{\gamma^*}{1-\gamma^*\theta_1}\Big)\Big((1-\theta_1)\alpha(H)+\alpha(L)\Big) = \frac{1-\gamma^*\theta_j}{1-\gamma^*\theta_1}\Big((1-\theta_1)\alpha(H)+\alpha(L)\Big).$$

According to constraint (3.3) that $(1-\theta_1)\alpha(H)+\alpha(L) \leq 1-\theta_1$, we have the following upper bound of the objective function (3.2):

$$(1-\theta_j)\alpha(H)+\alpha(L) \leq \frac{1-\gamma^*\theta_j}{1-\gamma^*\theta_1}\Big((1-\theta_1)\alpha(H)+\alpha(L)\Big) \leq (1-\gamma^*\theta_j)\frac{1-\theta_1}{1-\gamma^*\theta_1} = v_j^*.$$

The above upper bound is attained by the following distribution over action profiles:

$$\alpha(H) = \frac{(1-\theta_1)\gamma^*}{1-\gamma^*\theta_1}, \alpha(L) = \frac{(1-\theta_1)(1-\gamma^*)}{1-\gamma^*\theta_1}, \text{ and } \alpha(N) = \frac{\theta_1(1-\gamma^*)}{1-\gamma^*\theta_1},$$

which satisfies constraints (3.3) and (3.4). Therefore, the value of the optimization problem is v_j^*.

B Proof of Theorem 3

Statement 1: Suppose there exists type $\theta_i \neq \theta_m$ and a pure strategy $\widehat{\sigma}_{\theta_i}$ that is type θ_i's best reply to σ_2, such that

$$\frac{\mathbb{E}^{(\widehat{\sigma}_{\theta_i},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = H\}\Big]}{\mathbb{E}^{(\widehat{\sigma}_{\theta_i},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = L\}\Big]} = \frac{\gamma_i}{1-\gamma_i} \tag{B.1}$$

for some $\gamma_i < \gamma^*$. Let p_i be the discounted average frequency with which player 2 plays T under $(\widehat{\sigma}_{\theta_i}, \sigma_2)$.

Let $\widehat{\sigma}_{\theta_m}$ be an arbitrary pure-strategy best reply of type θ_m against σ_2. Let p_m be the discounted average frequency with which player 2 plays T under $(\widehat{\sigma}_{\theta_m}, \sigma_2)$ and let γ_m be pinned down via:

$$\frac{\mathbb{E}^{(\widehat{\sigma}_{\theta_m},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = H\}\Big]}{\mathbb{E}^{(\widehat{\sigma}_{\theta_m},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = L\}\Big]} = \frac{\gamma_m}{1-\gamma_m}. \tag{B.2}$$

The long-run player's ex ante incentive constraints, namely, first, type θ_i prefers $\widehat{\sigma}_{\theta_i}$ to $\widehat{\sigma}_{\theta_m}$, and second, type θ_m prefers $\widehat{\sigma}_{\theta_m}$ to $\widehat{\sigma}_{\theta_i}$ imply that $p_i \geq p_m$ and $\gamma_i \geq \gamma_m$. This further implies that according to type θ_m's equilibrium strategy σ_{θ_m},

$$\frac{\mathbb{E}^{(\sigma_{\theta_m},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = H\}\Big]}{\mathbb{E}^{(\sigma_{\theta_m},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = L\}\Big]} \leq \frac{\gamma_i}{1-\gamma_i},$$

or equivalently,

$$\gamma_i \mathbb{E}^{(\sigma_{\theta_m},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = L\}\Big] - (1-\gamma_i)\mathbb{E}^{(\sigma_{\theta_m},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = H\}\Big] > 0. \tag{B.3}$$

Since type θ_m's payoff from σ_{θ_m} is at least $v_m^* - \varepsilon$, which is strictly greater than $1-\theta_m$ when ε is small enough. This places a lower bound on p_m. If type θ_m plays according to σ_{θ_m}, then the learning arguments in Fudenberg and Levine (1992) and Gossner (2011) imply that for every $\varepsilon > 0$, there exists $\overline{\delta}$ such that when $\delta > \overline{\delta}$,

$$\gamma^* \mathbb{E}^{(\sigma_{\theta_m},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = L\}\Big] - (1-\gamma^*)\mathbb{E}^{(\sigma_{\theta_m},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{y_t = H\}\Big] < \varepsilon. \tag{B.4}$$

This contradicts (B.3) once we pick ε to be small enough, which establishes the lower bound on the relative frequencies of actions.

