

Information Acquisition and Reputation Dynamics

Qingmin Liu^{*†}

September 2010

Abstract

We study dynamic incentives and behavior in markets with costly discovery of past transactions. In our model, a sequence of short-lived customers interact over time with a single long-lived firm that privately knows its type (good or opportunistic). Customers must pay to observe the firm's past behavior. We characterize the equilibrium structure that features accumulation, consumption, and restoration of reputation. The opportunistic firm deliberately builds its reputation up to a point where the maximum periods of information acquired by customers do not reveal past opportunistic behavior, and exploits the customers who most trust the firm.

1 Introduction

Reputation is an important asset that can be accumulated, consumed, and restored. For example, many tourist destinations experience a phase of development with increasing popularity, followed by a phase of decline with reduced quality and reputation, and then a phase of rejuvenation.¹ Mohan and Blough (2008, p. 207) describe the following behavior often observed in online trading: an agent “behaves honestly for some time and then tries to milk his reputation by cheating. This reputation building and milking cycle continues. During the reputation

^{*}Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104. Email: qingmin@econ.upenn.edu

[†]This paper is based on my dissertation, submitted to Stanford Graduate School of Business. I am grateful to Yossi Feinberg, Andy Skrzypacz, Bob Wilson, and George Mailath for guidance and continuous help. I thank Bruno Biais, the editor, and two anonymous referees for their comments and suggestions that significantly improved the paper and gave the paper its current form. I thank Heski Bar-Isaac, Mehmet Ekmekci, Jihong Lee, John Roberts, Larry Samuelson, Renna Wen, Muhamet Yildiz, and many seminar participants for helpful comments, and Danielle Catambay and Michael Borns for professional proofreading.

¹See Butler (2006) for many careful case studies of tourism area life cycles, including the cases for the Atlantic City in the United States and Opatija Riviera in Croatia. I thank Denis Claude for drawing my attention to this literature and its relation to reputation.

milking phase, the user cheats on a number of transactions continuously until his reputation is very low; and then either enters the reputation building stage or reenters the system with a new pseudonym.”²

Reputation effects link past histories with future opportunities in repeated interactions with incomplete information. Intuitively, a firm’s lost reputation can only be restored if its opportunistic behavior does not have a long-lasting effect; however, the lack of long-term consequences will also encourage opportunistic behavior and discourage trust building. It is therefore not a priori clear whether long-run reputation dynamics can be an equilibrium phenomenon even with imperfect observations of past history.

Indeed, information about past transactions seldom comes for free. For instance, time and effort are required for a tourist to check other customers’ experiences of a hotel, and for a trader to verify the trading history of his counterpart. Moreover, information on transactions conducted in the distant past is more difficult to uncover.³ Costly discovery of history imposes a natural and serious constraint especially in new markets such as online trading, as the establishment of effective collective governance and institutional enforcements often lags behind the rapid development of the markets.

Despite its salience in many transactions, and its connection to reputation dynamics, the implications of costly information acquisition have been left unexplored. This paper develops a model of reputation with costly discovery of histories. Moreover, in this model, non-trivial long-run dynamics emerge as an intrinsic reputation phenomenon.

Standard repeated game models with perfect monitoring do not readily capture the richness of reputation dynamics. With imperfect monitoring, [Cripps, Mailath, and Samuelson \(2004\)](#) show that reputation effects are necessarily transient and the reputation dynamics degenerate over time. The feature of short-run reputation dynamics and long-run convergence was also investigated by [Bar-Isaac \(2003\)](#) and [Benabou and Laroque \(1992\)](#). In order to explain long-run reputation effects, the literature has explored the assumption that economic agents’ types secretly change over time; see, e.g., [Gale and Rosenthal \(1994\)](#) for a study of product quality cycles, [Holmström \(1999\)](#) for a theory of dynamic provision of managerial incentives, [Phelan \(2006\)](#) for a model of dynamic government policies, and [Mailath and Samuelson \(2001\)](#) for an investigation of reputation dynamics and ownership transactions.⁴ These models

²A large and fast-growing literature on reputation systems in electronic commerce has emerged aimed at designing various practically implementable mechanisms to identify and manage agents with dishonest conducts. This literature often takes these behavior patterns as exogenously given. Recent surveys are provided by [Hoffman, Zage, and Nita-Rotaru \(2009\)](#) and by [Marti and Garcia-Molina \(2006\)](#).

³In some environments, reviewing the entire transaction history is costly even when data on detailed transactions are accessible. This observation cost motivates some online transaction platforms such as Amazon.com to search for summary statistics of past histories.

⁴[Bar-Isaac \(2007\)](#) provides a mechanism that endogenizes the uncertainty of types: an old agent can sell his

have two shortcomings. First, their predictions crucially rely on the assumption concerning *unobservable* variables; the equilibrium unravels if type changes are public. Second, the equilibrium dynamics are driven by exogenous stochastic processes that govern type changes, while the aforementioned applications suggest that the long-run reputation dynamics are an intrinsic reputational phenomenon.

In our model, a single long-lived agent, say, a firm, interacts with a sequence of short-lived agents, say, the customers. The long-lived agent is either *good* or *opportunistic*, and privately knows his type. The good type always makes a choice preferred by the short-lived agents. The opportunistic type is forward-looking and acts strategically, even though he prefers an exploitive action in the short term. The long-lived agent's reputation is the public belief on the good type. There is a tension between reputation building and exploitation as the benefit of exploitation increases with reputation. We depart from the existing literature by postulating that the long-lived agent's past transactions are costly to observe, and observing longer history is more costly. Unlike bounded rational agents whose ability to process information is exogenously constrained, the uninformed agents in our model are Bayesian and rational, but are constrained on information. They acquire data optimally by taking into account the stake in a transaction, and the fact that the data are the outcomes of the strategic play of the informed agent. Therefore, the value of information *and* information acquisition are determined endogenously in equilibrium.

We obtain long-run dynamics consistent with the observations in the first paragraph. The equilibrium behavior of the opportunistic type features a recurrent process with gradual reputation building and a sudden crash. Reputation is indeed accumulated, milked, and regained. Such dynamics illustrate the persistent tension between the temporary gain from exploitation and its effect on continuation payoffs. Our analysis differs from [Azam, Bates, and Biais \(2009\)](#) who develop a model investigating reputation building and exploiting behavior of an opportunistic government. They predict that the collapse of reputation is durable unless a new government is drawn, and the opportunistic government will opt for predation even anticipating that its reputation will be permanently lost.⁵

The short-lived agents always randomize as to how much information to acquire, and the opportunistic long-lived agent builds his reputation up exactly to a point where the maximum periods of information acquired by a customer do not reveal past exploitation. A short-lived agent is thus fully exploited when he most trusts the long-lived agent, according to a correct

firm to a young agent who can maintain the firm's value, and the old agent needs to exert effort to improve the young agent's reputation because of *joint production*. See [Tadelis \(1999\)](#) for an OLG model of name trading, and [Bar-Isaac and Tadelis \(2008\)](#) for a comprehensive survey of this literature.

⁵We shall return to this point later, as it demonstrates a different but related driving force of reputation dynamics neglected in the literature.

posterior belief. This result captures the deliberate maneuver of the long-lived agent who possesses superior information. Additional reputation building beyond that limit is clearly unprofitable because unobservable good behavior will not be rewarded with additional trust, while reputation falling short of that limit will induce some of the agents to observe the past exploitation.

By taking the long-run reputation dynamics as exogenously given and ignoring economic agents' strategic responses, a number of researchers have argued that keeping the observable history short can prevent agents from building a high, exploitable reputation, and hence prohibit cheating behavior.⁶ Our analysis reveals the opposite: shortening the observable history merely makes the reputation building process shorter and exploitation occurs more frequently.

The rest of this paper is organized as follows. Section 2 sets up the model. Section 3 elaborates on the equilibrium definition. In the same section, we show that in the degenerate case of complete information the repetition of the static Nash equilibrium is the unique perfect equilibrium outcome. Section 4 uses a simple numerical example to illustrate the concepts and demonstrate the basic intuition. Section 5 contains the main results. Section 6 constructs the equilibria for linear cost functions. Section 7 discusses extensions and related papers on bounded memory.

2 The Model

2.1 The Underlying Game

Our model is built on the model of experienced good considered by Mailath and Samuelson (2001, 2006). Phelan (2006) and Azam, Bates, and Biais (2009, Section 4) study a similar game.

A long-lived firm (player 1) meets a sequence of short-lived customers (player 2), with a new short-lived customer entering in every period. In each period, player 1 chooses from two possible product or service qualities, high (H) and low (L). Player 2, without knowing the quality, chooses whether to take a trusting action h or a distrusting action l . One can interpret h as a high quantity or a high price.

High quality is more costly to produce, and therefore in a one-shot game player 1 strictly prefers the cheating action L . Anticipating a low quality L , player 2 will respond by taking the distrusting action l . Therefore, (L, l) is the unique static Nash equilibrium. However, if

⁶See, for example, Aperjis and Johari (2010) and Feldman, Lai, Stoica, and Chuang (2004). A recent survey of this literature is provided by Hoffman, Zage, and Nita-Rotaru (2009).

the quality is observable in advance, player 1 will choose H to induce player 2's trusting action h . That is, (H, h) is the *Stackelberg* outcome.⁷

For an action profile $a \in \{H, L\} \times \{h, l\}$, let $u_1(a)$ and $u_2(a)$ be player 1's and player 2's stage game payoffs. All payoffs are normalized to be nonnegative. Since action L is player 1's stage game dominant action, we have $u_1(L, h) > u_1(H, h)$ and $u_1(L, l) > u_1(H, l)$; moreover, since the profile (L, l) is the stage game Nash equilibrium, we have $u_2(L, l) > u_2(L, h)$; since the profile (H, h) is the Stackelberg outcome, we need $u_1(H, h) > u_1(L, l)$. It is assumed that a firm's short-term cheating benefit is higher when customers acquire a large quantity. Formally, $u_1(L, h) - u_1(H, h) > u_1(L, l) - u_1(H, l)$.⁸ This condition implies a growing tension between reputation building and exploitation: as player 2's trust increases, player 1's temptation to exploit its reputation becomes stronger. We shall demonstrate the role of this condition in Section 4.

2.2 Types

There are two types of player 1: the *good* type always chooses high quality H , and the *opportunistic* type maximizes his expected payoff with a discount factor $\delta \in (0, 1)$. The common prior probability on the good type is $\mu_0 > 0$. This information structure is common knowledge. In the sequel, whenever we refer to player 1 without mentioning his type, we shall mean the opportunistic type.

To focus on the interesting cases, assume $\mu_0 < \bar{\mu} = \frac{u_2(L, l) - u_2(L, h)}{u_2(L, l) - u_2(L, h) + u_2(H, h) - u_2(H, l)}$. Here $\bar{\mu}$ is such that player 2 is indifferent between h and l in a one-shot interaction. If $\mu_0 > \bar{\mu}$, player 2 will play h even though he knows that the opportunistic type plays L with probability 1.

⁷The tension between the static Nash outcome and the Stackelberg outcome differentiates our model from the main model of Azam, Bates, and Biais (2009, Sections 2 and 3). Our stage game corresponds to the limited predation model of Azam, Bates, and Biais (2009, Section 4), where they show that pooling is an equilibrium for large discount factors. This is not the case in our model. These differences reveal an interesting feature of reputation games overlooked in the previous literature. The tension between the static Nash outcome and the Stackelberg outcome creates the tradeoff between the temporary gain from exploitation and the loss of continuation payoff. If the former force (temporary gain) dominates, the short-lived players' distrust will eliminate any reputation building incentives. If the latter force dominates, exploitation will be fully deterred, and a pooling equilibrium will result from a trigger strategy argument. In Azam, Bates, and Biais (2009, Sections 2 and 3), the opportunistic government derives the same payoff from the static Nash and the Stackelberg outcomes. The balance of the two forces in their model lead to reputation dynamics. In our model, costly observation of histories endogenously balances the two forces. On the one hand, a lost reputation could be regained, and so exploitation cannot be fully deterred. On the other hand, exploitation incentive will not eliminate trusts: a partially observable history provides enough room for the opportunistic type to manipulate the public beliefs.

⁸Similar conditions also obtain in Azam, Bates, and Biais (2009) and Phelan (2006): the government's incentive for exploitation is higher if the public invests/produces more.

2.3 Information Acquisition

The long-lived player observes the full history of outcomes. The short-lived players, upon entering the game, observe neither the previous outcomes nor the number of transactions before them: they are symmetric *ex ante*. A short-lived player 2 can pay a cost, $C(n)$, upon entering, to observe player 1's actions in the previous n periods, $n \in \{0, 1, \dots\}$.⁹ Information acquisition is not observable to player 1. We make the following assumption on the cost function $C : \{0, 1, \dots\} \rightarrow \mathbb{R}_+$.

Assumption. (1) $C(n)$ is weakly increasing. (2) $C(0) = 0$ and there exists $N_C > 0$ such that $C(n) > \max\{u_2(H, h), u_2(L, l)\}$ for any $n > N_C$.

