Review of Economic Studies (2023) **90**, 1441–1469 https://doi.org/10.1093/restud/rdac052

Advance access publication 26 July 2022

Reputation Building under Observational Learning

HARRY PEI

Department of Economics, Northwestern University

First version received November 2020*; Editorial decision July* 2022*; Accepted July* 2022 (*Eds.*)

A patient seller interacts with a sequence of myopic consumers. Each period, the seller chooses the quality of his product, and a consumer decides whether to trust the seller after she observes the seller's actions in the last K periods (limited memory) and at least one previous consumer's action (observational learning). However, the consumer cannot observe the seller's action in the current period. With positive probability, the seller is a commitment type who plays his Stackelberg action in every period. I show that under limited memory and observational learning, consumers are concerned that the seller will not play his Stackelberg action when he has a positive reputation and will play his Stackelberg action after he has lost his reputation. Such a concern leads to equilibria where the seller receives a low payoff from building a reputation. I also show that my reputation failure result hinges on consumers' observational learning.

Key words: Reputation, Reputation failure, Imitation, Observational learning.

JEL Codes: C73, D82, D83

1. INTRODUCTION

Economists have long recognized that individuals and firms can benefit from good reputations. This idea was formalized by Fudenberg and Levine (1989), who show that a patient agent is guaranteed to receive a high payoff if he builds a good reputation. The intuition behind their result is that when other people observe the agent taking a particular action for a long time, they will be convinced that he will take the same action in the future. An insight from their result is that reputation concerns can significantly alleviate moral hazard problems, e.g., a seller may build a reputation for offering high-quality goods even when quality is hard to observe at the time of purchase.

However, Fudenberg and Levine (1989)'s result relies on an *unlimited record-keeping* assumption: the market can observe the entire history of the agent's behaviours. This assumption does not fit applications where consumers cannot observe the seller's records in the distant past. For example, in many informal markets, consumers have limited access to the seller's past records due to the lack of record-keeping institutions. The Better Business Bureau only reports complaints from the last 3 years, and some online platforms do not display reviews received by sellers in the distant past.

One may suspect that consumers' limited memory will lower the seller's returns from building reputations since each of the seller's actions is only observed by a bounded number of consumers.

The editor in charge of this paper was Adam Szeidl.

I show that the effects of limited memory on the seller's reputational incentives are more subtle. In particular, limited memory by themselves may not cause reputation effects to fail, but reputation effects will fail under limited memory when consumers can also observe other consumers' choices.

I study a repeated game between a patient seller and a sequence of myopic consumers, arriving one in each period and each plays the game only once. Players' stage-game payoffs satisfy a monotone-supermodularity condition. A leading example that satisfies my condition is the product choice game:[1]

Seller\consumer	**T**rust	**N**o trust
High effort	$1,2$	$-c_N,1$
Low effort	$1+c_T,-1$	$0,0$

where $c_N, c_T > 0$.

I use this example to illustrate my results throughout the introduction. The seller observes all past actions and is either a strategic type who maximizes his discounted average payoff or a commitment type who plays his Stackelberg action (in the example, it would be *high effort*) in every period.

My modelling innovation is in the monitoring structure. I assume that each consumer can only observe a bounded number of the seller's actions and can also observe her immediate predecessor's action. This assumption describes situations where consumers can learn about the seller's previous actions via word-of-mouth communication and can also learn about other consumers' choices via observational learning. I only require that each consumer can observe her immediate predecessor's action and can talk to at most a bounded number of other consumers before making her decision.

My main result shows that when the prior probability of commitment type is below some cutoff, there are equilibria where the patient seller receives his minmax payoff 0. This implies that even when the seller exerts high effort in every period, consumers may not trust him for a long time, causing inefficiencies in equilibrium. My result stands in contrast to Fudenberg and Levine (1989)'s result, which shows that when consumers can observe the entire history of the seller's actions, the patient seller receives at least his Stackelberg payoff 1 in *all equilibria*. My model also stands in contrast to Liu and Skrzypacz (2014) and existing reputation models with limited memory, which assume that consumers *cannot* observe other consumers' choices.

My reputation failure result is not obvious for two reasons. First, the strategic-type seller cannot exert low effort in every period in any low-payoff equilibrium. This is because, otherwise, consumers' posterior beliefs will assign probability one to the commitment type after they observe high effort, in which case the strategic type can obtain his Stackelberg payoff if he exerts high effort in every period. This suggests that the strategic type needs to exert high effort with a high enough probability to slow down consumers' learning, though the equilibrium probability with which he exerts high effort cannot be too high, since that will provide consumers a strict incentive to play T.

Second, my reputation failure result is not a direct consequence of limited memory. As a counterexample, I show that the seller receives at least his Stackelberg payoff in *all* equilibria when the seller's effort and consumers' trust are strategic complements (i.e. $c_N > c_T > 0$), and each consumer can only observe the seller's action in the period before but *cannot* observe other

1. Following Mailath and Samuelson (2015, page 168), I interpret "Trust" as purchasing a premium product or a customized product and "No Trust" as purchasing a standardized product. Under this interpretation, future consumers may observe the seller's effort even when the current-period consumer does not trust the seller.

consumers' choices. This implies that consumers' observational learning is not redundant for my reputation failure result.

I argue in two steps that my reputation failure result is driven by consumers' concern that the seller will exert low effort when he has a positive reputation (i.e. he will *milk* his reputation), and that after he loses his reputation, he will exert high effort until he has a positive reputation again (i.e. he will *rebuild* his reputation). Such a concern prompts consumers not to trust the seller even when they observe no low effort, and the seller receives his minmax payoff even if he always exerts high effort.

First, consumers have no such concern when they can observe the entire history of the seller's actions. This is because the seller loses his reputation forever after he exerts low effort, in which case he does not have the ability to rebuild his reputation. Consumers also have no such concern when they can only observe the seller's action in the period before, in which case the seller's incentive depends *only* on his action in the period before and the current-period consumer's action. This is because consumers trust the seller with a higher probability when the seller exerted high effort in the period before and the seller's effort and consumers' trust are strategic complements. Hence, the seller has a stronger incentive to exert high effort in the current period if he exerted high effort in the period before. Therefore, it is never optimal for the seller to milk his reputation and then rebuild his reputation.

Next, when each consumer can observe her immediate predecessor's action in addition to a bounded number of the seller's actions, I construct equilibria where (i) it is optimal for the seller to milk his reputation and then rebuild it, (ii) consumers may not trust the seller even when they observe no low effort, and (iii) the seller receives his minmax payoff. Such a construction is feasible since unlike the case without observational learning, each consumer's behaviour depends not only on the seller's action in the period before but also on her predecessor's action, so it is not necessarily true that consumers trust the seller with higher probability when the seller exerted high effort in the period before.

In the equilibrium I construct, the first consumer does not trust the seller. Each consumer imitates her predecessor with probability close to one, and with complementary probability, she trusts the seller when effort was high in the period before and vice versa. The strategic type exerts high effort when (H,T) was played in the period before. Otherwise, he plays a mixed action so that the unconditional probability of high effort is a half. The seller receives his minmax payoff when he exerts high effort in every period since the first consumer does not trust him and each consumer imitates her predecessor with probability close to one. Consumers have an incentive to play N when (H,N) was played in the period before, since (i) the strategic type exerts low effort with positive probability and (ii) the posterior probability of the commitment type is bounded above. This upper bound follows from the observation that in my equilibrium, the seller exerts high effort with a probability at least a half in each period, and each consumer only observes a limited number of the seller's actions.

My low-payoff equilibrium has two properties: When the seller plays H in every period, consumers never herd on action N and the seller receives a high *undiscounted average payoff*. These properties apply to *all* equilibria. First, consumers never herd on action N in any equilibrium for any prior belief and any discount factor. This stands in contrast to the canonical social learning results where inefficiencies are caused by herding. Second, if each consumer observes *all* previous consumers' choices and the seller's last K actions, then in *all* equilibria for any prior belief and any discount factor, the seller's undiscounted average payoff from exerting high effort in every period is at least $\frac{K}{K+1}$ times his Stackelberg payoff plus $\frac{1}{1+K}$ times his minimal stage-game payoff. When this guaranteed undiscounted average payoff is strictly greater than the seller's minmax payoff 0, the seller can eventually secure a positive payoff by building a reputation. This does not contradict my main result since the time it takes for the seller to secure this positive payoff is endogenous. For example, in my low-payoff equilibrium, it takes longer for a more patient seller to switch consumers' actions from N to T. The prolonged process of establishing a reputation wipes out the seller's benefit from that reputation.

I also study an extension where each consumer can observe a private signal about the seller's current action in addition to what she can observe in the baseline model. I focus on the case where each consumer observes all previous consumers' choices, which is reminiscent of the social learning models of Banerjee (1992), Bikhchandani, Hirshleifer and Welch (1992), and Smith and Sørensen (2000).

My reputation failure result extends to the case where the private signals are not very informative since the low-payoff equilibrium in the baseline model remains an equilibrium. However, if there is a realization of the private signal that is much more likely to occur when the seller exerts high effort compared to the case where he exerts low effort (i.e. the signal is *unboundedly informative*), then consumers have a strict incentive to trust the seller after they observe that signal realization. The low-payoff equilibrium in my baseline model unravels since consumers do not have an incentive to imitate their immediate predecessors. I formalize this logic and show that the seller can secure his Stackelberg payoff in *all* equilibria if and only if consumers' signals are unboundedly informative.

2. BASELINE MODEL

Time is indexed by $t=0,1,2,\ldots$ A long-lived Player 1 (e.g. a seller) interacts with an infinite sequence of short-lived player 2s (e.g. consumers). Each Player 2 plays the game only once. I use 2_t to denote the Player 2 who plays in Period t. Player 1 discounts future payoffs for two reasons. First, he exits the game after every period with probability $1-\delta_1$, and the game ends after he exits. Second, he is indifferent between receiving one unit of utility in period t and receiving δ_2 unit of utility in period $t-1$. I assume that $\delta_1,\delta_2\in(0,1)$, so that Player 1 discounts future payoffs by $\delta\equiv\delta_1\delta_2\in(0,1)$.

In Period t, Player 1 chooses $a_t\in A$ and Player 2_t chooses $b_t\in B$. I assume that both A and B are finite sets. Player $i\in\{1,2\}$'s stage-game payoff is $u_i(a_t,b_t)$. Let $\mathrm{BR}_2(a)\subset B$ denote Player 2's best reply to a. Player 1's (pure) Stackelberg action is $\arg\max_{a\in A}\left\{\min_{b\in \mathrm{BR}_2(a)} u_1(a,b)\right\}$.

Assumption 1. *Player 1 has a unique best reply to every pure action $b\in B$. Player 2 has a unique best reply to every pure action $a\in A$. Player 1 has a unique Stackelberg action.*

Since A and B are finite sets, Assumption 1 is satisfied for generic (u_1,u_2). Let a^* be player 1's Stackelberg action. I focus on games with monotone-supermodular payoffs, which have been studied in the reputation literature by Phelan (2006), Ekmekci (2011), and Liu (2011).

Assumption 2. *Players' stage-game payoffs (u_1,u_2) are monotone-supermodular if there exist a complete order on A, $\succ_A$, and a complete order on B, $\succ_B$, such that:*

1. *Player 1's payoff function $u_1(a,b)$ is strictly decreasing in a and is strictly increasing in b.*
2. *Player 2's payoff function $u_2(a,b)$ has strictly increasing differences in (a,b).*
3. *Player 1's Stackelberg action a^* is not the minimal element of A.*

The product choice game in the introduction satisfies Assumption 2 once players' actions are ranked according to $H\succ_A L$ and $T\succ_B N$. This is because consumers have stronger incentives to trust the seller when the seller exerts higher effort, the seller prefers to exert low effort but benefits from consumers' trust, and the seller's Stackelberg action H differs from his lowest-cost action L.

Before choosing a_t, Player 1 observes all past actions $(a_0,...,a_{t-1},b_0,...,b_{t-1})$, and his perfectly persistent type $\omega \in \{\omega_s,\omega_c\}$. Let ω_c stand for a *commitment type* who plays a^* in every period. Let ω_s stand for a *strategic type* who maximizes his discounted average payoff $\sum_{t=0}^{\infty}(1-\delta)\delta^t u_1(a_t,b_t)$. That is, Player 1's payoff is normalized so that the weight on his period-t payoff is $(1-\delta)\delta^t$ and the sum of the weights is 1. Let $\pi_0 \in (0,1)$ be the prior probability of the commitment type.

My modelling innovation is on Player 2's information structure. I assume that there exist $K \in \mathbb{N}$ and $M \in \mathbb{N} \cup \{+\infty\}$ such that for every $t \in \mathbb{N}$, Player 2_t can observe Player 1's actions in the last K periods $(a_{\max\{0,t-K\}},...,a_{t-1})$ and Player 2's actions in the last M periods $(b_{\max\{0,t-M\}},...,b_{t-1})$, where $M=+\infty$ means that every Player 2 can observe the entire history of her predecessors' choices.

1. I assume that K is finite. That is, every consumer observes a bounded number of the seller's actions. This stands in contrast to the reputation model of Fudenberg and Levine (1989) where every consumer observes the entire history of the seller's actions (i.e. $K=+\infty$).
2. I assume that $M \geq 1$. That is, every consumer can observe at least her immediate predecessor's action. This stands in contrast to existing reputation models with limited memory such as Liu (2011) and Liu and Skrzypacz (2014) where consumers cannot observe other consumers' choices.

I also assume that Player 2s cannot directly observe calendar time.[2] This assumption fits when consumers do not know when the game started but may infer it from their observations. Since the game ends after every period with probability $1-\delta_1$, the probability that the game *does not end* before period t is δ_1^t. This implies that for every $t \in \mathbb{N}$, the probability Player 2's belief assigns to calendar time being $t+1$ equals δ_1 times the probability her belief assigns to calendar time being t. Therefore, Player 2's belief assigns probability $\frac{\delta_1^t}{\sum_{s=0}^{+\infty}\delta_1^s}=(1-\delta_1)\delta_1^t$ to calendar time being t. A useful property is that when δ is close to 1, δ_1 is also close to 1. This implies that the game is likely to last for a long time. As a result, Player 2's belief assigns probability close to 0 to any particular calendar time. After observing $(a_{\max\{0,t-K\}},...,a_{t-1})$ and $(b_{\max\{0,t-M\}},...,b_{t-1})$, Player 2_t updates her belief about calendar time according to Bayes rule. For example, Player 2_t can perfectly infer calendar time if $t \leq \max\{K,M\}-1$ but may not be able to do so if $t \geq \max\{K,M\}$.

Let $\mathcal{H}_i$ be the set of player i's histories. A strategy of the strategic-type player 1 is $\sigma_1 : \mathcal{H}_1 \to \Delta(A)$ and a strategy of Player 2 is $\sigma_2 : \mathcal{H}_2 \to \Delta(B)$. The solution concept is Perfect Bayesian equilibrium.

3. MAIN RESULT

Recall that a^* is Player 1's Stackelberg action. Let $b^* \equiv \mathrm{BR}_2(a^*)$. Player 1's *Stackelberg payoff* is $u_1(a^*,b^*)$. Let a' be the minimal element of A. Let $b' \equiv \mathrm{BR}_2(a')$. The first two parts of Assumption 2 imply that (a',b') is the unique stage-game Nash equilibrium and that $u_1(a',b')$ is player 1's *minmax payoff* in the sense of Fudenberg, Kreps and Maskin (1990). The third part of Assumption 2 implies that $a^* \neq a'$. This together with Assumption 1 implies that $u_1(a',b') < u_1(a^*,b^*)$.

