Reputation Effects with Endogenous Records

Harry PEI*

November 12, 2023

Abstract: A patient player interacts with a sequence of short-run players. The patient player is either an *honest type* who always takes a commitment action and never erases any record, or an *opportunistic type* who decides which action to take and whether to erase that action from his record at a low cost. We show that the patient player will have an incentive to build a reputation in every equilibrium and can secure a payoff that is strictly greater than his commitment payoff after accumulating a long enough good record. However, as long as the patient player has a sufficiently long lifespan, his *equilibrium payoff* must be close to his minmax value. Although a small probability of opportunistic type can wipe out all of the patient player's returns from building reputations, it only has a negligible effect on the short-run players' welfare.

Keywords: record length, endogenous records, reputation effects, reputation failure.

1 Introduction

Most of the existing works on repeated games and reputations assume that the lengths of players' records are *exogenous*.[1] These include the models in Fudenberg and Maskin (1986) and Fudenberg and Levine (1989) where players' records contain the full history of play and the ones in Liu and Skrzypacz (2014), Bhaskar and Thomas (2019), Levine (2021), and Pei (2023) where the lengths of players' records are bounded.

However, in many situations, players' record lengths are *endogenous* and are affected by their strategic behaviors. For example, sellers in online platforms may bribe consumers for deleting negative reviews, and may even retaliate or threaten to sue them for defamation if the negative reviews are not removed.[2] Similarly, politicians may collude with media outlets to limit the coverage of bad news (Besley and Prat 2006).

*Northwestern University. Email: harrydp@northwestern.edu. I thank Jeff Ely, Nina Fluegel, Drew Fudenberg, Marina Halac, David Levine, Teddy Mekonnen, Alessandro Pavan, Larry Samuelson, Andrzej Skrzypacz, Egor Starkov, Yiman Sun, and Alex Wolitzky for helpful comments. I thank the NSF Grant SES-1947021 and the Cowles Foundation for financial support.

[1] A few exceptions include models where the monitoring structure is designed by social planners such as Ekmekci (2011), Vong (2022), and Wong (2023), as well as models where the uninformed players decide how much information to acquire about the informed player's history, such as Liu (2011). We explain the differences between our model and theirs in the literature review.

[2] According to Section 5 in Tadelis (2016), the lack of negative reviews due to seller reciprocity, retaliation, and harassment has caused significant biases in online reviews. A 2019 report in CNBC documents that many consumers who left negative reviews on Yelp were sued by firms in SLAPP lawsuits. Reports from CNET and the Guardian document that in the US and the UK, many Amazon sellers bribe consumers for deleting negative reviews. Bolton, Greiner, and Ockenfels (2013) provide empirical evidence for seller reciprocity and retaliation, making it more costly for buyers to post negative reviews than posting positive ones. Nosko and Tadelis (2015) document that only 0.07% of the reviews on eBay are negative despite a much larger fraction of the consumers complained to consumer service. Cai, et al (2014) and Tadelis (2016) document similar findings on EachNet and Airbnb.

This paper takes a first step to analyze reputation effects when players' record lengths are determined *endogenously* by their strategic behaviors. We analyze a novel reputation model in which a patient player can erase actions from his records at a low cost. We show that the patient player will build a reputation in *all* equilibria and can secure a high payoff after accumulating a long enough good record. However, his equilibrium payoff must be close to his minmax value when he has a sufficiently long lifespan, even if with *high probability*, he is an honest type who always takes some commitment action and never erases any action. The main takeaway is that the possibility of erasing records *cannot* eliminate the patient player's reputational incentives, but it can wipe out his returns from building reputations by slowing down the process of reputation building. Although a tiny probability of opportunistic type who may manipulate records can significantly lower the patient player's payoff, it only has a negligible effect on the short-run players' welfare.

We study a repeated game between a long-run player (e.g., a firm) and a sequence of short-run players (e.g., consumers). The long-run player discounts future payoffs and exits the game with some exogenous probability after each period. Players' stage-game payoffs are monotone-supermodular. The product choice game in Mailath and Samuelson (2001) satisfies our assumption, which we use to illustrate our results:

firm \ consumer	**L**arge Quantity	**S**mall Quantity
Good Products	$\mathbf{1}, 1$	$-g, x$
Bad Products	$1+g, -x$	$0, 0$

with $g > 0$ and $x \in (0, 1)$.

By the end of each period, the firm can erase its action in that period at a cost c.[3] We focus on the case in which the cost of erasing an action is strictly lower than the cost of supplying good products g.[4]

The long-run player has private information about his type: He is either an *honest type* who supplies good products in every period and never erases any action, or an *opportunistic type* who strategically decides which products to supply and whether to erase his actions. Each short-run player can observe the long-run player's *unerased actions* but cannot observe how many actions were erased. Hence, they cannot observe the long-run player's age in the game, i.e., calendar time.[5] Consistent with the literature on reputation effects with limited memories such as Liu and Skrzypacz (2014), the short-run players have a prior belief about the

[3]The firm in our model can only erase reviews but cannot modify the content of reviews. Arguably, it is harder to persuade dissatisfied consumers to write positive reviews than to ask them to stay silent. Our main result shows that reputation effects will fail when firms can manipulate their records, which is *stronger* when they can only erase reviews but cannot modify their content.

[4]We study the case where $c > g$ in Online Appendix A. Our assumption that $c < g$ seems reasonable since the consumers' losses from their bad experiences are sunk, so they might be willing to remove their negative reviews in exchange for a small bribe or to avoid a defamation lawsuit. The firms' costs of issuing a giftcard and making legal threats seem to be lower than the cost of supplying high quality. This is consistent with the empirical observation that negative reviews are rare in online marketplaces.

[5]Information about the seller's age on the market is not available or cannot be easily obtained by consumers in online platforms such as Yelp, Amazon, and TMall since they only disclose the number of reviews each seller received, the number times that he received each rating, together with some comments. Section 5.3 extends our results to settings in which either the consumers arrive *stochastically* or they *post reviews with probability less than* 1 after interacting with the seller, in which cases the seller's age on the market, the number of consumers that he has interacted with, and the number of reviews that he has received may not be the same.

long-run player's age, which is determined by the rate with which the long-run player leaves the game. The short-run players update their beliefs according to Bayes rule after observing the long-run player's record.

When the firm can erase its actions at a low cost, it will never supply good products *after* the consumers rule out the honest type. However, as long as the honest type occurs with positive probability, the opportunistic type will supply good products with positive probability in *all* equilibria until it has a sufficiently long good record at which point its continuation value will be strictly greater than its commitment payoff 1. The intuition is that the firm can signal its honesty via the *length* of its good record and in every equilibrium, both the firm's reputation and the consumers' willingness to play L *increase* in the firm's record length.

Although the firm can secure a high payoff in the long run, our main result, Theorem 1, shows that when the firm has a sufficiently long lifespan, (i) the opportunistic type's equilibrium payoff must be close to its minmax value 0 and (ii) the consumers' equilibrium welfare must be close to that in an auxiliary setting where they can observe the firm's type and the opportunistic type always supplies bad products. This result implies that the presence of *a small fraction of opportunistic types* who may supply bad products and may erase records can wipe out all of the firm's returns from building reputations but it only has a negligible effect on consumer welfare. It also implies that when a firm *commits* to the consumers that it will supply good products and will never erase any action, the value of such a commitment will be seriously compromised as long as the consumers entertain a grain of doubt about the firm's willingness to honor its commitment.

Theorem 1 is driven by two forces that are caused by the firm's ability to erase actions. On the one hand, the opportunistic type's ability to erase actions implies that it can sustain its current continuation value at a low cost. In order to motivate the opportunistic type to build a reputation for supplying good products, its continuation value has to increase fast enough with the length of its good record. This leads to *an upper bound* on the maximal length of good record that the opportunistic type may have in equilibrium, or equivalently, the minimal length of good record required for the firm to have a perfect reputation.

On the other hand, the opportunistic type may *not* lose its reputation after supplying bad products since it can erase that action. The firm's incentive to do so will slow down consumer learning and in equilibrium, firms with shorter records will have worse reputations. The upper bound on the speed of learning leads to *a lower bound* on the length of good record required for the firm to have a perfect reputation.

As the firm's expected lifespan increases, the honest type's expected record length increases, which also increases the lower bound on the length of good record required for the firm to have a perfect reputation. Once this lower bound exceeds the upper bound implied by the need to provide the opportunistic type incentives, the opportunistic type needs to separate from the commitment type with positive probability in the first period in order to boost its reputation. Since the opportunistic type's continuation value after separating from the honest type equals its minmax value 0, its equilibrium payoff must be close to 0.

In terms of consumer welfare, its equilibrium value is *no less* than the consumers' payoff when they play L if and only if the firm has a perfect reputation, and is *no more* than the consumers' payoff when they can observe the firm's *realized pure action* before choosing their actions. Although the opportunistic type will play G with high probability at some histories, we show that the probability with which a history occurs is bounded above zero *if and only if* the opportunistic type plays G at that history with a probability that is *at most proportional to its exit rate*. Since the opportunistic type's maximal length of good record is bounded, the *average* probability with which it plays G vanishes as its exit rate vanishes. If this is the case, then the upper and lower bounds on consumer welfare converge to the same value as the firm's exit rate vanishes.

Theorem 1 suggests that when the consumers expect that the honest type will have a long good record, firms with short records will have low reputations and will receive low payoffs. A natural question is that whether this reputation failure problem can be solved when the honest type does not reveal *all* its actions.

Theorem 2 studies a setting where the honest type discloses information about its past actions according to a *disclosure policy*. It shows that when the firm is sufficiently long-lived, regardless of the honest type's disclosure policy (i) the opportunistic type's payoff *cannot* exceed its equilibrium payoff in an auxiliary game where the honest type reveals no information and (ii) consumer welfare *cannot* exceed its equilibrium value in the baseline model where the honest type reveals *all* past actions. These conclusions also apply, for example, when a platform commits to reveal at most K of the firm's unerased actions to the consumers.

Theorem 2 implies that a long-lived firm *cannot* benefit from allowing the consumers to observe its history as long as the consumers suspect that with positive probability, it can erase actions from its records at a low cost. This stands in contrast to the standard lessons from the theories of repeated games such as Fudenberg, Kreps and Maskin (1990), that a patient player can obtain higher payoffs in *some equilibria* when his opponents can monitor his past actions relative to the case in which there is no monitoring at all.

Theorem 2 is driven by a conflict between motivating the opportunistic type to supply good products and persuading the consumers to buy a large quantity starting from period 0. As in Theorem 1, the opportunistic type's incentives lead to *an upper bound* on the longest good record that it may have in any equilibrium.

Persuading the consumers in our setting differs from the Bayesian persuasion problem in Kamenica and Gentzkow (2011). This is because the consumer's payoff in our model depends only on the firm's *endogenous actions* rather than on some *exogenous state of the world*, and the honest type's disclosure policy affects the opportunistic type's equilibrium behavior. Our proof uses the upper bound on the opportunistic type's *average behavior* that we derived earlier, which implies that the average probability with which it supplies good products vanishes when its expected lifespan goes to infinity. If this is the case, then it is impossible to persuade the consumers to play L in every period when (i) the honest type occurs with low probability and (ii) the opportunistic type erases B at every history and never separates from the honest type.

Related Literature: This paper takes a first step to analyze reputation effects when players' records, and in particular, their record lengths, are determined endogenously by their own behaviors. Our model stands in contrast to Ekmekci (2011), Vong (2022), Kovbasyuk and Spagnolo (2023), and Wong (2023) in which the record systems are designed by planners who do not participate in the game, as well as Liu (2011) in which the short-run players decide how much information to acquire about the long-run player's history.[6]

Our reputation model has several merits compared to our benchmark where the honest type occurs with probability 0 and to some of the existing reputation models such as the one in Fudenberg and Levine (1989).

First, our model leads to sharp predictions not only on the long-run player's payoff, but also on both players' behaviors and the short-run players' welfare. In contrast, the results in Fudenberg and Levine (1989) focus exclusively on the long-run player's equilibrium payoff but do *not* lead to sharp predictions on players' behaviors and the short-run players' welfare. The details are explained in Li and Pei (2021).