Part (2) says that it is too costly to acquire all the information. If the marginal cost of information is bounded below by some positive number, this condition does not restrict the scope of the analysis, for as we shall see later, player 2 will not buy all that he can afford because the marginal value of information decreases exponentially in the amount of information. This assumption is convenient as it allows us to define a finite state space consisting of finite histories. An example is the linear cost: $C(n) = cn$ for some constant $c > 0$. Another example is the “wholesale of information”: banks can obtain a borrower's credit history for a fixed number of periods at a fee.¹⁰

3 Strategies and Equilibria

The equilibrium notion used in this paper is perfect Bayesian equilibrium in stationary strategies tailored for our model without calendar times.

3.1 State Space and Stationary Strategies

For a given cost function C , individual rationality of player 2 implies he never buys more than N_C periods of information. Fix any $N \geq N_C$. Let us define the *state space* $S = \{H, L\}^N$ as the set of player 1's feasible plays in the last N periods. For a state/history $s = (s_N, s_{N-1}, \dots, s_1) \in$

⁹We assume that the information on the play of the short-lived players is unavailable or prohibitively expensive. This assumption is reasonable for a number of applications: a customer might complain about a restaurant's services without mentioning how he treated the waiters; a credit history records a borrower's default but not the terms set by the lender, and so on. This assumption helps us to get rid of the bootstrapping equilibria because player 2's actions cannot be used as a signal of past histories (Theorem 1). See Section 7.3 for an equilibrium when this assumption is relaxed.

¹⁰Miller (2003, p. 110) reports that in Italy, only credit files for the past 12 months of new customers are available and each month the oldest files are deleted from the database.

S , s_1 is the most recent outcome and s_N the oldest.¹¹

Player 1 knows the true state, while player 2 needs to acquire information to estimate it. Since player 2s are ex ante symmetric, we consider their symmetric strategies. Denote player 2's information acquisition strategy by a probability measure $\alpha \in \Delta\{1, 2, \dots, N\}$. With probability $\alpha(n)$, player 2 acquires n periods of information, and his information is represented by a partition \mathcal{P}^n on S . The partition element containing s , $\mathcal{P}^n(s)$, is the set of finite histories having the same most recent n entries as s . In particular, $\mathcal{P}^0(s) = S$ for each $s \in S$. Note that since \mathcal{P}^n becomes finer as n increases, more information is more informative *ex post*.

Player 2's strategy after acquiring n periods of information is a \mathcal{P}^n -measurable function $\sigma_2^n : S \rightarrow [0, 1]$ which specifies a probability of playing the trusting action h for each $s \in S$. Let us write $\sigma_2 = (\sigma_2^n)_{0 \leq n \leq N}$.

Since information acquisition activity is private, player 1 might not know what player 2 actually observes. Player 1 will respond to player 2's *expected strategy* weighted by the information acquisition probability α . Write player 2's expected strategy as $\bar{\sigma}_2(s) = \sum_n \alpha(n) \sigma_2^n(s)$. Consequently, player 1 always has a stationary best response that only depends on the finite state space S . Denote $\sigma_1 : S \rightarrow [0, 1]$ as player 1's stationary strategy that specifies a probability of playing H for each $s \in S$. We focus on these stationary strategies, allowing for arbitrary deviations on the part of player 1.

Remark 1. *We do not specify the strategies for the initial periods. For example, the first short-lived player will observe an empty history if he acquires information. However, the finite game consisting of the initial periods can be solved after the continuation play is identified and it does not affect the long-run dynamics.*

3.2 Belief Updating

Since in general short-lived players have no knowledge about when the game starts, their beliefs need to agree with the steady-state beliefs induced by the play of the game. In other words, their beliefs are updated according to Bayes' rule, as if they held an improper uniform prior over the number of transactions before them. This idea was first formulated by Rosenthal (1979) for general games with varying opponents.

To facilitate the reader's understanding, we elaborate on the equilibrium notion based on such belief consistency as is appropriate for our model. A reader familiar with this idea can skip to Section 3.3. For $N = 3$, all possible transitions induce a directed graph as depicted in Figure 1.

¹¹Defining a state space by invoking the individual rationality of player 2 facilitates the analysis, even though we neglect irrelevant off-equilibrium-path histories in which players observe more than N periods of information.

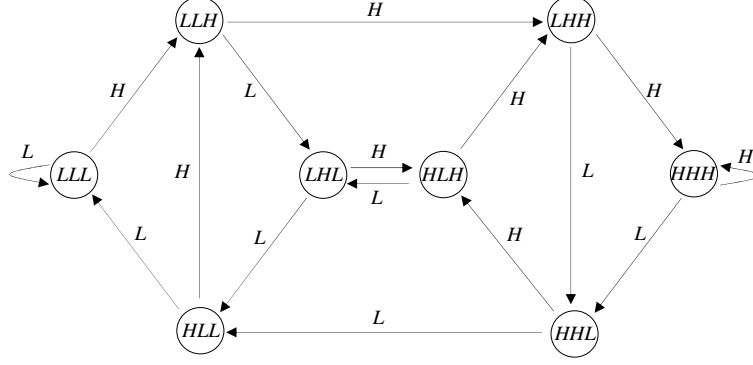


Figure 1: Transition on $S = \{H, L\}^3$ with 8 states

From player 2's point of view, nature chooses one of the two types of player 1 and hence one of the following two stochastic processes on S :

(1) The good type plays H constantly and the process is degenerate. Let λ^* be the steady-state distribution of this process, $\lambda^*(\underbrace{HH\dots H}_N) = 1$.

(2) The opportunistic type's strategy σ_1 induces a Markov process on the directed graph whose state space is S . Let $\lambda \in \Delta(S)$ be a steady-state distribution of this process.

After learning his information set $\mathcal{P}^n(s)$, player 2 updates his belief about the first process (the good type) according to Bayes' rule whenever possible:

$$\frac{\mu_0 \lambda^*(\mathcal{P}^n(s))}{\mu_0 \lambda^*(\mathcal{P}^n(s)) + (1 - \mu_0) \lambda(\mathcal{P}^n(s))}. \quad (1)$$

Consistent with this formulation, player 2 holds the prior belief μ_0 when he does not acquire information (i.e., $n = 0$), and he assigns probability 0 to the good type if he sees an L .

3.3 Equilibria and Preliminary Results

Table 1: Frequently used notations

$S = \{H, L\}^N$	the state space of finite histories with length $N \geq N_C$
$\sigma_1 : S \rightarrow [0, 1]$	player 1's stationary strategy (prob. of playing H)
$\alpha \in \Delta\{1, 2, \dots, N\}$	player 2's information acquisition strategy
$\sigma_2^n : S \rightarrow [0, 1]$	player 2's strategy (prob. on h) after observing n periods of data
$\bar{\sigma}_2 = \sum \alpha(n) \sigma_2^n$	the expected strategy of player 2
$\lambda \in \Delta S$	the steady-state distribution on S

A perfect Bayesian equilibrium requires that strategies be best responses to each other given beliefs and that the beliefs be consistent with the strategies.

Definition 1. A quintuple $(S, \sigma_1, \alpha, \sigma_2, \lambda)$ is a perfect Bayesian equilibrium under the cost function C if (1) $S = \{H, L\}^N$, $N \geq N_C$, (2) player 2 updates his belief using prior and λ , and σ_1 and (α, σ_2) are best responses given belief, and (3) λ is a steady-state distribution on S consistent with strategy σ_1 .

An equilibrium is still well defined when $\mu_0 = 0$. Note that the state space S is part of the equilibrium definition in the model of endogenous history discovery. There seems to be a flexibility in the choice of N . This is not essential because player 2 does not observe more than N_C periods of data, and hence additional periods of data simply serve as a randomizing device whose outcomes are observed only by player 1.

Lemma 1. An equilibrium exists.

The proof is omitted as it follows the usual fixed-point argument (Rosenthal, 1979, Theorem 1).¹² Despite the complicated transition on S , we obtain a sharp result for the degenerate case of $\mu_0 = 0$.

Theorem 1. In the complete information game ($\mu_0 = 0$), the infinite repetition of the static Nash equilibrium (L, l) is the unique perfect equilibrium outcome for any discount factor $\delta \in (0, 1)$. Consequently, player 2 never acquires information if $C(1) > 0$.

This result contrasts with the folk theorem proved by Fudenberg, Kreps, and Maskin (1990).¹³ The proof is in the Appendix.

4 An Example

	h	l
H	2,3	0,2
L	4,0	1,1

Figure 2: A version of the product-choice game

¹²Information acquisition does not pose a problem: we first prove the existence of a triple $(\sigma_1, \sigma_2, \lambda)_\alpha$ for any given α and then choose α to maximize player 2's payoff.

¹³For the game we considered, the folk theorem does not require knowledge about histories of the short-lived players.

We use a numerical example to demonstrate the equilibrium properties and the basic ideas. The payoffs in Figure 2 satisfy all of our assumptions. Consider the following cost structure: $C(1) = 0$ and $C(2) = 5$. The first period of information is free but additional information is prohibitively expensive. We consider the state space $S = \{H, L\}$ and assume that player 2 always observes player 1's previous-period play, i.e., $\alpha(1) = 1$. For simplicity, we write player 2's strategy as σ_2 instead of σ_2^1 .

4.1 $\mu_0 = 0$

Theorem 1 predicts that repetition of (L, l) is the unique perfect equilibrium outcome. The basic idea of the argument is that player 2 finds it irrelevant to observe information concerning player 1's previous play.¹⁴ This information is valuable to player 2 only if he plays differently after seeing different histories. By definition, player 2 plays the trusting action h with probability $\sigma_2(s)$ when the previous-period play is $s \in \{H, L\}$. Assume $\sigma_2(H) > \sigma_2(L)$ (the other case is symmetric). Since player 2 plays h with probability $\sigma_2(L) < 1$ upon observing L , it must be that player 1 plays the cheating action L with a positive probability upon the same history (because h is the best response to H).

We derive a contradiction based on the following intuition: the incentive of the reputation game is such that player 1's incentive for exploitation grows with player 2's trust; now we have just seen that player 1 cheats upon history L where he is trusted less (because $\sigma_2(L) < \sigma_2(H)$), and hence player 1 should strictly prefer to cheat upon history H – but then player 2 should respond with l upon this history; i.e., $\sigma_2(H) = 0$. However, by assumption, $\sigma_2(H) > \sigma_2(L) \geq 0$.

Formally, let $U_1(s)$ denote player 1's normalized discounted continuation payoff upon history $s \in \{H, L\}$. Then the IC condition for player 1 to play action L at history L is

$$\delta [U_1(H) - U_1(L)] \leq (1 - \delta)[2\sigma_2(L) + 1 - \sigma_2(L)], \quad (2)$$

where LHS is player 1's loss in continuation payoff from exploitation (which only depends on the current play because of one-period memory) and RHS is the immediate benefit from exploitation: when player 2 plays h , player 1 gets an extra 2 units of payoff by playing L instead of H ; when player 2 plays l , player 1 gets an extra 1 unit.

Since by assumption $\sigma_2(L) < \sigma_2(H)$, exploiting upon history H should be more profitable; i.e., the following inequality holds:

$$2\sigma_2(L) + 1 - \sigma_2(L) < 2\sigma_2(H) + 1 - \sigma_2(H). \quad (3)$$

¹⁴I am grateful to an anonymous referee for clarifying these arguments.

Combining (3) and (2), we have

$$\delta [U_1(H) - U_1(L)] < (1 - \delta) [2\sigma_2(H) + 1 - \sigma_2(H)]. \quad (4)$$

This inequality is the IC condition for exploitation to be strictly optimal for player 1 on history H . This implies that player 2 should not play h upon this history, i.e., $\sigma_2(H) = 0$, a contradiction. We conclude that observing player 1's previous play will not affect player 2's continuation play. Therefore, player 1 will play his stage-game dominant strategy L , and player 2 will respond with l .

As this argument demonstrates, besides the fact that the stage game has a unique equilibrium in which player 1 plays the dominant strategy, the driving force of this result is that in a reputation game player 1's incentive to play a cheating action is higher if player 2 trusts him more. This property eliminates the incentive for player 2 to play differently upon different histories.

In this simple example, player 1 has the ability to hide the past cheating actions. This results from the assumption of one-period memory and the fact that the short-lived players are prevented from transmitting information to future players using their own actions.¹⁵

4.2 $\mu_0 > 0$

In the non-degenerate case we focus on, the intertemporal incentive will not fully collapse, if the long-lived player 1 is patient enough. The reasoning is as follows. If player 1 plays L repeatedly in a perfect equilibrium, player 1's discounted average payoff is 1. If, however, player 1 defects to playing H forever, future player 2 will assign probability 1 to the good type (except in the period immediately following the deviation) and respond by h , and so his payoff will be at least 2δ and the deviation will be profitable for $\delta > \frac{1}{2}$.