2. My theorems extend to the case where Player 2 can directly observe calendar time. However, Proposition 1 requires Players 2 not being able to observe calendar time. I provide a counterexample in Supplementary Appendix D.

Theorem 1. *For every memory length $K \in \mathbb{N}$ and every stage-game payoff (u_1, u_2) that satisfies Assumptions 1 and 2, there is a cutoff discount factor $\underline{\delta} \in (0,1)$ and an upper bound on the prior probability of the commitment type $\overline{\pi}_0 > 0$,[3] such that for every $\pi_0 < \overline{\pi}_0$ and $\delta > \underline{\delta}$, there is a PBE in which player 1's discounted average payoff equals his minmax payoff $u_1(a', b')$.*

The proof is in Appendix A. According to Theorem 1, there exist equilibria in which the patient seller receives his minmax payoff when the probability of commitment type is below some cutoff, each consumer observes a limited number of the seller's actions, and can observe other consumers' choices.

The existence of low-payoff equilibria stands in contrast to the reputation result in Fudenberg and Levine (1989), which shows that the patient seller receives at least his Stackelberg payoff in *all* equilibria when each consumer observes the entire history of the seller's actions (i.e. $K = +\infty$). This holds for all π_0 and regardless of how many of their predecessors' actions each consumer can observe.

My reputation failure result is not immediately obvious for two reasons. First, repeated play of the stage-game Nash equilibrium (a', b') cannot lead to a low-payoff equilibrium in the reputation game. This is because under such a strategy profile, Player 2's posterior belief assigns probability 1 to the commitment type after observing a^*, which implies that Player 1 can secure payoff approximately $u_1(a^*, b^*)$ by playing a^* in every period. Hence, in every low-payoff equilibrium, the strategic type needs to play a^* with positive probability in order to slow down Player 2's learning, but he cannot play a^* with probability close to 1 since that will provide Player 2 a strict incentive to play b^*. This suggests the need to leverage more subtle forces in order to obtain a low-payoff equilibrium.

Second, Theorem 1 is not a direct consequence of limited memory. For a counterexample, take the product choice game with an additional supermodularity assumption that $0 < c_T < c_N$, I show that the patient player receives at least his Stackelberg payoff in *all* equilibria when $(K, M) = (1, 0)$. The comparison between this example and Theorem 1 implies that the short-run players' observational learning is not redundant for my reputation failure result.

Proposition 1. *Suppose $0 < c_T < c_N$ and $(K, M) = (1, 0)$. For every $\pi_0 > 0$, there exists $\underline{\delta} \in (0,1)$, such that when $\delta > \underline{\delta}$, player 1's payoff is at least $\delta - (1-\delta)c_N$ in every equilibrium.*

The proof is in Appendix B. Proposition 1 is related to Liu and Skrzypacz (2014), who study a reputation model where consumers observe the seller's last K actions but *cannot* observe previous consumers' choices and *cannot* observe calendar time. In contrast to Proposition 1, they assume that the seller's payoff is *submodular*, which translates into $c_T > c_N > 0$ in the product choice game. They show that for every $\varepsilon > 0$ and $\pi_0 \in (0,1)$, there exists $K(\varepsilon, \pi_0) \in \mathbb{N}$ such that when $K > K(\varepsilon, \pi_0)$, the patient seller's payoff at every history of every stationary equilibrium is at least $(1-\delta^K)u_1(a^*, b') + \delta^K u_1(a^*, b^*) - \varepsilon$.[4] That is, their reputation result requires a large K. This stands in contrast to Proposition 1, which shows that the patient seller can secure his Stackelberg

3. Player 1's discount factor needs to be above some cutoff $\underline{\delta}$, which ensures that the strategic type has an incentive to play the Stackelberg action although doing so gives him a lower stage-game payoff. In the product choice game, $\underline{\delta} = \max\{\frac{c_T}{c_T+1}, \frac{c_N}{c_N+1}\}$. By the end of Appendix A, I discuss how large $\underline{\delta}$ needs to be in general.

4. A *stationary equilibrium* is a PBE in which Player 1's action in period t depends only on Player 2_t's history. The equilibria I construct in the proof of Theorem 1 are stationary equilibria. This is because Player 1's action in period t depends only on (a_{t-1}, b_{t-1}) and his reputation in period t, both of which are measurable with respect to Player 2_t's information. This implies that my reputation failure result is also true when we restrict attention to stationary equilibria.

payoff in all equilibria even when π_0 is close to 0 and $K=1$, provided that the seller's payoff is supermodular.

Proof Sketch of Theorem 1: I explain how the proof of Theorem 1 works using the product choice game. I focus on the case where $(K,M)=(1,1)$, $\pi_0 \leq \frac{1}{9}$, and $\delta \geq \max\{\frac{c_T}{c_T+1}, \frac{c_N}{c_N+1}\}$. I construct a class of equilibria called *imitation equilibria* where Player 1's discounted average payoff is 0.

In every imitation equilibrium, player 2_t's action depends only on (a_{t-1}, b_{t-1}), and Player 1's action in period t depends only on (a_{t-1}, b_{t-1}) as well as his *reputation* π_t, which is the probability Player 2_t's belief assigns to the commitment type after observing (a_{t-1}, b_{t-1}). Play consists of four phases:

1. **Mistrust phase:** When $t=0$ or $(a_{t-1}, b_{t-1})=(L,N)$, Player 2_t plays N and the strategic-type Player 1 plays H with probability $\frac{1-2\pi_0}{2-2\pi_0}$ if $t=0$ and plays H with probability $\frac{1}{2}$ if $t \geq 1$.
2. **Doubting phase:** When $(a_{t-1}, b_{t-1})=(H,N)$, Player 2_t plays T with probability $r_1 \equiv \frac{1-\delta}{\delta} c_N$ and the strategic-type Player 1 plays H with probability $p_t \equiv \frac{1-2\pi_t}{2-2\pi_t}$.
3. **Testing phase:** When $(a_{t-1}, b_{t-1})=(L,T)$, Player 2_t plays T with probability $r_2 \equiv 1-\frac{1-\delta}{\delta} c_T$ and the strategic-type Player 1 plays H with probability $\frac{1}{2}$.
4. **Trusting phase:** When $(a_{t-1}, b_{t-1})=(H,T)$, Player 2_t plays T and Player 1 plays H.

I depict the phase transitions in Figure 1. My imitation equilibrium has two features. First, consumers do not trust the seller in period 0, which is interpreted as saying that consumers do not trust sellers who newly arrive and have no past record. Second, for every $t \geq 1$, consumer t plays N with probability 1 or close to 1 when $b_{t-1}=N$, and plays T with probability 1 or close to 1 when $b_{t-1}=T$. Hence, imitation equilibria describe situations where the first consumer does not trust the seller and every consumer imitates her predecessor with high probability. If the seller plays H in every period, then consumers first take action N and then switch to T at some random time. Since the probability of switching is strictly positive conditional on the seller's effort being high, the seller's asymptotic payoff equals his Stackelberg payoff 1. However, the probability with which each consumer takes a different action compared to her immediate predecessor is close to 0, so switching from N to T takes a long time in expectation. This explains why the seller's discounted average payoff is 0.

Next, I check players' incentive constraints. Every consumer best replies to the seller's action at every (a_{t-1}, b_{t-1}). The seller's continuation value depends only on (a_{t-1}, b_{t-1}), which is denoted by $V(a_{t-1}, b_{t-1})$. The seller's incentive constraints and promise-keeping constraints are satisfied when $V(H,T)=1$, $V(L,N)=0$, $V(H,N)=\frac{1-\delta}{\delta} c_N$, and $V(L,T)=1-\frac{1-\delta}{\delta} c_T$. To see this, note that

1. When $t=0$ or $(a_{t-1}, b_{t-1})=(L,N)$, the seller's discounted average payoff from playing L is 0, and his discounted average payoff from playing H is $(1-\delta)(-c_N)+\delta V(H,N)=V(L,N)=0$.
2. When $(a_{t-1}, b_{t-1})=(H,N)$, the seller's discounted average payoff from playing L is

$$(1-\delta)u_1(L, r_1 T+(1-r_1)N)+\delta\{r_1 V(L,T)+(1-r_1)V(L,N)\}=\frac{1-\delta}{\delta} c_N = V(H,N),$$

and his discounted average payoff from playing H is

$$(1-\delta)u_1(H, r_1 T+(1-r_1)N)+\delta\{r_1 V(H,T)+(1-r_1)V(H,N)\}=\frac{1-\delta}{\delta} c_N = V(H,N).$$

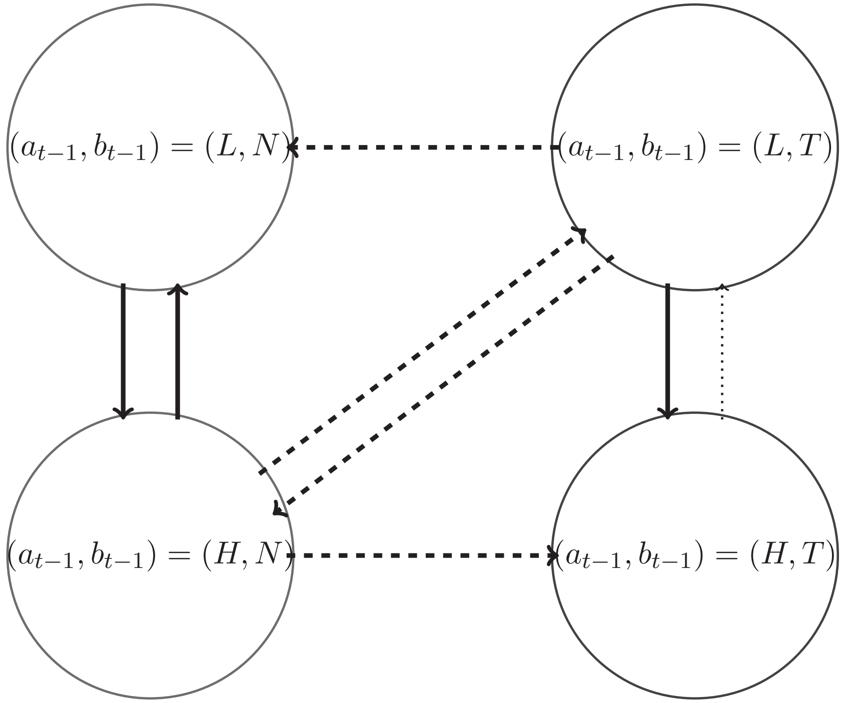

FIGURE 1

Phase transitions in imitation equilibria. Play starts from (L,N). The dashed lines denote events that occur with probability proportional to $1-\delta$ (i.e. close to 0). The solid lines denote events that occur with probability bounded away from 0. The dotted line denotes an event that occurs with zero probability in equilibrium but occurs with positive probability when the strategic seller deviates. In every phase, play stays in the same phase with probability bounded away from 0, which I did not draw explicitly.

3. When $(a_{t-1},b_{t-1})=(L,T)$, the seller's discounted average payoff from playing L is

$$(1-\delta)u_1(L,r_2T+(1-r_2)N)+\delta\{r_2V(L,T)+(1-r_2)V(L,N)\}=1-\frac{1-\delta}{\delta}c_T=V(L,T),$$

and his discounted average payoff from playing H is

$$(1-\delta)u_1(H,r_2T+(1-r_2)N)+\delta\{r_2V(H,T)+(1-r_2)V(H,N)\}=1-\frac{1-\delta}{\delta}c_T=V(L,T).$$

4. When $(a_{t-1},b_{t-1})=(H,T)$, the seller's discounted average payoff from playing H is 1, and his discounted average payoff from playing L is no more than 1.

Note that conditional on $b_t=T$, the seller's continuation value is 1 regardless of his action, and conditional on $b_t=N$, the seller's continuation value is 0 regardless of his action.

I verify that when $\pi_0\leq\frac{1}{9}$, p_t is a well-defined probability for every $t\in\mathbb{N}$. In Appendix A, I use a fixed point argument to show that $\pi_t\leq\frac{1}{3}$ whenever $(a_{t-1},b_{t-1})=(H,N)$,[5] which implies that

5. My fixed point argument applies when calendar time is not observed and M is finite. When $M=+\infty$ or when Player 2s can directly observe calendar time, I bound π_t from above using an induction argument. See Appendix A.

$p_t \in [\frac{1}{4}, 1]$. Intuitively, this is because (i) observing $a_{t-1} = H$ only provides consumer t limited information about the seller's type given that the strategic-type seller mixes between high and low effort, so consumer t's posterior belief about the commitment type cannot increase by too much compared to her prior; (ii) observing $b_{t-1} = N$ lowers the seller's reputation since N is more likely to occur under the strategy of the strategic type compared to the strategy of the commitment type. □

Discussions. I explain why my construction breaks down in cases where $K = +\infty$ and $(K, M) = (1, 0)$. The comparison between these two cases and the case where $(K, M) = (1, 1)$ highlights the roles of limited memory and observational learning in my reputation failure result.

The above comparison also unveils the mechanism behind Theorem 1: When K is finite and $M \geq 1$, consumers have a justified concern that the seller will milk his reputation and then rebuild his reputation after losing it. Such a concern prompts consumers not to trust the seller even when they do not observe any low effort, causing reputation effects to fail. By contrast, the seller does not have the ability to rebuild his reputation when $K = +\infty$, and he has no incentive to milk his reputation and then rebuild it when $(K, M) = (1, 0)$ and players' actions are strategic complements. In these cases, the seller can secure his Stackelberg payoff in all equilibria since consumers are free of such concerns.

The role of bounded memory. When $K = +\infty$, the seller loses his reputation forever after he exerts low effort, in which case consumers have no concerns about him rebuilding his reputation.

The imitation equilibria I constructed break down when $K = +\infty$ since it is no longer rational for consumers to imitate their immediate predecessors. To see this, note that for every $t \in \mathbb{N}$, either consumer t believes that $a_t = H$ with probability more than $\frac{1}{2}$, in which case she has a strict incentive to play T, or the probability consumer $t+1$ assigns to the commitment type after she observes $a_t = H$ is at least two times the probability consumer t assigns to the commitment type. When the strategic-type seller deviates and plays H in every period, there can be at most a bounded number of consumers who have incentives to imitate a predecessor who played N, which unravels my imitation equilibria.

By contrast, consumers' imitation behaviours *can* be rationalized when each consumer observes only a bounded number of the seller's actions and the probability of the commitment type is low enough:

1. Consumers may not be convinced that H will be played in the future after observing H in at most K periods. This is because even when consumer t believes that H will be played with probability less than $\frac{1}{2}$, consumer $t+1$'s posterior belief may not be greater than consumer t's since she cannot observe a_{t-K}. When the seller plays H in every period, consumers after period K obtain the same information from the seller's actions. Unlike the case where $K = +\infty$, there can be infinitely many consumers who are concerned that the seller will play L in the future.
2. Although consumers may learn from other consumers' choices, the information each consumer obtains from them never discourages her from imitating her immediate predecessor. Intuitively, when the seller is the commitment type, consumers never play N after they have played T. If $b_{t-1} = N$ and consumer t's posterior belief assigns positive probability to the commitment type after observing $(a_{\max\{0,t-K\}}, \ldots, a_{t-1})$ and $(b_{\max\{0,t-M\}}, \ldots, b_{t-1})$, it must be the case that $(b_{\max\{0,t-M\}}, \ldots, b_{t-1}) = (N, \ldots, N)$. Since the probability with which previous consumers playing N is greater when the seller is strategic, observing $(b_{\max\{0,t-M\}}, \ldots, b_{t-1}) = (N, \ldots, N)$ lowers the seller's reputation, which encourages consumer t to imitate consumer $t-1$.