Second, our predictions are consistent with a number of empirical findings in online marketplaces. Livingston (2005) finds that negative reviews are rare on eBay and sellers' sales depend mostly on the lengths of their good records. Nosko and Tadelis (2015) document that 99.3% of the reviews on eBay are positive despite a significant fraction of the consumers are dissatisfied and complained to customer service. These are consistent with our finding that in *all* equilibria, (i) the opportunistic firm supplies good products with positive probability in many periods, (ii) the number of negative reviews is much smaller than the number of times that the seller supplied bad products, and (iii) the consumers play L with higher probability when the firm has a longer good record. Meanwhile, in the benchmark without any honest type, the firm supplies bad products and the consumers play S in every period. In Fudenberg and Levine (1989), the relationship between the short-run player's action and the length of the long-run player's good record is sensitive to the selection of equilibrium and may not be monotone. In Liu and Skrzypacz (2014), the short-run player's action depends only on the timing of the latest bad review rather than on the length of good records.

Third, our results point out that when record lengths are endogenous, the qualitative features of the equilibria depend not only on the long-run player's discount factor, but also on the reason for why he discounts future payoffs, that is, the extent to which discounting is driven by his *time preference* versus his *survival probability*. In contrast, there is no distinction between the long-run player's time preference and survival probability in existing reputation models, which include models where players cannot observe

[6] Ekmekci, Gorno, Maestri, Sun and Wei (2022) study a continuous-time stopping game where an informed long-run player can manipulate the *content* of public signals rather than the *length* of his records. This stands in contrast to our model.

calendar time such as Liu (2011), Liu and Skrzypacz (2014), Levine (2021), and Pei (2023).[7]

We show that the patient player will receive a low payoff in *all* equilibria even when he is the opportunistic type with low probability. This is related to the literature on bad reputation, most notably Ely and Välimäki (2003) and Ely, Fudenberg and Levine (2008). They show that the patient player will receive a low payoff in all equilibria when there is a lack-of-identification problem and *bad commitment types* occur with positive probability. Their papers focus on *participation games* in which the short-run players can unilaterally shut down learning by taking a non-participating action. This stands in contrast to our model where no player can unilaterally stop learning and reputation failure is caused by the low rate of learning.[8]

Our model can be interpreted as a continuum of firms and consumers being randomly matched in each period and each consumer observing the *record* of the firm she is matched with. This is related to the literature on community enforcement with a continuum of players, such as Takahashi (2010), Heller and Mohlin (2018), Bhaskar and Thomas (2019), and Clark, Fudenberg and Wolitzky (2021). One of our contributions to this literature is to introduce *endogenous record length* and *reputations*. We show that even a small fraction of opportunistic types who may manipulate records can significantly lower social welfare.[9]

Our work is also related to several recent papers on dynamic information censoring such as Smirnov and Starkov (2022), Sun (2023), and Hauser (2023). Unlike those papers in which the uninformed player's payoff depends only on the informed player's *type*, the uninformed player's payoff depends only on the informed player's *action* in our model. Our formulation is standard in models of repeated games and reputations, which fits markets where quality provision is subject to moral hazard. Compared to their works, we highlight the roles of *record length* and *the informed player's expected lifespan* on the value of reputations.

2 The Baseline Model

Time is discrete, indexed by $k = 0, 1, 2, ...$ A long-lived player 1 interacts with an infinite sequence of short-lived player 2s, arriving one in each period and each plays the game only in the period she arrives.

Player 1 discounts future payoffs for two reasons. First, by the end of each period, he exits the game with probability $1 - \overline{\delta}$ where $\overline{\delta} \in (0, 1)$, after which the game ends and players' payoffs are 0. Second, he is

[7] The results on steady state learning in Fudenberg and Levine (1993) and Fudenberg and He (2018) require players' expected lifespans to be much longer relative to their patience, under which players spend most of their lives playing their best replies in the steady state. In contrast, a longer expected lifespan in our model lowers the long-run player's reputation when he has a short record.

[8] The conclusion that delays are necessary for the patient player to signal his type also appears in repeated signaling games with *interdependent values* where the receiver's payoff depends on the sender's type, such as in the model of Kaya (2009). In contrast, our model has *private values* and costly delays are caused by the patient player's ability to erase actions from his records.

[9] Sugaya and Wolitzky (2020) show that players' payoffs are arbitrarily close to their minmax values in all equilibria of a community enforcement model with *bad commitment types*. However, their model has a finite number of players and focuses on *symmetric* stage games with a pairwise dominant action (e.g., the prisoner's dilemma). Both features stand in contrast to our model.

indifferent between $\widehat{\delta} \in (0,1)$ unit of utility in period k and 1 unit in period $k+1$. Therefore, the long-run player discounts future payoffs by $\delta \equiv \overline{\delta} \cdot \widehat{\delta}$, which we call his *effective discount factor*. His time preference $\widehat{\delta}$ and his survival probability $\overline{\delta}$ play different roles since only $\overline{\delta}$ affects his *expected lifespan* $(1-\overline{\delta})^{-1}$.

In period $k \in \mathbb{N}$, players simultaneously choose their actions $a_{1,k} \in A_1$ and $a_{2,k} \in A_2$ from finite sets A_1 and A_2. Players' stage-game payoffs are $u_1(a_{1,k}, a_{2,k})$ and $u_2(a_{1,k}, a_{2,k})$. We make two assumptions on (u_1, u_2) throughout the paper. The first one is a monotone-supermodularity assumption, which is standard in the literature and is also assumed in Mailath and Samuelson (2001), Ekmekci (2011), and Liu (2011).

Assumption 1. *There exist a complete order on A_1 and a complete order on A_2 such that $u_1(a_1, a_2)$ is strictly decreasing in a_1 and is strictly increasing in a_2, and $u_2(a_1, a_2)$ has strictly increasing differences.*

The product choice game in the introduction satisfies Assumption 1 once we rank the row player's actions according to $G \succ B$ and the column player's actions according to $L \succ S$. We make another assumption, which is *generically* satisfied as long as A_1 and A_2 are finite sets.

Assumption 2. *Player 1 has a strict best reply to every $a_2 \in A_2$. Player 2 has a strict best reply to every $a_1 \in A_1$. For every $a_2 \in A_2$, $a_1, \widetilde{a}_1 \in A_1$, and $\lambda \in [0,1]$, if a_2 best replies to player 1's mixed action $\lambda a_1 + (1-\lambda)\widetilde{a}_1$, then there exists $\widetilde{\lambda} \in [0,1] \backslash \{\lambda\}$ such that a_2 also best replies to $\widetilde{\lambda} a_1 + (1-\widetilde{\lambda})\widetilde{a}_1$.*

The first part of Assumption 2 requires each player to have a strict best reply to each of their opponent's *pure* actions. The second part of Assumption 2 rules out situations in which some a_2 *only* best replies to a knife-edge distribution over player 1's actions: It allows some of player 2's actions to be strictly dominated and also allows some actions to best reply to *an open set* of player 1's mixed actions. We explain why this part of Assumption 2 is *generically* satisfied by the end of Appendix C, after showing Lemma 2.

Let $\underline{a}_1$ denote the *lowest* action in A_1. Assumption 1 implies that $\underline{a}_1$ is strictly dominant in the stage game. Let $\underline{a}_2 \in A_2$ denote player 2's best reply to $\underline{a}_1$, which is unique under Assumption 2. By definition, $u_1(\underline{a}_1, \underline{a}_2)$ is player 1's *minmax value* in the sense of Fudenberg, Kreps and Maskin (1990), which requires player 2 to play a best reply to some $\alpha_1 \in \Delta(A_1)$. We normalize player 1's payoff so that $u_1(\underline{a}_1, \underline{a}_2) \equiv 0$.

By the end of period k but before period $k+1$, player 1 decides whether to erase his period-k action $a_{1,k}$ at cost $c > 0$. We denote this decision by $c_k \in \{0, c\}$, where $c_k = 0$ stands for $a_{1,k}$ not being erased and $c_k = c$ stands for $a_{1,k}$ being erased. All of our results except for Proposition 3 extend to the case where the cost of erasing actions is 0. Except for an extension in Online Appendix A, our analysis focuses on the case where the cost of erasing actions c is lower than the cost of taking high actions. Formally, let a_1' denote the lowest action in A_1 such that $\underline{a}_2$ does *not* best reply to a_1'. By definition, a_1' is strictly greater than $\underline{a}_1$.

We assume that:

$$c < \overline{c} \equiv \min_{\beta \in \Delta(A_2)} \Big\{ u_1(\underline{a}_1, \beta) - u_1(a_1', \beta) \Big\}. \tag{2.1}$$

In the product choice game, $\overline{c}$ equals the cost of supplying good products g.[10] This restriction seems reasonable since consumers' losses from their bad experiences are sunk.[11] Therefore, they might be willing to remove their negative reviews in exchange for a small bribe, or to avoid a defamation lawsuit. Paying bribes (e.g., by issuing a giftcard) and making threats are usually not that costly from the firm's perspective.

Player 1 has persistent private information about his type $\omega \in \{\omega_o, \omega_h\}$, where ω_o stands for an *opportunistic type* who chooses his actions as well as whether to erase them in order to maximize his *discounted average payoff* $\sum_{k=0}^{+\infty} (1-\delta)\delta^k \{u_1(a_{1,k}, a_{2,k}) - c_k\}$, and ω_h stands for an *honest type* who takes action $a_1^* \neq \underline{a}_1$ in every period and never erases any action. Let a_2^* denote player 2's unique best reply to a_1^*. Player 1's *commitment payoff* is $u_1(a_1^*, a_2^*)$. We focus on the interesting case where $u_1(a_1^*, a_2^*) > u_1(\underline{a}_1, \underline{a}_2) \equiv 0$.

Before choosing $a_{1,k}$, player 1 observes his type ω and the *full history* of the game up to period k, which we denote by $h^k \equiv \{a_{1,s}, a_{2,s}, c_s\}_{s=0}^{k-1}$. Player 1 observes ω, h^k, and $(a_{1,k}, a_{2,k})$ before choosing c_k.

Player 2 observes the sequence of player 1's *unerased actions* but *cannot* observe which actions were erased and how many actions were erased. Formally, her history in period k is a sequence $\{a_{1,\tau_0}, ..., a_{1,\tau_{m(k)}}\}$ where $0 \leq \tau_0 < \tau_1 < ... < \tau_{m(k)} \leq k-1$ such that for every $s \in \{0, 1, ..., k-1\}$, $c_s = 0$ if and only if there exists $i \in \{0, 1, ..., m(k)\}$ such that $s = \tau_i$.[12] Let $\mathcal{H}_2$ denote the set of player 2's histories, or *player 1's records*. Our results extend to any monitoring structure as long as player 2 observes (i) the number of a_1^* among player 1's unerased actions and (ii) whether the set of player 1's unerased actions contains any action other than a_1^*. One example is that player 2_k only observes *the number of times* that each $a_1 \in A_1$ occurred in the sequence $\{a_{1,\tau_0}, ..., a_{1,\tau_{m(k)}}\}$, that is, the *summary statistics* of player 1's unerased actions.

We also assume that the short-run players *cannot* directly observe the long-run player's age in the game, or equivalently, calendar time. This assumption is common in reputation models with finite memories such as Liu and Skrzypacz (2014), Levine (2021), and Pei (2023), reputation models with incomplete records such as Liu (2011), and community enforcement models such as Heller and Mohlin (2018). It is reasonable on platforms such as Yelp, Amazon, and TMall which reveal the number of each rating the seller has received but do *not* disclose in a salient place the seller's time on the market. It is consistent with our interpretation

[10]Even when the long-run player has *a continuum of actions*, the value of $\overline{c}$ remains bounded above 0. Take the product choice game example. If the firm chooses its effort level from the unit interval $[0, 1]$ and the consumers have an incentive trust the firm (i.e., play L) only when its expected effort is more than $x \in (0, 1)$, then $\overline{c}$ equals the firm's cost of choosing effort level x.

[11]An interpretation is that the seller decides whether to supply high quality experienced goods and the consumers decide how much to purchase *without* knowing product quality. The consumers can observe the seller's action, i.e., quality, *after* purchase. Then they post a review that honestly reflects the seller's action. The seller then decides whether to erase that review at cost $c \geq 0$.

[12]Our definition of player 2's history implies that she *cannot* observe previous short-run players' actions. This assumption is standard in reputation models with limited memories, which is also made in Liu (2011), Liu and Skrzypacz (2014), and Pei (2023).

that firms can erase actions from their records and consumers cannot observe the extent to which records were erased. Sections 5.3 and 5.4 extend our results to settings where consumers either *arrive stochastically* or *do not post reviews with positive probability*, in which cases the seller's age may not coincide with the number of consumers that he has interacted with and the number of reviews that he has received.