While the intertemporal incentive does not fully collapse, the Stackelberg outcome (H, h) cannot be enforced. To see this, suppose player 1 plays H and hence obtains an average discounted payoff of 2 on the equilibrium path. The worst off-the-equilibrium-path payoff player 1 could ever receive is 0.¹⁶ Following a history H , player 1 has a profitable deviation of alternating between L and H , which results in a sequence of payoffs of at least $4, 0, 4, 0, \dots$ with a discounted average larger than 2.¹⁷ Why doesn't a grim trigger strategy argument work here?

¹⁵This is where the unobservability of the actions short-lived players are used. See Footnote 9 and Section 7.3.

¹⁶Player 1 can actually get strictly more than 0. The reason is that player 1 gets 0 from outcome (H, l) but the short-lived player 2 is unwilling to "punish" player 1 by playing l which is not a best response to H . The lower bound will be too loose for general payoffs.

¹⁷The alternative repeated game strategy involves a series of deviations from the constant play of H . This

The answer has to do with off-path histories. A grim trigger strategy profile in standard games specifies the worst punishment after a single deviation even when the deviation occurs in the remote past. There is no escape from this punishment once deviation occurs. With one-period monitoring, however, an immediate deviation can be detected, but deviations in the remote past cannot. For example, an off-path history with a deviation from 3 periods earlier cannot be distinguished from an on-path history. Therefore, player 1 can elude the punishment after a past deviation.

The argument above shows that we cannot have an equilibrium with degenerate long-run dynamics. The following graph describes an equilibrium when $\delta > \frac{1}{3}$ and $\mu_0 > \frac{1}{4}$.¹⁸

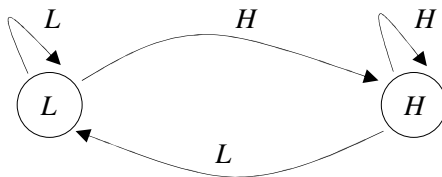


Figure 3: In state H : (L, h) is played, and hence the transition back to state H is off the equilibrium path. In state L : Player 1 plays H with probability $\sigma_1(L) = \frac{1}{2}$; player 2 plays h with probability $\sigma_2(L) = \frac{3\delta-1}{3\delta+1}$.

In this equilibrium, the opportunistic type plays a completely mixed strategy in state L . When H is observed, player 2 updates his belief on the good type to such an extent that he is willing to play the trusting action h , while the opportunistic player 1 exploits this trust by playing L . State L is a reputation-building state and state H is a reputation-exploitation state.

Let us verify that the prescribed strategy profile is indeed an equilibrium. First, in state L , player 2 knows player 1's type; we need $\sigma_1(L) = \frac{1}{2}$ to make player 2 indifferent. Second, in state H , the posterior belief on the good type is given by $\mu_H = \frac{\mu_0}{\mu_0 + (1-\mu_0)\frac{\sigma_1(L)}{1+\sigma_1(L)}} = \frac{\mu_0}{\mu_0 + \frac{1}{3}(1-\mu_0)}$, where $\frac{\sigma_1(L)}{1+\sigma_1(L)}$ in the denominator is player 2's steady-state belief on state H ; to induce player 2 to play the trusting action h , we need $\mu_H \geq \frac{1}{2}$, i.e., $\mu_0 \geq \frac{1}{4}$.

We still need to check player 1's incentives upon the two histories. For player 1 to be indifferent in state L , we need $\sigma_2(L) = \frac{3\delta-1}{3\delta+1}$, which requires $\delta \geq \frac{1}{3}$. It's straightforward to check that action L is player 1's best response in state H .

allows for an easy computation of payoffs for this specific example. With general payoff and cost structures, such an alternative strategy will be difficult to construct. Note that the one-deviation principle still holds. Our game has “perfect recall.” A short-lived player takes a one-shot action based on *all* the information he acquires. Mailath and Samuelson (2006, p. 397) present a general result for the one-deviation principle.

¹⁸In this example $\bar{\mu} = \frac{1}{2}$. If $0 < \mu_0 < \frac{1}{4}$, one obtains similar equilibrium dynamics in which $\sigma_2(L) = 0$, $\sigma_2(H) \in (0, 1)$, $\sigma_1(L) \in (0, 1)$, and $\sigma_1(H) = 1$.

4.3 Generalization

In this example, the maximum amount of information that can be acquired is exogenously assumed to be 1. We must address two critical features of the equilibrium in Figure 3 in order to generalize the idea to endogenous information acquisition with general payoffs and cost structures. First, in the equilibrium, the trusting action h is a best response for player 2 regardless of the states. If $C(1) = \varepsilon > 0$, the equilibrium will unravel even when ε is small, because player 2 should play the pure strategy h without paying the cost, and hence player 1 should always cheat. Second, conditional on player 1 being the opportunistic type, the equilibrium path never contains two consecutive H 's, even though there may be an arbitrary number of consecutive L 's. If $C(1) = 0$ and $C(2) = \varepsilon$, then player 2 could acquire two periods of information to find out player 1's type and to avoid being exploited.

We show that the dynamics of reputation building and exploitation will not unravel for general cost structures and payoffs. Generically, player 1 will build his reputation up to an endogenous point where the number of consecutive H 's coincides with the maximum periods of information observed by player 2. The behavior of player 2 changes from the example. Player 2 will not trust the firm at all upon observing a bad action after paying a positive cost. However, this is not enough to prevent player 1 from exploitation because player 2 plays mixed strategies in information acquisition, while continuing to play trusting actions in response to good actions.

5 Reputation Dynamics

The state space $S = \{H, L\}^N$ grows exponentially with N , and *a priori* the direct graph on the 2^N states lacks an obvious, tractable transition (See Figure 1 for $N = 3$). Notice that a history reveals player 1's type as long as it contains a single cheating action L and the number of H 's since the most recent L measures how far player 1 has gone toward a clean record with straight H 's. This observation motivates a natural index which cuts through the graph on S .

Definition 2. *The clean history index of s , $I(s)$, is the number of good actions H since the most recent cheating action, L , in state s . Formally,*

$$I(s) = \begin{cases} N & \text{if } \forall i \in \{1, \dots, N\}, s_i = H \\ \min \{i : s_i = L\} - 1 & \text{otherwise} \end{cases}.$$

For example, $I(\dots LHH) = 2$, $I(\dots L) = 0$, and $I(H \dots HH) = N$. Note that I partitions S . By definition, an additional H will increase the index from i to $\min\{i + 1, N\}$ and an L will

reduce the index down to 0. If player 2 buys $n < N$ periods of information, he is able to tell whether the index is $0, 1, \dots, n$ or larger, even though he does not know the true history s . It is easier to study the linear transition using dynamic programming techniques. We will focus on the class of equilibria that are characterized by the index set \mathbf{I} . That is, if two histories have the same index, then the players play the same strategy in the two histories. Indeed, all equilibria on S can be “transformed” into equilibria on \mathbf{I} .¹⁹

Remark 2. *With some abuse of notation, write an equilibrium characterized by the index as $(\mathbf{I}, \sigma_1, \alpha, \sigma_2, \lambda)$. We refer to \mathbf{I} as the new state space. For a history with index i , write $\sigma_1(i)$ as player 1’s strategy and $\sigma_2^n(i)$ as player 2’s strategy after acquiring n periods of information.²⁰ In addition, $\lambda(i)$ and $U_1(i)$ are defined as the belief and the expected discounted average payoff in state i . Write $\bar{\sigma}_2(i) = \sum_n \alpha(n)\sigma_2^n(i)$ as player 2’s expected strategy in state i .*

If player 2 never acquires information, the opportunistic type of player 1 plays L . Player 2 is tempted to acquire one period of information to check player 1’s type. The benefit of such information is $\mu_0 [u_2(H, h) - u_2(H, l)]$ because player 2 should play h instead of l if he finds out that player 1 is the good type. The cost is $C(1)$. Information acquisition is beneficial if $C(1) < \mu_0 [u_2(H, h) - u_2(H, l)]$. We summarize this observation below.

Lemma 2. *Not acquiring information is an equilibrium if and only if $C(1) \geq \mu_0 [u_2(H, h) - u_2(H, l)]$. In this equilibrium, player 2 plays l and the opportunistic player 1 plays L .*

Remark 3. *We shall assume that $C(1) < \mu_0 [u_2(H, h) - u_2(H, l)]$ in the sequel, because it makes sense to talk about equilibrium dynamics only when the cost is small enough that player 2 is willing to discover the past play. Furthermore, for the reputation effect to work, player 1 has to sufficiently care about continuation payoff. Let $\bar{\delta} = \frac{u_1(L, h) - u_1(H, h)}{u_1(L, h) - u_1(L, l)} \in (0, 1)$. In what follows, we characterize the structure of all equilibria $(\mathbf{I}, \sigma_1, \alpha, \sigma_2, \lambda)$ for $\delta > \bar{\delta}$.*

5.1 Long-Run Reputation Dynamics

Theorem 2. *If $\delta > \bar{\delta}$, then every equilibrium $(\mathbf{I}, \sigma_1, \alpha, \sigma_2, \lambda)$ exhibits a stochastic reputation cycle: there exists an integer n^* , $0 < n^* \leq N$, such that the set $\{0, 1, \dots, n^*\}$ consists of all states on the equilibrium path and it can be broken down into the following two phases:*

¹⁹See the Supplementary Materials for details.

²⁰Let $\mathcal{I}(s) = \{s' \in S : I(s') = I(s)\}$ be the partition element containing s and $\mathbf{I} = \{0, 1, \dots, N\}$. Player 2 might not learn the index i if he acquires $n < i + 1$ periods of information. Therefore, $\sigma_2^n(i)$ is measurable with respect to the finest common coarsening of $\mathcal{I}(s)$ and \mathcal{P}^n given by $(\mathcal{I} \vee \mathcal{P}^n)(s) = \begin{cases} \mathcal{P}^n(s) & \text{if } n \leq I(s) \\ \mathcal{I}(s) & \text{otherwise} \end{cases}$.

(1) the set $\{0, 1, \dots, n^* - 1\}$ forms a reputation-building phase in which player 1 plays strict mixed strategies: $0 < \sigma_1(i) \leq \bar{\mu} < 1$ for each state i in this set.²¹

(2) the state n^* is a reputation-exploitation phase where player 1 plays L with probability 1.

The reputation dynamics are characterized by a recurrent stochastic process that “starts afresh” in state 0, as illustrated in Figure 4. In the reputation-building phase, player 1’s strategies are completely mixed, and the probabilities assigned to the good action are uniformly bounded above by the constant $\bar{\mu} < 1$. This property gives us a bound on the likelihood of longer clean histories, and hence on the value of information. In equilibrium, player 1 builds his reputation by manipulating public beliefs. As long as player 2 does not see an L , the posterior belief on the good type will increase in the index. Once he has played H ’s for n^* consecutive periods, player 1 depletes his reputation completely and hence all states higher than n^* are off the equilibrium path. Therefore, n^* is the maximum number of consecutive good actions, and it is also the minimum length of time that a long-run player needs to spend before reaching the peak of his reputation.

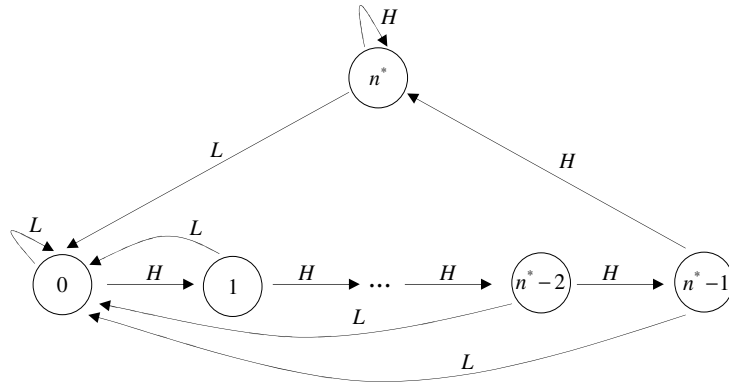


Figure 4: $\sigma_1(i) \in (0, \bar{\mu}]$ if $i < n^*$ and $\sigma_1(n^*) = 0$. Therefore, reputation collapses completely at n^* and the transition from n^* back to itself is off the equilibrium path.

The proof of Theorem 2 utilizes “rolling induction” extensively. We sketch the three steps below. Details can be found in the Appendix.

Step 1. Unique Ergodic Set. The opportunistic type’s play in an equilibrium $(\mathbf{I}, \sigma_1, \alpha, \sigma_2, \lambda)$ must induce at least one ergodic set on \mathbf{I} .²² Because of the linear transition on \mathbf{I} , an ergodic set must take the form of $\{m, m + 1, \dots, n\}$ where $0 \leq m \leq n \leq N$. We show that the equilibrium ergodic set is unique with $m = 0$, $n > 0$, and $\sigma_1(n) = 0$. The uniqueness also guarantees that

²¹Recall that $\bar{\mu}$ is the probability on action H such that player 2 is indifferent between h and l in a one-shot game.