Statement 2: Suppose towards a contradiction that according to one of type θ ($\neq \theta_1$)'s pure-strategy best reply to σ_2, denoted by $\widehat{\sigma}_\theta$,

$$\frac{\mathbb{E}^{(\widehat{\sigma}_\theta,\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t=H\}\Big]}{\mathbb{E}^{(\widehat{\sigma}_\theta,\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t=L\}\Big]}=\frac{\gamma}{1-\gamma} \tag{B.5}$$

where $\gamma>\gamma^*$. Let

$$p\equiv \mathbb{E}^{(\widehat{\sigma}_\theta,\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t=H\}\Big]+\mathbb{E}^{(\widehat{\sigma}_\theta,\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t=L\}\Big].$$

If type θ_1 plays according to $\widehat{\sigma}_\theta$, her payoff is $p(1-\gamma\theta_1)$. According to Theorem 1,

$$p(1-\gamma\theta_1)\leq 1-\theta_1. \tag{B.6}$$

If type θ plays according to $\widehat{\sigma}_\theta$, she receives her equilibrium payoff, which is $p(1-\gamma\theta)$. The equilibrium payoff is within ε of v^* implies that:

$$p(1-\gamma\theta)\geq \frac{1-\theta_1}{1-\gamma^*\theta_1}(1-\gamma^*\theta)-\varepsilon. \tag{B.7}$$

Inequalities (B.6) and (B.7) together imply that:

$$\varepsilon>(1-\theta_1)\Big\{\frac{1-\gamma^*\theta}{1-\gamma^*\theta_1}-\frac{1-\gamma\theta}{1-\gamma\theta_1}\Big\}. \tag{B.8}$$

The RHS is strictly positive since $\gamma>\gamma^*$ and $\theta>\theta_1$. As a result, inequality (B.8) cannot hold for ε smaller than the RHS. For every $\gamma>\gamma^*$, take ε to be smaller than

$$\min_{\theta\neq\theta_1}(1-\theta_1)\Big\{\frac{1-\gamma^*\theta}{1-\gamma^*\theta_1}-\frac{1-\gamma\theta}{1-\gamma\theta_1}\Big\},$$

we obtain a contradiction. This establishes the upper bound on the relative frequencies.

C Payoff Types that Behave Like Stackelberg Commitment Type

I explore when can a payoff type mixes between H and L at every history, i.e., behaves like the Stackelberg commitment type in canonical reputation models. Players' stage-game payoffs are shown in Figure 6. Player 1's stage-game payoff is a function of her type $\theta\in\Theta$, where Θ is finite. Different from the baseline model which requires all types of player 1 sharing the same ordinal preferences over stage-game outcomes, the current setup allows for arbitrary preferences for player 1, and only requires one type of player 1 having the ordinal preference in the baseline model, which I call the *normal type*.

Proposition C.1. *Suppose there exists $\theta^*\in\Theta$, such that $y(\theta^*)>x(\theta^*)>z(\theta^*)$. For every small enough $\varepsilon>0$, there exists $\underline{\delta}\in(0,1)$, such that when $\delta>\underline{\delta}$, if there exists BNE such that:*

- *type θ^* attains within ε of her Stackelberg payoff,*
- *there exists $\theta'\in\Theta$ such that type θ' mixes between H and L at every history,*

then it must be the case that $x(\theta')=y(\theta')=z(\theta')$.

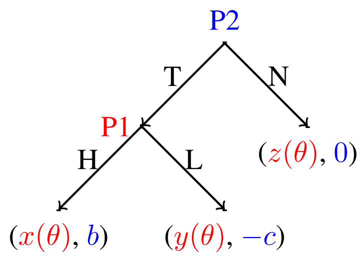

Figure 6: Generalized Stage Game

Proof. Suppose there exists such an equilibrium σ. For every $\varepsilon > 0$, there exists $\underline{\delta}$ such that when $\delta > \underline{\delta}$, such that

$$\frac{\mathbb{E}^{(\sigma_{\theta^*},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t = H\}\Big]}{\mathbb{E}^{(\sigma_{\theta^*},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t = L\}\Big]} \geq \frac{\gamma^* - \varepsilon}{1-\gamma^*+\varepsilon}. \tag{C.1}$$