The role of observational learning. When $(K,M)=(1,0)$, consumers' strategies in imitation equilibria are no longer feasible since consumer t cannot observe b_{t-1}. Since consumers cannot observe calendar time, consumer t's probability of playing T depends only on a_{t-1}, which I denote by $\beta(a_{t-1})\in[0,1]$. The seller's period-t continuation value depends only on a_{t-1}, which I denote by $V(a_{t-1})\in\mathbb{R}$.

The first observation is that $\beta(H)>\beta(L)$ in all equilibria. This is because when $\beta(H)\leq\beta(L)$, the strategic seller has no incentive to play H since playing H is costly in the stage game, does not increase the probability of being trusted in the next period, and has no impact on consumers' behaviour after two periods. Hence, each consumer's posterior belief assigns probability 1 to the commitment type after observing $a_{t-1}=H$. As a result, $1=\beta(H)\leq\beta(L)=1$. However, consumer t strictly prefers N when $a_{t-1}=L$ since the seller will play L in period t. This contradicts $\beta(L)=1$.

Under a generic condition on the seller's stage-game payoff function that $c_T\neq c_N$, it is impossible to keep the seller indifferent both when $a_{t-1}=H$ and when $a_{t-1}=L$. This is because the seller's indifference requires both $(1-\delta)u_1(H,\beta(H))+\delta V(H)=(1-\delta)u_1(L,\beta(H))+\delta V(L)$ and $(1-\delta)u_1(H,\beta(L))+\delta V(H)=(1-\delta)u_1(L,\beta(L))+\delta V(L)$, which together imply that

$$V(H)-V(L)=\frac{1-\delta}{\delta}\Big(u_1(L,\beta(H))-u_1(H,\beta(H))\Big)=\frac{1-\delta}{\delta}\Big(u_1(L,\beta(L))-u_1(H,\beta(L))\Big). \quad (3.1)$$

When $c_T\neq c_N$, $u_1(L,\beta(H))-u_1(H,\beta(H))=u_1(L,\beta(L))-u_1(H,\beta(L))$ if and only if $\beta(H)=\beta(L)$. Since I have shown that $\beta(H)>\beta(L)$, the seller must have a strict incentive at some histories.

I use this observation to show Proposition 1 that when $c_N>c_T>0$, the patient seller can secure his Stackelberg payoff in all equilibria. Intuitively, since $\beta(H)>\beta(L)$ and $c_N>c_T>0$, the strategic type has a stronger incentive to play H in period t when $a_{t-1}=H$. This implies that either

- *The seller has no incentive to milk his reputation* when $a_{t-1}=H$. In this case, as long as $a_{t-1}=H$, both types of the seller will play H in period t and consumer t will play T.
- Or *the seller has no incentive to rebuild his reputation*, i.e., the seller never plays H after playing L. I show in Appendix B that if this is the case and δ is close to 1, consumer t prefers T after observing $a_{t-1}=H$. Intuitively, if the strategic-type seller has no incentive to rebuild his reputation, then there is at most one period where he plays L given that he played H before. Since the commitment type plays H in every period and consumers do not observe calendar time, consumers' posterior belief after observing $a_{t-1}=H$ assigns significantly higher probability to *the seller being the commitment type* relative to the event that *the seller will milk his reputation in the current period*. The last step uses the observation that consumers' belief assigns probability close to 0 to any particular calendar time when $\delta\to 1$.[6]

By contrast, in the case *with* observational learning (e.g. when $(K,M)=(1,1)$), consumers' ability to observe their predecessors' choices enriches their strategy space, under which it can be optimal for the seller to milk his reputation and then rebuild his reputation. When a consumer is concerned that the seller will behave in such a way, she has a rationale for not trusting the seller

6. In Supplementary Appendix D, I present a counterexample which shows that Proposition 1 fails when consumers can directly observe calendar time. Intuitively, when consumers' strategy can depend on calendar time, although at any given calendar time, the seller either has no incentive to milk his reputation or has no incentive to rebuild his reputation, *he may have an incentive to milk his reputation at some calendar time and then rebuild his reputation at another calendar time*. This provides consumers a rationale for playing N even after they observe the seller played H in the period before.

despite not observing any low effort. This is because upon observing $a_{t-1}=H$, consumers are not convinced that the strategic seller will exert high effort in the current period due to the seller's incentives to milk his reputation. They are also not convinced that the seller is likely to be the commitment type due to the seller's incentives to rebuild his reputation. This is because when the strategic type has an incentive to rebuild his reputation, there can be infinitely many periods where the strategic type exerts low effort despite having exerted high effort in the period before. Unlike the case where the strategic type has no incentive to rebuild his reputation, consumer t's posterior belief after observing $a_{t-1}=H$ can assign probability more than $\frac{1}{2}$ to $a_t=L$, which provides her an incentive to play N.

My proof of Theorem 1 confirms this intuition by showing that consumers' concerns about the seller's milking-and-then-rebuilding behaviour can arise in equilibrium and can provide them an incentive to play N even when they do not observe any low effort. This is the case when every consumer imitates her predecessor with probability close to 1 so that the seller is indifferent at every history. In the equilibrium I construct, the seller milks his reputation with positive probability when $(a_{t-1},b_{t-1})=(H,N)$ and rebuilds his reputation with positive probability when $a_{t-1}=L$. This provides consumer t a rationale for not trusting the seller even when H was played in the last K periods. The seller receives his minmax payoff when he plays H in every period, since play gets stuck at (H,N) for a long time.

Practical relevance. My imitation equilibria have two qualitative features: (i) consumers do not trust newly arrived sellers who have no past record and (ii) consumers imitate their predecessors. Both features are plausible and are supported by empirical evidence.

On consumer imitation, Cai, Chen and Fang (2009) find that consumers imitate each other in the Chinese food market. Zhang (2010) finds that patients are more likely to refuse a kidney that has been refused by earlier patients, even conditional on the objective quality of kidneys. Cai, De Janvry and Sadoulet (2015) find that farmers in rural China are more likely to purchase weather insurance when they were told that other farmers had purchased the insurance.[7]

My *no initial trust* condition fits some informal markets in developing countries. For example, Michelson, Fairbairn, Ellison, Maertens and Manyong (2021) find that many farmers in Tanzania suspect the fertilizers sold in local markets are adulterated and their pessimistic beliefs about the seller's integrity persists over time. Such persistent mistrust contributes to the under-adoption of fertilizers.

Although details about farmers' information structures are not available, three characteristics of this market make my model a plausible fit. First, farmers' payoffs depend on the seller's action: namely, whether the seller has adulterated products currently sold on the market. Second, farmers are myopic; that is, they won't trust the seller if they believe that his products are adulterated and won't punish the seller if they believe that his products are authentic. Although some farmers may buy multiple times, they are unlikely to sacrifice their current-period profits, since most of them have low incomes and cannot afford to do so. Third, I require that every farmer observes the choice of her predecessor and a limited number of the seller's actions. This is plausible when farmers live close to each other, so that it is easy to observe other farmers' recent choices, and farmers have limited memory about the seller's actions. My result suggests a rationale for persistent mistrust when farmers do not trust the seller in the beginning due to a pessimistic prior and are unwilling to trust the seller even after they observe him supplying authentic products in a bounded number of periods.

7. Cai *et al.* (2015) write on page 82 that "*...when we told farmers about other villagers decisions, these decisions strongly influenced their own take-up choices...*", and "*...if information on other villagers decisions can be revealed in complement to the performance of the network, it can have a large impact on adoption decisions...*"

3.1. *Connections with canonical social learning models*

The imitation equilibria constructed in the proof of Theorem 1 are reminiscent of the canonical results on social learning. In Banerjee (1992), Bikhchandani, Hirshleifer and Welch (1992), and Smith and Sørensen (2000), a sequence of players choose their actions sequentially after observing all predecessors' actions and a private signal of some exogenous state. Inefficiencies take the form of *herding* in the sense that myopic players ignore their private signals and imitate their immediate predecessors.

My model is analogous once we view $(a_{\max\{0,t-K\}},...,a_{t-1})$ as player 2_t's private signal. The conceptual difference is that in my model, the myopic players' payoffs depend only on the patient player's endogenous actions not on the patient player's exogenous type. The myopic players never herd on action N in imitation equilibria since their actions are responsive to the seller's action in the period before. Proposition 2 shows that the *no bad herd* conclusion applies more generally. Formally, I say that Player 2s herd on action b at $h^t \equiv (a_s,b_s)_{s\leq t-1}$ if Player 2s play b at h^t and at every successor of h^t. Let $\pi(h^t)\in[0,1]$ be the probability Player 2_t's belief at h^t assigns to the commitment type.

Proposition 2. *Suppose players' payoffs satisfy Assumption 1. Then for every $(\delta,\pi_0)\in(0,1)^2$, every $b\neq b^*$, and every equilibrium, Player 2s cannot herd on b at any history h^t with $\pi(h^t)>0$.*

The proof is in Supplementary Appendix A, which considers separately the case where M is finite and M is infinite. Proposition 2 implies that as long as Player 1 imitates the commitment type, Player 2s can never herd on any action that does not best reply to a^* regardless of Player 1's discount factor, Player 2's prior belief, and the equilibrium we focus on. This implies that reputation failure cannot be caused by myopic players herding on actions that give the patient player a low payoff.

For a heuristic explanation, once Player 2s herd on action $b\neq b^*$, the strategic-type Player 1 cannot affect Player 2s' future actions, so he has no intertemporal incentives. As a result, the strategic-type Player 1 will not play a^* when a^* is not a best reply to b in the stage game.[8] This implies that Player 2 will learn that Player 1 is the commitment type upon observing a^*, and hence, will have a strict incentive to play b^*. This contradicts the hypothesis that Player 2s herd on action $b\neq b^*$.

3.2. *Connections with canonical reputation models*

Fudenberg and Levine (1992) show that a patient player can secure his Stackelberg payoff in all equilibria if (i) with positive probability, he is a commitment type who plays his Stackelberg action in every period and (ii) every short-run player can observe the entire history of a noisy signal that can statistically identify the patient player's action. An elegant proof of their result is provided by Gossner (2011). The key is to show that for any $\delta\in(0,1)$ and any equilibrium (σ_1,σ_2),

$$\mathbb{E}^{(a^*,\sigma_2)}\Big[\sum_{t=0}^{\infty} d\Big(y_t(\cdot|a^*)\Big|\Big|y_t(\cdot)\Big)\Big] \leq -\log\pi_0, \tag{3.2}$$

where $y_t(\cdot)$ is the equilibrium distribution of Player 2's signals about a_t, $y_t(\cdot|a^*)$ is the distribution of Player 2's signals about a_t conditional on Player 1 being the commitment type, $d(\cdot||\cdot)$ is

8. In the case where a^* is Player 1's myopic best reply to b, both types of player 1 play a^* in equilibrium after Player 2s herd on action b. When both types of Player 1 play a^*, Player 2 has a strict incentive to play b^*, which is not b.

the Kullback–Leibler divergence between two distributions, and $\mathbb{E}^{(a^*,\sigma_2)}[\cdot]$ is the expectation operator when player 1 plays a^* in every period and Player 2 plays σ_2. When Player 2s' signals can identify Player 1's actions, $d\big(y_t(\cdot|a^*)\big|\big|y_t(\cdot)\big)$ is bounded away from 0 whenever Player 2_t does not have a strict incentive to play b^*. Inequality (3.2) implies that in expectation, there can be at most a bounded number of periods in which Player 2s do not have strict incentives to play b^*. Importantly, this upper bound does not depend on δ. This explains why Player 1's equilibrium payoff is at least $u_1(a^*,b^*)$ when $\delta \to 1$.

Fudenberg and Levine (1992)'s model is analogous to mine when $M=+\infty$, i.e., every consumer can observe the entire history of her predecessors' actions. This is because each consumer's action can be viewed as an informative signal about the seller's past actions, so observing the entire history of consumers' choices can be viewed as observing the entire history of some noisy signal about the seller's actions. Inequality (3.2) applies to imitation equilibria of my model once we take $y_t(\cdot)$ as the equilibrium distribution of b_{t+1} and $y_t(\cdot|a^*)$ as the distribution of b_{t+1} conditional on Player 1 being the commitment type. Consumer $t+1$'s action can statistically identify the seller's action in period t, so $d\big(y_t(\cdot|a^*)\big|\big|y_t(\cdot)\big)>0$ when Player 2_t does not have a strict incentive to play b^*.

However, the distribution of b_{t+1} in imitation equilibria is such that $d\big(y_t(\cdot|a^*)\big|\big|y_t(\cdot)\big)\to 0$ as $\delta \to 1$. This stands in contrast to Fudenberg and Levine (1992)'s model in which $d\big(y_t(\cdot|a^*)\big|\big|y_t(\cdot)\big)$ is bounded away from zero whenever Player 2_t does not have a strict incentive to play b^*. As a result, inequality (3.2) cannot rule out situations where the expected number of periods in which Player 2 has no incentive to play b^* grows without bound as $\delta \to 1$. This is indeed the case in imitation equilibria, where the prolonged process of reputation building cancels out the positive effects of increased patience.

The above discussion unveils an interesting feature of imitation equilibria: Although the patient player can eventually guarantee a high continuation value by exerting high effort in every period, his *discounted average payoff* equals his minmax payoff. Intuitively, each action of Player 2 is informative about her observations of Player 1's past actions, and every Player 2 can observe the entire history of Player 2s' actions. As a result, either Player 2_t strictly prefers to play b^*, or all future Player 2s learn something about Player 1's type from b_t. The arguments in Gossner (2011) imply that there exist at most a finite number of periods where Player 2 does not have a strict incentive to play b^*. Therefore, the patient Player 1 can eventually secure a high continuation value by playing a^* in every period. This logic generalizes to all equilibria when every consumer can observe all of her predecessors' choices.

Proposition 3. *Suppose $M=+\infty$ and players' payoffs satisfy Assumptions 1 and 2. For every $\delta\in(0,1)$, $\pi_0\in(0,1)$, and every strategy profile (σ_1,σ_2) that is part of a PBE, we have:*[9]

$$\liminf_{t\to\infty}\frac{1}{t}\mathbb{E}^{(a^*,\sigma_2)}\Big[\sum_{s=0}^{t-1}u_1(a_s,b_s)\Big]\geq\frac{K}{K+1}u_1(a^*,b^*)+\frac{1}{K+1}u_1(a^*,b'). \tag{3.3}$$

When π_0 is small and δ is large, there exists an equilibrium such that (3.3) holds with equality.