As in Liu and Skrzypacz (2014), the short-run players have a prior belief about calendar time and update their beliefs using Bayes rule after they observe their histories. Since the long-run player exits the game with probability $1-\overline{\delta}$ after each period, the probability that the short-run players' prior belief assigns to the long-run player's age being $k+1$ should equal $\overline{\delta}$ times the probability that her prior belief assigns to the long-run player's age being k. The unique prior belief that satisfies this condition for every $k \in \mathbb{N}$ is the one that assigns probability $(1-\overline{\delta})\overline{\delta}^k$ to the long-run player's age being k.[13] In Appendix I, we explain how to calculate the short-run player's posterior belief about calendar time after they observe their history.

Let $\pi \in (0,1)$ denote the prior probability of the honest type, which can be different from player 2's *posterior belief* after she observes a history with length zero. Player 1's *reputation* at a history $h_2 \in \mathcal{H}_2$ is the probability that player 2's posterior belief assigns to the honest type after observing history h_2.

Players' strategies σ_1 and σ_2 are mappings from their histories to their actions. A *Nash equilibrium* is a strategy profile (σ_1, σ_2) such that σ_i best replies to σ_{-i} for every $i \in \{1,2\}$. A *stationary equilibrium* (σ_1^*, σ_2) is a Nash equilibrium such that the opportunistic type of player 1's strategy σ_1^* is measurable with respect to player 2's history. Let $\mathcal{H}(\sigma_1,\sigma_2) \subset \mathcal{H}_2$ denote the set of player 2's histories that occur with positive probability under (σ_1,σ_2), which we refer to as the set of *on-path histories*.

Lemma 1. *If (σ_1,σ_2) is a Nash equilibrium, then there exists a stationary equilibrium (σ_1^*,σ_2) such that $\mathcal{H}(\sigma_1,\sigma_2) = \mathcal{H}(\sigma_1^*,\sigma_2)$, player 2's beliefs about player 1's action and type at every history that belongs to $\mathcal{H}(\sigma_1,\sigma_2)$ are the same under (σ_1,σ_2) and (σ_1^*,σ_2), and the expected values of $\sum_{k=0}^{+\infty}(1-\delta)\delta^k\{u_1(a_{1,k},a_{2,k})-c_k\}$ and $U_2 \equiv \sum_{k=0}^{+\infty}(1-\overline{\delta})\overline{\delta}^k u_2(a_{1,k},a_{2,k})$ are the same under (σ_1,σ_2) and (σ_1^*,σ_2).*

The proof is in Appendix A. Lemma 1 shows that for every Nash equilibrium, there exists an *equivalent* stationary equilibrium in the sense that (i) the short-run players' strategy σ_2 and their beliefs about the long-run player's action and type at every on-path history remain the same and (ii) the long-run player's discounted average payoff and the short-run players' welfare, measured by the *sum* of their payoffs U_2,[14] remain the same. Our subsequent analysis focuses on the common properties of *all* stationary equilibria (or *equilibria* for short). According to Lemma 1, the properties we derived on players' payoffs, the short-run

[13] Hu (2020) provides a microfoundation for models in which calendar time is not observed. He constructs a probability measure over the order of entry and shows that a common posterior belief axiom is satisfied if and only if the short-run players entertain an exponential prior belief about calendar time. His result justifies the exponential prior belief about calendar time in our model.

[14] The sum of player 2's payoffs is $\sum_{k=0}^{+\infty}(1-\overline{\delta})\overline{\delta}^k u_2(a_{1,k},a_{2,k})$ since the game ends with probability $1-\overline{\delta}$ after each period.

players' actions, beliefs, and learning, as well as the long-run player's reputation apply to *all* Nash equilibria.

2.1 Alternative Interpretations of the Baseline Model

We spell out two alternative interpretations of our model, which will be used to discuss the implications of our results. Our model can be interpreted as one of *imperfect commitment*: Player 1 commits to player 2 that he will always play a_1^* and will never erase any action. Player 2 is concerned that player 1 will *not* honor his commitment with probability $1-\pi$, in which case player 1 may take other actions and may erase actions.[15] Section 4 explores alternative forms of commitment, in which the honest type commits to play a_1^* in every period but reveals information about his history according to an arbitrary information disclosure policy.

Our model also has a *population-level interpretation*. As in Heller and Mohlin (2018) and Clark, Fudenberg and Wolitzky (2021), consider a doubly infinite repeated game between a unit mass of player 1 and a unit mass of player 2. A fraction π of player 1s are honest types and the rest are opportunistic types. Each period, each player 1 is matched with a player 2 uniformly at random to play the stage game. Player 2 observes the *record* of the player she is matched with before choosing her action. Player 1 cannot observe player 2's record. After the interaction, player 1 decides whether to erase the action that he has just taken at a cost c. Then a fraction $1-\overline{\delta}$ of player 1s exit the game and are replaced by new ones without any record, a fraction π of which are honest types. Whether a player will exit is independent of his type, record, and past actions. As a result, player 2's *prior belief* assigns probability $(1-\overline{\delta})\overline{\delta}^t$ to player 1's age being t.

Each player's *strategy* is a mapping from their histories to their actions. A *steady state* consists of a strategy for each player and a distribution over records such that the record distribution is consistent with the strategy profile and players' strategies best reply to each other given the record distribution. One can show that every steady state of this model corresponds to a stationary equilibrium of our baseline model.

2.2 Discussions on the Modeling Assumptions and Extensions

We assume that player 1 can erase actions from his records but *cannot* modify the content of his records. In practice, whether firms can modify the reviews as well as their costs of doing so depend on the institutional details. Our assumption is motivated by the observation in Tadelis (2016) that most of the consumers post reviews because they are motivated to share their opinions, to reward firms' good behaviors and to punish bad ones, or to provide future consumers useful information. If this is the case, then it seems more challenging for firms to convince consumers to lie about their experiences than to ask them to stay silent. We show that firms will receive low payoffs even when only a small fraction of them can manipulate their records. This

[15] This form of imperfect commitment is studied by Lipnowski, Ravid and Shishkin (2022) in a one-shot communication model.

conclusion is stronger when opportunistic firms can only erase reviews but cannot modify their content.[16]

Our analysis focuses on the case where the cost of erasing actions is low. In the product choice game, it translates into $c < g$. In Online Appendix A, we show that when $c > g$, for every $\pi \in (0,1)$ and $\varepsilon > 0$, there exists $\delta^* \in (0,1)$ such that when $\delta > \delta^*$, the opportunistic type's payoff in every equilibrium is ε-close to his commitment payoff and that he will never erase any action whenever he has a positive reputation. Therefore, the ability to erase actions affects the equilibrium outcomes only when the cost of erasing is low.

Our baseline model focuses on the case in which there is only one honest type and one opportunistic type and assumes that the honest type *cannot* erase actions. Section 5.2 studies an extension where the honest type commits to take action a_1^* in every period and *strategically* decides whether to erase his action at a strictly positive cost $c > 0$ in order to maximize his discounted average payoff. In Online Appendix B, we study an extension where (i) there are multiple honest types taking different pure actions and (ii) there are multiple opportunistic types with different payoff functions and different costs of erasing actions.

Our baseline model focuses on the case in which (i) the short-run players' best reply does not depend on whether the long-run player will erase his action and (ii) the cost of erasing actions does not depend on the action profile being played. In practice, if we interpret erasing actions as offering a partial refund or a giftcard in exchange for deleting the review, then the consumers' demands for refunds or giftcards may depend on the quantity they purchased and on the firm's action (e.g., product quality). Our theorems can be extended to cases where (i) each consumer's best reply depends also on the probability with which the firm will erase its action, and (ii) the firm's cost of erasing actions depends on the action profile being played.

Our model assumes that the short-run players can perfectly observe the long-run player's *unerased actions*. This rules out situations in which they can only observe noisy signals about those actions. Analyzing repeated games with incomplete information, imperfect monitoring, and limited observations is challenging. This explains why most of the existing analysis on reputation games with finite record lengths such as Liu (2011), Liu and Skrzypacz (2014), and Heller and Mohlin (2018) all focus on the case with perfect monitoring.[17] The case with *endogenous* record length and imperfect monitoring is left for future work.

3 Analysis and Results

Section 3.1 examines two benchmarks, which includes the case where the honest type occurs with probability 0 and the case where player 1 cannot erase actions. Then we analyze the case of interest in which

[16]If the opportunistic type can modify the content of his records at a low cost, e.g., he can change his record to a_1^* when his action was $\underline{a}_1$, then he will never play a_1^* since doing so is strictly dominated by playing $\underline{a}_1$ and then modifying his record to a_1^*.

[17]Bhaskar and Thomas (2019) allow for imperfect monitoring, but they focus on a *complete information game* in which the long-run player has only one type. Levine (2021) allows for imperfect monitoring but assumes that players' record length is 1.

$\pi \in (0,1)$ and player 1 can erase records. Section 3.2 establishes some properties of players' behaviors and beliefs that apply to *all* equilibria. Sections 3.3 and 3.4 state and prove our main result, Theorem 1, which provides a sharp characterization of players' equilibrium payoffs when player 1 is sufficiently long-lived.

3.1 Benchmarks

First, suppose the opportunistic type *cannot* erase actions. When $\pi = 0$, that is, the honest type occurs with zero probability, the folk theorem in Fudenberg, Kreps and Maskin (1990) implies that there exists an equilibrium in which player 1's payoff is $u_1(a_1^*, a_2^*)$ and there also exist equilibria in which player 1 receives his minmax value $u_1(\underline{a}_1, \underline{a}_2) \equiv 0$. When $\pi > 0$, the result in Fudenberg and Levine (1989) implies that player 1's payoff in *every* equilibrium is approximately his commitment payoff $u_1(a_1^*, a_2^*)$ when δ is sufficiently close to 1. In Online Appendix A, we show that a similar conclusion holds when c is large enough. However, as shown in Li and Pei (2021), there is little robust prediction on players' behaviors and the short-run players' welfare. For example, in the product choice game, there exist equilibria in which the firm takes action G in every period and the consumers obtain their first-best payoff 1, and there also exist equilibria in which the firm takes action G with low frequency and the consumers receive low payoffs.

Second, suppose the opportunistic type *can* erase his actions but the honest type occurs with probability 0. Recall that $\mathcal{H}(\sigma_1, \sigma_2)$ is the set of player 2's histories that occur with positive probability under (σ_1, σ_2).

Proposition 1. *If $c < \overline{c}$ and $\pi = 0$, then in every equilibrium (σ_1, σ_2), action profile $(\underline{a}_1, \underline{a}_2)$ is played at every history that belongs to $\mathcal{H}(\sigma_1, \sigma_2)$.*

The proof is in Appendix B. Intuitively, player 1's ability to erase actions implies that he can take $\underline{a}_1$ and then erase it, by which he can *sustain his current continuation value*. Fix any equilibrium (σ_1, σ_2) and consider any history h_2 where player 1's continuation value is *close* to his highest continuation value among the histories in $\mathcal{H}(\sigma_1, \sigma_2)$. When $c < \overline{c}$, player 1 strictly prefers taking action $\underline{a}_1$ and then erasing it to taking any action that is no less than a_1'. This is because taking any action at h_2 will lead to at most a negligible increase in player 1's continuation value. Once player 2 anticipates this, she will have a strict incentive to take action $\underline{a}_2$ at h_2. This will unravel any equilibrium where player 1 takes any action other than $\underline{a}_1$.

3.2 Preliminary Observations

Starting from this section, we analyze the case in which $\pi \in (0,1)$ and player 1 can erase records at cost $c < \overline{c}$. It is straightforward that player 1 will never take any action other than a_1^* and $\underline{a}_1$ at any on-path history in any equilibrium. This is because if in equilibrium, he takes any action $a_1 \notin \{a_1^*, \underline{a}_1\}$ with positive probability and then not erase it, then he will be separated from the honest type and Proposition 1 implies that his continuation value will be 0. In order to obtain a strictly positive continuation value after taking action a_1, he needs to erase a_1. But *taking action* a_1 *and then erasing it* is strictly dominated by *taking action* $\underline{a}_1$ *and then erasing it*. This is because they lead to the same history for player 2 in the next period but the latter results in a higher payoff in the current period. Since player 1's action at every on-path history belongs to $\{a_1^*, \underline{a}_1\}$, player 2's (potentially mixed) action at every $h_2 \in \mathcal{H}(\sigma_1, \sigma_2)$ belongs to

$$\mathcal{B} \equiv \Big\{\beta \in \Delta(A_2) \Big| \beta \text{ best replies to } \lambda a_1^* + (1-\lambda)\underline{a}_1 \text{ for some } \lambda \in [0,1]\Big\}. \tag{3.1}$$

Lemma 2 shows that every pair of player 2's mixed actions in $\mathcal{B}$ can be ranked according to FOSD and that one can generate a rich set of payoffs for player 1 by varying player 2's actions in $\mathcal{B}$.