Since type θ^*'s payoff is within ε of her Stackelberg payoff $\gamma^* x(\theta^*) + (1-\gamma^*)y(\theta^*)$, we have:

$$\mathbb{E}^{(\sigma_{\theta^*},\sigma_2)}\Big[\sum_{t=0}^{\infty}(1-\delta)\delta^t\mathbf{1}\{y_t = N\}\Big] \leq \underbrace{\frac{1+y(\theta^*)-x(\theta^*)}{(1-\gamma^*+\varepsilon)y(\theta^*)+(\gamma^*-\varepsilon)x(\theta^*)-z(\theta^*)}}_{\equiv C}\varepsilon. \tag{C.2}$$

For type θ' to mix at every history, she is indifferent between σ_{θ^*}, the strategy of playing H at every history, denoted by σ_H, and the strategy of playing L at every history, denoted by σ_L. Inequality (C.2) implies that unless $x(\theta') = y(\theta')$, type θ' receives equilibrium payoff *strictly* between $x(\theta')$ and $y(\theta')$. If type θ' plays σ_H, then her payoff is between $x(\theta')$ and $z(\theta')$. If type θ' plays σ_L, then her payoff is between $y(\theta')$ and $z(\theta')$, and it cannot be equal $y(\theta')$. These three payoffs coincide, which I denote by $\Pi(\theta')$. I consider two cases separately.

First, suppose $x(\theta') = y(\theta') \neq z(\theta')$. Then $\Pi(\theta')$ is a convex combination of $y(\theta')$ and $z(\theta')$, with the convex weight on $z(\theta')$ being less than the RHS of (C.2). Type θ' being indifferent between σ_{θ^*} and σ_L implies that type θ^*'s payoff by playing σ_L is at least:

$$(1-C\varepsilon)y(\theta^*) + C\varepsilon z(\theta^*),$$

which is strictly greater than his Stackelberg payoff when ε is small enough. This implies that type θ^* has a strict incentive to deviate to σ_L, which leads to a contradiction.

Next, suppose $x(\theta') \neq y(\theta')$. Consider three subcases. First, if $z(\theta') \geq \max\{y(\theta'), x(\theta')\}$, then any convex combination between $z(\theta')$ and the larger one among $x(\theta')$ and $y(\theta')$ is strictly larger than a number strictly between $x(\theta')$ and $y(\theta')$. This leads to a contradiction. Similarly, if $z(\theta') \leq \min\{y(\theta'), x(\theta')\}$, then any convex combination between $z(\theta')$ and the smaller one among $x(\theta')$ and $y(\theta')$ is strictly smaller than a number strictly between $x(\theta')$ and $y(\theta')$. This leads to a contradiction. Third, if $z(\theta')$ is strictly between $x(\theta')$ and $y(\theta')$. Then for a convex combination between $z(\theta')$ and $x(\theta')$ to equal a convex combination between $z(\theta')$ and $y(\theta')$, both of which attach convex weight 1 to $z(\theta')$, i.e., unless player 1 has played different actions in the past, otherwise, player 2 never plays T. On the other hand, if player 1 plays according to σ_{θ^*}, then the discounted average frequency of outcome N is close to 0. Hence, there exists h^t that occurs with positive probability under $(\sigma_{\theta^*}, \sigma_2)$, such that T is played with positive probability under h^t but not at any predecessor of h^t. This leads to a contradiction since by using strategy σ_H, player 1 can also reach history h^t, which implies that the discounted average frequency of N under (σ_H, σ_2) is strictly smaller than 1. $\square$

D Generalizations & Robustness

I state generalizations of my findings to a class of monotone-supermodular games. I setup the model and state my results when players move *simultaneously* in the stage game, although analogous results hold in sequential-move stage games. Proofs of these generalized results are available upon request.

By allowing the reputation-building player to have any finite number of actions, any number of rational types, and non-separability in players' payoffs, my results surpass the generality of many existing reputation papers that focus on equilibrium behavior. For example, Mailath and Samuelson (2001), Phelan (2006), Ekmekci (2011), Liu (2011), Jehiel and Samuelson (2012), and section 4 of Mathevet, Pearce, and Stacchetti (2019) focus on 2×2 games with only one type of player 1 being rational when they construct equilibrium and study the long-run player's equilibrium behavior.