²²See Stokey and Lucas (1989) for the definition of ergodic sets.

initial periods do not affect the long-run dynamics. The uniqueness characterization is not trivial: *a priori* both $\{0\}$ and $\{N\}$ can be ergodic sets on the equilibrium path conditional on the opportunistic type. If $\{N\}$ is ergodic, then there is a “permanent reputation” at the absorbing state N . For the simple example in Section 4.2, we find a specific profitable deviation for player 1 with a simple payoff bound; see footnotes 16 and 17. This is difficult to do for general payoff and cost structures. We show *by induction* that the exploitation incentive is so strong that the cheating action L must eventually be played in high states, save this case. If $\{0\}$ is absorbing, then there is no reputation building. Player 2 will find out player 1’s type if he buys one period of information. This induces player 1 to play H and leave the absorbing state in order to get at least $u_1(H, l)$ now and $u_1(H, h)$ later, which is better than the absorbing state if $(1 - \delta)u_1(H, l) + \delta u_1(H, h) > u_1(L, l)$. A sufficient condition for this inequality to hold is $\delta > \bar{\delta}$.

This step does not complete the characterization. Player 1 could still play H in pure strategies in lower states *a priori*. This is not sufficient to support the claim that player 1’s reputation is player 2’s *belief about types*.

Step 2. Increasing Incentive. Step 1 shows that player 1 must go from state 0 to higher states in equilibrium. Therefore, player 1 must be compensated for “climbing” toward higher states by playing the static *dominated* action H . We show that player 2’s expected strategies $\bar{\sigma}_2(i) = \sum_n \alpha(n)\sigma_2^n(i)$ must be non-decreasing on the equilibrium path in order to provide reputation-building incentives.

Step 3. Uniform Bound on Mixing Probabilities. Using Step 2, we provide a uniform upper bound on player 1’s mixing probabilities. The proof is by contradiction. If $\sigma_1(i) > \bar{\mu}$ for some i on the equilibrium path, then player 2 will respond with h whenever he observes this state by the definition of $\bar{\mu}$. Consequently, player 2’s expected strategies in higher states cannot assign larger probabilities to h ; they must all be equal by the weak monotonicity property in Step 2. As a result, player 1 has no incentive to play H in state i now and L later: he prefers to play L as soon as possible as he discounts his payoff.

We discuss several additional interesting properties. The result obtained in Step 3 can help us strengthen the monotonicity property obtained in Step 2.

Lemma 3. *On the equilibrium path, the probability that player 2 plays the trusting action h , $\bar{\sigma}_2(i) = \sum_n \alpha(n)\sigma_2^n(i)$, and player 1’s expected payoff, $U_1(i)$, are strictly increasing in reputation index i .*

However, $\sigma_1(i)$ may not be monotonic in i . Cost structures can be constructed such that $\sigma_1(i)$ increases or decreases in i . Intuitively, player 1 is more willing to mimic the good type in a state when the information about that state is less expensive.

Corollary 1. *If the cost structure is $C(N) = 0$ and $C(N + 1) = \infty$ as in the example in Section 4, then $n^* = N$.*

The reasoning is as follows. If $n^* < N$, then player 2 will know player 1's type by observing L . So player 2 will play l at n^* , which contradicts Lemma 3. Together with Theorem 2, this result says that imposing a shorter memory record will only make the exploitation more frequent.

Corollary 2. *Fix an equilibrium $(\mathbf{I}, \sigma_1, \alpha, \sigma_2, \lambda)$. Then $\lambda(i)$, the probability that state i is reached in equilibrium, is strictly decreasing in i and $0 < \lambda(i) < \bar{\mu}^i$ for any $i \in \{1, 2, \dots, n^*\}$, where n^* is characterized in Theorem 2.*

This result is an immediate consequence of the uniform bound on player 1's mixing probabilities. A uniform bound on player 1's mixing probabilities also implies a bound on the equilibrium value of information. To see this, suppose player 2 acquires i periods of information. An additional period of information is helpful only when it reveals a cheating action, which the first i periods do not. In this case, the additional information may change player 2's action from h to l . The posterior belief on the opportunistic type upon observing i straight H 's is bounded above by $\frac{(1-\mu_0)\bar{\mu}^i}{\mu_0+(1-\mu_0)\bar{\mu}^i}$. The value of an additional period of information is therefore at most $\frac{(1-\mu_0)\bar{\mu}^i}{\mu_0+(1-\mu_0)\bar{\mu}^i} [u_2(L, l) - u_2(L, h)]$, which decreases exponentially in i .

5.2 The Short-lived Players

Theorem 2 shows that the opportunistic type plays H for at most n^* consecutive periods. Therefore, player 2 will learn player 1's type and avoid being exploited if he acquires $n^* + 1$ periods of information. But the value of additional information decreases exponentially. The following result demonstrates player 2's behavior under the delicate manipulation of the informed player 1.

Theorem 3. *Suppose $\delta > \bar{\delta}$ and the cost function is strictly increasing. Consider any equilibrium $(\mathbf{I}, \sigma_1, \alpha, \sigma_2, \lambda)$ and let n^* be the highest reputation index in this equilibrium characterized in Theorem 2. Then the following properties hold:*

(1) n^* is the maximum number of periods of information that player 2 acquires for generic cost functions.²³

²³There is always a knife-edge case in which a player is exactly indifferent between acquiring n^* and $n^* + 1$

(2) *Player 2 plays completely mixed strategies: $\alpha(i) > 0, i \in \{0, 1, \dots, n^*\}$.*

(3) *Upon acquiring information, player 2 plays pure strategy l whenever he sees an L and he plays pure strategy h whenever he sees straight H 's.*

We provide the key intuition behind the results. For (1), if player 2 observes at most $n^* - 1$ periods, player 1 should exploit his reputation when the index is $n^* - 1$ because player 1, who discounts his payoff, prefers to reap the benefit sooner rather than later; if player 2 acquires $n^* + 1$ periods or more, player 1 will have an incentive to wait: his payoff is higher if he exploits his reputation when player 2 trusts him more. Genericity is used to break the indifference. Part (2) is a consequence of a free-rider problem. Absent direct information sharing, the free-rider problem arises from the disciplining effect that information discovery has on the opportunistic type.²⁴ The somewhat surprising part is that the randomization has full support. Part (3) is intuitive: if player 2 is indifferent between his two actions upon obtaining the information, then he should play a pure strategy without paying for the information. Note that Part (3) does not predict player 2's behavior when he does not acquire information. He will be more likely to play the trusting action h if the observation cost is lower because the short-lived players on average acquire more information which disciplines the long-lived player more. In the next section, we demonstrate this intuition with linear costs.

From Part (2) of Theorem 3, the short-lived player 2's indifference conditions characterize player 1's strategy, and hence player 1's strategy is independent of discount factors. Its economic interpretation is clear. Player 1 plays H merely to gain enough trust from player 2 and then exploits it; the opportunistic type has no other long-term incentives. It is in this sense that our model, though somewhat complicated, is parsimonious. We summarize the result below.

Proposition 1. *If $\delta > \bar{\delta}$, then n^* and σ_1 are independent of δ .*

6 Linear Cost

The characterization in the previous section enables us to construct equilibria explicitly. First, for any n^* (its exact value will be pinned down later), indifference conditions of both players determine player 1's strategy σ_1 and player 2's expected strategy $\bar{\sigma}_2$. We then pin down α from $\bar{\sigma}_2(i) = \sum \alpha(n)\sigma_2^n(i), i = 0, 1, \dots, n^*$. Finally, we pin down n^* from player 2's IC condition for information acquisition.

periods of information. See Section E.2 in Appendix for details. For the family of linear cost functions $C(n) = cn$, Section 6 shows that the non-generic c 's are countable.

²⁴If others acquire lots of information, the opportunistic player 1 will be more likely to refrain from exploitation and this will in turn reduce player 2's incentive to acquire information.

Here we demonstrate the construction for the family of linear cost functions: $C(n) = cn$, where $0 < c < \bar{c} = \mu_0[u_2(H, h) - u_2(H, l)]$. We focus on the characterization of n^* and player 2's action when he does not buy information. The proof, together with the computation of σ_1, σ_2 , and α , involves routine algebra and can be found in the Supplementary Materials.

Proposition 2. *If $C(n) = cn$, the equilibrium on \mathbf{I} is unique for generic $c \in (0, \bar{c})$ when $\delta > \bar{\delta}$. Moreover, there exist a strictly decreasing sequence $\{c_n\}_{n=1}^\infty$ and a cutoff $c^* > 0$ that are independent of δ , such that*

- (1) *if $c \in (c_n, c_{n+1})$, then $n^* = n$;*
- (2) *if player 2 does not acquire information, he plays h if $c \in (0, c^*)$ and l if $c \in (c^*, \bar{c})$.*

When c belongs to the countable set of cutoffs, there are multiple equilibria. The proposition implies that if the cost is higher, then the reputation-building process is shorter, and hence reputation exploitation occurs more frequently.

7 Extensions and Related Literature

Costly discovery of history is a salient ingredient in repeated interactions with incomplete information. Reputation models provide a natural framework to study the connection between endogenous discovery of history and long-term strategic behavior of asymmetrically informed parties. As an initial step in this direction, our setting is perhaps the simplest in which the relevant issues can be addressed. We discuss several features of our model and also relate our model to the literature on bounded memory.

7.1 Other Information Acquisition Schemes

For tractability we have assumed that the short-lived players observe the recent periods in sequence. One may also consider a sampling procedure where the short-lived players observe a random subset of past transactions. While similar long-run reputation dynamics might still hold, the analysis certainly becomes cumbersome. Let us consider the simplest case in which each short-lived player randomly samples only one period. Using a similar example and argument in Section 4.2, it is not difficult to see that we cannot have a long-run convergence to the play of H or L , as before. However, it is no longer true that we can define stationary strategies on a finite state space, because any transaction in the past may be sampled by a short-lived player and as a result the long-lived player will condition his strategy on his entire history, and the number of H 's since the most recent L is no longer a sufficient statistic of a history. This poses a technical difficulty. Various *exogenous* sampling procedures have

been considered in social learning environments; see Ellison and Fudenberg (1995), Banerjee and Fudenberg (2004), and Smith and Sørensen (2008). Information decay in which sampling probability for past transactions decreases exponentially along the history is of particular interest to this literature which mainly concentrates on the possibility of long-run convergence. *Endogenous* sampling in a *repeated game with an informed player* would be an interesting extension. To characterize the long-run dynamics, one may want to search for an alternative tractable sufficient statistic of histories.

7.2 Multiple Types

Our model has two types. This captures the tradeoff between reputation building and exploitation in a clear way. A possible extension is multiple types. We conjecture that the nature of reputation dynamics will not change qualitatively if in addition there is a third type who always plays the cheating action L . In this case, however, a history with straight cheating actions is of particular interest. It contains different information from a history with index 0. As a result, the space \mathbf{I} has to be enlarged to incorporate “cheating history indices.” One may also consider types that commit to mixed strategies. This introduces imperfect monitoring into the model. This case is harder unless a mixed strategy assigns a very high probability to a pure action. In general, a sufficient statistic for histories should capture the proportion of high actions on top of the clean history index introduced in this paper.

7.3 Observability

Our reputation model is clean in the sense that information acquisition incentives are created by uncertainty about types. This is demonstrated by the no information acquisition result in the degenerate complete information case. In our model, the long-lived player has the ability to clean his past history, and we have assumed away the possibility that short-lived players transmit information using their own actions. Unobservability of short-lived players’ actions is valid in many applications (see Footnote 9).

If, instead, the history of short-lived players is observable, then the short-lived players could coordinate their actions. We first look at the degenerate complete information case. Consider the game in Section 4.1. The repetition of the static Nash equilibrium is still an equilibrium. It is easy to check that the following trigger strategy profile is also an equilibrium: on the equilibrium path, the profile (H, h) is played; if anyone plays differently in the previous period, (L, l) is played forever. This trigger strategy profile is still an equilibrium even when $\mu_0 > 0$. The reasoning in Section 4.1 does not work here. Once player 1 plays L , he is not able to

cover it up by playing H because player 2 will switch to l , which tells the next player 2 that L has been played before. In general, since player 2's action is a function of his observation, the current player 2 will infer something (from the previous player 2) not directly observable to him.

When $C(1)$ is positive but small, we can construct the following equilibrium for $\mu_0 \geq 0$ based on the trigger strategy equilibrium above provided that δ is large.

Player 1 randomizes between playing L forever and playing H forever, with probability $\frac{C(1)}{(1-\mu_0)[u_2(L,l)-u_2(L,h)]}$ on the former; the first short-lived player plays h ; subsequent short-lived players randomize between acquiring 1 period of information and not acquiring information at all, with probability $\frac{u_1(L,h)-u_1(H,h)}{\delta[u_1(L,h)-u_1(L,l)]}$ on the former; they play h if they do not acquire information; after acquiring information, they play l unless both H and h are played.