Lemma 2. *Each pair of elements in $\mathcal{B}$ can be ranked according to FOSD. For every $a_1 \in A_1$ and $v \in [u_1(a_1, \underline{a}_2), u_1(a_1, a_2^*)]$, there exists a unique $\beta \in \mathcal{B}$ such that $u_1(a_1, \beta) = v$.*

The proof is in Appendix C. Let h_*^k denote player 2's history where she observes k actions, all of which are a_1^*. Let $\mathcal{H}_* \equiv \{h_*^k | k \in \mathbb{N}\}$, which is the set of histories where player 1 has a strictly positive reputation. Let $p_k^* \in [0,1]$ denote the probability with which the opportunistic type of player 1 plays a_1^* at h_*^k. Let $\beta_k \in \Delta(A_2)$ denote player 2's mixed action at h_*^k. Let $\pi_k \in [0,1]$ denote player 1's reputation at h_*^k. Proposition 2 establishes some properties of players' behaviors and beliefs that apply to *all* equilibria.

Proposition 2. *For every $\pi \in (0,1)$ and $c \in (0, \overline{c})$, there exists a constant $\lambda > 0$ such that in every equilibrium under (π, c, δ), there exist $t_0, t \in \mathbb{N}$ that satisfy $0 \leq t_0 \leq t < +\infty$ and $t > \frac{\lambda}{1-\delta}$ such that:*

- *The opportunistic type plays $\underline{a}_1$ with positive probability at every on-path history. The opportunistic type plays a_1^* with positive probability at h_2 if and only if there exists $k < t-1$ such that $h_2 = h_*^k$.*
- *The opportunistic type's continuation value at h_*^{t-1} is at least $(1-\delta)u_1(a_1^*, \underline{a}_2) + \delta(u_1(\underline{a}_1, a_2^*) - c)$.*
- *The opportunistic type never erases a_1^* at any history and never erases any action unless the history belongs to $\mathcal{H}_*$. At history h_*^k, player 1 never erases $\underline{a}_1$ if $k < t_0$ and erases $\underline{a}_1$ for sure if $k > t_0$.*
- *For every $k < t$, player 1's reputation π_k at h_*^k is strictly less than 1. For every $k < t-1$, player 2's action β_k is strictly less than a_2^*. For every $k \geq t$, we have $\beta_k = a_2^*$ and $\pi_k = 1$.*
- *When $k < t$, β_k is strictly increasing in k in the sense of FOSD, π_k is strictly increasing in k, and p_k^* is strictly decreasing in k.*

The proof is in Appendix D. Compared to the game *without* honest type, the opportunistic type will build a reputation for playing a_1^* even though he can play $\underline{a}_1$ and then erase it at a low cost. When his effective discount factor δ is close to 1, he has an incentive to take action a_1^* *for a long time* in the sense that the number of such periods is bounded below by a linear function of $(1-\delta)^{-1}$. The intuition is that the long-run player can signal his honesty via the *length* of his good record. By the time he stops playing a_1^*, his continuation value at h_*^{t-1} is at least $(1-\delta)u_1(a_1^*, \underline{a}_2) + \delta(u_1(\underline{a}_1, a_2^*) - c)$, which is *strictly bounded above* his commitment payoff $u_1(a_1^*, a_2^*)$ given that $c < \overline{c}$. The opportunistic type can secure this payoff by (i) taking action a_1^* at h_*^{t-1} and (ii) in every subsequent period, taking action $\underline{a}_1$ and then erasing it at cost c.

In terms of the equilibrium dynamics, Proposition 2 shows that when the long-run player has a longer good record, he will have a stronger incentive to erase $\underline{a}_1$, and both his reputation and the short-run player's action will increase. The intuition is that the long-run player faces decreasing returns from generating longer good records, faces a greater loss from losing his reputation once he has a longer good record, and needs to be rewarded for taking the commitment action given that he has the option to take the lowest action and then erase it at a low cost. The first part is reminiscent of reputation models with changing types such as Phelan (2006) and the second part is reminiscent of the bad news model of Board and Meyer-ter-Vehn (2013).

3.3 Main Result

Although Proposition 2 implies that both players will obtain high payoffs after the long-run player has accumulated a long enough good record, it remains silent on the *time* it takes for players to obtain the high payoffs. The latter is also crucial both for player 1's equilibrium payoff and for player 2's welfare.

Our main result, Theorem 1, answers this question and provides a sharp characterization of the long-run player's payoff and the short-run players' welfare in the case where the long-run player has *a long lifespan*. Recall that $\widehat{\delta}$ measures player 1's patience, $\overline{\delta}$ stands for the probability that the game continues after each period, and $\delta \equiv \widehat{\delta} \cdot \overline{\delta}$. Recall the definition of player 2's welfare U_2 in Lemma 1 and that of $\overline{c}$ in (2.1).

Theorem 1. *For every $c < \overline{c}$, $\pi \in (0,1)$, $\widehat{\delta} \in (0,1)$, and $\varepsilon > 0$, there exists $\delta^* \in (0,1)$ such that in every equilibrium when $\overline{\delta} \in (\delta^*, 1)$,*

1. *The opportunistic type of player 1 receives a payoff no more than $(1-\delta)c/\delta$.*
2. *Player 2's welfare U_2 belongs to an ε-neighborhood of $\pi u_2(a_1^*, a_2^*) + (1-\pi)u_2(\underline{a}_1, \underline{a}_2)$ under an additional assumption that $u_2(a_1, a_2)$ is weakly increasing in a_1.*

According to Theorem 1, as long as player 1 has a sufficiently long lifespan, even when his discount factor δ is close to 1 and the probability of opportunistic type is close to 0, his equilibrium payoff must be close to his minmax value 0.[18] The honest type's payoff is also no more than $(1-\delta)c/\delta$ when the two types share the same stage-game payoff function since the opportunistic type can imitate the honest type.

If in addition that $u_2(a_1, a_2)$ is weakly increasing in a_1, which in the product choice game translates into the consumers receiving higher payoffs when the firm exerting higher effort, then player 2's welfare in every equilibrium is arbitrarily close to their payoff in an auxiliary setting where (i) they can observe the long-run player's type and (ii) the opportunistic type always takes his lowest-cost action $\underline{a}_1$. In fact, the additional assumption is only needed to show that player 2's welfare is at least $\pi u_2(a_1^*, a_2^*) + (1-\pi)u_2(\underline{a}_1, \underline{a}_2)$ when $\overline{\delta}$ is close to 1, since Assumption 1 only requires player 2's payoff to be supermodular but does not say anything about player 2's preference over player 1's actions. Without this additional assumption, we can show that player 2's welfare is at most $\pi u_2(a_1^*, a_2^*) + (1-\pi)u_2(\underline{a}_1, \underline{a}_2)$ and is at least $\pi u_2(a_1^*, a_2^*) + (1-\pi)\min_{(\alpha,\beta)\in\mathbf{B}} u_2(\alpha, \beta)$ where $\mathbf{B} \subset \Delta\{\underline{a}_1, a_1^*\} \times \mathcal{B}$ such that $(\alpha, \beta) \in \mathbf{B}$ if and only if β best replies to α.

One caveat is that we measure welfare via $U_2 \equiv \sum_{k=0}^{+\infty}(1-\overline{\delta})\overline{\delta}^k u_2(a_{1,k}, a_{2,k})$, which is the *expected sum* of player 2s' payoffs. This means that *conditional on* the game will continue to the next period, the planner who evaluates welfare does *not* value the current player's payoff more than the next one's. Take the product choice game, the first t consumers receive their minmax value 0, where t is bounded above and below by linear functions of $(1-\delta)^{-1}$. Therefore, when the firm is patient, consumer welfare is low once it is evaluated by a planner who weights the current consumer's payoff much more than the ones in the future.

In terms of implications, under the standard interpretation of reputation models as well as the population-level interpretation in Section 2.1, Theorem 1 implies that a small fraction of opportunistic types who may erase actions from their records can significantly lower the long-run player's payoff, although it has a negligible effect on the short-run players' welfare. Nevertheless, both players will receive low payoffs in all equilibria when the honest type is rare or occurs with low probability. Under the interpretation that the long-run player *commits* to the short-run players that it will always take the commitment action a_1^* and will never erase any action, Theorem 1 implies that the long-run player's benefit from commitment is entirely wiped out as long as his opponents entertain a grain of doubt about his willingness to honor his commitment.

We provide some intuition for Theorem 1 before presenting the proof. In the benchmark where the long-run player *cannot* erase actions, the opportunistic type separates from the honest type right after he takes any action other than a_1^*. The Bayesian learning argument in Fudenberg and Levine (1989) implies that there exist at most a *bounded* number of 'bad periods' where (i) the short-run player's belief assigns a strictly positive probability to the honest type and (ii) the short-run player believes that the long-run player will take action a_1^* with probability lower than some cutoff $\gamma \in (0,1)$. Since the upper bound on the number of bad

[18] Our result can accommodate the usual order of limits in Fudenberg and Levine (1989) and others, which fix the type distribution π and send $\delta \equiv \widehat{\delta} \cdot \overline{\delta}$ to 1. Our Theorem 1 applies when $\delta \to 1$ as long as $\overline{\delta}$ goes to 1 at a faster rate than $\widehat{\delta}$ goes to 1.

periods does not depend on the long-run player's discount factor δ, a sufficiently patient player can secure approximately his commitment payoff $u_1(a_1^*, a_2^*)$ by taking his commitment action a_1^* in every period.

In our model, the opportunistic type may not separate from the honest type after he takes actions other than a_1^* since he can erase those actions. This makes it hard for the short-run players to distinguish the honest type from the opportunistic type. As a result, they may not be convinced that the long-run player will play a_1^* with high probability in the future even after many periods where (i) they believe that the long-run player will play a_1^* with low probability but (ii) observe that the long-run player has actually played a_1^*.

Hence, the opportunistic type's ability to erase actions *slows down* the process of reputation building. In order to convince the short-run players to play a_2^*, the long-run player needs to have a longer good record compared to that in the benchmark scenario where he cannot erase actions. This will lower his returns from building reputations. The extent to which the short-run players' learning is slowed down depends on player 1's incentive to erase actions, which is a key object to examine in our proof.

Recall the definition of h_*^k and recall that $\beta_k \in \Delta(A_2)$ denotes player 2's action at h_*^k. Proposition 2 implies that the opportunistic type has an incentive to play $\underline{a}_1$ at every on-path history of every equilibrium. Therefore, it is optimal for the opportunistic type to play $\underline{a}_1$ and then not erase it at history h_*^k *if and only if*

$$\underbrace{(1-\delta)u_1(\underline{a}_1, \beta_k)}_{\text{player 1's payoff from playing } \underline{a}_1 \text{ and not erasing it}} \geq \underbrace{u_1(\underline{a}_1, \beta_k) - c}_{\text{player 1's payoff from playing } \underline{a}_1 \text{ in every period and then erasing it}},$$

or equivalently,

$$u_1(\underline{a}_1, \beta_k) \leq c/\delta. \tag{3.2}$$

In any equilibrium where the opportunistic type finds it optimal *not* to erase $\underline{a}_1$ in period 0, his equilibrium payoff is no more than $(1-\delta)u_1(\underline{a}_1, \beta_0)$, which according to (3.2), is no more than $(1-\delta)c/\delta$.

In order to show the first part of Theorem 1, we only need to rule out equilibria in which the opportunistic type has a *strict* incentive to erase action $\underline{a}_1$ after taking it in period 0. Our proof uses a conflict caused by the long-run player's ability to erase actions, which is between (i) motivating the opportunistic type to take the commitment action a_1^* and (ii) motivating the short-run players to take actions greater than $\underline{a}_2$. The same conflict is also used to show the second part of the result, since it provides an upper bound on the maximal length of good record that the opportunistic type may have in any equilibrium. This in turn leads to a uniform upper bound on the *expected probability* with which the opportunistic type takes action a_1^*.