According to Proposition 3, Player 1's undiscounted average payoff from playing the Stackelberg action is at least a fraction $\frac{K}{K+1}$ of his Stackelberg payoff plus a fraction $\frac{1}{K+1}$ of some low payoff

9. Supplementary Appendix B.3 shows that Proposition 3 is not true when M is finite, in the sense that there exist equilibria where Player 1's undiscounted average payoff from imitating the commitment type equals his minmax payoff $u_1(a',b')$.

$u_1(a^*, b')$. This is true in all equilibria for all discount factors and for all prior beliefs. This lower bound is tight in the sense that it can be attained by some equilibria when π_0 is small and δ is large.

When the right-hand side of (3.3) is strictly greater than $u_1(a', b')$, the patient Player 1 can guarantee an asymptotic payoff that is strictly greater than his minmax payoff by playing a^* in every period. The only way to reconcile this conclusion and Theorem 1 is that when Player 1 plays a^* in every period, it takes more time for him to secure this high asymptotic payoff when δ is larger. It is exactly this prolonged process of reputation building that cancels out the direct effects of increased δ.

The proof of Proposition 3 is in Supplementary Appendix B.1 and the proof for the tightness of my lower bound is in Supplementary Appendix B.2. For a heuristic explanation, Assumption 2 implies that a^* is suboptimal for Player 1 in the stage game. Therefore, for every $t \in \mathbb{N}$, either the strategic type has no incentive to play a^* in period t, or $(b_{t+1}, ..., b_{t+K})$ is informative about a_t. In the first case, players 2_{t+1} to 2_{t+K} learn that Player 1 is committed after observing $a_t = a^*$. By playing a^* in every period, Player 1's average payoff from period t to $t+K$ is at least a fraction $\frac{K}{K+1}$ of his Stackelberg payoff plus $\frac{1}{K+1}$ times his minimal stage-game payoff. In the second case, all future Player 2s observe an informative signal about a_t since $M = +\infty$. According to the arguments in Fudenberg and Levine (1992) and Gossner (2011), Player 2s' posterior beliefs assign probability close to 1 to the commitment type after a finite number of periods with learning. The two parts together imply that Player 1's asymptotic payoff is no less than the right-hand side of (3.3).

4. EXTENSION: REPUTATION WITH CONTEMPORANEOUS INFORMATION

Motivated by the social learning models of Banerjee (1992), Bikhchandani, Hirshleifer, and Welch (1992), and Smith and Sørensen (2000), I study an extension where each Player 2 observes Player 1's actions in the last K periods, the entire history of past Player 2s' actions (i.e. $M = +\infty$), and a private signal s_t about Player 1's current-period action a_t. Whether player 1 can observe s_t is irrelevant for my results. Let $s_t \in S$, where S is a countable set. Let $f(s_t|a_t)$ be the probability of s_t when Player 1's action is a_t. I restrict attention to signal distributions that satisfy a *monotone likelihood ratio property* (MLRP), under which Player 2 is more likely to observe a higher signal when Player 1's action is higher.

MLRP. *The distribution of Player 2's private signal satisfies MLRP if there exists a complete order on S, $\succ_S$, such that $f(s|a)f(s'|a') \geq f(s'|a)f(s|a')$ for every $a \succ_A a'$ and $s \succ_S s'$.*

I replace $\succ_A$, $\succ_B$, and $\succ_S$ with $\succ$. The imitation equilibrium I constructed in the proof of Theorem 1 remains an equilibrium when s_t is not very informative about player 1's Stackelberg action a^* and the prior probability of the commitment type is below some cutoff. Later on, I show that my reputation failure result extends to this case. By contrast, if there exists a realization of s_t that is much more likely to occur when $a_t = a^*$ compared to the case when $a_t \neq a^*$, then regardless of previous short-run players' actions, the current-period short-run player will have a strict incentive to play b^* after observing that signal realization. This will unravel the imitation equilibrium, and the patient player can potentially secure a high payoff by establishing a reputation.

My results in this section formalize the above intuition. I show that whether the patient player can secure his Stackelberg payoff in all equilibria depends on whether the distribution of the short-run players' private signals satisfies the following *unbounded informativeness* condition:

Unbounded informativeness. *Player 2's private signal is unboundedly informative about a^* if for every $L > 0$, there exists $s \in S$ such that $f(s|a^*) > Lf(s|a)$ for every $a \neq a^*$.*

My notion of unbounded informativeness is similar to that in Smith and Sørensen (2000).[10] When S is a finite set, unbounded informativeness requires the existence of an $s^* \in S$ such that $f(s^*|a) > 0$ if and only if $a = a^*$. When S is countably infinite, $f(\cdot|a)$ can have full support for every $a \in A$, as long as there exists a sequence $\{s_n\}_{n\in\mathbb{N}} \subset S$ such that $\lim_{n\to+\infty} \frac{f(s_n|a^*)}{f(s_n|a)} = +\infty$ for every $a \neq a^*$.

For an interpretation of s_t and my unbounded informativeness condition, consider a regulator who only has the budget to inspect a fraction ε of the sellers in each period and can issue certificates to the ones that are inspected. The certificate can be taken as s_t, which the current-period consumer can observe before deciding what to buy. MLRP implies that the seller is more likely to obtain a good certificate when he exerts higher effort. If S is a finite set, then consumers' private signal is unboundedly informative about a^* when the seller can obtain a good certificate only if he plays a^*. This is the case, for example, when the regulator can observe the seller's action after her inspection.

Theorem 2 shows that the patient player receives at least his Stackelberg payoff in all equilibria when the short-run players' private signals are unboundedly informative.

Theorem 2. *Suppose players' payoffs satisfy Assumptions 1 and 2, every Player 2 can observe all previous Player 2s' choices, Player 2's private signal satisfies MLRP, and is unboundedly informative about a^*. Then for every prior belief $\pi_0 > 0$ and constant $\varepsilon > 0$, there exists $\delta^* \in (0,1)$ such that player 1's payoff is at least $u_1(a^*, b^*) - \varepsilon$ in all equilibria when $\delta > \delta^*$.*

My next result establishes a partial converse of Theorem 2. For every $a \neq a^*$, I say that a^* is *not strongly separable* from a if there exists $\varepsilon > 0$ such that $f(s|a) \geq \varepsilon f(s|a^*)$ for every $s \in S$. If Player 2's private signal is unboundedly informative about a^*, then there exists no $a \neq a^*$ such that a^* is not strongly separable from a. However, Player 2's private signal not being unboundedly informative about a^* does not imply that a^* is not strongly separable from some $a \neq a^*$.

Theorem 3. *Suppose players' payoffs satisfy Assumptions 1 and 2, every Player 2 can observe all previous Player 2s' choices, Player 2's private signal satisfies MLRP, and a^* is not strongly separable from a'. For every $K \in \mathbb{N}$, there exists $\overline{\pi}_0 \in (0,1)$ such that for every $\pi_0 \in (0, \overline{\pi}_0)$ and δ large enough, there exists an equilibrium where Player 1's payoff is $u_1(a', b')$.*

The proofs of these theorems are in Supplementary Appendix C. Theorem 2 implies that the patient player can guarantee his Stackelberg payoff in all equilibria when each of his opponents can observe the entire history of their predecessors' choices and an unboundedly informative private signal about his current-period action.[11] Theorem 3 extends the reputation failure result of Theorem 1 to situations where K is finite, M is infinite, and Player 2_t observes a private signal about a_t before choosing b_t.

When $|A| = 2$, every signal distribution satisfies MLRP. Since Assumption 2 requires that $a^* \neq a'$, we have $A = \{a^*, a'\}$. The private signal is not unboundedly informative about a^* *if and only if* a^* is not strongly separable from a'. Hence, the private signal being unboundedly

10. First, when S is infinite, I allow for, but does not require, signal realizations that can perfectly rule out some of Player 1's actions, while Smith and Sørensen (2000) require the signal distribution to have full support conditional on every state. Second, I restrict attention to S that is countable while Smith and Sørensen (2000) allow S to be uncountable.

11. Theorem 2 only establishes a common property of all equilibria but does not establish the existence of equilibrium. When S is infinite, the existence of equilibrium does not follow from the canonical result of Fudenberg and Levine (1983). I provide a constructive proof for the existence of equilibrium in Supplementary Appendix C.2.

informative about a^* is both necessary and sufficient for Player 1 to secure his Stackelberg payoff in all equilibria.[12]

The conclusion of Theorem 2 is reminiscent of a well-known result in Bikhchandani, Hirshleifer and Welch (1992) and Smith and Sørensen (2000). They show that in canonical social learning models where there are two states, every myopic player has a finite number of actions, and all players share the same payoff function, the myopic players' actions are asymptotically efficient if and only if their private signals are unboundedly informative about the payoff-relevant state.[13]

My model differs from Smith and Sørensen (2000), since the myopic players' payoffs depend only on the action profile but do not depend on the persistent state—which is Player 1's type in my model. My analysis focuses on the patient player's *discounted average payoff* instead of his *asymptotic payoff* or the game's *asymptotic outcome*. In fact, the myopic players asymptotically learning about the persistent state is *neither necessary nor sufficient* for the patient player to receive a high discounted average payoff. It is not necessary since player 2's payoff depends only on the action profile but not on player 1's type. For example, suppose Player 2s believe that the strategic-type Player 1 plays a^* in every period. They cannot learn Player 1's type, but Player 1 can receive his Stackelberg payoff $u_1(a^*,b^*)$ by playing a^* in every period. It is not sufficient since in imitation equilibria, Player 1's asymptotic payoff is $u_1(a^*,b^*)$ but his discounted average payoff is $u_1(a',b')$.

In what follows, I sketch the proof of Theorem 2 in the case where S is finite, under which there exists $s^* \in S$ such that $f(s^*|a)>0$ if and only if $a=a^*$.[14]

A rough intuition is that Player 2_t observing an unboundedly informative private signal s_t about a_t guarantees a positive lower bound on the informativeness of Player 2_t's action b_t about a_t. Unlike imitation equilibria where the informativeness of b_t about a_{t-1} converges to 0 as δ goes to 1, the informativeness of b_t about a_t is bounded away from zero for all $\delta \in (0,1)$. Since every Player 2 can observe all of her predecessors' actions, the arguments in Fudenberg and Levine (1992) and Gossner (2011) imply that the patient player receives at least his Stackelberg payoff in all equilibria.

A more detailed explanation proceeds in two steps, which highlights the role of unbounded informativeness and MLRP. First, I examine whether Player 2_t's action is informative about her private signal s_t. Intuitively, b_t can be uninformative about s_t for two reasons: (i) Player 2_t is unwilling to play b^* no matter which s_t she observes or (ii) Player 2_t is willing to play b^* no matter which s_t she observes. Since s_t is unboundedly informative about a^*, Player 2 has a strict incentive to play b^* when she observes $s_t=s^*$. This rules out the first possibility. When Player 2_t is willing to play b^* no matter which s_t she observes, Player 1's stage-game payoff is $u_1(a^*,b^*)$ when he plays a^* in period t.

Second, I examine whether Player 2_t's action is informative about Player 1's type. When Player 1's action choice is binary, i.e., $A \equiv \{a^*,a'\}$, Player 2_t is willing to play b^* if and only if $\frac{f(s_t|a^*)}{f(s_t|a')}$ is above some cutoff. This implies that $\Pr(b_t=b^*|a_t=a^*)-\Pr(b_t=b^*|a_t=a')\geq 0$: since

12. When $|A|\geq 3$, MLRP cannot be dropped and the condition in Theorem 3 cannot be replaced by "*the private signal not being unboundedly informative about* a^*", or "a^* *is not strongly separable from* $a^\dagger$ *for some* $a^\dagger \notin \{a^*,a'\}$".

13. Lee (1993) shows that asymptotic efficiency can be achieved under boundedly informative signals when players have a rich set of actions (e.g. a continuum). When the states, actions, and signals can be ordered such that players' payoffs satisfy single-crossing differences, Kartik, Lee and Rappoport (2021) show that asymptotic efficiency can be achieved as long as the signal distribution satisfies directionally unbounded beliefs, which is weaker than unbounded informativeness.

14. When S is infinite and the signal is unboundedly informative about a^*, there exists a non-empty subset $S(\pi)\subset S$ for every $\pi\in(0,1)$ such that when the prior probability of commitment type is at least π before player 2_t observes s_t, she has a strict incentive to play b^* after observing any $s_t \in S(\pi)$. See Supplementary Appendix C.1 for details.

Player 2_t plays b^* after observing s^*, which occurs if and only if Player 1 plays a^*, there exists $c>0$ such that

$$\Pr(b_t=b^*|a_t=a^*)-\Pr(b_t=b^*|a_t=a')\geq c(1-\Pr(b_t=b^*|a_t=a^*)), \tag{4.1}$$

i.e., the informativeness of b_t about a_t is bounded below by some positive function of $1-\Pr(b_t= b^*|a_t=a^*)$. For every $\nu\in(0,1)$, when $\Pr(b_t=b^*|a_t=a^*)\leq 1-\nu$, the strategic type plays a^* with probability bounded away from 1, so the informativeness of b_t about player 1's type is bounded below by a strictly positive function of ν.

When player 1 has three or more actions, Player 2_t's incentive to play b^* can no longer be summarized by a likelihood ratio. As a result, Player 2_t's action can be uninformative about Player 1's type even when the private signal is unboundedly informative about a^* and b_t is informative about s_t. I provide a counterexample in Section 4.1. Nevertheless, when the private signal satisfies MLRP, b_t is informative about Player 1's type in every period where $\Pr(b_t=b^*|a_t=a^*)\neq 1$.

Formally, for every $\alpha\in\Delta(A)$ and $\beta:S\to\Delta(B)$, let $\gamma(\alpha,\beta)\in\Delta(B)$ be the distribution of b induced by (α,β). I show in Supplementary Appendix C.1 that there exists $c>0$ such that for every $\nu\in(0,1)$, every $\alpha\in\Delta(A)$ such that a^* belongs to the support of α, and every β that best replies to α, if the probability of b^* under $\gamma(a^*,\beta)$ is less than $1-\nu$, then the Kullback–Leibler divergence between $\gamma(\alpha,\beta)$ and $\gamma(a^*,\beta)$ is at least $c\nu^2$. This implies that when Player 1 imitates the commitment type, either b^* occurs with probability at least $1-\nu$ under (a^*,β), or the informativeness of b_t about Player 1's type, measured by the Kullback–Leibler divergence between the distribution induced by the equilibrium strategy and the distribution induced by the commitment type, is bounded away from 0.

Back to the discussion on the connections between my results and the canonical reputation results in Section 3.2. Inequality (3.2) applies to my model with contemporaneous private signals as well once we view $y_t(\cdot)$ as the equilibrium distribution of b_t and $y_t(\cdot|a^*)$ as the distribution of b_t conditional on Player 1 being the commitment type. The above discussion implies that when Player 2's private signal is unboundedly informative about a^* and satisfies MLRP, there exists a strictly increasing function $g:[0,1]\to\mathbb{R}_+$ such that $g(0)=0$ and $d\big(y_t(\cdot|a^*)\big|\big|y_t(\cdot)\big)>g(\nu)$ when player 2_t plays b^* with probability less than $1-\nu$. Inequality (3.2) implies that for every $\nu\in(0,1)$, the expected number of periods where $\Pr(b_t=b^*|a_t=a^*)<1-\nu$ is bounded above and this upper bound depends only on ν and is independent of δ. Hence, for every $\nu>0$, there exists $\underline{\delta}\in(0,1)$ such that when $\delta>\underline{\delta}$, player 1 receives at least a fraction $1-\nu$ of $u_1(a^*,b^*)$ when he plays a^* in every period.