It is routine algebra to verify that this is an equilibrium. Player 2's randomization probability is chosen so that player 1 is initially indifferent between H and L .²⁵ In this equilibrium, player 2 acquires information to find out the outcome of player 1's initial randomization.

7.4 Related Literature on Bounded Memory

Our result on the degenerate case is related to the literature on bounded recall in repeated games with complete information, especially the anti-folk theorems obtained in games with *two long-lived players* by Jéheil (1995), Bhaskar (1998), and Cole and Kocherlakota (2005). Bhaskar and Vega-Redondo (2002) also provide a foundation for Markov equilibria with memory cost. Their setups and assumptions are very different from ours.

Our main result on reputation dynamics is related to the recent literature on bounded memory with incomplete information. Wilson (2002) proposed a model of biased information processing with a finite automaton. A non-Bayesian decision maker receives a sequence of exogenous signals and updates his belief using a given set of memory states. The decision maker chooses a transition rule for going from one memory state to another, knowing that he will forget his current memory in the future. Our model differs in several respects. First, we have a repeated game instead of a decision problem, and the signals in our model are the outcomes of strategic play. Second, agents in our model are rational and Bayesian. The fact that uninformed agents are sometimes exploited is not a consequence of bounded rationality, but the maneuver of the informed long-lived agent. The Bayesian agents in our model have an unlimited information-processing capacity, but are constrained on data. They acquire

²⁵If player 1 decides to take L initially, he has no incentive to take H afterwards because player 2 will find out the initial L with positive probability and therefore switch to l ; this l will reveal to the next player 2 that a deviation has occurred. If player 1 decides to take H initially, he has no incentive to take L , which is absorbing according to the prescribed strategy.

data optimally by taking into account the stake in a transaction, and the fact that data are manipulated.

Ekmecki (2010) takes a mechanism design approach to study *payoff* effects of reputation. A third party can use a finite automaton (a rating system) to record information in the spirit of Wilson (2002). He shows that there exists a carefully designed automaton and an equilibrium with high payoffs for the long-lived agent in any history. This model shares with our model the feature that the observations of short-lived agents are restricted. However, in our model, there is a real tradeoff between information cost and the payoff from the game, and the short-lived agents optimize on information acquisition. Our focus is on how this tradeoff leads to different equilibrium *behavior without third-party interventions*, and we demonstrate that this tradeoff endogenously determines the reputation-building process *and* information acquisition.

Wilson and Ekmecki (2006) combine the ideas of Wilson (2002) and Ekmecki (2010), obtaining a high payoff equilibrium with the assumption that consumers could use a finite automaton to update their beliefs and also truthfully report their memory states to future consumers. Apart from the differences addressed previously regarding the two papers, in our model short-lived agents act independently and a free-riding problem naturally arises. Monte (2010) studies reputation effects in zero-sum games with two long-lived agents and the uninformed agent has a finite memory in the spirit of Wilson (2002).

Liu and Skrzypacz (2010) consider a model with exogenous finite memory and establish a payoff bound uniformly for *all* perfect Bayesian equilibria after any history, not only for a particular equilibrium. Without the complication of endogenous information acquisition, the paper also derives sharper behavior characterizations for a larger class of reputation games. Ekmecki and Wilson (2009) obtain a uniform payoff bound in a model of switching types à la Phelan (2006).

Appendix

A Preliminaries

Recall that on space S , player 1's strategy $\sigma_1 : S \rightarrow [0, 1]$ assigns a probability $\sigma_1(s)$ on action H at state $s \in S$; player 2's strategy after acquiring n periods of information is a measurable function $\sigma_2^n : S \rightarrow [0, 1]$ that assigns a probability $\sigma_2^n(s)$ on action h at state $s \in S$; player 2's expected action at state s is $\bar{\sigma}_2(s) = \sum \alpha(n)\sigma_2^n(s)$. If player 2 plays h with probability $\bar{\sigma}_2(s)$ in expectation, and player 1 plays pure strategy H , we write player 1's expected payoff at state s as $u_1(H, \bar{\sigma}_2(s))$ which is obtained by the usual multilinear extension of an expected utility

function. Similarly, we write $u_1(L, \bar{\sigma}_2(s))$ when player 1 plays L , and $u_1(\sigma_1(s), \bar{\sigma}_2(s))$ when player 1 plays a mixed strategy denoted by $\sigma_1(s)$. We shall follow the same convention when the state space is \mathbf{I} .

Recall that $u_1(L, h) - u_1(H, h) > u_1(L, l) - u_1(H, l)$. This property extends to player 2's mixed strategies. We summarize this observation below for later reference.

Lemma 4. *If $\bar{\sigma}_2(s) > \bar{\sigma}_2(s')$, then*

$$u_1(L, \bar{\sigma}_2(s)) - u_1(H, \bar{\sigma}_2(s)) > u_1(L, \bar{\sigma}_2(s')) - u_1(H, \bar{\sigma}_2(s')).$$

B Proof of Theorem 1

Proof. The proof generalizes the idea used in Section 4.1. We show inductively that the uninformed player 2's equilibrium play does not depend on the information he acquired if any. As player 1's play has no effect on player 2's action choices, player 1 will take his myopic optimal action L , and player 2 should respond with l . Since there is no uncertainty, there is no information acquisition for positive costs.

By definition, N periods of information is the most that can be paid for. We show that data of the N th period in the past is useless; i.e., player 2's play does not depend on data of that period. Suppose that the claim is not true. Then it must be that player 2 acquires N periods of information with positive probability and player 2 plays differently at two histories s and s' that differ only in the N th period in the past, $\sigma_2^N(s) \neq \sigma_2^N(s')$. Assume without loss of generality that $\sigma_2^N(s) > \sigma_2^N(s')$ (the other case is symmetric). Hence

$$\bar{\sigma}_2(s) > \bar{\sigma}_2(s'). \tag{5}$$

Since $\sigma_2^N(s) > 0$, i.e., player 2 plays h with positive probability at s , it follows that player 1 plays H with positive probability at this history. The IC condition for player 1 is therefore

$$(1 - \delta) u_1(H, \bar{\sigma}_2(s)) + \delta U_1(s \wedge H) \geq (1 - \delta) u_1(L, \bar{\sigma}_2(s)) + \delta U_1(s \wedge L), \tag{6}$$

where $s \wedge H$ denotes the history of length N after H is played at history s .

Since $\sigma_2^N(s') < 1$, i.e., player 2 plays l with a positive probability at s' , it must be that player 1 plays L with a positive probability at this history. The IC condition for player 1 is

$$(1 - \delta) u_1(L, \bar{\sigma}_2(s')) + \delta U_1(s' \wedge L) \geq (1 - \delta) u_1(H, \bar{\sigma}_2(s')) + \delta U_1(s' \wedge H). \tag{7}$$

Note that since s and s' differ only in the play of the N th period in the past, after a play of the same action at s and s' , player 2 cannot distinguish the continuation finite histories of length N : $s \wedge L = s' \wedge L$ and $s \wedge H = s' \wedge H$. Hence the continuation payoff is the same: $U_1(s \wedge L) = U(s' \wedge L)$ and $U_1(s \wedge H) = U(s' \wedge H)$. Adding the above two equalities together, we have

$$u_1(L, \bar{\sigma}_2(s')) - u_1(H, \bar{\sigma}_2(s')) \geq u_1(L, \bar{\sigma}_2(s)) - u_1(H, \bar{\sigma}_2(s)). \quad (8)$$

This contradicts Lemma 4. Therefore, $\sigma_2^N(s) = \sigma_2^N(s')$.

We have just shown that data of the N th period in the past is useless. Suppose this property holds for data of periods $N, N-1, \dots, k+1$ in the past, for $k > 0$. We shall show that this property holds for k th period data as well. Suppose that this is not true. Then without loss of generality assume that $\sigma_2^k(s) > \sigma_2^k(s')$ for two histories s and s' that differ only in the k th period in the past and player 2 acquires at least k periods of data with positive probability. Therefore, $\bar{\sigma}_2(s) > \bar{\sigma}_2(s')$. Applying previous arguments line by line, we again obtain inequalities (6) and (7). Now, the continuation histories $s \wedge L$ and $s' \wedge L$ differ only in period $k+1$. The same is true for $s \wedge H$ and $s' \wedge H$. Note that by the induction hypothesis, player 2's action does not depend on the play in periods $N, N-1, N-2, \dots, k+1$ in the past. Hence, $\bar{\sigma}_2(s \wedge L) = \bar{\sigma}_2(s' \wedge L)$, and $\bar{\sigma}_2(s \wedge H) = \bar{\sigma}_2(s' \wedge H)$; furthermore, $U_1(s \wedge L) = U(s' \wedge L)$ and $U_1(s \wedge H) = U(s' \wedge H)$. Adding the two inequalities (6) and (7), we again obtain inequality (8) and a contradiction. ■

C Proof of Theorem 2

We proceed according to the three steps outlined in the main text.

C.1 Unique Ergodic Set

The linear transition on \mathbf{I} implies that an ergodic set induced by player 1's strategy σ_1 must consist of consecutive integers. We shall show that the unique ergodic set is of the form $\{0, 1, \dots, n^*\}$ with $n^* > 0$ and $\sigma_2(n^*) = 0$. This result is summarized below.

Proposition 3. *Suppose $\delta > \bar{\delta}$. For each equilibrium $(\mathbf{I}, \sigma_1, \alpha, \sigma_2, \lambda)$, there exists an n^* , $0 < n^* \leq N$, such that (1) $\sigma_1(n^*) = 0$ and $\sigma_1(i) > 0$ if $0 \leq i < n^*$, and (2) states higher than n^* are off the equilibrium path.*

We first show that it is impossible that $\sigma_1(N) > 0$. Let $n_1 := \max\{n : p(n) > 0\}$. Because $C(1) < \mu_0[u_2(H, h) - u_2(H, l)]$, we have $n_1 > 0$.

Lemma 5. *If player 1 weakly prefers to play H in some state $i \in \{n_1, \dots, N\}$, then player 1 must weakly prefer to play H in state N .*

Proof. Suppose to the contrary that player 1 strictly prefers to play L in state N . Consider

$$i^* := \max\{i : \text{player 1 weakly prefers to play } H \text{ in state } i\}.$$

Note that $n_1 \leq i^* < N$. In state $i > i^*$, player 1 strictly prefers to play L , and hence

$$U_1(i) = (1 - \delta)u_1(L, \bar{\sigma}_2(n)) + \delta U_1(0) \quad (9)$$

is constant. In state $i > i^*$,

$$U_1(i) = (1 - \delta)u_1(L, \bar{\sigma}_2(n)) + \delta U_1(0) > (1 - \delta)u_1(H, \bar{\sigma}_2(n_1)) + \delta U_1(\min\{i + 1, N\}). \quad (10)$$

In state i^* ,

$$U_1(i^*) = (1 - \delta)u_1(H, \bar{\sigma}_2(n_1)) + \delta U_1(i^* + 1) \geq (1 - \delta)u_1(L, \bar{\sigma}_2(n)) + \delta U_1(0). \quad (11)$$

Adding the two, we have

$$U_1(i^* + 1) > U_1(\min\{i + 1, N\}),$$

which contradicts (9). ■

Lemma 6. *If player 1 weakly prefers to play H in some state $i \in \{n_1, \dots, N\}$, then for each $j \in \{n_1, \dots, N\}$, $\bar{\sigma}_2(j) = \bar{\sigma}_2(n_1)$, $U_1(j) = u_1(H, \bar{\sigma}_2(n_1))$ and player 1 weakly prefers to play H in state j .*

Proof. By the definition of n_1 , $\bar{\sigma}_2(j) = \bar{\sigma}_2(n_1)$ if $n_1 \leq j \leq N$. We prove the rest by induction. By Lemma 5, player 1 weakly prefers H in state N and $U_1(N) = u_1(H, \bar{\sigma}_2(n_1))$. Assume $U_1(j) = u_1(H, \bar{\sigma}_2(n_1))$ and player 1 weakly prefers H in state j for $j = k + 1, \dots, N$, where $k \geq n_1$. In these states,

$$U_1(j) = u_1(H, \bar{\sigma}_2(n_1)) \geq (1 - \delta)u_1(L, \bar{\sigma}_2(n)) + \delta U_1(0). \quad (12)$$

Consider the state k . If $\sigma_1(k) < 1$,

$$u_1(H, \bar{\sigma}_2(n_1)) \leq (1 - \delta)u_1(L, \bar{\sigma}_2(n)) + \delta U_1(0) = U_1(k). \quad (13)$$

By comparing (13) and (12), we conclude that the equality holds in (13). Therefore, player

1 weakly prefers H in state k and $U_1(k) = u_1(H, \bar{\sigma}_2(n_1))$. The conclusion is immediate if $\sigma_1(k) = 1$. ■