Consider a repeated game in discrete time between an informed player 1 with discount factor $\delta_1 \in (0, 1)$, and an uninformed player 2 with discount factor $\delta_2 \in [0, 1)$.

- Player 1 has a perfectly persistent type $\theta \in \Theta$. Player 2's full support prior is $\pi \in \Delta(\Theta)$.
- Player 1's action $a_1 \in A_1$, player 2's action $a_2 \in A_2$. A pure action profile is $a \in A \equiv A_1 \times A_2$, and player 1's mixed action is denoted by $\alpha_1 \in \Delta(A_1)$.
- Players' stage game payoffs are $u_1(\theta, a_1, a_2)$ and $u_2(a_1, a_2)$, i.e., values are private.
- P2's stage-game best reply to $\alpha_1 \in \Delta(A_1)$ is $\text{BR}_2(\alpha_1)$, which is a nonempty subset of A_2.
- For every $\theta \in \Theta$, the set of type θ's pure Stackelberg action is $\arg\max_{a_1 \in A_1} \min_{a_2 \in \text{BR}_2(a_1)} u_1(\theta, a_1, a_2)$.
- Players' pure actions are perfectly monitored. Players *cannot* observe each other's mixed actions.
- I assume that Θ and A_1 are finite sets, and $|A_2| = 2$.

I start from Assumption 1 that is satisfied for generic u_1 and u_2:

Assumption 1. *Players' stage-game payoff functions u_1 and u_2 satisfy:*

1. *For every pure action $a_1 \in A_1$, $BR_2(a_1)$ is a singleton.*
2. *For every $\theta \in \Theta$, type θ has a unique pure Stackelberg action.*

Assumption 2 is called *monotone-supermodularity* (MSM), which captures the lack-of-commitment problem.

Assumption 2 (MSM). *Θ and A_2 are fully ordered sets, and A_1 is a lattice, such that:*

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,*
   $u_1(\theta, a_1, a_2)$ has weakly increasing differences in θ and a_2.
3. *$u_2(a_1, a_2)$ has strictly increasing differences in a_1 and a_2.*

To map Assumption 2 into the seller-buyer application, let a_1 be the quality of good the seller supplies, a_2 represents whether the buyer makes the purchase or not, or whether she purchases the customized or standardized version, and θ measures the efficiency of the seller's production technology. Assumption 2 requires that:

1. supply high quality is strictly costly for the seller, but he strictly benefits from buyers' trusting action;
2. a higher type faces lower cost to supply high quality, and values buyers' trust weakly more;
3. a buyer has stronger incentive to trust when her expectation of product quality is higher.

That being said, my general framework allows the seller to have

1. any finite number of rational types,

2. any finite number of actions, and the seller's effort can be multi-dimensional.

My framework surpasses the generality of Mailath and Samuelson (2001), Phelan (2006), Ekmekci (2011), Liu (2011), Jehiel and Samuelson (2012), section 4 of Mathevet, Pearce, and Stacchetti (2019) that focus on 2×2 stage-games with player 1 having only one rational type. It also allows for non-separability between θ and a_2, and between a_1 and a_2 in seller's payoff, i.e., my framework can accommodate, but is not limited to the case in which player 1's benefit from player 2s' trust being independent of her type and her action.

Let $\Theta \equiv \{\theta_1, \theta_2, ..., \theta_m\}$ with $\theta_1 \succ \theta_2 \succ ... \succ \theta_m$. For $i \in \{1, 2\}$, let $\overline{a}_i \equiv \max A_i$ and $\underline{a}_i \equiv \min A_i$. Under Assumption 2, $(\underline{a}_1, \underline{a}_2)$ is all types of player 1's minmax outcome.

Let $a_1^*(\theta) \in A_1$ be type θ's *pure Stackelberg action*, namely, her optimal commitment *if she can only commit to pure actions*. This is uniquely defined under the second statement of Assumption 1. I assume that the most efficient type of seller finds it optimal to commit to supply the highest quality.

Assumption 3. $a_1^*(\theta_1) = \overline{a}_1$.

Assumption 3 allows some types of the seller to be *inefficient*, i.e, their cost of supplying high quality is so high that they strictly prefer the minmax outcome $(\underline{a}_1, \underline{a}_2)$ to the highest outcome $(\overline{a}_1, \overline{a}_2)$.