3.4 Proof of Theorem 1

Fix any $c < \overline{c}$, $\pi \in (0,1)$, and $\widehat{\delta} \in (0,1)$. Suppose by way of contradiction that for every $\delta^* \in (0,1)$, there exist $\overline{\delta} \in (\delta^*, 1)$ as well as an equilibrium (σ_1, σ_2) under $(c, \pi, \widehat{\delta}, \overline{\delta})$ such that the opportunistic type of player 1 has a *strict* incentive to erase action $\underline{a}_1$ after taking it in period 0.

First, we use the opportunistic type's incentive to take action a_1^* to derive a *lower bound* on the rate with which the opportunistic type's continuation value needs to increase as his good record becomes longer.

According to Proposition 2, there exists $t \in \mathbb{N}$ such that player 1's reputation is strictly less than 1 at h_*^{t-1} but reaches 1 at h_*^t. By definition, the opportunistic type takes action a_1^* with positive probability at h_*^k for every $k < t-1$. Let V_k denote player 1's continuation value at h_*^k. If the opportunistic type strictly prefers to erase $\underline{a}_1$ after taking it at h_*^0, then by Proposition 2, he will also have a strict incentive to do so at every h_*^k. This implies that at every h_*^k with $k < t-1$, the opportunistic type is indifferent between *playing* a_1^* and *playing* $\underline{a}_1$ *and then erasing it*. This leads to the following indifference condition:

$$V_k = u_1(\underline{a}_1, \beta_k) - c = (1-\delta) u_1(a_1^*, \beta_k) + \delta V_{k+1} \text{ for every } k < t-1. \tag{3.3}$$

Plugging $V_{k+1} = u_1(\underline{a}_1, \beta_{k+1}) - c$ into (3.3) and using the conclusion that $\beta_{k+1} \succ_{FOSD} \beta_k$, we have

$$u_1(\underline{a}_1, \beta_{k+1}) - u_1(\underline{a}_1, \beta_k) = (1-\delta)\Big(u_1(\underline{a}_1, \beta_{k+1}) - c - u_1(a_1^*, \beta_k)\Big) > (1-\delta)\Big(u_1(\underline{a}_1, \beta_k) - c - u_1(a_1^*, \beta_k)\Big).$$

Assumption 1 implies that $u_1(\underline{a}_1, \beta_{t-1}) \leq u_1(\underline{a}_1, a_2^*)$ and $u_1(\underline{a}_1, \beta_0) \geq u_1(\underline{a}_1, \underline{a}_2) \equiv 0$. Therefore,

$$t \leq \overline{T} \equiv \frac{u_1(\underline{a}_1, a_2^*)}{(1-\delta)\Delta}, \tag{3.4}$$

where $\Delta \equiv \min_{\beta \in \mathcal{B}} \big\{ u_1(\underline{a}_1, \beta) - u_1(a_1^*, \beta) - c \big\} > 0$. This upper bound $\overline{T}$ is linear in $(1-\delta)^{-1}$.

The second step uses player 2's incentives to derive a lower bound on t, which is strictly greater than the upper bound $\overline{T}$ as $\overline{\delta} \to 1$. One caveat is that player 2's *posterior belief* about player 1's type after observing player 1's record length being 0 may *not* be the same as her prior belief π. This is because the opportunistic type may erase actions from his records, so the length of player 1's record is a signal about his type.

We compute player 2's belief about player 1's action. For every $k \leq t-1$, let μ_k^* denote the probability of history h_*^k *conditional on* player 1 being the opportunistic type. Let q_k^* denote the probability that the opportunistic type plays $\underline{a}_1$ at h_*^k and then erases it. Recall that p_k^* is the probability that the opportunistic type plays a_1^* at h_*^k. Since player 1 exits the game with probability $1-\overline{\delta}$ after each period, we have:

$$\mu_0^* = (1-\overline{\delta}) + \overline{\delta}\mu_0^* q_0^* \quad \text{and} \quad \mu_k^* = \overline{\delta}(\mu_{k-1}^* p_{k-1}^* + \mu_k^* q_k^*) \text{ for every } k \in \{1,...,t-1\},$$

or

$$\mu_0^* = \frac{1-\overline{\delta}}{1-\overline{\delta}q_0^*} \quad \text{and} \quad \frac{\mu_k^*}{\mu_{k-1}^*} = \frac{\overline{\delta}p_{k-1}^*}{1-\overline{\delta}q_k^*} \text{ for every } k \in \{1,...,t-1\}. \tag{3.5}$$

Our hypothesis that the opportunistic type strictly prefers to erase action $\underline{a}_1$ after taking it at h_*^0 implies that $q_k^* = 1-p_k^*$ for every $k \in \{0,1,...,t-1\}$. Let x_k denote the probability with which player 2's belief at h_*^k assigns to player 1's current-period action being a_1^*. Recall that π is the probability that player 2's prior belief assigns to the honest type and let $l \equiv \frac{\pi}{1-\pi}$. According to Bayes rule, we have:

$$\frac{x_k}{1-x_k} = \frac{\pi(1-\overline{\delta})\overline{\delta}^k + (1-\pi)\mu_k^* p_k^*}{(1-\pi)\mu_k^*(1-p_k^*)},$$

or equivalently,

$$l(1-\overline{\delta})\overline{\delta}^k = \mu_k^*\Big\{\frac{x_k}{1-x_k}(1-p_k^*) - p_k^*\Big\} = \mu_k^* \frac{x_k - p_k^*}{1-x_k}. \tag{3.6}$$

Applying equation (3.6) to both h_*^k and h_*^{k-1}, we obtain that

$$\frac{\mu_k^*}{\mu_{k-1}^*} = \overline{\delta}\frac{x_{k-1}-p_{k-1}^*}{x_k - p_k^*} \cdot \frac{1-x_k}{1-x_{k-1}} \leq \overline{\delta}\frac{x_k - p_{k-1}^*}{x_k - p_k^*}, \tag{3.7}$$

where the last inequality comes from Proposition 2 that β_k is strictly increasing in k in the sense of FOSD, and by Assumption 1, x_k is weakly increasing in k. Equation (3.5) and inequality (3.7) together imply that

$$\overline{\delta}\frac{x_k - p_{k-1}^*}{x_k - p_k^*} \geq \frac{\mu_k^*}{\mu_{k-1}^*} = \frac{\overline{\delta}p_{k-1}^*}{1-\overline{\delta}q_k^*} = \frac{\overline{\delta}p_{k-1}^*}{1-\overline{\delta}(1-p_k^*)},$$

or equivalently,

$$p_{k-1}^* - p_k^* \leq (1-\overline{\delta})\frac{x_k - p_{k-1}^*}{x_k}(1-p_k^*) \leq 1-\overline{\delta}. \tag{3.8}$$

Since the opportunistic type plays a_1^* with zero probability at h_*^{t-1}, we have $p_{t-1}^* = 0$. This implies that

$$t \geq p_0^*(1-\overline{\delta})^{-1}. \tag{3.9}$$

We show that there exists $p^* \in (0,1)$ such that $p_0^* > p^*$ for all $\overline{\delta}$ close to 1. In order to derive this uniform lower bound for p_0^*, take $k = 0$ in (3.6) and replace μ_0^* with its expression in (3.5), we have

$$\frac{\pi}{1-\pi} = l = \frac{1}{1-\overline{\delta} q_0^*} \cdot \frac{x_0 - p_0^*}{1-x_0} = \frac{1}{1-\overline{\delta}(1-p_0^*)} \cdot \frac{x_0 - p_0^*}{1-x_0}. \tag{3.10}$$

The value of x_0 is bounded above 0 since $u_1(\underline{a}_1, \beta_0) > c/\delta$ requires β_0 to assign positive probability to some action strictly greater than $\underline{a}_2$. The RHS of (3.10) converges to $\frac{1}{1-\overline{\delta}} \cdot \frac{x_0}{1-x_0}$ as $p_0^* \to 0$, and there exists $\delta^* \in (0,1)$ such that $\frac{1}{1-\overline{\delta}} \cdot \frac{x_0}{1-x_0} > \frac{\pi}{1-\pi}$ for every $\overline{\delta} > \delta^*$. This implies that p_0^* must be bounded above 0.

Therefore, the lower bound on t, implied by player 2's incentive constraints, is a linear function of $(1-\overline{\delta})^{-1}$. The upper bound on t in (3.4), implied by the opportunistic type's incentive constraints, is a linear function of $(1-\delta)^{-1}$. The lower bound exceeds the upper bound once we fix any $\widehat{\delta}$ and let $\overline{\delta} \to 1$. This leads to a contradiction. Therefore, in every equilibrium, the opportunistic type has an incentive *not* to erase action $\underline{a}_1$ after taking it at h_*^0, so his equilibrium payoff must be no more than $(1-\delta)c/\delta$.

Next, we bound player 2's equilibrium welfare in three steps. In the first step, we show that there exist a constant $m \in \mathbb{N}$ that depends only on (π, u_1, u_2) as well as a constant $\lambda > 0$ that depends only on (u_1, u_2) such that $t \leq m + \lambda(1-\delta)^{-1}$. Recall the definitions of t_0 and t in Proposition 2. Since player 1's continuation value at $h_*^{t_0}$ is at least 0, the opportunistic type's incentive to take action a_1^* from $h_*^{t_0}$ to h_*^{t-2} implies that $t - t_0$ is bounded above by a linear function of $(1-\delta)^{-1}$. At every h_*^k with $k < t_0$, player 1 has no incentive to erase actions at h_*^k and player 2 takes actions other than a_2^* with strictly positive probability. The learning argument in Fudenberg and Levine (1989) implies that for every (π, u_1, u_2), there exists $m \in \mathbb{N}$ such that $t_0 \leq m$. The two parts together imply that $t \leq m + \lambda(1-\delta)^{-1}$.

In the second step, we bound the short-run players' welfare from below, which uses the additional assumption that $u_2(a_1, a_2)$ is weakly increasing in a_1. Notice that our reputation game is equivalent to one in which a hypothetical planner chooses actions on behalf of the short-run players in order to maximize $\sum_{k=0}^{+\infty}(1-\overline{\delta})\overline{\delta}^k u_2(a_{1,k}, a_{2,k})$, the planner's action in period k depends only on short-run player k's history, and the long-run player cannot observe any action chosen by the planner. Since the honest type reaches h_*^t in period t and the opportunistic type never reaches h_*^t, the planner can secure the following payoff for the short-run players when he takes action a_2^* at every h_*^k with $k \geq t$ and takes action $\underline{a}_2$ at any other history:

$$\Big\{(1-\overline{\delta}^t)\pi + (1-\pi)\Big\} u_2(\underline{a}_1, \underline{a}_2) + \overline{\delta}^t \pi u_2(a_1^*, a_2^*). \tag{3.11}$$

Since $t \leq m + \lambda(1-\delta)^{-1}$, we know that for every $\widehat{\delta} \in (0,1)$ and $\varepsilon > 0$, there exists $\delta^* \in (0,1)$ such that when $\overline{\delta} > \delta^*$, the right-hand-side of (3.11) is at least $\pi u_2(a_1^*, a_2^*) + (1-\pi)u_2(\underline{a}_1, \underline{a}_2) - \varepsilon$.