Lemma 7. *If player 1 weakly prefers to play H in some state $i \in \{n_1, \dots, N\}$, then for each $j \in \{0, 1, \dots, n_1\}$, $\bar{\sigma}_2(j) = \bar{\sigma}_2(n_1)$, $U_1(j) = u_1(H, \bar{\sigma}_2(n_1))$, and player 1 weakly prefers to play H in state j .*

Proof. We prove by induction again. When $i = n_1$, the conclusion follows from Lemma 6. Suppose the conclusion holds for $j = k + 1, \dots, n_1$, $k < n_1$. Consider $j = k$. We first establish the following property:

$$\bar{\sigma}_2(k) \leq \bar{\sigma}_2(k + 1) = \bar{\sigma}_2(n_1). \quad (14)$$

To see this, suppose to the contrary that $\bar{\sigma}_2(k) > \bar{\sigma}_2(n_1)$. By playing L or H for one period in state k and then following the equilibrium strategy, player 1 gets the expected payoff

$$r_k(L) = (1 - \delta)u_1(L, \bar{\sigma}_2(k)) + \delta U_1(0), \text{ or}$$

$$r_k(H) = (1 - \delta)u_1(H, \bar{\sigma}_2(k)) + \delta u_1(H, \bar{\sigma}_2(n_1)),$$

respectively. Analogously for state $k + 1$, we have

$$r_{k+1}(L) = (1 - \delta)u_1(L, \bar{\sigma}_2(n_1)) + \delta U_1(0),$$

$$r_{k+1}(H) = (1 - \delta)u_1(H, \bar{\sigma}_2(n_1)) + \delta u_1(H, \bar{\sigma}_2(n_1)).$$

Let us consider

$$\Delta = [r_{k+1}(L) - r_{k+1}(H)] - [r_k(L) - r_k(H)]. \quad (15)$$

Simple computation shows that

$$\Delta/(1 - \delta) = [u_1(L, \bar{\sigma}_2(n_1)) - u_1(H, \bar{\sigma}_2(n_1))] - [u_1(L, \bar{\sigma}_2(k)) - u_1(H, \bar{\sigma}_2(k))]. \quad (16)$$

By our supposition $\bar{\sigma}_2(k) > \bar{\sigma}_2(n_1)$. It follows from Lemma 4 that $\Delta < 0$. Since player 1 weakly prefers to play H in state $k + 1$ by the induction hypothesis, $r_{k+1}(L) \leq r_{k+1}(H)$. We have two cases to consider.

(i) $r_{k+1}(L) = r_{k+1}(H)$.

Then, $r_k(L) > r_k(H)$ by (15), i.e., $\sigma_1(k) = 1$. Thus, $\sigma_2^n(k) = 0$ for $n \geq k + 1$. Therefore,

$$\begin{aligned}\bar{\sigma}_2(k) &= \sum_{n \leq k} \alpha(n) \sigma_2^n(k) + \sum_{n \geq k+1} \alpha(n) 0 = \sum_{n \leq k} \alpha(n) \sigma_2^n(k+1) + \sum_{n \geq k+1} \alpha(n) 0 \\ &\leq \sum_{n \leq k} \alpha(n) \sigma_2^n(k+1) + \sum_{n \geq k+1} \alpha(n) \sigma_2^n(k+1) \\ &= \bar{\sigma}_2(k+1).\end{aligned}$$

The second equality follows because player 2 sees straight H 's after acquiring n periods of information in states k and $k + 1$, $n \leq k$. But, we have derived $\bar{\sigma}_2(n_1) = \bar{\sigma}_2(k+1) \geq \bar{\sigma}_2(k)$, which contradicts the assumption.

(ii) $r_{k+1}(L) < r_{k+1}(H)$.

By the induction hypothesis, $\bar{\sigma}_2(m) = \bar{\sigma}_2(n_1) = \bar{\sigma}_2(k+1)$ if $m > k$ and hence $r_m = r_{k+1}$. Thus $r_m(L) < r_m(H)$ and $\sigma_1(m) = 1$ if $m > k$. Therefore, player 2 will play h for sure whenever he buys $k + 1$ or more periods of information and then sees straight H 's. But, we will have a contradiction:

$$\begin{aligned}\bar{\sigma}_2(k) &= \sum_{n \leq k} \alpha(n) \sigma_2^n(k) + \sum_{n \geq k+1} \alpha(n) \sigma_2^n(k) \\ &\leq \sum_{n \leq k} \alpha(n) \sigma_2^n(k+1) + \sum_{n \geq k+1} \alpha(n) 1 \\ &= \sum_{n \leq k} \alpha(n) \sigma_2^n(k+1) + \sum_{n \geq k+1} \alpha(n) \sigma_2^n(k+1) \\ &= \bar{\sigma}_2(k+1).\end{aligned}$$

The property (14) is established.

Now, let us continue the main induction on j . We have two cases to consider when $j = k$.

Case 1: $\sigma_1(k) = 1$.

Then, $\sigma_2^n(k) = 1$ for $n > k$, and hence $\bar{\sigma}_2(k) \geq \bar{\sigma}_2(n_1)$. It follows from Claim 2 above that $\bar{\sigma}_2(k) = \bar{\sigma}_2(n_1)$. Player 1 prefers H in state k since $\sigma_1(k) = 1$.

Case 2: $\sigma_1(k) < 1$.

Then, $r_k(L) - r_k(H) \geq 0$. If $\bar{\sigma}_2(n_1) > \bar{\sigma}_2(k)$, then $\Delta > 0$ by inspecting Equation (16). Thus, $r_{k+1}(L) - r_{k+1}(H) > 0$, which contradicts the induction hypothesis: player 1 weakly prefers H in state $k + 1$. Therefore, $\bar{\sigma}_2(k) = \bar{\sigma}_2(k)$ by the claim above and hence $\Delta = 0$ by inspecting Equation (16). Hence,

$$r_k(L) - r_k(H) = r_{k+1}(L) - r_{k+1}(H) \leq 0.$$

That is, player 1 weakly prefers H in state k .

To conclude, Cases 1 and 2 imply that $\sigma_{n_1} = \sigma_k$ and player 1 weakly prefers H in k . Hence,

$$U_1(k) = (1 - \delta)u_1(H, \bar{\sigma}_2(n_1)) + \delta U_1(k + 1) = u_1(H, \bar{\sigma}_2(n_1)).$$

The induction is complete. ■

Corollary 3. *Player 1 strictly prefers to play L in each state $i \in \{n_1, \dots, N\}$.*

Proof. If player 1 weakly prefers to play H in some state $i \in \{n_1, \dots, N\}$, then Lemmas 6 and 7 imply that player 2's expected behavioral strategies will be constant in all states in $\{0, 1, \dots, N\}$. Therefore player 1 should always strictly prefer to play the stage game dominant strategy L , a contradiction. ■

We now complete the proof of Proposition 3.

Proof of Proposition 3. If an equilibrium does *not* take the form of Proposition 3, there are three cases to consider.

- (1) $\{N\}$ is an ergodic set (there could be other ergodic sets). In this case $\sigma_1(N) = 1$.
- (2) $\{0, 1, \dots, N\}$ is the unique ergodic set and $\sigma_1(N) > 0$.
- (3) $\{0\}$ is the unique ergodic set. That is, player 1 plays L with probability 1 in state 0 and the probability of reaching state 0 from any other state is 1.

The first two cases contradict Corollary 3. The argument against case (3) is as follows. By Lemma 2, player 2 will strictly prefer to buy 1 period of information. If player 2 commits to buying 1 period of information and state 0 is absorbing, then player 1 ends up with $u_1(L, l)$. But, if he is patient enough, player 1 can “escape the trap” by playing H for 1 period, in which his average payoff is at least $u_1(H, l)$ and after which his average payoff is at least $u_1(H, h)$. The deviation is profitable if $(1 - \delta)u_1(H, l) + \delta u_1(H, h) > u_1(L, l)$. This holds when $\delta > \bar{\delta}$.²⁶ If $C(n) = C(1)$ for some $n > 1$, it is possible that player 2 buys more than 1 period of information. The analysis for this case is similar.

Since $\sigma_1(n_1) = 0$, there exists $n^* \in \{1, 2, \dots, n_1\}$ such that $\sigma_1(i) > 0$ for each $i \in \{0, 1, \dots, n^* - 1\}$ and $\sigma_1(n^*) = 0$. ■

C.2 Increasing Incentive

Lemma 8. $\bar{\sigma}_2(n^* - 1) \leq \bar{\sigma}_2(n^*)$ and $U_1(n^* - 1) \leq U_1(n^*)$.

$$\frac{26 \frac{u_1(L, h) - u_1(H, h)}{u_1(L, h) - u_1(L, l)} - \frac{u_1(L, l) - u_1(H, l)}{u_1(H, h) - u_1(H, l)}}{\frac{u_1(L, l) - u_1(H, l)}{u_1(H, h) - u_1(H, l)}} = \frac{[u_1(H, h) - u_1(L, l)][u_1(L, h) - u_1(H, h) - u_1(L, l) + u_1(H, l)]}{[u_1(H, h) - u_1(H, l)][u_1(L, h) - u_1(L, l)]} > 0. \quad \text{Thus, } \delta > \frac{u_1(L, l) - u_1(H, l)}{u_1(H, h) - u_1(H, l)} \text{ whenever } \delta > \frac{u_1(L, h) - u_1(H, h)}{u_1(L, h) - u_1(L, l)}.$$

Proof. Note that

$$U_1(n^* - 1) = (1 - \delta)u_1(H, \bar{\sigma}_2(n^* - 1)) + \delta U_1(n^*) \geq (1 - \delta)u_1(L, \bar{\sigma}_2(n^* - 1)) + \delta U_1(0),$$

$$U_1(n^*) = (1 - \delta)u_1(L, \bar{\sigma}_2(n^*)) + \delta U_1(0) \geq (1 - \delta)u_1(H, \bar{\sigma}_2(n^*)) + \delta U_1(\min\{n^* + 1, N\}).$$

Adding the two, we have

$$\begin{aligned} & \delta(U_1(n^*) - U_1(\min\{n^* + 1, N\})) \geq \\ & (1 - \delta)[u_1(L, h) - u_1(H, h) - u_1(L, l) + u_1(H, l)][\bar{\sigma}_2(n^* - 1) - \bar{\sigma}_2(n^*)]. \end{aligned} \quad (17)$$

Notice that $U_1(\min\{n^* + 1, N\}) \geq (1 - \delta)u_1(L, \bar{\sigma}_2(n^* + 1)) + \delta U_1(0)$ and $U_1(n^*) = (1 - \delta)u_1(L, \bar{\sigma}_2(n^*)) + \delta U_1(0)$. From (17),

$$\begin{aligned} & \delta(u_1(L, h) - u_1(L, l))(\bar{\sigma}_2(n^*) - \bar{\sigma}_2(\min\{n^* + 1, N\})) \geq \\ & (1 - \delta)[u_1(L, h) - u_1(H, h) - u_1(L, l) + u_1(H, l)][\bar{\sigma}_2(n^* - 1) - \bar{\sigma}_2(n^*)]. \end{aligned} \quad (18)$$

We claim that $\bar{\sigma}_2(n^*) \leq \bar{\sigma}_2(\min\{n^* + 1, N\})$. The argument is as follows. If $n^* < n_1$, then $\sigma_1(n^*) = 0$, and hence $\sigma_2^k(n^*) = 0 \leq \sigma_2^k(n^* + 1)$ when $k > n^*$. If $n^* = n_1$, then $\bar{\sigma}_2(n^*) = \bar{\sigma}_2(\min\{n^* + 1, N\})$. Therefore, $\bar{\sigma}_2(n^* - 1) \leq \bar{\sigma}_2(n^*)$ from (18).

Since $\sigma_1(n^*) = 0$, $\sigma_2^k(n^*) = 0 \leq \sigma_2^k(\min\{n^* + m, N\})$ for any $k > n^*$ and $m > 0$. Therefore, $\bar{\sigma}_2(\min\{n^* + m, N\}) \geq \bar{\sigma}_2(n^*) \geq \bar{\sigma}_2(n^* - 1)$ for any $m > 0$. In state n^* , player 1's payoff is at least the payoff from always playing H . That is,

$$\begin{aligned} U_1(n^*) & \geq (1 - \delta)[u_1(H, \bar{\sigma}_2(n^*)) + \delta u_1(H, \bar{\sigma}_2(\min\{n^* + 1, N\})) + \dots] \\ & \geq u_1(H, \bar{\sigma}_2(n^*)) \\ & \geq u_1(H, \bar{\sigma}_2(n^* - 1)). \end{aligned}$$

Therefore, $U_1(n^*) \geq (1 - \delta)u_1(H, \bar{\sigma}_2(n^* - 1)) + \delta U_1(n^*) = U_1(n^* - 1)$. ■

Now we show that player 2's expected strategies must be weakly increasing.

Lemma 9. $\bar{\sigma}_2(i)$ must be weakly increasing in i , $i \in \{0, 1, \dots, n^*\}$.