I generalize Theorem 1 by characterizing every type of P1's *highest equilibrium payoff* when δ_1 is close to 1 and δ_2 is close to or equal to 0. For every $j \in \{1, 2, ..., m\}$, let v_j^* be the value of the following problem:

$$\max_{\alpha \in \Delta(A_1 \times A_2)} \sum_{a \in A} \alpha(a) u_1(\theta_j, a) \tag{D.1}$$

subject to:

$$\sum_{a \in A} \alpha(a) u_1(\theta_1, a) \leq u_1(\theta_1, \overline{a}_1, \overline{a}_2), \tag{D.2}$$

and for every $a_2^* \in A_2$ such that the marginal distribution of α on A_2 attaches positive probability to a_2^*,

$$a_2^* \in \arg\max_{a_2 \in A_2} \sum_{a_1 \in A_1} \alpha_1(a_1 | a_2^*) u_2(a_1, a_2), \tag{D.3}$$

with $\alpha_1(\cdot | a_2^*) \in \Delta(A_1)$ the distribution over player 1's actions conditional on a_2^* induced by joint distribution α.

Theorem 1'. *If players' stage-game payoffs satisfy Assumptions 1, 2 and 3, then for every $\varepsilon > 0$, there exist $\underline{\delta}_1 \in (0, 1)$ and $\overline{\delta}_2 \in (0, 1)$ such that for every $\delta_1 \in (\underline{\delta}_1, 1)$ and $\delta_2 \in [0, \overline{\delta}_2)$,*

1. *There exists no BNE in which type θ_1's payoff is strictly more than v_1^*. There exists no BNE in which type θ_j's payoff is strictly more than $v_j^* + \varepsilon$ for some $j \in \{2, 3, ..., m\}$.*

2. *There exists sequential equilibrium in which P1 attains payoff within ε of $v^* \equiv (v_1^*, ... v_m^*)$.*

According to Theorem 1', v_j^* is type θ_j's highest equilibrium payoff. Similar to the baseline model, the most efficient type θ_1 cannot strictly benefit from incomplete information and her maximal payoff in the repeated incomplete information game coincides with her highest equilibrium payoff in the repeated complete information game. Second, all types except for type θ_1 can strictly benefit from incomplete information. Moreover, every type's highest equilibrium payoff depends only on their own type and the most efficient type in the support of player 2s' prior belief. Third, v_j^* is pinned down by two constraints, which generalize constraints (3.3) and (3.4) in the baseline model. Constraint (D.2) says that the most efficient type P1 receives payoff no more than $u_1(\theta_1, \overline{a}_1, \overline{a}_2)$ under the distribution over stage-game outcome induced by type θ_j. Similar to the baseline model, this arises due to type θ_1's incentive constraint and that he cannot benefit from private information. This constraint is absent in commitment-type models as commitment types face no incentive constraint. Constraint (D.3) says that conditional on every a_2 that occurs with positive probability under α, P2 has an incentive to play a_2 against the conditional distribution over player 1's actions. This comes from P2's learning, namely, her prediction about P1's action is arbitrarily close to P1's action in the true state in all except for a bounded number of periods.

Let Θ^* be the set of types whose *mixed Stackelberg payoffs* are strictly greater than their minmax payoffs:

$$\Theta^* \equiv \Big\{\theta \Big| \text{ there exist } \alpha_1 \in \Delta(A_1) \text{ and } a_2 \in \text{BR}_2(\alpha_1) \text{ such that } u_1(\theta, \alpha_1, a_2) > u_1(\theta, \underline{a}_1, \underline{a}_2)\Big\}, \tag{D.4}$$

Given $\sigma_2 : \mathcal{H} \to \Delta(A_2)$, I say that $\sigma_\theta : \mathcal{H} \to \Delta(A_1)$ is *stationary* with respect to σ_2 if it takes the same value for every h^t that occurs with positive probability under $(\sigma_\theta, \sigma_2)$. I say that $\sigma_\theta : \mathcal{H} \to \Delta(A_1)$ is *completely mixed* with respect to σ_2 if $\sigma_\theta(h^t)$ has full support for every h^t that occurs with positive probability under $(\sigma_\theta, \sigma_2)$.