4.1. *Conditions in Theorems 2 and 3*

Bounded informativeness. I provide an example in order to explain why "a^* is not strongly separable from a'" in Theorem 3 cannot be replaced by a weaker condition that "consumers' private signal s_t is not unboundedly informative about a^*". Suppose players' stage-game payoffs are

-	b^*	b'
$\overline{a}$	$1,4$	$-2,0$
a^*	$2,1$	$-1,0$
$\underline{a}$	$3,-2$	$0,0$

Let $S\equiv\{\overline{s},s^*,\underline{s}\}$, with $f(\overline{s}|\overline{a})=2/3$, $f(s^*|\overline{a})=1/3$, $f(\overline{s}|a^*)=1/3$, $f(s^*|a^*)=2/3$, and $f(\underline{s}|\underline{a})=1$. Players' stage-game payoffs are monotone-supermodular when Player 1's actions are ranked according to $\overline{a}\succ a^*\succ\underline{a}$ and Player 2's actions are ranked according to $b^*\succ b'$. When signal realizations are ranked according to $\bar{s} \succ s^* \succ \underline{s}$, the signal distribution satisfies MLRP, and is not unboundedly informative about a^*. Player 1's payoff is at least 2 in every equilibrium. This is because when he plays a^*, Player 2 observes either s^* or $\bar{s}$, and has a strict incentive to play b^*.

Not strongly separable from other actions. I use an example to explain why in Theorem 3, "a^* is not strongly separable from a'" cannot be replaced by "a^* is not strongly separable from $a^\dagger$ for some $a^\dagger \notin \{a^*, a'\}$". Suppose Player 1's stage-game payoff is given by the following matrix:

-	b^*	$b^\dagger$	b''	b'
a^*	5	-2	-3	-4
$a^\dagger$	6	-1	-2	-3
a''	7	2	1	-1
a'	8	3	2	0

Player 2's stage-game payoff function is such that b^* is a strict best reply to a^*, $b^\dagger$ is a strict best reply to $a^\dagger$, b'' is a strict best reply to a'', and b' is a strict best reply to a'.

Suppose $S \equiv \{s^*, s'', s'\}$ such that $f(s'|a')=1$ and $f(s'|a)=0$ for every $a \neq a'$. For every $s \in \{s^*, s''\}$ and $a \in \{a^*, a^\dagger, a''\}$, we have $f(s|a)>0$, and $\frac{f(s^*|a)}{f(s''|a)}$ is strictly increasing in a.

Players' stage-game payoffs satisfy Assumptions 1 and 2 once we rank Player 1's actions according to $a^* \succ a^\dagger \succ a'' \succ a'$ and Player 2's actions according to $b^* \succ b^\dagger \succ b'' \succ b'$. The signal distribution satisfies MLRP once we rank the signal realizations according to $s^* \succ s'' \succ s'$. Player 1's Stackelberg action is a^*, which is not strongly separable from $a^\dagger$.

Player 1's commitment payoff from $a^\dagger$ is -1, which is strictly less than his minmax payoff 0. Hence, there exists no equilibrium in which Player 1's payoff equals his commitment payoff from $a^\dagger$.

Next, I show there is no equilibrium where Player 1's payoff equals his minmax payoff 0. Since Player 1 is the commitment type with positive probability, both s^* and s'' occur with positive probability in period 0. Since both s^* and s'' occur on the equilibrium path, Player 2's action is supported in $\{b^*, b^\dagger, b''\}$ after she observes s^* or s''. Suppose Player 1 plays a'' in period 0, and Player 2_0 observes either s^* or s'', so that her action is supported in $\{b^*, b^\dagger, b''\}$. This implies that Player 1's stage-game payoff in period 0 is at least 1 and his expected continuation value after playing a'' is at least 0 in any Perfect Bayesian equilibrium. Hence, Player 1's discounted average payoff is strictly greater than his minmax payoff 0 in all Perfect Bayesian equilibria.

One can obtain a higher payoff lower bound under the following refinement of PBE: For every history h^t, no matter whether it is on-path or off-path, Player 2_t's posterior belief about a_t after observing s_t is supported in $A(s_t) \equiv \{a \in A | f(s_t|a) > 0\}$. In every PBE that satisfies this refinement, suppose Player 1 deviates and plays a'' in every period. Then at every history, Player 2 must be playing some mixed action supported in $\{b^*, b^\dagger, b''\}$. Hence, Player 1's discounted average payoff from playing a'' in every period is at least 1, so his equilibrium payoff in every refined PBE must be no less than 1.

MLRP. In order to demonstrate that the MLRP condition is not redundant, consider the game in the following matrix. Let $S \equiv \{\bar{s}, s^*, \underline{s}\}$. The signal distribution is given by $f(s^*|a^*)=2/3$, $f(\underline{s}|a^*)=1/3$, $f(\bar{s}|\bar{a})=1$, $f(\bar{s}|\underline{a})=1/3$, and $f(\underline{s}|\underline{a})=2/3$.

-	b^*	b'
$\bar{a}$	$1,4$	$-2,0$
a^*	$2,1$	$-1,0$
$\underline{a}$	$3,-2$	$0,0$

Players' payoffs satisfy Assumptions 1 and 2 when plzayer 1's actions are ranked according to $\overline{a} \succ a^* \succ \underline{a}$ and Player 2's actions are ranked according to $b^* \succ b'$. Player 1's Stackelberg action is a^*, his Stackelberg payoff is 2, and s_t is unboundedly informative about a^*. However, MLRP is violated.

I construct an equilibrium where Player 1's payoff is 1, which is bounded below his Stackelberg payoff 2. The strategic-type Player 1 plays a mixed action that depends only on Player 2's posterior belief about his type. If Player 2's posterior belief assigns probability π to the commitment type, then the strategic-type Player 1 plays $\alpha(\pi) \in \Delta(A)$ such that $(1-\pi)\cdot\alpha(\pi)+\pi\cdot a^* = 0.5\cdot a^* + 0.25\cdot \overline{a} + 0.25\cdot \underline{a}$. Player 2_t plays b^* if $s_t \in \{s^*, \overline{s}\}$, and plays b' if $s_t = \underline{s}$.

This strategy profile is an equilibrium since Player 1's expected stage-game payoff is 1 no matter which action he plays, and his continuation value is independent of his current-period action. Player 2 has a strict incentive to play b^* after observing $\overline{s}$ or s^*, and has an incentive to play b' after observing $\underline{s}$. Regardless of Player 1's type, the probability with which Player 2 plays b^* in each period is $2/3$.

In the above example, b_t is uninformative about Player 1's type despite the probability of $b_t = b^*$ being bounded away from 1. As a result, even when Player 1 builds a reputation for playing a^*, Player 2 can still play b' with significant probability in an unbounded number of periods. This explains why the patient player's equilibrium payoff is bounded below his Stackelberg payoff in some equilibria.

5. CONCLUDING REMARKS

This article examines a patient seller's returns from building a reputation when consumers have limited access to his past records and can also learn from other consumers' choices.

My main result shows that limited memory and observational learning lead to reputation failures. This is because consumers' abilities to observe their predecessors' choices enrich their strategy space, rationalizing a larger set of the seller's behaviours. When consumers are concerned that the strategic seller will milk his reputation and then rebuild his reputation, they have a rationale for not trusting the seller despite having observed high effort in the previous period, or more generally, in the last K periods. This is because (i) consumers are not convinced that the seller is the commitment type and (ii) they are also not convinced that the strategic-type seller will exert high effort. I construct equilibria in which the above forces come into play and the patient seller receives his minmax payoff. Consumers' strategies in this equilibrium have a natural interpretation: the first consumer does not trust the seller and every subsequent consumer imitates her predecessor with a probability close to one.

By contrast, when consumers can observe the entire history of the seller's actions, the seller cannot rebuild his reputation after he loses it, in which case concerns about the seller rebuilding his reputation are irrelevant. When each consumer only observes the seller's action in the period before but cannot observe other consumers' choices and cannot observe calendar time, consumers' strategy space is more restricted. Under an additional assumption that players' actions are strategic complements, the restriction in consumers' strategy space implies that the strategic seller either has no incentive to milk his reputation or has no incentive to rebuild his reputation. When consumers are not concerned that the seller will milk-and-then-rebuild his reputation, they have a strict incentive to trust the seller after they observe high effort in the period before. As a result, the seller receives his Stackelberg payoff in all equilibria. I conclude by reviewing the related literature on social learning and reputation formation.

Social learning. The special case of my model where each consumer observes all previous consumers' choices is analogous to a social learning model: A sequence of myopic players observe their predecessors' choices as well as some private signals (e.g. the long-run player's last K actions) in order to predict the long-run player's action in the current period. This stands in contrast to the social learning models in Banerjee 1992, Bikhchandani, Hirshleifer and Welch (1992), Lee (1993), Smith and Sørensen (2000), Bose, Orosel, Ottaviani and Vesterlund (2006), and Kartik *et al.* (2021). In those models, a sequence of myopic players learn about *an exogenous state* instead of *the endogenous actions* of a long-run player.

Due to differences in the object to learn, the myopic players asymptotically learning about the patient player's type is neither sufficient nor necessary for the long-run player to receive a high discounted average payoff, which I have explained earlier. The differences in the learning objective also leads to different forms of inefficiencies. In canonical social learning models, inefficiencies arise when the myopic players ignore their private signals and herd on some inefficient action. In contrast, the myopic players can never herd on any action other than b^* in any equilibrium of my baseline model.[15]

My main result examines the effects of observational learning on a patient player's discounted average payoff. This stands in contrast to existing results that focus on players' asymptotic beliefs, their asymptotic rates of learning (e.g. Gale and Kariv, 2003; Hann-Caruthers, Martynov and Tamuz, 2018; Harel, Mossel, Strack and Tamuz, 2021), and their asymptotic payoffs (e.g. Rosenberg and Vieille, 2019).[16] As demonstrated by the imitation equilibria in the constructive proof of Theorem 1, the patient player's discounted average payoff can be low even when his asymptotic payoff is high.

My article is also related to social learning models with bounded memories. For example, Drakopoulos, Ozdaglar and Tsitsiklis (2012) study a model where a sequence of myopic players learns about an exogenous state. Each player observes a private signal and the actions of her last M predecessors. They show that learning is possible when $M \geq 2$ but not when $M = 1$. In contrast, the myopic players in my model are learning about the endogenous behaviour of a strategic long-run player instead of an exogenous state. As a result, the informativeness of their private signal (which is the patient player's actions in the last K periods in my model) is also endogenous. In contrast to their conclusion which highlights the distinction between the case where $M = 1$ and the case where $M \geq 2$, the values of K and M do not play an important role in my model as long as K is finite and M is at least one.

Reputation failure. Theorem 1 is related to the literature on reputation failures. Schmidt (1993), Cripps and Thomas (1997), and Chan (2000) assume that the uninformed player is forward looking. They show that reputation effects fail in the sense that there *exist equilibria* in which the informed player receives a low payoff. The takeaway from their analysis is that the informed player's patience helps reputation building while the uninformed player's patience hurts reputation building. In contrast, my analysis highlights another effect: the informed player's patience makes it hard for his opponents to distinguish between the commitment type and the strategic type. This effect does not affect the patient player's payoff when his opponents can observe his entire history but plays an important role when each of his opponents can only observe a bounded number of his

15. Logina, Lukyanov and Shamruk (2019) study a social learning model in which every myopic player observes a private signal about a patient player's action. They show that the patient player exerts high effort only when the myopic players' beliefs are intermediate. Their logic is similar to the one in Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992). Board and Meyer-ter-Vehn (2021) study a model of innovation adoption in which players learn about a persistent exogenous state, and characterize the rate of learning under different network structures.

16. In social learning models where a sequence of myopic players learn about an *exogenous state*, Rosenberg and Vieille (2019) calculate the discounted sum of the myopic players' payoffs, and Che and Hörner (2018) and Smith, Sørensen and Tian (2021) design mechanisms that maximize the myopic players' discounted average payoff.

actions. When each uninformed player receives limited information, there is a rationale for her to imitate her immediate predecessor, and her imitation behaviour wipes out the seller's returns from building reputations.

Ely and Välimäki (2003), Ely, Fudenberg and Levine (2008), and Deb, Mitchell and Pai (2022) focus on *participation games* where the uninformed player(s) can take an action under which the informed player receives his minmax payoff and future uninformed players cannot learn about his current-period action.[17] This lack-of-identification problem leads to an equilibrium in which the patient player receives a low payoff. Pei (2020) and Deb and Ishii (2021) show that lack-of-identification occurs when uninformed players' signals cannot identify the state or the uninformed players do not know the monitoring structure. In contrast, the uninformed players cannot shut down learning in my model and consumers' actions in imitation equilibria can statistically identify the seller's past actions.

Bai (2021) assumes that the seller is either a low-cost type who may exert effort or a high-cost type who never exerts effort. Every consumer observes a noisy signal of the seller's effort and communicates the realized signal to all future consumers. She shows that the low-cost type has no incentive to exert effort when δ is low, when his initial reputation is low, and when the fixed cost of establishing a reputation is high. In contrast, reputation effects fail in my model since consumers have limited observations of the seller's past actions and can also observe previous consumers' choices.

Reputation models with limited memory. Liu (2011) and Liu and Skrzypacz (2014) study reputation models where the seller's payoff is strictly submodular and each consumer can observe a bounded number of the seller's actions but cannot observe other consumers' choices and cannot observe calendar time.[18] By contrast, I show that consumers' ability to observe other consumers' choices leads to qualitatively different predictions. First, my reputation failure result requires consumers to observe other consumers' choices. Second, the reputation cycles described in Liu (2011) and Liu and Skrzypacz (2014) cannot arise in my model due to consumers' ability to observe other consumers' choices.

Kaya and Roy (2022) study a model where a long-lived seller decides whether to sell to a myopic consumer in each period. Their model has interdependent values since the seller has persistent private information about his product quality, and quality affects both his production cost and consumers' willingness to pay. When each consumer observes a bounded number of the seller's past actions but *cannot* observe previous consumers' price offers, longer memories encourage the low-quality seller to imitate the high-quality seller, making it harder for the market to screen the seller's type. In contrast, the consumers' payoffs in my model do not directly depend on the seller's type and each consumer observes at least one other consumer's action in addition to a bounded number of the seller's actions. My main result shows that bounded memories by themselves may not lower the seller's returns from building a good reputation, but bounded memories together with consumers' ability to observe other consumers' choices can lower the seller's returns from building a good reputation.

17. Levine (2021) studies a model where signals are less informative when the uninformed players do not participate.

18. Heller and Mohlin (2018) study repeated games with anonymous random matching in which all players are long-lived and each player only observes a finite sample of his opponent's past play. Section 4 of Sperisen (2018) numerically computes the seller's equilibrium payoff set in the product choice game when $c_T = c_N > 0$. Bhaskar and Thomas (2019) study a repeated trust game between a patient player who has *no* reputation concern and a sequence of short-run players, each of them has a finite memory.