Proof. Suppose $\bar{\sigma}_2(i)$ is not weakly increasing in i . Define $i^* = \max\{i : \bar{\sigma}_2(i - 1) > \bar{\sigma}_2(i)\}$. By the definition of i^* , $\bar{\sigma}_2(i^* - 1) > \bar{\sigma}_2(i^*)$. By Lemma 8, $i^* < n^*$. Since $\sigma_1(i)$ for each

$i \in \{0, 1, \dots, n^* - 1\}$, we have

$$\left(\begin{array}{l} u_1(H, \bar{\sigma}_2(i^* - 1)) + \delta u_1(H, \bar{\sigma}_2(i^*)) + \dots \\ + \delta^{n^* - i^* - 1} u_1(H, \bar{\sigma}_2(n^* - 2)) + \delta^{n^* - i^*} \frac{U_1(n^* - 1)}{1 - \delta} \end{array} \right) \geq u_1(L, \bar{\sigma}_2(i^* - 1)) + \delta \frac{U_1(0)}{1 - \delta}. \quad (19)$$

Claim: $0 < \sigma_1(i^*) \leq \bar{\mu}$.

To see this, first note that $\sigma_1(i^*) > 0$ by Proposition 3. We suppose to the contrary that $\sigma_1(i^*) > \bar{\mu}$. Therefore, $\sigma_2^k(i^*) = 1$ for each $k > i^*$. By the definition of i^* , $\bar{\sigma}_2(i) \geq \bar{\sigma}_2(i^*)$ for each $i > i^*$. On the other hand, for $i > i^*$,

$$\begin{aligned} \bar{\sigma}_2(i) &= \sum_{k \leq i^*} \alpha(k) \sigma_2^k(i) + \sum_{k > i^*} p(k) \sigma_2^k(i) \\ &= \sum_{k \leq i^*} \alpha(k) \sigma_2^k(i^*) + \sum_{k > i^*} \alpha(k) \sigma_2^k(i) \\ &\leq \sum_{k \leq i^*} \alpha(k) \sigma_2^k(i^*) + \sum_{k > i^*} \alpha(k) \\ &= \sum_{k \leq i^*} \alpha(k) \sigma_2^k(i^*) + \sum_{k > i^*} \alpha(k) \sigma_2^k(i^*) \\ &= \bar{\sigma}_2(i^*). \end{aligned}$$

The second equality follows from the fact that $\sigma_2^k(i) = \sigma_2^k(i^*)$ when $i > i^* \geq k$. Therefore, if $\bar{\sigma}_1(i^*) > \bar{\mu}$, then

$$\bar{\sigma}_1(i) = \bar{\sigma}_1(i^*) \text{ for each } i > i^*. \quad (20)$$

In particular, $\bar{\sigma}_1(n^*) = \bar{\sigma}_1(i^*)$. This is impossible when $n^* < n_1$, since $\sigma_2^k(i^*) = 1$ for $k > i^*$ but $\sigma_2^{n_1}(n^*) = 0$. Let us consider the case $n^* = n_1$. In state $n^* - 1$,

$$U_1(n^* - 1) = (1 - \delta)u_1(H, \bar{\sigma}_2(n^* - 1)) + \delta U_1(n^*) \geq (1 - \delta)u_1(L, \bar{\sigma}_2(n^* - 1)) + \delta U_1(0).$$

In state $n^* = n_1$, by Corollary 3,

$$U_1(n^*) = (1 - \delta)u_1(L, \bar{\sigma}_2(n^*)) + \delta U_1(0) > (1 - \delta)u_1(H, \bar{\sigma}_2(n^*)) + \delta U_1(n^* + 1).$$

Therefore,

$$\begin{aligned} 0 &= \delta(U_1(n^*) - U_1(n^* + 1)) \\ &> (1 - \delta)[u_1(L, h) - u_1(H, h) - u_1(L, l) + u_1(H, l)][\bar{\sigma}_2(n^* - 1) - \bar{\sigma}_2(n^*)]. \end{aligned} \quad (21)$$

This will lead to $\bar{\sigma}_2(n^*) > \bar{\sigma}_2(n^* - 1)$, contradicting Equation (20). Therefore, $0 < \sigma_1(i^*) \leq \bar{\mu}$. The claim is established.

From this claim, we have

$$\left(\begin{array}{l} u_1(H, \bar{\sigma}_2(i^*)) + \delta u_1(H, \bar{\sigma}_2(i^* + 1)) + \dots \\ + \delta^{n^* - i^* - 1} u_1(H, \bar{\sigma}_2(n^* - 1)) + \delta^{n^* - i^*} \frac{U_1(n^*)}{1 - \delta} \end{array} \right) = u_1(L, \bar{\sigma}_2(i^*)) + \delta \frac{U_1(0)}{1 - \delta}. \quad (22)$$

From (19), we have

$$\begin{aligned} & [u_1(L, h) - u_1(H, h) - u_1(L, l) + u_1(H, l)][\bar{\sigma}_2(i^* - 1) - \bar{\sigma}_2(i^*)] \\ & \leq \sum_{i=i^*}^{n^* - 2} \delta^{i - i^* + 1} (u_1(H, \bar{\sigma}_2(i)) - u_1(H, \bar{\sigma}_2(i + 1))) + \delta^{n^* - i^*} \frac{U_1(n^* - 1) - U_1(n^*)}{1 - \delta}. \end{aligned} \quad (23)$$

For $i \geq i^*$, $u_1(H, \bar{\sigma}_2(i)) - u_1(H, \bar{\sigma}_2(i + 1)) \leq 0$ since $u_1(H, h) > u_1(H, l)$ and $\bar{\sigma}_2(i) \leq \bar{\sigma}_2(i + 1)$. Note that $U_1(n^* - 1) \leq U_1(n^*)$ by Lemma 8. Thus, the right-hand side of (23) is non-positive. Therefore,

$$[u_1(L, h) - u_1(H, h) - u_1(L, l) + u_1(H, l)][\bar{\sigma}_2(i^* - 1) - \bar{\sigma}_2(i^*)] \leq 0.$$

It follows from $u_1(L, h) - u_1(H, h) > u_1(L, l) - u_1(H, l)$ that $\bar{\sigma}_2(i^* - 1) \leq \bar{\sigma}_2(i^*)$, contradicting the definition of i^* . Therefore, $\bar{\sigma}_2(i)$ is non-decreasing in i when $i \in \{0, 1, \dots, n^*\}$. This completes the proof of monotonicity. ■

C.3 Uniform Bound on Player 1's Mixing Probabilities

Lemma 10. *If $i \in \{0, 1, \dots, n^* - 1\}$, then $0 < \sigma_1(i) \leq \bar{\mu}$.*

Proof. Suppose to the contrary that $\sigma_1(i) > \bar{\mu}$ for some $i < n^*$. Let $i^* := \min\{i : \sigma_1(i) > \bar{\mu}\}$. By definition $i^* < n^*$. Therefore, $\sigma_2^m(i^*) = 1$ for $m > i^*$. It follows from Lemma 9 that $\bar{\sigma}_2(i^*) = \bar{\sigma}_2(i^* + 1) = \dots = \bar{\sigma}_2(n^*)$. The argument between Equation (20) and Inequality (21) then leads to the required contradiction. ■

Summary of Sections C.1–C.3. It follows from Lemma 10 that n^* defined in Proposition 3 satisfies all the requirements of Theorem 2. The proof of Theorem 2 is now complete.

D Proof of Lemma 3

Proof. In Section C.2, we have shown that $\bar{\sigma}_2(i)$ is weakly increasing on $\{0, 1, \dots, n^*\}$. Suppose there is a $j \in \{0, 1, \dots, n^* - 1\}$ such that $\bar{\sigma}_2(j) = \bar{\sigma}_2(j + 1)$. By Lemma 10,

$$U_1(j) = (1 - \delta)u_1(L, \bar{\sigma}_2(j)) + \delta U_1(0) = U_1(j + 1).$$

We claim that $U_1(j) > u_1(H, \bar{\sigma}_2(j))$. To see this, notice that (1) $\bar{\sigma}_2(j) \leq \bar{\sigma}_2(j + k)$ for each k (if $j + k \leq n^*$, the result follows from monotonicity; otherwise, $\bar{\sigma}_2(j) \leq \bar{\sigma}_2(n^*) \leq \bar{\sigma}_2(n^* + k)$ for each k), (2) in state j , player 1 can achieve at least a payoff from playing H forever, and (3) in a state greater than or equal to n_1 , player 1 strictly prefers to play L . Thus,

$$U_1(j) > (1 - \delta)(u_1(H, \bar{\sigma}_2(j)) + \delta u_1(H, \bar{\sigma}_2(\min\{j + 1, N\})) + \dots) \geq u_1(H, \bar{\sigma}_2(j)).$$

Hence, $U_1(j) > (1 - \delta)u_1(H, \bar{\sigma}_2(j)) + \delta U_1(j + 1)$. This contradicts the fact that player 1 is indifferent in state j . Therefore, $\bar{\sigma}_2(i)$ is strictly increasing. It follows immediately that $U_1(j) = (1 - \delta)u_1(L, \bar{\sigma}_2(j)) + \delta U_1(0)$ is also strictly increasing. ■

E Proof of Theorem 3

We shall first prove properties (2) and (3), as they do not require generic cost functions.

E.1 Proof of Properties (2) and (3)

Let n_1 and n_2 be the largest and the second largest amounts of information that are ever acquired in equilibrium, respectively. The next two lemmas investigate the relationships of n^* , n_1 and n_2 .

Lemma 11. *If the cost function is strictly increasing and $\delta > \bar{\delta}$, then $n^* = n_2$ or $n^* = n_1$.*

Proof. Suppose $n_2 < n^* < n_1$. Since $\sigma_1(n^*) = 0$ and $n^* < n_1$, we have $\sigma_2^{n_1}(n^*) = 0$. Note that $\sigma_2^{n_2}(n^* - 1) = \sigma_2^{n_2}(n^*)$ by $n_2 < n^*$ and $\bar{\sigma}_2(n^* - 1) < \bar{\sigma}_2(n^* - 1)$ (Lemma 3), and hence $\sigma_2^{n_2}(n^* - 1) < 0$. This is impossible. Furthermore, note that $n^* < n_2$ is impossible because if $\sigma_1(n^*) = 0$, then states $\{n_2, n_2 + 1, \dots, N\}$ are off the equilibrium path (Proposition 3). Hence player 2 needs only to buy at most n_2 periods of information. This contradicts the definition of n_1 . ■

Lemma 12. *If $n^* = n_2$, then $n_2 = n_1 - 1$.*

Proof. If $n^* = n_2 < n_1 - 1$, then states $n_2 + 1, \dots, n_1$ are off the equilibrium path. Player 2 will be better off buying $n^* = n_2 + 1 < n_1$ periods of information. ■

Part (2) and (3) of Theorem 3 are immediate implications of Lemma 11 and the following.

Proposition 4. *If $C(\cdot)$ is strictly increasing and $\delta > \bar{\delta}$, then $\alpha(i) > 0$ and $\sigma_1(i) < \bar{\mu}$ for each $i \in \{0, 1, \dots, n_1\}$. Upon acquiring information, player 2 plays pure strategy l whenever he sees an L and he plays pure strategy h whenever he sees straight H 's.*

Proof. $\alpha(n_1) > 0$ by definition and $\sigma_1(n_1) = 0$ by Theorem 2. Since $\bar{\sigma}_2(i)$ is strictly increasing in i , we have $\bar{\sigma}_2(n_1) > \bar{\sigma}_2(n_1 - 1)$. That is, $\sum \alpha(k)\sigma_2^k(n_1) > \sum \alpha(k)\sigma_2^k(n_1 - 1)$. Since $\sigma_2^k(n_1) = \sigma_2^k(n_1 - 1)$ when $k \leq n_1 - 1$, we have $\sigma_2^{n_1}(n_1) > \sigma_2^{n_1}(n_1 - 1)$.

Claim 1: $\sigma_2^{n_1}(n_1) = 1$ and $\sigma_2^{n_1}(n_1 - 1) = 0$.

To see this, suppose to the contrary that one of them is in $(0, 1)$; i.e., player 2 plays a strictly mixed strategy in that state upon observing n_1 periods of information. Suppose that a_2 is played with positive probability in both states n_1 and $n_1 - 1$, $a_2 \in \{h, l\}$. Then, we can construct a profitable deviation $(\alpha, \tilde{\sigma}_2)$ for player 2 as follows. Set $\alpha(n_1 - 1) := \alpha(n_1 - 1) + \alpha(n_1)$ and $\alpha(k) = \alpha(k)$ when $k < n_1 - 1$. That is, player 2 buys $n_1 - 1$ instead of n_1 periods of information. Set $\tilde{\sigma}_2^{n_1-1}(n_1) = \tilde{\sigma}_2^{n_1-1}(n_1 - 1) = 1$ if $a_2 = h$ and $\tilde{\sigma}_2^{n_1-1}(n_1) = \tilde{\sigma}_2^{n_1-1}(n_1 - 1) = 0$ if $a_2 = l$. Set $\tilde{\sigma}_2^{n_1-1}(i) := \sigma_2^{n_1}(i)$ if $i < n_1 - 1$. Let $\tilde{\sigma}_2^k := \sigma_2^k$ if $k \neq n_1, n_1 - 1$. $\tilde{\sigma}_2^{n_1-1}(n_1)$ and $\tilde{\sigma}_2^{n_1-1}(n_1 - 1)$ are optimal since a_2 is optimal in both states n_1 and $n_1 - 1$. The optimality of $\tilde{\sigma}^k$ follows from the optimality of σ^k for $k \neq n_1, n_1 - 1$. By using this alternative strategy, player 2 achieves the same payoff as before but saves $(C(n_1) - C(n_1 - 1))\alpha(n_1)$. We have derived a contradiction. Therefore, $\sigma_2^{n_1}(n_1) = 1$ and $\sigma_2^{n_1}(n_1 - 1) = 0$.