In the third step, we bound the short-run players' welfare from above by deriving an upper bound on the opportunistic type's *expected behavior*. For every $k \in \mathbb{N}$, we define event $\mathcal{E}^k$ as

$$\mathcal{E}^k \equiv \{\text{the current history is } h_*^k \text{ and player 1 plays } a_1^* \text{ in the current period}\}.$$

Lemma 3. *In every equilibrium and for every $k \in \mathbb{N}$, the probability of event $\mathcal{E}^k$ is no more than $(1-\overline{\delta})\overline{\delta}^k$ conditional on player 1 being the opportunistic type.*

Proof. By definition, the probability of event $\mathcal{E}^k$ conditional on player 1 being the opportunistic type is $\mu_k^* p_k^*$. We show by induction that $\mu_k^* p_k^* \leq (1-\overline{\delta})\overline{\delta}^k$ for every $k \in \mathbb{N}$. The first part of (3.5) implies that:

$$\mu_0^* p_0^* = \frac{(1-\overline{\delta})p_0^*}{1-\overline{\delta}q_0^*} \leq \frac{(1-\overline{\delta})p_0^*}{1-\overline{\delta}(1-p_0^*)} \leq 1-\overline{\delta} \text{ for every } p_0^* \in [0,1].$$

Next, suppose $\mu_{j-1}^* p_{j-1}^* \leq (1-\overline{\delta})\overline{\delta}^{j-1}$ for some $j \geq 1$, then the second part of (3.5) implies that

$$\mu_j^* p_j^* = \frac{\overline{\delta} p_j^* p_{j-1}^* \mu_{j-1}^*}{1-\overline{\delta}q_j^*} \leq \frac{\overline{\delta} p_j^* p_{j-1}^* \mu_{j-1}^*}{1-\overline{\delta}(1-p_j^*)} \leq \overline{\delta}^j \frac{(1-\overline{\delta})p_j^*}{1-\overline{\delta}(1-p_j^*)} \leq (1-\overline{\delta})\overline{\delta}^j \text{ for every } p_j^* \in [0,1].$$

This completes the proof of the lemma. ☐

Conditional on player 1 being the opportunistic type, the probability of event $\mathcal{E}^k$ is 0 for every $k \geq t$ since the opportunistic type never reaches h_*^t. The *ex ante* probability that the opportunistic type takes action a_1^* is $\sum_{k=0}^{+\infty} \Pr(\mathcal{E}^k)$, which by Lemma 3, is no more than $1-\overline{\delta}^t$ in any equilibrium. Fix any equilibrium (σ_1, σ_2) as well as the resulting distribution over player 1's actions. Player 2's payoff is no more than her payoff when she can observe player 1's *realized pure action* before choosing her action, which is at most

$$(1-\pi)\Big\{1-(1-\overline{\delta}^t)\Big\}u_2(\underline{a}_1, \underline{a}_2) + \Big\{\pi + (1-\overline{\delta}^t)(1-\pi)\Big\}u_2(a_1^*, a_2^*). \tag{3.12}$$

Since t is bounded above by a linear function of $(1-\delta)^{-1}$, for every $\widehat{\delta} \in (0,1)$ and $\varepsilon > 0$, there exists $\delta^* \in (0,1)$ such that when $\overline{\delta} > \delta^*$, the value of (3.12) is no more than $\pi u_2(a_1^*, a_2^*) + (1-\pi)u_2(\underline{a}_1, \underline{a}_2) + \varepsilon$.

3.5 Additional Implications of Theorem 1

Theorem 1 implies the following two corollaries. The first one examines the case where $\widehat{\delta}$ and $\overline{\delta}$ are fixed.

Corollary 1. *For every $c < \overline{c}$, $\widehat{\delta} \in (0,1)$, and $\overline{\delta} \in (0,1)$, there exists $\overline{\pi} \in (0,1)$ such that when $\pi \in (0, \overline{\pi})$, player 1's payoff is no more than $(1-\delta)c/\delta$ in every equilibrium.*

The proofs of this corollary and the next can be found in Appendix E. Corollary 1 shows that the patient player's payoff is close to his minmax value when the probability of honest type is lower than some cutoff $\overline{\pi}$, even when his discount factor δ is arbitrarily close to 1. This stands in contrast to the main result in Fudenberg and Levine (1989), which shows that for any $\pi > 0$, the patient player can secure his commitment payoff $u_1(a_1^*, a_2^*)$ in all equilibria when δ approaches 1. More generally, one can show that whether reputation effects will fail *when π is close to* 1 depends on the comparison between (i) the rate with which $\overline{\delta} \to 1$ and (ii) the rate with which $\delta \equiv \overline{\delta} \cdot \widehat{\delta} \to 1$: Once we fix the patient player's time preference $\widehat{\delta}$ and increase his expected lifespan, or more generally, when $\frac{1-\delta}{1-\overline{\delta}}$ diverges to infinity, the patient player receives (approximately) his minmax value in all equilibria even when π is arbitrarily close to 1. If $\frac{1-\delta}{1-\overline{\delta}}$ does *not* diverge to infinity, then the patient player receives his minmax value when π is lower than some cutoff.

The next corollary provides sufficient conditions under which player 1's equilibrium payoff is uniquely pinned down as well as sufficient conditions under which there is a unique equilibrium outcome. Let π^* denote the lowest $\widetilde{\pi} \in [0, 1]$ such that player 2 has an incentive to play a_2^* when player 1 takes a mixed action $\widetilde{\pi} a_1^* + (1-\widetilde{\pi})\underline{a}_1$. Let $\underline{\delta} \equiv \frac{c}{u_1(\underline{a}_1, a_2^*)}$, which is strictly between 0 and 1 given that $u_1(\underline{a}_1, a_2^*) > u_1(a_1^*, a_2^*) > 0$ and $c < u_1(\underline{a}_1, a_2^*) - u_1(a_1^*, a_2^*)$. In the product choice game, we have $\pi^* = x$ and $\underline{\delta} = \frac{c}{1+g}$.

Corollary 2 shows that when the fraction of honest types π is more than π^* and player 1's effective discount factor $\delta \equiv \overline{\delta} \cdot \widehat{\delta}$ is greater than $\underline{\delta}$, player 1's payoff is exactly $(1-\delta)c/\delta$ in every equilibrium when he has a sufficiently long lifespan, and that generically, there is a unique equilibrium outcome.

Corollary 2. *For every $c < \overline{c}$, $\pi > \pi^*$, and $\widehat{\delta} \in (0, 1)$, there exists $\delta^* \in (0, 1)$ such that player 1's payoff in every equilibrium equals $\frac{(1-\delta)c}{\delta}$ for every $\overline{\delta} \in \big(\max\{\delta^*, \underline{\delta}/\widehat{\delta}\}, 1\big)$ and all equilibria lead to the same outcome for generic $\overline{\delta} \in \big(\max\{\delta^*, \underline{\delta}/\widehat{\delta}\}, 1\big)$.*

The proof is in Appendix E. Although the equilibrium outcome is not necessarily unique under every parameter configuration, all equilibria share the same qualitative features (Proposition 2 and Theorem 1), so in some sense, they are all *close*. Section 5.1 discusses the sources of multiplicity in our model.

4 Extension: Limiting the Honest Type's Record

In our baseline model, when player 1 has a sufficiently long lifespan, the honest type will have a long record in expectation. This implies that player 1 will have a low reputation when he has a short record and hence will receive a low payoff. One natural question is that whether player 1 can obtain higher payoffs when the honest type does *not* reveal all his past actions to the short-run players, either because he commits to reveal a shorter record or his record length is limited by some third party, such as an online platform.

In the spirit of commitment types in the reputation literature, we assume that the honest type takes action a_1^* in every period, and in addition, he commits to disclose information about his past actions according to a *disclosure policy*. Our analysis and conclusion also apply to the case where each short-run player can observe at most $\widetilde{K} \in \mathbb{N}$ actions, where $\widetilde{K}$ can be any random variable that takes values in $\mathbb{N} \cup \{+\infty\}$. This is relevant, for example, when a platform commits to reveal at most $\widetilde{K}$ unerased actions from each seller.[19]

We restrict attention to *disclosure policies* that take the form of a mapping $\mathbf{q} \equiv (q_0, q_1, q_2, ...)$ such that if the honest type's record length is $m \in \mathbb{N}$, then (i) with probability $q_m(n)$, he reveals a record with length $n \in \mathbb{N}$ consisting only of a_1^*, and (ii) with probability $q_m(s^*)$, he reveals a public signal s^* that is infeasible for the opportunistic type to generate. This class of disclosure policies includes disclosing the last K actions, randomizing between disclosing all past actions and disclosing no past action, and so on.

Compared to general disclosure policies that map the honest type's action history to a distribution over public signals, we only rule out disclosure policies which require the honest type to *fabricate records*, by which we mean revealing records that contain actions that he has never taken in the past. The motivation is that fabricating records and lying to the consumers about their past actions are against the spirit of *honest* types. It also sounds unrealistic for firms to tell consumers that it has supplied low quality while it has actually supplied high quality. We leave the study of such disclosure policies for future research.

There are other reasonable restrictions on the set of disclosure policies that the honest type can commit to, such as (i) the length of record he reveals cannot exceed the number of actions that he has taken and (ii) the honest type cannot generate any signal that the opportunistic type cannot generate, such as signal s^*. We do not include any additional restriction since our result applies to *all* disclosure policies that belong to our class. That is to say, the result is stronger when it is stated without any additional restriction.

We start from two useful benchmarks. First, in an auxiliary game where player 2 *cannot* observe any information about player 1's past actions, the opportunistic type will never take any action that is strictly greater than $\underline{a}_1$ and will never erase any action. For every $\pi \in (0, 1)$, let $\overline{a}_2^\pi$ denote player 2's *highest* best reply to player 1's mixed action $\pi a_1^* + (1 - \pi)\underline{a}_1$ and let $\underline{a}_2^\pi$ denote player 2's *lowest* best reply to $\pi a_1^* + (1 - \pi)\underline{a}_1$. The opportunistic type's equilibrium payoff is between $u_1(\underline{a}_1, \underline{a}_2^\pi)$ and $u_1(\underline{a}_1, \overline{a}_2^\pi)$.

Next, suppose player 2 can observe the signal revealed by the honest type as well as the history of the opportunistic type's unerased actions, but the honest type commits to reveal the null history regardless of his past actions, that is, $q_m(0) = 1$ for every $m \in \mathbb{N}$. Lemma 4 derives lower and upper bounds on the opportunistic type's equilibrium payoff in this benchmark.

Lemma 4. *If* $q_m(0) = 1$ *for every* $m \in \mathbb{N}$*, then the opportunistic type's payoff in every equilibrium is at least* $\max\{0, u_1(\underline{a}_1, \underline{a}_2^\pi) - c\}$ *and is at most* $\max\{\frac{1-\delta}{\delta}c, u_1(\underline{a}_1, \overline{a}_2^\pi) - c\}$.

[19]This is because when the honest type's record length is at most K, either due to his commitment or due to some physical constraints imposed by the platform or by other planners, the opportunistic type has no incentive to take action a_1^* for more than K times in any equilibrium, even when it is feasible for the opportunistic type to do that.

The proof is in Appendix F. The intuition is that the opportunistic type will take action $\underline{a}_1$ in every period, and whether he will erase that action depends on the short-run player's action when they observe a record of length 0. For generic π, player 2 has a unique best reply to $\pi a_1^* + (1-\pi)\underline{a}_1$, in which case $\underline{a}_2^\pi = \overline{a}_2^\pi$ and the payoff lower and upper bounds derived in Lemma 4 coincide as $\delta \to 1$.

Theorem 2 shows that as long as player 1 is sufficiently long-lived, regardless of the honest type's disclosure policy (i) the opportunistic type's payoff in any equilibrium cannot be significantly greater than his highest equilibrium payoff when the honest type discloses no information and (ii) the short-run players' welfare cannot be significantly greater than that when the honest type discloses all his actions.

Theorem 2. *For every $c < \overline{c}$, π, $\widehat{\delta}$, and $\varepsilon > 0$, there exists $\delta^* \in (0,1)$ such that for every $\overline{\delta} > \delta^*$, the opportunistic type's payoff in any equilibrium under any of the honest type's disclosure policy is at most*

$$\max\Big\{\frac{1-\delta}{\delta}c,\ u_1(\underline{a}_1, \overline{a}_2^\pi) - c\Big\}. \tag{4.1}$$

If in addition that $u_2(a_1, a_2)$ is weakly increasing in a_1, then player 2's welfare U_2 in any equilibrium under any disclosure policy is no more than $\pi u_2(a_1^*, a_2^*) + (1-\pi)u_2(\underline{a}_1, \underline{a}_2) + \varepsilon$.

Theorem 2 implies that as long as the short-run players suspect that the long-run player can erase his actions at a low cost and believe that the honest type will never fabricate records, allowing them to observe the game's history can no longer benefit the long-run player relative to the two benchmarks where either the honest type reveals no information, or the short-run players cannot monitor the long-run player's behavior. This result neither implies nor is implied by Theorem 1, since Theorem 1 derives a stronger payoff upper bound under full disclosure while Theorem 2 derives a weaker payoff upper bound but allows for a larger class of disclosure policies. Our conclusion stands in contrast to usual intuition in the theories of repeated games and reputations, that the patient player can obtain higher payoffs when his opponents can monitor his actions compared to the case in which his actions cannot be monitored at all.

Compared to the baseline model where the honest type discloses his entire history to the short-run players, Theorem 2 implies that policies that limit the honest type's record length *cannot* improve the short-run players' welfare. Nevertheless, they can *strictly* benefit the opportunistic type *if and only if* the short-run players assign a high enough probability to the honest type so that the opportunistic type's payoff is greater than his minmax value in the auxiliary game where the honest type discloses no information.