A. PROOF OF THEOREM 1

I show Theorem 1 both in the case where calendar time is directly observed and in the case where calendar time is not directly observed. Since $a^* \neq a'$, a' is Player 1's lowest action, and a^* is the unique Stackelberg action, we know that $u_1(a',b') < u_1(a^*,b^*)$. I normalize Player 1's payoff function by setting $u_1(a',b')=0$ and $u_1(a^*,b^*)=1$. Assumption 2 implies that $u_1(a,b')<0$ for every $a \succ a'$ and $u_1(a,b^*)>1$ for every $a \prec a^*$.

Let π_t be the probability Player 2_t's posterior belief assigns to the commitment type after observing $(a_{\max\{0,t-K\}},...,a_{t-1})$ and $(b_{\max\{0,t-M\}},...,b_{t-1})$. Note that when the short-run players do not directly observe calendar time, π_t does not depend on t when $t \geq \max\{M,K\}$. Let $\underline{q}$ be the largest $q \in [0,1]$ such that b' is not Player 2's strict best reply to mixed action $qa^*+(1-q)a'$. Let $\overline{q}$ be the smallest $q \in [0,1]$ such that b^* is not Player 2's strict best reply to mixed action $qa^*+(1-q)a'$. Assumption 1 implies that b^* is a strict best reply to a^* and b' is a strict best reply to a'. Hence $0<\underline{q}<\overline{q}<1$ and there exist $b^{**} \neq b'$ and $b'' \neq b^*$ such that $\{b^{**},b'\} \subset \mathrm{BR}_2(\underline{q}a^*+(1-\underline{q})a')$ and $\{b^*,b''\} \subset \mathrm{BR}_2(\overline{q}a^*+(1-\overline{q})a')$. Assumption 2 implies that $b^* \succ b''$, $b^{**} \succ b'$, and $b^* \succ b'$. I consider the following three cases:

- Case 1: $b^*=b^{**}$ and $b'=b''$.
- Case 2: $b^* \succ b'' \succ b^{**} \succ b'$.
- Case 3: $b^* \succ b''=b^{**} \succ b'$.

I construct equilibria in which (i) Player 1's *ex ante* payoff is 0, (ii) Player 2_t's action depends only on (a_{t-1},b_{t-1}), (iii) Player 1's action in period t depends only on (a_{t-1},b_{t-1}) and Player 2's posterior belief about player 1's type, (iv) Player 1 plays either a^* or a' on the equilibrium path, and (v) if $a_{t-1} \notin \{a',a^*\}$, then the continuation play proceeds as if $(a_{t-1},b_{t-1})=(a',b_{t-1})$. Since $u_1(a,b)$ is strictly in a and a' is Player 1's lowest action, the strategic-type Player 1 strictly prefers a' to actions other than a^* and a' at any private history. I comment on $\underline{\delta}(u_1,u_2)$ by the end of this section.

Case 1: $b^*=b^{**}$ and $b'=b''$. In this case, $\underline{q}=\overline{q}\equiv q$. The construction resembles that in the product choice game after replacing H with a^*, L with a', T with b^*, and N with b'.

1. When $(a_{t-1},b_{t-1})=(a',b')$ or $\varnothing$. Player 2 plays b'. The strategic type player 1 mixes between a^* and a'. His probability of playing a^* is $\frac{q-\pi_0}{1-\pi_0}$ when $t=0$, and is q when $t \geq 1$.

2. When $(a_{t-1},b_{t-1})=(a^*,b')$. Player 2 plays b^* with probability $-\frac{1-\delta}{\delta}u_1(a^*,b')$ and plays b' with complementary probability. The strategic-type Player 1 mixes between a^* and a'. His probability of playing a^*, denoted by p_t, satisfies $\pi_t+(1-\pi_t)p_t=q$.

3. When $(a_{t-1},b_{t-1})=(a',b^*)$. Player 2 plays b^* with probability $\frac{1-(1-\delta)u_1(a',b^*)}{\delta}$ and plays b' with complementary probability. The strategic type Player 1 plays a^* with probability q and plays a' with complementary probability.

4. When $(a_{t-1},b_{t-1})=(a^*,b^*)$, Player 2 plays b^* and player 1 plays a^*.

Players' incentive constraints are satisfied, and in particular, the strategic-type player 1 is indifferent between a^* and a' at every history. In what follows, I show that when π_0 is small enough such that

$$\frac{\pi_0}{1-\pi_0} \leq \Big(\frac{q}{2-q}\Big)^{K+1}, \tag{A.1}$$

$\pi_t \leq \frac{q}{2}$ at every history where $(a_{t-1},b_{t-1})=(a^*,b')$. This implies that the strategic-type's mixed strategy p_t is a well-defined probability and that he plays a^* with probability at least $\frac{q}{2}$ at every history. This conclusion applies both when Player 2s can and cannot observe calendar time.

If $(a_{t-1},b_{t-1})=(a^*,b')$, then Player 2_t's belief assigns positive probability to the commitment type only when $(a_{\max\{0,t-K\}},...,a_{t-1})=(a^*,...,a^*)$ and $(b_{\max\{0,t-M\}},...,b_{t-1})=(b',b',...,b')$. In what follows, I bound Player 2's belief from above when $(a_{\max\{0,t-K\}},...,a_{t-1})=(a^*,...,a^*)$ and $(b_{\max\{0,t-M\}},...,b_{t-1})=(b',b',...,b')$ by considering two cases separately.

First, I study cases where either Player 2s can directly observe calendar time or $M=+\infty$ (in which case Player 2s can perfectly infer calendar time). Let π_t^* be player 2's belief about the commitment type after observing $(a_{\max\{0,t-K\}},...,a_{t-1})=(a^*,...,a^*)$, $(b_{\max\{0,t-M\}},...,b_{t-1})=(b',b',...,b')$, and calendar time. I show that $\pi_t^* \leq \frac{q}{2}$ for every $t \in \mathbb{N}$ by induction on $t \in \mathbb{N}$.

First, $\pi_0^*=\pi_0 \leq \frac{q}{2}$ according to (A.1). Suppose $\pi_s^* \leq \frac{q}{2}$ for every $s \leq t-1$. The induction hypothesis implies that in every period before t, the probability that the strategic type plays a^* is at least $\frac{q}{2}$. Let $P^{\omega_s}(\cdot)$ be the probability measure induced by the equilibrium strategy of the strategic type. Let $P^{\omega_c}(\cdot)$ be the probability measure induced by the commitment type. Let E_t be the event that $(a_{\max\{0,t-K\}},...,a_{t-1})=(a^*,...,a^*)$. Let F_t be the event that $(b_{\max\{0,t-M\}},...,b_{t-1})=(b',...,b')$. According to Bayes rule,

$$\frac{\pi_t^*}{1-\pi_t^*}\Big/\frac{\pi_0}{1-\pi_0}=\frac{P^{\omega_c}(E_t \cap F_t)}{P^{\omega_s}(E_t \cap F_t)}=\frac{P^{\omega_c}(E_t)}{P^{\omega_s}(E_t)}\cdot\frac{P^{\omega_c}(F_t|E_t)}{P^{\omega_s}(F_t|E_t)}. \tag{A.2}$$

Since the strategic type plays a^* with probability at least $\frac{q}{2}$ in every period before t, and b' occurs with lower probability under the strategy of type ω_c compared to that under type ω_s, we have

$$\frac{P^{\omega_c}(E_t)}{P^{\omega_s}(E_t)} \leq \Big(\frac{q}{2}\Big)^{-K} \quad \text{and} \quad \frac{P^{\omega_c}(F_t|E_t)}{P^{\omega_s}(F_t|E_t)} \leq 1. \tag{A.3}$$

When $\pi_0 \leq (\frac{q}{2})^{-K-1}$, (A.2) and (A.3) together imply that $\pi_t^* \leq \frac{q}{2}$.

Next, I study the case where M is finite and Player 2s cannot directly observe calendar time. Let $T \equiv \max\{K, M\}$. Any Player 2 who arrives before period T can perfectly infer calendar time, in which case we can use the same induction argument to bound π_t^* from above by $\frac{q}{2}$. Any Player 2 who arrives after period T cannot perfectly infer calendar time. Let π^* be Player 2_t's belief after observing $(a_{t-K}, ..., a_{t-1}) = (a^*, ..., a^*)$ and $(b_{t-M}, ..., b_{t-1}) = (b', b', ..., b')$ for $t \geq T$.

One complication is that the strategic type's probability of playing a^* depends on π^*, and Player 2's posterior belief after observing $(a_{t-K}, ..., a_{t-1}) = (a^*, ..., a^*)$ and $(b_{t-M}, ..., b_{t-1}) = (b', b', ..., b')$ depends on the strategic type's probability of playing a^*. This leads to a fixed point problem. I show that there exists a fixed point π^* that is less than $\frac{q}{2}$ when π_0 satisfies (A.1).

Let P^{ω_s, π^*} be the probability measure induced by the strategic type when he plays a^* with probability $\frac{q-\pi^*}{1-\pi^*}$ when $(a_{t-K}, ..., a_{t-1}) = (a^*, ..., a^*)$ and $(b_{t-M}, ..., b_{t-1}) = (b', b', ..., b')$. Recall P^{ω_c}, E_t, and F_t defined earlier in this proof. For every $\pi^* \in [0, \frac{q}{2}]$, let $\Pi(\pi^*)$ be defined as:

$$\frac{\Pi(\pi^*)}{1-\Pi(\pi^*)} = \frac{\pi_0}{1-\pi_0} \cdot \frac{\sum_{t=T}^{+\infty}(1-\delta_1)\delta_1^t P^{\omega_c}(E_t) P^{\omega_c}(F_t|E_t)}{\sum_{t=T}^{+\infty}(1-\delta_1)\delta_1^t P^{\omega_s,\pi^*}(E_t) P^{\omega_s,\pi^*}(F_t|E_t)}. \tag{A.4}$$

By definition, $\Pi(\pi^*)$ is a continuous function of π^* for every $\pi^* \in [0, \frac{q}{2}]$. When $\pi^* = 0$, we have $\Pi(\pi^*) > 0 = \pi^*$. The right-hand side of (A.4) is no more than

$$\frac{\pi_0}{1-\pi_0} \cdot \max_{t \geq T} \Big\{ \frac{P^{\omega_c}(E_t) P^{\omega_c}(F_t|E_t)}{P^{\omega_s,\pi^*}(E_t) P^{\omega_s,\pi^*}(F_t|E_t)} \Big\}.$$

Under probability measure P^{ω_s,π^*}, the strategic-type plays a^* with probability at least $\frac{q-\pi^*}{1-\pi^*}$ at every history, which implies that

$$\frac{P^{\omega_c}(E_t) P^{\omega_c}(F_t|E_t)}{P^{\omega_s,\pi^*}(E_t) P^{\omega_s,\pi^*}(F_t|E_t)} \leq \Big(\frac{q-\pi^*}{1-\pi^*}\Big)^{-K}.$$

When $\pi^* = \frac{q}{2}$,

$$\frac{\Pi(\pi^*)}{1-\Pi(\pi^*)} \leq \frac{\pi_0}{1-\pi_0} \cdot \Big(\frac{2-q}{q}\Big)^K \leq \frac{q}{2-q},$$

where the last inequality comes from (A.1). Since $\Pi(\pi^*) > \pi^*$ when $\pi^* = 0$, $\Pi(\pi^*) \leq \pi^*$ when $\pi^* = \frac{q}{2}$, and $\Pi(\pi^*)$ is continuous, there exists a fixed point $\pi^* \in (0, \frac{q^*}{2}]$ such that $\Pi(\pi^*) = \pi^*$.

Case 2: $b^* \succ b'' \succ b^{**} \succ b'$. Consider the following strategy profile, which is parameterized by $r(a^*, b')$, $r(a^*, b'')$, $r(a', b^*)$, and $r(a', b^{**})$, all of them belong to $(0, 1)$ and will be specified later on. Recall that π_t is Player 2_t's belief about the commitment type.

1. When $(a_{t-1}, b_{t-1}) = (a', b')$ or (a', b'') or $\varnothing$. Player 2 plays b'. The strategic type player 1 mixes between a^* and a'. He plays a^* with probability p_t such that $\pi_t + (1-\pi_t)p_t = \underline{q}$.
2. When $(a_{t-1}, b_{t-1}) = (a^*, b')$. Player 2 plays b^{**} with probability $r(a^*, b')$ and b' with complementary probability. The strategic type Player 1 mixes between a^* and a'. He plays a^* with probability p_t such that $\pi_t + (1-\pi_t)p_t = \underline{q}$.
3. When $(a_{t-1}, b_{t-1}) = (a^*, b'')$. Player 2 plays b^{**} with probability $r(a^*, b'')$ and b' with complementary probability. The strategic type Player 1 mixes between a^* and a'. He plays a^* with probability p_t such that $\pi_t + (1-\pi_t)p_t = \underline{q}$.
4. When $(a_{t-1}, b_{t-1}) = (a', b^*)$. Player 2 plays b^* with probability $r(a', b^*)$ and b'' with complementary probability. The strategic type Player 1 plays a^* with probability $\overline{q}$ and plays a' with complementary probability.
5. When $(a_{t-1}, b_{t-1}) = (a', b^{**})$. Player 2 plays b^* with probability $r(a', b^{**})$ and b'' with complementary probability. The strategic type Player 1 plays a^* with probability $\overline{q}$ and plays a' with complementary probability.
6. When $(a_{t-1}, b_{t-1}) = (a^*, b^*)$ or (a^*, b^{**}). Player 2 plays b^* and player 1 plays a^*.

Player 2's incentive constraint at every history is satisfied. Next, I compute Player 1's continuation value in period t for every (a_{t-1}, b_{t-1}), which I denote by $V(a_{t-1}, b_{t-1})$. Then, I verify Player 1's incentive constraints. From the descriptions of players' strategies from (1) to (6), we know that $V(\varnothing) = V(a', b') = V(a', b'') = 0$ and $V(a^*, b^{**}) = V(a^*, b^*) = 1$. Player 1's indifference at $(a_{t-1}, b_{t-1}) = (a', b')$ implies that

$$V(a^*, b') = -\frac{1-\delta}{\delta} u_1(a^*, b'). \tag{A.5}$$

Since $(1-\delta)u_1(a^*,b')+\delta V(a^*,b')=(1-\delta)u_1(a',b')+\delta u_1(a',b')=0$, Player 1 is indifferent when $(a_{t-1},b_{t-1})\in\{(a',b''),(a^*,b'),(a^*,b'')\}$ if and only if

$$(1-\delta)u_1(a',b^{**})+\delta V(a',b^{**})=(1-\delta)u_1(a^*,b^{**})+\delta V(a^*,b^{**})=(1-\delta)u_1(a^*,b^{**})+\delta, \tag{A.6}$$

which implies that

$$V(a',b^{**})=1-\frac{1-\delta}{\delta}\Big(\underbrace{u_1(a',b^{**})-u_1(a^*,b^{**})}_{>0}\Big). \tag{A.7}$$

Let $V(a',b^*)$ be such that Player 1 is indifferent when $(a_{t-1},b_{t-1})=(a^*,b^*)$. This yields:

$$V(a',b^*)=\frac{1-(1-\delta)u_1(a',b^*)}{\delta}. \tag{A.8}$$

According to (A.8), Player 1 is indifferent when $(a_{t-1},b_{t-1})\in\{(a^*,b^{**}),(a',b^*),(a',b^{**})\}$ if and only if

$$(1-\delta)u_1(a^*,b'')+\delta V(a^*,b'')=(1-\delta)u_1(a',b'')+\delta V(a',b'')=(1-\delta)u_1(a',b''). \tag{A.9}$$

This yields:

$$V(a^*,b'')=\frac{1-\delta}{\delta}\Big(\underbrace{u_1(a',b'')-u_1(a^*,b'')}_{>0}\Big). \tag{A.10}$$

Next, I pin down variables $r(a^*,b')$, $r(a^*,b'')$, $r(a',b^*)$, and $r(a',b^{**})$.