Theorem 2'. *If $|\Theta^*| \geq 2$ and stage-game payoffs satisfy Assumptions 1, 2, and 3, then for every small enough $\varepsilon > 0$, there exist $\underline{\delta}_1 \in (0,1)$ and $\overline{\delta}_2 \in (0,1)$ such that when $\delta_1 \in (\underline{\delta}_1, 1)$ and $\delta_2 < [0, \overline{\delta}_2)$, in any BNE $\sigma \equiv \big((\sigma_\theta)_{\theta\in\Theta}, \sigma_2\big)$ that attains payoff within ε of v^*:*

1. *for every $\theta \in \Theta^*$ and every $\widehat{\sigma}_\theta$ that is type θ's best reply against σ_2, $\widehat{\sigma}_\theta$ is not completely mixed.*
2. *for every $\theta \in \Theta^*$, σ_θ is not stationary.*

References

[1] Aumann, Robert and Michael Maschler (1965) *Repeated Games with Incomplete Information*, MIT Press.

[2] Bar-Isaac, Heski and Steven Tadelis (2008) "Seller Reputations," *Foundations and Trends in Micro*.

[3] Barro, Robert (1986) "Reputation in a Model of Monetary Policy with Incomplete Information," *Journal of Monetary Economics*, 17, 3-20.

[4] Benabou, Roland and Guy Laroque (1992) "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *Quarterly Journal of Economics*, 107(3), 921-958.

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

[6] Bohren, Aislinn (2018) "Using Persistence to Generate Incentives in a Dynamic Moral Hazard Problem," *Theoretical Economics*, forthcoming.

[7] Bower, Anthony, Steven Garber and Joel Watson (1997) "Learning About a Population of Agents and the Evolution of Trust and Cooperation," *International Journal of Industrial Organization*, 15(2), 165-190.

[8] Caruana, Guillermo and Liran Einav (2008) "A Theory of Endogenous Commitment," *Review of Economic Studies*, 75(1), 99-116.

[9] Chassang, Sylvain (2010) "Building Routines: Learning, Cooperation, and the Dynamics of Incomplete Relational Contracts," *American Economic Review*, 100(1), 448-465.

[10] Cripps, Martin and Jonathan Thomas (2003) "Some Asymptotic Results in Discounted Repeated Games of One-Sided Incomplete Information," *Mathematics of Operations Research*, 28, 433-462.

[11] Dellarocas, Chrysanthos (2006) "Reputation Mechanisms," *Handbook on Information Systems and Economics*, T. Hendershott edited, Elsevier Publishing, 629-660.

[12] Dilmé, Francesc (2018) "Reputation Building through Costly Adjustment," *Journal of Economic Theory*, forthcoming.

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

[14] 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.

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

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

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

[18] Hart, Sirgiu (1985) “Nonzero-Sum Two-Person Repeated Games with Incomplete Information,” *Mathematics of Operations Research*, 10(1), 117-153.

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

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

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

[25] Mailath, George and Larry Samuelson (2006) “Repeated Games and Reputations: Long-Run Relationships,” Princeton University Press.

[26] Mathevet, Laurent, David Pearce, and Ennio Stacchetti (2019) “Reputation and Information Design,” Working Paper, NYU.

[27] Milgrom, Paul and John Roberts (1982) “Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis,” *Econometrica*, 50(2), 443-459.

[28] Pei, Harry (2018) “Reputation Effects under Interdependent Values,” Working Paper.

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

[30] Pęski, Marcin (2014) “Repeated Games with Incomplete Information and Discounting,” *Theoretical Economics*, 9(3), 651-694.

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

[32] Shalev, Jonathan (1994) “Nonzero-Sum Two-Person Repeated Games with Incomplete Information and Known-Own Payoffs,” *Games and Economic Behavior*, 7(2), 246-259.

[33] Sobel, Joel (1985) "A Theory of Credibility," *Review of Economic Studies*, 52(4), 557-573.

[34] Tirole, Jean (1996) "A Theory of Collective Reputations (with applications to the persistence of corruption and to firm quality)," *Review of Economic Studies*, 63(1), 1-22.

[35] Weinstein, Jonathan and Muhamet Yildiz (2016) "Reputation without Commitment in Finitely Repeated Games," *Theoretical Economics*, 11(1), 157-185.

[36] Wiseman, Thomas (2005) "A Partial Folk Theorem for Games with Unknown Payoff Distributions," *Econometrica*, 73(2), 629-645.