Under our imperfect commitment interpretation, Theorem 2 implies that the long-run player may benefit from committing to non-full disclosure *only when* his opponents believe that he will honor his commitment with a high enough probability. As we mentioned earlier, our conclusion also applies to policies imposed by third parties, such as an online platform committing to reveal at most K unerased actions from each seller to the consumers, in which case it is infeasible for any type to reveal more than K actions.

The proof is in Appendix G. Theorem 2 does not follow from the standard arguments in the Bayesian persuasion literature, such as the one in Kamenica and Gentzkow (2011). This is because unlike those models where the uninformed player's payoff depends only on some *exogenous state*, their payoff in our model depends only on the patient player's *endogenous action*.[20] Since Proposition 2 implies that the opportunistic type will take action a_1^* with positive probability and that the probability with which he takes action a_1^* depends on the disclosure policy, it is not straightforward that the opportunistic type cannot obtain higher payoffs under any of the honest type's disclosure policy relative to the no disclosure benchmark.

Our proof uses Lemma 3 in Section 3.3 that *conditional on* player 1 being the opportunistic type, the probability of event $\mathcal{E}^k$ is no more than $1-\overline{\delta}$ for every $k \leq t-1$. This lemma applies under *all* disclosure policies of the honest type since it is a statement about the probabilities *conditional on player* 1 *being the opportunistic type* and the proof of that result does *not* rely on any player's incentive constraint.

As in the baseline model, the opportunistic type's incentive to play a_1^* implies that (i) the short-run player's action increases in FOSD with the length of the long-run player's good record, and (ii) the maximal length of good record that the opportunistic type may have in any equilibrium is bounded above by a linear function of $(1-\delta)^{-1}$. If the opportunistic type's equilibrium payoff is strictly more than (4.1), then (i) the short-run players need to take actions strictly greater than $\overline{a}_2^{\pi}$ with positive probability at every on-path history that belongs to $\mathcal{H}_*$, and (ii) the opportunistic type will never take $\underline{a}_1$ and then not erase it, since his payoff in any equilibrium where he does so with positive probability is no more than $(1-\delta)c/\delta$. The second requirement is satisfied *only if* the opportunistic type never reaches histories that do not belong to $\mathcal{H}_*$.

Therefore, a *necessary* condition for the existence of such an equilibrium is that player 2 has an incentive to take an action that is strictly greater than $\overline{a}_2^{\pi}$ under her *prior belief about player* 1*'s action*. The definition of $\overline{a}_2^{\pi}$ implies that player 2 has no incentive to take actions strictly greater than $\overline{a}_2^{\pi}$ under her prior belief about player 1's action *unless* in expectation, the opportunistic type will take action a_1^* with probability bounded above 0. However, conditional on player 1 being the opportunistic type, the probability of every $\mathcal{E}^k$ is no more than $(1-\overline{\delta})\overline{\delta}^k$ for every $k \leq t$ and the probability of $\mathcal{E}^k$ is 0 for every $k > t$. Therefore, persuading the short-run players to take actions greater than $\overline{a}_2^{\pi}$ requires t to be at least proportional to $(1-\overline{\delta})^{-1}$.

Fix any $\widehat{\delta} \in (0,1)$ and as $\overline{\delta} \to 1$, the upper bound on t implied by the need to motivate the opportunistic type will be strictly lower than the lower bound on t implied by the need to persuade the short-run players to

[20] More closely related is Bhaskar and Thomas (2019), who study the design of disclosure policies in a repeated *complete information* game, under the restriction that the short-run players only have information about the long-run player's last K actions. They examine whether cooperation can be sustained in *one* purifiable equilibrium while the current paper focuses on the common properties of *all* equilibria. As a result, the techniques used in their paper cannot be directly applied to our setting and vice versa.

take actions strictly greater than $\overline{a}_2^{\pi}$. This contradiction rules out equilibria in which player 1's payoff being strictly greater than (4.1), regardless of the disclosure policy that the honest type commits to.

The short-run players' welfare is no more than their welfare in the case where the honest type reveals all his past actions. This is because regardless of the honest type's disclosure policy, the expected probability with which the opportunistic type takes the commitment action vanishes to 0 as $\overline{\delta}$ goes to 1.

5 Concluding Remarks

This paper takes a first step to analyze reputation effects when a patient player's record length is determined endogenously by his own behavior. Since the patient player can signal his honesty via the length of his good record, he will have an incentive to build a reputation even when he can erase actions from his records at a low cost and that he can secure a high payoff after accumulating a long enough good record.

However, as long as the patient player has a sufficiently long lifespan, his equilibrium payoff must be close to his minmax value and the short-run players' welfare must be close to their payoff when they can observe the patient player's type and the opportunistic type always takes the lowest action. Even when the honest type of the patient player can disclose information about his history according to any disclosure policy, the opportunistic type's payoff is no more than his equilibrium payoff in the auxiliary game where the honest type discloses no information, and the one in which the short-run players receive no information.

Our results imply that (i) the possibility of erasing actions *cannot* eliminate patient players' incentives to build reputations, (ii) a small probability of opportunistic type can entirely wipe out the patient player's returns from building reputations, although it has a negligible effect on the short-run players' welfare, (iii) when a sufficiently long-lived player commits to his opponents that he will always take some cooperative action and will never erase any action, the value of such a commitment will be seriously compromised as long as his opponents entertain a grain of doubt about his willingness to honor this commitment, and (iv) policies that limit the long-run player's record length cannot improve the short-run players' welfare and can improve the long-run player's welfare only when he is believed to be honest with a high enough probability.

We conclude the paper with a discussion of our baseline model as well as a list of extensions. Section 5.1 discusses the sources of multiplicity in our baseline model. Section 5.2 studies an extension where the honest type can strategically decide whether to erase his action at a strictly positive cost. Sections 5.3 and 5.4 show that our results are robust when the consumers arrive stochastically, or post reviews with probability strictly between 0 and 1 after interacting with the seller. The extension to multiple honest types and multiple opportunistic types are relegated to Online Appendix B.

5.1 The Uniqueness and Multiplicity of Equilibrium Outcomes

Corollary 2 provides sufficient conditions under which the equilibrium outcome is unique in our baseline model. We comment on the sources of multiplicity in our model when these conditions are violated.

Recall the definition of $\mathcal{B}$ in (3.1) and that by Lemma 2, every pair of elements in $\mathcal{B}$ can be ranked according to FOSD. Fix any $\beta_0 \in \mathcal{B}$, player 2's actions when player 1 has a positive reputation, $\beta_1, \beta_2, ..., \beta_t$, are pinned down by player 1's indifference condition at h_*^k for $k \in \{0, 1, 2, ..., t-1\}$:

$$\underbrace{\max\Big\{u_1(\underline{a}_1, \beta_k) - c, (1-\delta)u_1(\underline{a}_1, \beta_k)\Big\}}_{\text{player 1's continuation value at } h_*^k} = (1-\delta)u_1(a_1^*, \beta_k) + \delta \underbrace{\max\Big\{u_1(\underline{a}_1, \beta_{k+1}) - c, (1-\delta)u_1(\underline{a}_1, \beta_{k+1})\Big\}}_{\text{player 1's continuation value at } h_*^{k+1}}.$$

This recursive process also pins down the value of t since β_{t-1} must be weakly lower than a_2^* but must be high enough so that player 1 does not have a strict incentive to play a_1^* at h_*^{t-1} even when $\beta_t = a_2^*$.

Let $\beta^\dagger \in \mathcal{B}$ be such that $u_1(\underline{a}_1, \beta^\dagger) = c/\delta$. When c satisfies (2.1), such an action exists when $\delta > \frac{c}{u_1(\underline{a}_1, a_2^*)}$ and is unique by Lemma 2. One can also show that $\beta^\dagger$ is nontrivially mixed when δ is large enough. We consider two cases. If there exists a pure action $\beta \in \mathcal{B}$ such that $u_1(\underline{a}_1, \beta) = c$, then $\beta^\dagger$ must be nontrivially mixed when δ is close to 1. If the unique $\beta \in \mathcal{B}$ that satisfies $u_1(\underline{a}_1, \beta) = c$ is nontrivially mixed, then a continuity argument implies that $\beta^\dagger$ is also nontrivially mixed for every δ close enough to 1.

When $\overline{\delta}$ and $\widehat{\delta}$ are bounded below 1, player 2's action in the first period can be bounded below $\beta^\dagger$. If her action in period 0 is pure, then there are multiple probabilities with which player 1 can play $\underline{a}_1$ in period 0, leading to a multiplicity of equilibrium outcomes.

Fix any π. When δ is close enough to 1, it must be the case that $\beta_0 = \beta^\dagger$ or β_0 is close to $\beta^\dagger$. This is because the speed with which β increases in t is proportional to $1-\delta$ and similar to Fudenberg and Levine (1989), the speed with which player 1's reputation increases when $\beta < \beta^\dagger$ is bounded above 0. If it takes too many periods for β to reach $\beta^\dagger$, then player 1's reputation will exceed 1 before β reaches $\beta^\dagger$, which will lead to a contradiction. If π is small enough such that β_0 is strictly lower than $\beta^\dagger$, then even when both $\beta^\dagger$ and β_0 are nontrivially mixed, there may exist multiple values of β_0 in equilibrium, which is another source of multiplicity. However, as long as δ is close to 1, β_0 must be close to $\beta^\dagger$, and the values of $\beta_1^*, ..., \beta_{t-1}^*$, $p_0^*, ..., p_{t-1}^*$, and $\mu_0^*, ..., \mu_{t-1}^*$ are also close across different equilibria.

When π is above some cutoff and δ is close to 1, we can show that $\beta_0 = \beta^\dagger$ in all equilibria, from which we can pin down the values of t as well as $\beta_1, \beta_2, ..., \beta_{t-1}$. If all of $\beta_0, ..., \beta_{t-1}$ are nontrivially mixed, which happens under generic δ, then the conclusion that $p_{t-1}^* = 0$ as well as player 2's indifference conditions pin down the values of $p_0^*, ..., p_{t-1}^*$ and $\mu_0^*, ..., \mu_{t-1}^*$. When some actions in $\{\beta_0, \beta_1, ..., \beta_{t-1}\}$ are pure, there are multiple actions of player 1's under which player 2 has an incentive to play that pure action. This implies that there are multiple $p_0^*, ..., p_{t-1}^*$ and $\mu_0^*, ..., \mu_{t-1}^*$ that satisfy player 2's incentive constraints, leading to multiple equilibrium outcomes. However, even at these degenerate values of δ where multiple equilibrium outcomes occur, the equilibrium values of $p_0^*, ..., p_{t-1}^*$ and $\mu_0^*, ..., \mu_{t-1}^*$ are pinned down except for periods in which player 2 takes a pure action in equilibrium.

5.2 Honest Types Erasing Actions

This section studies an extension where the honest type mechanically takes action a_1^* in every period but can strategically decide whether to erase his actions in order to maximize his discounted average payoff

$$\sum_{k=0}^{+\infty}(1-\delta)\delta^k\Big\{u_1(a_1^*, a_{2,k}) - c_k\Big\}.$$

Proposition 3. *If $c > 0$, then the honest type never erases his action in any equilibrium.*

The proof is in Appendix H. This result does not rely on the honest type sharing the same stage-game payoff as the opportunistic type. It remains valid when the honest type's stage-game payoff $\widetilde{u}_1(a_1, a_2)$ is different from that of the opportunistic type's $u_1(a_1, a_2)$, as long as $\widetilde{u}_1(a_1, a_2)$ is strictly increasing in a_2.

5.3 Stochastic Arrivals

This section discusses an extension in which the short-run players arrive *stochastically* over time. Suppose in each period, a short-run player arrives with some exogenous probability $p \in (0, 1)$. If a short-run player arrives, then players play the stage game, and by the end of that period, the long-run player decides whether to erase his action. If no short-run player arrives in a given period, then the long-run player's record remains the same regardless of his behavior in that period and his stage-game payoff is normalized to 0.