1. $r(a^*,b')$ is pinned down by:

$$\underbrace{V(a^*,b')}_{\text{positive but close to 0}}=r(a^*,b')\Big((1-\delta)u_1(a^*,b^{**})+\delta\underbrace{V(a^*,b^{**})}_{=1}\Big).$$

Such $r\in[0,1]$ exists since $0<V(a^*,b')<(1-\delta)u_1(a^*,b^{**})+\delta V(a^*,b^{**})$ when δ is large enough.

2. $r(a^*,b'')$ is pinned down by:

$$\underbrace{V(a^*,b'')}_{\text{positive but close to 0}}=r(a^*,b'')\Big((1-\delta)u_1(a^*,b^{**})+\delta V(a^*,b^{**})\Big).$$

Such $r\in[0,1]$ exists since $0<V(a^*,b'')<(1-\delta)u_1(a^*,b^{**})+\delta V(a^*,b^{**})$ when δ is large enough.

3. $r(a',b^*)$ is pinned down by:

$$\underbrace{V(a',b^*)}_{\text{less than but close to 1}}=r(a',b^*)+(1-r(a',b^*))\Big((1-\delta)u_1(a^*,b'')+\delta\underbrace{V(a^*,b'')}_{\text{positive but close to 0}}\Big).$$

Such $r\in[0,1]$ exists since $(1-\delta)u_1(a^*,b'')+\delta V(a^*,b'')<V(a',b^*)<1$ when δ is large enough.

4. $r(a',b^{**})$ is pinned down by:

$$\underbrace{V(a',b^{**})}_{\text{less than but close to 1}}=r(a',b^{**})+(1-r(a',b^{**}))\Big((1-\delta)u_1(a^*,b'')+\delta\underbrace{V(a^*,b'')}_{\text{positive but close to 0}}\Big).$$

Such $r\in[0,1]$ exists since $(1-\delta)u_1(a^*,b'')+\delta V(a^*,b'')<V(a',b^*)<1$ when δ is large enough.

When the prior probability of commitment type is less than $\overline{\pi}_0$, where $\overline{\pi}_0$ is given by

$$\frac{\overline{\pi}_0}{1-\overline{\pi}_0}=\Big(\frac{\underline{q}}{2}\Big)^{K+1}, \tag{A.11}$$

one can show using the same argument as the first case that Player 2's posterior belief assigns probability less than $\underline{q}/2$ to the commitment type at every history where $(a_{t-1},b_{t-1})\notin\{(a^*,b^*),(a^*,b^{**})\}$. This implies that the strategic type Player 1 plays a^* with probability at least $\underline{q}/2$ at every history, and that his mixed action at every history is well defined.

Case 3: $b^*\succ b''=b^{**}\succ b'$. I write b'' instead of b^{**}. Consider the following strategy profile, parameterized by $s(a^*,b')$, $s(a^*,b'')$, $s(a',b^*)$, and $s(a',b^{**})$.

1. When $(a_{t-1},b_{t-1})=(a',b')$ or $\varnothing$. Player 2 plays b'. If $t=0$, the strategic type player 1 plays a^* with probability $\frac{q-\pi_0}{1-\pi_0}$ and plays a' with complementary probability. If $t\geq 1$, the strategic type Player 1 plays a^* with probability $\underline{q}$ and plays a' with complementary probability.

2. When $(a_{t-1}, b_{t-1}) = (a^*, b')$. Player 2 plays b'' with probability $s(a^*, b')$ and b' with complementary probability. The strategic type Player 1 mixes between a^* and a'. He plays a^* with probability p_t such that $\pi_t + (1-\pi_t)p_t = \underline{q}$.

3. When $(a_{t-1}, b_{t-1}) = (a', b'')$. Player 2 plays b'' with probability $s(a', b'')$ and b' with complementary probability. The strategic type Player 1 plays a^* with probability $\underline{q}$ and plays a' with complementary probability.

4. When $(a_{t-1}, b_{t-1}) = (a^*, b'')$. Player 2 plays b^* with probability $s(a^*, b'')$ and b'' with complementary probability. The strategic type Player 1 mixes between a^* and a'. He plays a^* with probability p_t such that $\pi_t + (1-\pi_t)p_t = \overline{q}$.

5. When $(a_{t-1}, b_{t-1}) = (a', b^*)$. Player 2 plays b^* with probability $s(a', b^{**})$ and b'' with complementary probability. The strategic type Player 1 plays a^* with probability $\overline{q}$ and plays a' with complementary probability.

6. When $(a_{t-1}, b_{t-1}) = (a^*, b^*)$. Player 2 plays b^* and player 1 plays a^*.

According to (1) and (6), $V(\varnothing) = V(a', b') = 0$ and $V(a^*, b^*) = 1$. Player 1's indifference at (a', b') implies that $V(a^*, b') = -\frac{1-\delta}{\delta} u_1(a^*, b')$. Let $V(a', b^*) = \frac{1-(1-\delta)u_1(a', b^*)}{\delta}$, under which Player 1 is indifferent between a^* and a' when $(a_{t-1}, b_{t-1}) = (a^*, b^*)$.

Since $(1-\delta)u_1(a^*, b') + \delta V(a^*, b') = (1-\delta)u_1(a', b') + \delta V(a', b')$ and $(1-\delta)u_1(a^*, b^*) + \delta V(a^*, b^*) = (1-\delta)u_1(a', b^*) + \delta V(a', b^*)$ under these continuation values, the strategic type of Player 1 is indifferent at (a^*, b'), (a', b''), (a^*, b''), and (a', b^*) if and only if

$$(1-\delta)u_1(a^*, b'') + \delta V(a^*, b'') = (1-\delta)u_1(a', b'') + \delta V(a', b''). \tag{A.12}$$

Assumption 2 implies that $u_1(a', b'') > u_1(a^*, b'')$, $u_1(a^*, b'') < u_1(a^*, b^*)$ and $u_1(a', b'') > u_1(a', b')$.

Lemma 1. *There exists $\gamma \in (0,1) \cap (u_1(a^*, b''), u_1(a', b''))$ such that*

$$\gamma(1 - u_1(a^*, b'')) \geq (1-\gamma)u_1(a', b''). \tag{A.13}$$

Proof. Consider two cases separately. First, suppose $u_1(a', , b'') \leq 1$. By setting $\gamma = u_1(a', b'')$,

$$\gamma(1 - u_1(a^*, b'')) = u_1(a', b'')(1 - u_1(a^*, b'')) > u_1(a', b'')(1 - u_1(a', b'')).$$

The intermediate value theorem implies that (A.13) holds for some γ that is strictly less than $u_1(a', b'')$ but is strictly greater than $u_1(a^*, b'')$. Second, suppose $u_1(a', b'') > 1$. By setting $\gamma = 1$, the left-hand side of (A.13) is strictly positive while the right-hand side of (A.13) is 0. The intermediate value theorem implies that (A.13) holds for some γ that is strictly less than 1 but is strictly greater than $u_1(a^*, b'')$ □

Pick $\gamma \in (0,1) \cap (u_1(a^*, b''), u_1(a', b''))$ that satisfies (A.13) and set player 1's continuation values at (a^*, b'') and (a', b'') to be

$$V(a^*, b'') = \frac{1}{\delta}\big(\gamma - (1-\delta)u_1(a^*, b'')\big) \tag{A.14}$$

and

$$V(a', b'') = \frac{1}{\delta}\big(\gamma - (1-\delta)u_1(a', b'')\big). \tag{A.15}$$

These continuation values satisfy player 1's incentive constraint (A.12), and moreover,

$$V(a^*, b'') > (1-\delta)u_1(a^*, b'') + \delta V(a^*, b'') = \gamma = (1-\delta)u_1(a', b'') + \delta V(a', b'') > V(a', b'').$$

When δ is close to 1, both $V(a^*, b'')$ and $V(a', b'')$ are bounded away from 0 and 1, and moreover, $V(a', b'') < u_1(a', b'')$ and $V(a^*, b'') > u_1(a^*, b'')$.

Next, I pin down the values of $s(a^*, b')$, $s(a^*, b'')$, $s(a', b^*)$, and $s(a', b'')$ so that player 1 receives these continuation values. Recall that $V(a^*, b') = -\frac{1-\delta}{\delta} u_1(a^*, b')$ and $V(a', b^*) = \frac{1-(1-\delta)u_1(a', b^*)}{\delta}$, and the values of $V(a^*, b'')$ and $V(a', b'')$ are given by (A.14) and (A.15).

1. $s(a^*, b')$ is pinned down by:

$$\underbrace{V(a^*, b')}_{\text{positive but close to 0}} = s(a^*, b')\Big((1-\delta)u_1(a^*, b'') + \delta \underbrace{V(a^*, b'')}_{\text{bounded away from 0}}\Big).$$

Such $s \in [0,1]$ exists since $0 < V(a^*, b') < (1-\delta)u_1(a^*, b'') + \delta V(a^*, b'')$ when δ is large enough.

2. $s(a', b'')$ is pinned down by:

$$V(a', b'') = s(a', b'')\Big((1-\delta)u_1(a', b'') + \delta V(a', b'')\Big).$$

Such $s \in [0,1]$ exists since $0 < V(a', b'') < (1-\delta)u_1(a', b'') + \delta V(a', b'')$ when δ is large enough.

3. $s(a^*,b'')$ is pinned down by:

$$V(a^*,b'')=s(a^*,b'')+(1-s(a^*,b''))\Big((1-\delta)u_1(a^*,b'')+\delta V(a^*,b'')\Big).$$

Such $s\in[0,1]$ exists since $(1-\delta)u_1(a^*,b'')+\delta V(a^*,b'')<V(a^*,b'')<1$ when δ is large enough.

4. $s(a',b^*)$ is pinned down by:

$$\underbrace{V(a',b^*)}_{\text{close to but less than 1}}=s(a',b^*)+(1-s(a',b^*))\Big((1-\delta)u_1(a^*,b'')+\delta \underbrace{V(a^*,b'')}_{\text{bounded away from 1}}\Big)$$

Such $s\in[0,1]$ exists since $(1-\delta)u_1(a^*,b'')+\delta V(a^*,b'')<V(a',b^*)<1$ when δ is large enough.

Next, I show that Player 2's posterior belief assigns probability less than $\underline{q}/2$ to the commitment type at every history where $(a_{t-1},b_{t-1})\neq(a^*,b^*)$. The key step is Lemma 2.

Lemma 2. *If γ satisfies (A.13), then $s(a',b'')+s(a^*,b'')\geq 1$.*

Proof. According to the expressions of Player 1's continuation value, we have

$$s(a^*,b'')=\frac{V(a^*,b'')-\gamma}{1-\gamma}\quad\text{and}\quad s(a',b'')=\frac{V(a',b'')}{\gamma}. \tag{A.16}$$

Therefore, $s(a',b'')+s(a^*,b'')\geq 1$ if and only if

$$\frac{V(a^*,b'')-\gamma}{1-\gamma}+\frac{V(a',b'')}{\gamma}\geq 1,$$

which is equivalent to $(1-\gamma)V(a',b'')\geq\gamma(1-V(a^*,b''))$. Plugging in (A.14) and (A.15), this inequality is equivalent to $\gamma(1-u_1(a^*,b''))\geq(1-\gamma)u_1(a',b'')$, which is (A.13). □

Since Player 2 plays b'' with probability $1-s(a^*,b'')$ when $(a_{t-1},b_{t-1})=(a^*,b'')$ and plays b'' with probability $s(a',b'')$ when $(a_{t-1},b_{t-1})=(a',b'')$, Lemma 2 implies that

$$\Pr(b_{t+1}=b''|b_t=b'',a_t=a')\geq\Pr(b_{t+1}=b''|b_t=b'',a_t=a^*). \tag{A.17}$$

Therefore, the likelihood ratio between the commitment type and the strategic type does not increase when Player 2 observes $b_{t+1}=b''$ conditional on $b_t=b''$. Back to the proof of $\pi_t\leq\underline{q}/2$ whenever $(a_{t-1},b_{t-1})\neq(a^*,b^*)$, we only need to consider histories such that $a_{t-1}=a^*$. Assume $\pi_0<\overline{\pi}_0$ where $\overline{\pi}_0$ is given by

$$\frac{\overline{\pi}_0}{1-\overline{\pi}_0}=\Big(\frac{\underline{q}}{2}\Big)^{K+1}\frac{\underline{q}}{2-\underline{q}}. \tag{A.18}$$

1. At histories where $(a_{t-1},b_{t-1})=(a^*,b')$, then the same argument as that in the first case implies that when π_0 is no more than $\overline{\pi}_0$ defined in (A.18), Player 2's posterior belief assigns probability less than $\underline{q}/2$ at every such history.

2. At histories where $(a_{t-1},b_{t-1})=(a^*,b'')$, then Player 2's posterior belief assigns strictly positive probability to the commitment type *only if* $(a_{t-K},...,a_{t-1})=(a^*,...,a^*)$ and there exists $s\leq t-1$ such that $b_\tau=b'$ for every $\tau<s$ and $b_\tau=b''$ for every $t-1\geq\tau\geq s$.

I show this claim when $M=+\infty$ or when Player 2s can directly observe calendar time. The case where M being finite and Player 2s cannot directly observe calendar time can be shown using a fixed point argument similar to the one in Case 1, which I omit in order to avoid repetition.

Let E_t be the event that $(a_{t-K},...,a_{t-1})=(a^*,...,a^*)$, let $F_{s,t}$ be the event that $(b_0,...,b_{t-1})=(b',...,b',b'',b'',...,b'')$ where the first b'' occurs in period s. Let $\pi^*_{s,t}$ be the posterior probability of commitment type conditional on $E_t\cap F_t$. According to Bayes rule,

$$\frac{\pi^*_{s,t}}{1-\pi^*_{s,t}}\Big/\frac{\pi_0}{1-\pi_0}=\frac{P^{\omega_c}(E_t\cap F_t)}{P^{\omega_s}(E_t\cap F_t)}=\frac{P^{\omega_c}(E_t)}{P^{\omega_s}(E_t)}\cdot\frac{P^{\omega_c}(F_t|E_t)}{P^{\omega_s}(F_t|E_t)}. \tag{A.19}$$

The first term on the right-hand side of (A.19) is no more than $(\underline{q}/2)^{-K}$. For every $n<s$, let

$$l_n\equiv\frac{P^{\omega_c}(a_n=a'|E_t,(b_0,...,b_{n-1})=(b',...,b'))}{P^{\omega_s}(a_n=a'|E_t,(b_0,...,b_{n-1})=(b',...,b'))} \tag{A.20}$$

and for every $n\geq s$, let

$$l_n\equiv\frac{P^{\omega_c}(a_n=a''|E_t,(b_0,...,b_{n-1})=(b',...,b'))}{P^{\omega_s}(a_n=a''|E_t,(b_0,...,b_{n-1})=(b',...,b'))}. \tag{A.21}$$

According to Bayes rule, the second term on the right-hand side of (A.19) equals $\Pi_{i=0}^{t-1}l_i$. According to Lemma 2, $l_n\leq 1$ for every $n\neq s$. Since $\pi_0\leq\overline{\pi}_0$, we have $\pi^*_{s,t}\leq\underline{q}/2$ for every $t\leq s$. Since $\pi_t\leq\max_{s\leq t}\pi^*_{s,t}$, we have $\pi_t\leq\underline{q}/2$ for every $t\leq s$. Since the unconditional probability with which Player 1 plays a^* is at least $\underline{q}$ in every period and $\pi^*_{s,s}\leq\underline{q}/2$, we have $l_s\leq(\underline{q}/2)^{-1}$. This implies that $\pi_t\leq\underline{q}/2$ for every $t\in\mathbb{N}$, which concludes the proof.