Claim 2: $\sigma_1(n_1 - 1) < \bar{\mu}$.

To see this, suppose to the contrary that $\sigma_1(n_1 - 1) = \bar{\mu}$. Then, player 2 is indifferent after buying n_1 periods of information in state $n_1 - 1$. The identical argument that establishes Claim 1 also applies here.

We now prove the proposition by induction. Suppose that for $i = n_1, n_1 - 1, \dots, m + 1$, $m > 0$, we have $p(i) > 0$, $\sigma_1(i - 1) < \bar{\mu}$, $\sigma_2^i(i) = 1$ and $\sigma_2^k(i - 1) = 0$ if $k \geq i$.

Consider $i = m$. Suppose $\alpha(m) = 0$. Then

$$\begin{aligned} \bar{\sigma}_2(m) - \bar{\sigma}_2(m - 1) &= \sum_{k \geq m+1} \alpha(k)(\sigma_2^k(m) - \sigma_2^k(m - 1)) \\ &= - \sum_{k \geq m+1} \alpha(k)\sigma_2^k(m - 1) \\ &\leq 0. \end{aligned}$$

This is impossible because $\bar{\sigma}_2(m) > \bar{\sigma}_2(m - 1)$ by strict monotonicity. Therefore, $\alpha(m) \neq 0$

and

$$\begin{aligned}\bar{\sigma}_2(m) - \bar{\sigma}_2(m-1) &= \sum_{k \geq m} \alpha(k)(\sigma_2^k(m) - \sigma_2^k(m-1)) \\ &= \alpha(m)\sigma_2^m(m) - \sum_{k \geq m} \alpha(k)\sigma_2^k(m-1).\end{aligned}$$

It must be that $\sigma_2^m(m) = 1$, $\sigma_2^m(m-1) = 0$, and $\sigma_1(m-1) < \bar{\mu}$, by arguments similar to those used in the proof of Claims 1 and 2. Since $\sigma_1(m-1) < \bar{\mu}$, we have $\sigma_2^k(m-1) = 0$ if $k \geq m$. We therefore have that for each $i \in \{1, 2, \dots, n_1\}$, $\alpha(i) > 0$, $\sigma_1(i-1) < \bar{\mu}$, $\sigma_2^k(i-1) = 0$ if $k \geq i$ and $\sigma_2^i(i) = 1$.

We still need to show that $\alpha(0) > 0$. Suppose to the contrary $\alpha(0) = 0$. Then we will have $\bar{\sigma}_2(0) = 0$ and $\bar{\sigma}_2(n_1) = 1$. If $n^* = n_2$, then $n_2 = n_1 - 1$ by Lemma 12. Player 1 plays pure strategy L in state n_2 ,

$$\begin{aligned}(1 - \delta)u_1(H, \bar{\sigma}_2(n_2)) + \delta[(1 - \delta)u_1(L, \bar{\sigma}_2(n_1)) + \delta u_1(L, \bar{\sigma}_2(0))] \\ \leq (1 - \delta)u_1(L, \bar{\sigma}_2(n_2)) + \delta u_1(L, \bar{\sigma}_2(0)).\end{aligned}$$

Since $\bar{\sigma}_2(0) = 0$ and $\bar{\sigma}_2(n_1) = 1$, we have

$$\begin{aligned}\delta(u_1(L, h) - u_1(L, l)) &\leq u_1(L, \bar{\sigma}_2(n_2)) - u_1(H, \bar{\sigma}_2(n_2)) \\ &< u_1(L, h) - u_1(H, h).\end{aligned}$$

We reach an immediate contradiction if $\delta > \bar{\delta}$. The same argument shows that $n^* = n_1$ is impossible. ■

E.2 Proof of Property (1)

Proof. By Lemma 11, we only need to show that $n^* = n_2 \neq n_1$ is the non-generic case. Note that by Proposition 4, $n_1 = n_2 + 1$ if $n_1 \neq n_2$. By Proposition 4, we have two cases to consider.

Case 1: Player 2 plays l with positive probability when he does not acquire information.

If player 2 does not acquire information, then his payoff is

$$U_2^0 = \mu_0 u_2(H, l) + (1 - \mu_0) \sum_{j=0}^{n_1} \lambda(j) u_2(\sigma_1(j), l). \quad (24)$$

If player 2 acquires n_1 periods of information, we have

$$\begin{aligned}U_2^{n_1} &= \mu_0 u_2(H, h) + (1 - \mu_0) \sum_{j=0}^{n_1-1} \lambda(j) u_2(\sigma_1(j), l) \\ &\quad + (1 - \mu_0) \lambda(n_1) u_2(\sigma_1(n_1), h) - C(n_1).\end{aligned}$$

Since $U_2^0 = U_2^{n_1}$, we have

$$\mu_0(u_2(H, h) - u_2(H, l)) - (1 - \mu_0)\lambda(n_1)(u_2(L, l) - u_2(L, h)) = C(n_1).$$

If $n^* = n_2 \neq n_1$, $\lambda(n_1) = 0$. Then, $C(n_1) = \mu_0[u_2(H, h) - u_2(H, l)]$. This is clearly non-generic.

Case 2: Player 2 plays h when he does not acquire information.

If player 2 acquires i periods of information, $0 \leq i \leq n_1$,

$$\begin{aligned} U_2^i &= \mu_0 u_2(H, h) + (1 - \mu_0) \sum_{j=0}^{i-1} \lambda(j) u_2(\sigma_1(j), l) \\ &\quad + (1 - \mu_0) \sum_{j=i}^{n_1} \lambda(j) u_2(\sigma_1(n_1), h) - C(i). \end{aligned}$$

From $U_2^0 = \dots = U_2^{n_1}$, we have n_1 equations:

$$(1 - \mu_0)\lambda(i)[u_2(\sigma_1(i), l) - u_2(\sigma_1(i), h)] = C(i+1) - C(i), \quad 0 \leq i \leq n_1 - 1. \quad (25)$$

From the definition of invariant distribution, we have $n_1 + 1$ equations:

$$\lambda(i+1) = \lambda(i)\sigma_1(i) \text{ for } 0 \leq i \leq n_1, \quad (26)$$

$$\sum \lambda(i) = 1. \quad (27)$$

If $n^* = n_2 \neq n_1$, then $\sigma_1(n_2) = 0$. This together with $\sigma_1(n_1) = 0$ gives two initial conditions. Therefore, we have $n_1 + (n_1 + 1) + 2 = 2(n_1 + 1) + 1$ equations, but only $2(n_1 + 1)$ unknowns: $\sigma_1(i)$ and $\lambda(i)$ for $0 \leq i \leq n_1$.²⁷ This case is non-generic. ■

F Proof of Proposition 1

Proof. In both cases we considered in the proof in Section E.2, the discount factor does not enter the equations that characterize n^* , σ_1 and λ . The claim follows. ■

References

APERJIS, C., AND R. JOHARI (2010): “Optimal windows for aggregating ratings in electronic marketplaces,” *Forthcoming, Management Science*.

²⁷A simple computation by plugging (26) into (25) leads to a system of independent linear equations.

- AZAM, J., R. BATES, AND B. BIAIS (2009): “Political predation and economic development,” *Economics & Politics*, 21(2), 255–277.
- BANERJEE, A., AND D. FUDENBERG (2004): “Word-of-mouth learning,” *Games and Economic Behavior*, 46, 1–22.
- BAR-ISAAC, H. (2003): “Reputation and survival: Learning in a dynamic signalling model,” *Review of Economic Studies*, 70(2), 231–251.
- (2007): “Something to prove: reputation in teams,” *RAND Journal of Economics*, 38(2), 495–511.
- BAR-ISAAC, H., AND S. TADELIS (2008): “Seller reputation,” *Foundation and Trends in Microeconomics*, Now Publishers Inc.
- BENABOU, R., AND G. LAROQUE (1992): “Using privileged information to manipulate markets: Insiders, gurus, and credibility,” *The Quarterly Journal of Economics*, 107(3), 921–958.
- BHASKAR, V. (1998): “Informational constraints and the overlapping generations model: Folk and anti-folk theorems,” *Review of Economic Studies*, 65(1), 135–149.
- BHASKAR, V., AND F. VEGA-REDONDO (2002): “Asynchronous choice and Markov equilibria,” *Journal of Economic Theory*, 103(2), 334–350.
- BUTLER, R. (2006): *The tourism area life cycle: Applications and modifications*. Cromwell Press.
- COLE, H., AND N. KOCHERLAKOTA (2005): “Finite memory and imperfect monitoring,” *Games and Economic Behavior*, 53(1), 59–72.
- CRIPPS, M., G. MAILATH, AND L. SAMUELSON (2004): “Imperfect monitoring and impermanent reputations,” *Econometrica*, 72(2), 407–432.
- EKMEKCI, M. (2010): “Sustainable reputations with rating systems,” *Forthcoming, Journal of Economic Theory*.
- EKMEKCI, M., AND A. WILSON (2009): “Maintaining a permanent reputation with replacements,” *working paper, Northwestern University*.
- ELLISON, G., AND D. FUDENBERG (1995): “Word-of-mouth communication and social learning,” *The Quarterly Journal of Economics*, 110(1), 93–125.

- FELDMAN, M., K. LAI, I. STOICA, AND J. CHUANG (2004): “Robust incentive techniques for peer-to-peer networks,” in *Proceedings of the 5th ACM Conference on Electronic Commerce*, pp. 102–111.
- FUDENBERG, D., D. KREPS, AND E. MASKIN (1990): “Repeated games with long-run and short-run players,” *The Review of Economic Studies*, 57(4), 555–573.
- GALE, D., AND R. ROSENTHAL (1994): “Price and quality cycles for experience goods,” *The RAND Journal of Economics*, 25(4), 590–607.
- HOFFMAN, K., D. ZAGE, AND C. NITA-ROTARU (2009): “A survey of attack and defense techniques for reputation systems,” *ACM Computing Surveys (CSUR)*, 42(1), 1–31.
- HOLMSTRÖM, B. (1999): “Managerial incentive problems: A dynamic perspective,” *Review of Economic Studies*, 66(1), 169–182.
- JÉHEIL, P. (1995): “Limited horizon forecast in repeated alternate games,” *Journal of Economic Theory*, 67(2), 497–519.
- LIU, Q., AND A. SKRZYPACZ (2010): “Limited records and reputation,” *Mimeo, Stanford University*.
- MAILATH, G., AND L. SAMUELSON (2001): “Who wants a good reputation?,” *Review of Economic Studies*, 68(2), 415–441.
- (2006): *Repeated Games and Reputations: Long-run Relationships*. Oxford University Press, USA.
- MARTI, S., AND H. GARCIA-MOLINA (2006): “Taxonomy of trust: Categorizing P2P reputation systems,” *Computer Networks*, 50(4), 472–484.
- MILLER, M. (2003): *Credit Reporting Systems and the International Economy*. The MIT Press.
- MOHAN, A., AND D. BLOUGH (2008): “AttributeTrust? A framework for evaluating trust in aggregated attributes via a reputation system,” in *Sixth Annual Conference on Privacy, Security and Trust*, pp. 201–212.
- MONTE, D. (2010): “Bounded memory and permanent reputations,” *working paper, Simon Fraser University*.
- PHELAN, C. (2006): “Public trust and government betrayal,” *Journal of Economic Theory*, 130(1), 27–43.

- ROSENTHAL, R. (1979): “Sequences of games with varying opponents,” *Econometrica*, 47(6), 1353–1366.
- SMITH, L., AND P. SØRENSEN (2008): “Rational social learning with random sampling,” *working paper, University of Michigan*.
- STOKEY, N., AND R. LUCAS (1989): *Recursive methods in economic dynamics*. Harvard University Press.
- TADELIS, S. (1999): “What’s in a name? Reputation as a tradeable asset,” *American Economic Review*, 89(3), 548–563.
- WILSON, A. (2002): “Bounded memory and biases in information processing,” *NAJ Economics*, 5(3).
- WILSON, A., AND M. EKMEKCI (2006): “A note on reputation effects with finite memory,” *Discussion paper, Yale University*.