Remark. I provide sufficient conditions for the cutoff discount factor $\underline{\delta}(u_1,u_2)$. Recall we adopt the normalization that $u_1(a^*,b^*)=1$ and $u_1(a',b')=0$. In Case 1, the cutoff discount factor is:

$$\underline{\delta}(u_1,u_2)=\max\left\{\frac{-u_1(a^*,b')}{1-u_1(a^*,b')},1-\frac{1}{u_1(a',b^*)}\right\}.$$

In Case 2, the cutoff discount factor is pinned down by $V(a^*,b')\leq(1-\delta)u_1(a^*,b^{**})+\delta$, $V(a^*,b'')\leq(1-\delta)u_1(a^*,b^{**})+\delta$, $(1-\delta)u_1(a^*,b'')+\delta V(a^*,b'')\leq V(a',b^*)$, and $(1-\delta)u_1(a^*,b'')+\delta V(a^*,b'')\leq V(a',b^*)$, where $V(a^*,b')$, $V(a',b^{**})$, $V(a',b^*)$, and $V(a^*,b'')$ are given by (A.5), (A.7), (A.8), and (A.10). In Case 3, the cutoff discount factor is pinned down by $V(a^*,b')\leq(1-\delta)u_1(a^*,b'')+\delta V(a^*,b'')$, $V(a',b'')\leq(1-\delta)u_1(a',b'')+\delta V(a',b'')$, $(1-\delta)u_1(a^*,b'')+\delta V(a^*,b'')\leq V(a^*,b'')$, and $(1-\delta)u_1(a^*,b'')+\delta V(a^*,b'')\leq V(a',b^*)$, where $V(a^*,b')=-\frac{1-\delta}{\delta}u_1(a^*,b')$, $V(a',b^*)=\frac{1-(1-\delta)u_1(a',b^*)}{\delta}$, and the values of $V(a^*,b'')$ and $V(a',b'')$ are given by (A.14) and (A.15).

B. PROOF OF PROPOSITION 1

Since $(K,M)=(1,0)$, consumers' strategy is represented by a triple $(r_\varnothing,r_H,r_L)$, where r_x is the probability with which she plays T when $a_{t-1}=x$ for $x\in\{\varnothing,H,L\}$.

First, I show that $r_H>r_L$. Suppose by way of contradiction that $r_H\leq r_L$, then the strategic-type seller has no incentive to play H. After consumer t observes that $a_{t-1}=H$, she infers that the seller is the commitment type for sure and hence, she has a strict incentive to play T. This implies that $r_H=1$. Since $r_H\leq r_L$, we have $r_L=1$. However, since consumer t knows that the seller is the strategic type after observing $a_{t-1}=L$ and the strategic-type seller has no incentive to play H, which implies that $r_L=0$. This contradicts the previous conclusion that $r_L=1$.

Since consumer t's strategy depends only on a_{t-1}, starting from period 1, the seller's continuation value depends only on whether $a_{t-1}=L$ or $a_{t-1}=H$. Let $V(L)$ and $V(H)$ be these continuation values, respectively. The strategic-type seller has an incentive to play H when $a_{t-1}=H$ if and only if $(1-\delta)(r_H+(1-r_H)(-c_N))+\delta V(H)-(1-\delta)(1+c_T)r_H-\delta V(L)\geq 0$, or equivalently,

$$\frac{\delta}{1-\delta}(V(H)-V(L))\geq c_Tr_H+c_N(1-r_H). \tag{B.1}$$

Similarly, the seller has an incentive to play H when $a_{t-1}=L$ if and only if

$$\frac{\delta}{1-\delta}(V(H)-V(L))\geq c_Tr_L+c_N(1-r_L). \tag{B.2}$$

Since $r_H>r_L$ and $c_N>c_T$, the right-hand side of (B.2) is strictly greater than the right-hand side of (B.1), which implies that

- If the strategic-type seller is indifferent between H and L when $a_{t-1}=L$, then the strategic-type seller has a strict incentive to play H when $a_{t-1}=H$.
- If the strategic-type seller is indifferent between H and L when $a_{t-1}=H$, then the strategic-type seller has a strict incentive to play L when $a_{t-1}=L$.

I consider three cases. First, suppose the strategic-type seller has a strict incentive to play L when $a_{t-1}=H$, then he also has a strict incentive to play L when $a_{t-1}=L$. Then after she observes $a_{t-1}=H$, consumer t believes that the seller is the commitment type for sure and hence, she has a strict incentive to play T. This implies that $r_H=1$. As a result, the seller can guarantee discounted average payoff at least $\delta-(1-\delta)c_N$ by playing H in every period.

Next, suppose the strategic-type seller has a strict incentive to play H when $a_{t-1}=H$, then after consumer t observes that $a_{t-1}=H$, she knows that the seller will play H regardless of his type and therefore, she will have a strict incentive to play T. As a result, $r_H=1$, which implies that the seller can guarantee discounted average payoff at least $\delta-(1-\delta)c_N$ by playing H in every period.

The above reasoning implies that in every equilibrium where the strategic-type seller receives a payoff strictly less than $\delta-(1-\delta)c_N$, the strategic-type seller must be indifferent when $a_{t-1}=H$ and must strictly prefer to play L when $a_{t-1}=L$, and moreover, $r_H<1$. I show that there is no such equilibria when δ is close to 1. Let p_t be the probability of the event:

$$E_t\equiv\big\{\text{The seller is the strategic type and plays } H \text{ in period } t\big\}.$$

Since the strategic-type seller strictly prefers to play L in period t when $a_{t-1}=L$, we have $1-\pi_0\geq p_0\geq p_1\geq p_2\geq\ldots$ Since consumers' prior belief assigns probability π_0 to the commitment type and assigns probability $\delta_1^t(1-\delta_1)$ to the calendar time being t, she prefers N to T after observing $a_{t-1}=H$ only if

$$\frac{\sum_{t=1}^{+\infty}(1-\delta_1)\delta_1^t\pi_0}{\sum_{t=1}^{+\infty}(1-\delta_1)\delta_1^t(p_{t-1}-p_t)}\leq 1. \tag{B.3}$$

Since $\sum_{t=1}^{+\infty}(1-\delta_1)\delta_1^t\pi_0=\delta_1\pi_0$ and $\sum_{t=1}^{+\infty}(1-\delta_1)\delta_1^t(p_{t-1}-p_t)\leq(1-\delta_1)\sum_{t=1}^{\infty}(p_{t-1}-p_t)=1-\delta_1$, we have

$$\frac{\sum_{t=1}^{+\infty}(1-\delta_1)\delta_1^t\pi_0}{\sum_{t=1}^{+\infty}(1-\delta_1)\delta_1^t(p_{t-1}-p_t)}\geq\frac{\delta_1\pi_0}{1-\delta_1}$$

the right-hand side is strictly greater than 1 when δ_1 is close to 1, in which case (B.3) cannot be true. The above contradiction implies that the seller's payoff is at least $\delta-(1-\delta)c_N$ in every equilibrium.

Acknowledgments. I thank Daron Acemoglu, S. Nageeb Ali, Alp Atakan, Jie Bai, Dhruva Bhaskar, Sushil Bikhchandani, Joyee Deb, Eddie Dekel, Mira Frick, Drew Fudenberg, Olivier Gossner, Daniel Hauser, Kevin He, Ju Hu, Yuhta Ishii, Navin Kartik, Elliot Lipnowski, George Lukyanov, Daniel Luo, George Mailath, Moritz Meyer-ter-Vehn, Wojciech Olszewski, Mallesh Pai, Evan Sadler, Peter Norman Sørensen, Bruno Strulovici, Adam Szeidl, Caroline Thomas, Juuso Toikka, Chris Udry, Udayan Vaidya, Juuso Välimäki, Alex Wolitzky, Tomer Yehoshua-Sandak, and four anonymous referees for helpful comments. I thank NSF Grant SES-1947021 for financial support.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

BAI, J. (2021), "Melons as Lemons: Asymmetric Information, Consumer Learning and Quality Provision" (Working Paper).

BANERJEE, A. (1992), "A Simple Model of Herd Behavior", *Quarterly Journal of Economics*, **107**, 797–817.

BHASKAR, V. and THOMAS, C. (2019), "Community Enforcement of Trust with Bounded Memory", *Review of Economic Studies*, **86**, 1010–1032.

BIKHCHANDANI, S., HIRSHLEIFER, D. and WELCH, I. (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades", *Journal of Political Economy*, **100**, 992–1026.

BOARD, S. and MEYER-TER-VEHN, M. (2021), "Learning Dynamics in Social Networks", *Econometrica*, **89**, 2601–2635.

BOSE, S., OROSEL, G., OTTAVIANI, M. and VESTERLUND, L. (2006), "Dynamic Monopoly Pricing and Herding", *RAND Journal of Economics*, **37**, 910–928.

CAI, H., CHEN, Y. and FANG, H. (2009), "Observational Learning: Evidence from a Randomized Natural Field Experiment", *American Economic Review*, **99**, 864–882.

CAI, J., DE JANVRY, A. and SADOULET, E. (2015), "Social Networks and the Decision to Insure", *American Economic Journal: Applied Economics*, **7**, 81–108.

CHAN, J. (2000), "On the Non-Existence of Reputation Effects in Two-Person Infinitely-Repeated Games" (Working Paper).

CHE, Y.-K. and HÖRNER, J. (2018), "Recommender Systems as Mechanisms for Social Learning", *Quarterly Journal of Economics*, **133**, 871–925.

CRIPPS, M. and THOMAS, J. (1997), "Reputation and Perfection in Repeated Common Interest Games", *Games and Economic Behavior*, **28**, 433–462.

DEB, J. and ISHII, Y. (2021), "Reputation Building under Uncertain Monitoring" (Working Paper).

DEB, R., MITCHELL, M. and PAI, M. (2022), "(Bad) Reputation in Relational Contracting", *Theoretical Economics*, **17**, 763–800.

DRAKOPOULOS, K., OZDAGLAR, A. and TSITSIKLIS, J. (2012), "On Learning with Finite Memory", *IEEE Journal of Selected Topics in Signal Processing*, **6**, 780–794.

ELY, J. and VÄLIMÄKI, J. (2003), "Bad Reputation", *Quarterly Journal of Economics*, **118**, 785–814.

ELY, J., FUDENBERG, D. and LEVINE, D. (2008), "When is Reputation Bad?", *Games and Economic Behavior*, **63**, 498–526.

EKMEKCI, M. (2011), "Sustainable Reputations with Rating Systems", *Journal of Economic Theory*, **146**, 479–503.

FUDENBERG, D., KREPS, D. and MASKIN, E. (1990), "Repeated Games with Long-Run and Short-Run Players", *Review of Economic Studies*, **57**, 555–573.

FUDENBERG, D. and LEVINE, D. (1983), "Subgame Perfect Equilibria of Finite and Infinite Horizon Games", *Journal of Economic Theory*, **31**, 251–268.

FUDENBERG, D. and LEVINE, D. (1989), "Reputation and Equilibrium Selection in Games with a Patient Player", *Econometrica*, **57**, 759–778.

FUDENBERG, D. and LEVINE, D. (1992), "Maintaining a Reputation when Strategies are Imperfectly Observed", *Review of Economic Studies*, **59**, 561–579.

GALE, D. and KARIV, S. (2003), "Bayesian Learning in Social Networks", *Games and Economic Behavior*, **45**, 329–346.

GOSSNER, O. (2011), "Simple Bounds on the Value of a Reputation", *Econometrica*, **79**, 1627–1641.

HANN-CARUTHERS, W., MARTYNOV, V. and TAMUZ, O. (2018), "The Speed of Sequential Asymptotic Learning", *Journal of Economic Theory*, **173**, 383–409.

HAREL, M., MOSSEL, E., STRACK, P. and TAMUZ, O. (2021), "Rational Groupthink", *Quarterly Journal of Economics*, **136**, 621–668.
HELLER, Y. and MOHLIN, E. (2018), "Observations on Cooperation", *Review of Economic Studies*, **85**, 2253–2282.
KARTIK, N., LEE, S. and RAPPOPORT, D. (2021), "Observational Learning with Ordered States" (Working Paper).
KAYA, A. and ROY, S. (2022), "Market Screening with Limited Records", *Games and Economic Behavior*, **132**, 106–132.
LEE, I. H. (1993), "On the Convergence of Information Cascades", *Journal of Economic Theory*, **61**, 395–411.
LEVINE, D. (2021), "The Reputation Trap", *Econometrica*, **89**, 2659–2678.
LIU, Q. (2011), "Information Acquisition and Reputation Dynamics", *Review of Economic Studies*, **78**, 1400–1425.
LIU, Q. and SKRZYPACZ, A. (2014), "Limited Records and Reputation Bubbles", *Journal of Economic Theory*, **151**, 2–29.
LOGINA, E., LUKYANOV, G. and SHAMRUK, K. (2019), "Reputation and Social Learning" (Working Paper).
MAILATH, G. and SAMUELSON, L. (2015), "Reputations in Repeated Games", *Handbook of Game Theory with Economic Applications*, **4**, 165–238.
MICHELSON, H., FAIRBAIRN, A., ELLISON, B., MAERTENS, A. and MANYONG, V. (2021), "Misperceived Quality: Fertilizer in Tanzania", *Journal of Development Economics*, **148**, 102579.
PEI, H. (2020), "Reputation Effects under Interdependent Values", *Econometrica*, **88**, 2175–2202.
PHELAN, C. (2006), "Public Trust and Government Betrayal", *Journal of Economic Theory*, **130**, 27–43.
ROSENBERG, D. and VIEILLE, N. (2019), "On the Efficiency of Social Learning", *Econometrica*, **87**, 2141–2168.
SCHMIDT, K. (1993), "Reputation and Equilibrium Characterization in Repeated Games with Conflicting Interests", *Econometrica*, **61**, 325–351.
SMITH, L. and SØRENSEN, P. N. (2000), "Pathological Outcomes of Observational Learning", *Econometrica*, **68**, 371–398.
SMITH, L., SØRENSEN, P. N. and TIAN, J. (2021), "Informational Herding, Optimal Experimentation, and Contrarianism", *Review of Economic Studies*, **88**, 2527–2554.
SPERISEN, B. (2018), "Bounded Memory and Incomplete Information", *Games and Economic Behavior*, **109**, 382–400.
ZHANG, J. (2010), "The Sound of Silence: Observational Learning in the U.S. Kidney Market", *Management Science*, **29**, 315–335.