In this setting, the opportunistic type maximizes $p\sum_{k=0}^{+\infty}(1-\delta)\delta^k\Big(u_1(a_{1,k}, a_{2,k}) - c_k\Big)$. The short-run players' prior belief assigns probability $(1-\overline{\delta})\overline{\delta}^k$ to the long-run player's *age in the game* being k and assigns probability $(1-\widetilde{\delta})\widetilde{\delta}^k$ to the honest type having interacted with k short-run players, where

$$\widetilde{\delta} \equiv 1 - \frac{1-\overline{\delta}}{1-\overline{\delta}(1-p)}. \tag{5.1}$$

Using the same method as in the proof of Theorem 1, one can show that for every $c < \overline{c}$, $p \in (0, 1)$, $\pi \in (0, 1)$, $\widehat{\delta} \in (0, 1)$, and $\varepsilon > 0$, there exists $\delta^* \in (0, 1)$ such that when $\overline{\delta} > \delta^*$, the opportunistic type's payoff is no more than

$$u(p) \equiv \frac{(1-\delta)(1-\delta+\delta p)}{\delta}c \tag{5.2}$$

in every equilibrium. This payoff converges to his minmax value 0 as $\delta \to 1$. Under an additional assumption that $u_2(a_1, a_2)$ is weakly increasing in a_1, the short-run players' welfare, measured by the expected sum of their payoffs, is ε-close to $p\pi u_2(a_1^*, a_2^*) + p(1-\pi)u_2(\underline{a}_1, \underline{a}_2)$ in every equilibrium.

When the honest type also commits to an information disclosure policy in addition to committing to play a_1^* in every period, or when a third party (e.g., an online platform) limits the length of record the long-run player can disclose to the short-run players, one can show a generalized version of Theorem 2 that when $c < \overline{c}$ and $\widehat{\delta} \in (0,1)$, there exists $\delta^* \in (0,1)$ such that when $\overline{\delta} > \delta^*$, the opportunistic type's payoff in any equilibrium under any disclosure policy of the honest type is at most $\max\{u(p), p(u_1(\underline{a}_1, \overline{a}_2^\pi) - c)\}$, where $\overline{a}_2^\pi$ is player 2's highest best reply to mixed action $\pi a_1^* + (1-\pi)\underline{a}_1$. Under the additional assumption that $u_2(a_1, a_2)$ is weakly increasing in a_1, the short-run players' welfare is no more than their equilibrium welfare when the honest type reveals all of his past actions.

We briefly explain how to extend the proofs of Theorems 1 and 2 to the case with stochastic arrivals, with details available upon request. To show the generalized version of Theorem 1, suppose by way of contradiction that there exists an equilibrium in which player 1 strictly prefers to erase action $\underline{a}_1$ after taking it at every h_*^k. On the one hand, player 1 has an incentive to take action a_1^* at history h_*^k only if

$$(1-\delta)u_1(a_1^*, \beta_k) + \delta p(u_1(\underline{a}_1, \beta_{k+1}) - c) \geq (1-\delta)u_1(\underline{a}_1, \beta_k) + \delta p(u_1(\underline{a}_1, \beta_k) - c).$$

A necessary condition for the above inequality is that player 1's continuation value increases by something linear in $1-\delta$ when his record length increases from k to $k+1$. This leads to an upper bound on the number of periods with which the opportunistic type can take action a_1^*, which is a linear function of $(1-\delta)^{-1}$.

On the other hand, the argument in the proof of Theorem 1 implies that the rate with which player 1's reputation increases is bounded above by something proportional to $1-\widetilde{\delta}$. This implies that the number of periods with which the opportunistic type needs to take action a_1^* is at least a linear function of $(1-\widetilde{\delta})^{-1}$. As $\overline{\delta} \to 1$, expression (5.1) implies that $\widetilde{\delta}$ also goes to 1, and the lower bound on t will exceed the upper bound. Therefore, in every equilibrium, the opportunistic type has an incentive not to erase $\underline{a}_1$ at h_*^0.

We bound player 1's payoff in equilibria where he has an incentive *not to erase* $\underline{a}_1$ at history h_*^0. Suppose a short-run player arrives at history h_*^0 and that the opportunistic type took action $\underline{a}_1$ at that history, the opportunistic type prefers not to erase $\underline{a}_1$ if

$$(1-\delta)u_1(\underline{a}_1,\beta_0) \geq (1-\delta)(u_1(\underline{a}_1,\beta_0)-c)+\delta p(u_1(\underline{a}_1,\beta_0)-c),$$

or equivalently,

$$u_1(\underline{a}_1,\beta_0) \leq \frac{(1-\delta)(1-\delta+\delta p)}{\delta p}c = u(p). \tag{5.3}$$

In order to bound the short-run players' payoffs and to show the generalized version of Theorem 2, we establish a generalized version of Lemma 3 that the probability of event $\mathcal{E}_k$ is no more than $1-\widetilde{\delta}$.

The short-run players' equilibrium welfare is no less than their payoff when they take action a_2^* if and only if the length of player 1's good record exceeds t, and takes action $\underline{a}_2$ otherwise. This lower bound converges to $p\pi u_2(a_1^*,a_2^*)+p(1-\pi)u_2(\underline{a}_1,\underline{a}_2)$ as $\widetilde{\delta}\to 1$, since the *expected* number of periods for the honest type to obtain a record length t is a linear function of $(1-\delta)^{-1}$, which is lower than the decay rate of their belief $\widetilde{\delta}$. Their equilibrium welfare is no more than their payoff when they can observe the realized pure action of player 1. This upper bound converges to $p\pi u_2(a_1^*,a_2^*)+p(1-\pi)u_2(\underline{a}_1,\underline{a}_2)$ as $\widetilde{\delta}\to 1$, since the average probability with which the opportunistic type takes action a_1^* vanishes as $\widetilde{\delta}\to 1$. The lower and the upper bounds coincide, which pins down the short-run players' equilibrium welfare.

For the generalization of Theorem 2, in order for the opportunistic type to obtain payoff strictly greater than $\max\{u(p),p(u_1(\underline{a}_1,\overline{a}_2^{\pi})-c)\}$ in an equilibrium, he must strictly prefer to erase $\underline{a}_1$ at every on-path history and the short-run players must have an incentive to take some action that is strictly greater than $\overline{a}_2^{\pi}$ at h_*^0 as well as at every other on-path history. The necessity to provide the opportunistic type an incentive to take action a_1^* implies that the length of the opportunistic type's good record must be bounded above by some linear function of $(1-\delta)^{-1}$. Since the probability of event $\mathcal{E}_k$ is no more than $1-\widetilde{\delta}$, the need to persuade the short-run players to take actions greater than $\overline{a}_2^{\pi}$ leads to a lower bound on t, which is a linear function of $(1-\widetilde{\delta})^{-1}$. As $\overline{\delta}\to 1$, one cannot simultaneously provide incentives to the opportunistic type and to the short-run players. This rules out equilibria in which player 1's payoff exceeds $\max\{u(p),p(u_1(\underline{a}_1,\overline{a}_2^{\pi})-c)\}$ under any of the honest type's disclosure policy.

5.4 Stochastic Records

This section discusses an extension in which the short-run players post reviews with probability less than 1. Suppose after interacting with the patient player, the short-run player does not leave any review with probability $p\in[0,1)$, in which case the patient player's record does not change regardless of his action. Conditional on the short-run player posts a review, the patient player decides whether to erase it at cost c.

Our baseline model assumes that $p = 0$. We explain how to extend our main result to any arbitrary $p < 1$.

The opportunistic type's continuation value equals his minmax value 0 after separating from the honest type. After taking action $\underline{a}_1$ at history h, he prefers to erase it if and only if his continuation value $V(h)$ satisfies $V(h) \geq \frac{(1-\delta)c}{\delta}$. Hence, in any equilibrium where the opportunistic type has an incentive not to erase $\underline{a}_1$ after taking it in period 0, player 1's equilibrium payoff is no more than $\frac{(1-\delta)c}{\delta}$.

Suppose by way of contradiction that for every $\delta^* \in (0,1)$, there exist $\overline{\delta} > \delta^*$ and an equilibrium under which the opportunistic type strictly prefers to erase $\underline{a}_1$ after taking it in period 0. Let $t \in \mathbb{N}$ be such that the opportunistic type plays a_1^* with positive probability at h_*^k if and only if $k < t-1$. Let V_k denote player 1's continuation value at h_*^k and let β_k denote player 2's action at h_*^k. Since at every h_*^k with $k < t-1$, the opportunistic type is indifferent between playing a_1^* and playing $\underline{a}_1$ and then erasing it, we have:

$$(1-\delta)u_1(a_1^*, \beta_k) + \delta\Big\{pV_k + (1-p)V_{k+1}\Big\} = (1-\delta)\Big\{u_1(\underline{a}_1, \beta_k) - (1-p)c\Big\} + \delta V_k,$$

which implies that

$$V_{k+1} - V_k = \frac{1-\delta}{\delta(1-p)}\Big\{u_1(\underline{a}_1, \beta_k) - u_1(a^*, \beta_k) - c(1-p)\Big\}. \tag{5.4}$$

Since $u_1(\underline{a}_1, \beta_k) - u_1(a^*, \beta_k) - c(1-p) > 0$, there exists a constant $\lambda > 0$ such that $t \leq \frac{\lambda}{1-\delta}$.

Recall the definitions of μ_k^* and p_k^* in the proof of Theorem 1. When the short-run players do not leave reviews with probability p, we have

$$\mu_0^* = (1-\overline{\delta}) + \overline{\delta}\Big\{(1-p_0^*) + p_0^* p\Big\}$$

and

$$\mu_k^* = \overline{\delta}\Big\{\mu_{k-1}^* p_{k-1}^*(1-p) + \mu_k^*(1 - p_k^* + p_k^* p)\Big\}$$

or equivalently,

$$\mu_0^* = \frac{1-\overline{\delta}}{1-\overline{\delta}(1-p_0^*+p_0^* p)} \tag{5.5}$$

and

$$\frac{\mu_k^*}{\mu_{k-1}^*} = \frac{\overline{\delta} p_{k-1}^*(1-p)}{1-\overline{\delta}(1-p_k^*+p_k^* p)}. \tag{5.6}$$

Let x_k denote player 2's belief about the probability of a_1^* at history h_*^k. As in the baseline model, we have

$$\frac{\pi}{1-\pi}(1-\overline{\delta})\overline{\delta}^k = \mu_k^*\Big\{\frac{x_k}{1-x_k}(1-p_k^*) - p_k^*\Big\} = \mu_k^* \frac{x_k - p_k^*}{1-x_k} \tag{5.7}$$

and

$$\frac{\mu_k^*}{\mu_{k-1}^*} = \overline{\delta}\frac{x_{k-1} - p_{k-1}^*}{x_k - p_k^*} \cdot \frac{1 - x_k}{1 - x_{k-1}} \leq \overline{\delta}\frac{x_k - p_{k-1}^*}{x_k - p_k^*}. \tag{5.8}$$

Plugging $k = 0$ into (5.7) and applying equation (5.5), we know that p_0^* is bounded above 0 as $\overline{\delta} \to 1$. Equations (5.6) and (5.8) together imply that

$$\frac{x_k - p_{k-1}^*}{x_k - p_k^*} \geq \frac{\overline{\delta}p_{k-1}^*(1-p)}{1 - \overline{\delta}(1 - p_k^* + p_k^* p)}, \tag{5.9}$$

which is equivalent to

$$p_{k-1}^* - p_k^* \leq (1 - \overline{\delta})\frac{x - p_{k-1}^*}{x(1-p)}(1 - p_k^*(1-p)). \tag{5.10}$$

This leads to an upper bound on t, which is proportional to $(1 - \overline{\delta})^{-1}$.

When $\overline{\delta} \to 1$, the lower bound on t exceeds the upper bound on t driven by the opportunistic type's incentives to take action a_1^*, which rules out equilibria in which the opportunistic type has a strict incentive to erase $\underline{a}_1$ in period 0 and implies that player 1's payoff in every equilibrium is no more than $\frac{(1-\delta)c}{\delta}$.

The short-run players' welfare is arbitrarily close to $\pi u_2(a_1^*, a_2^*) + (1-\pi)u_2(\underline{a}_1, \underline{a}_2)$ in every equilibrium. This is because the probability of event $\mathcal{E}^k$, defined before the statement of Lemma 3, is $\mu_k^* p_k^*$ and satisfies

$$\mu_k^* p_k^* \leq (1 - \overline{\delta})\overline{\delta}^k \frac{1}{1 - \overline{\delta}p}. \tag{5.11}$$

This bound coincides with the one in Lemma 3 when $p = 0$. This can be shown using the same induction argument as in the proof of Lemma 3. Since t is bounded above by a linear function of $(1 - \delta)^{-1}$, the average probability with which the opportunistic type takes action a_1^* is close to 0 when $\overline{\delta}$ is close to 1 in every equilibrium. This conclusion can be used to extend the proof of Theorem 2 to any $p < 1$.

A Proof of Lemma 1

Let $\mathcal{H}_i$ denote the set of player $i \in \{1,2\}$'s histories. The opportunistic type's strategy is $\sigma_1 : \mathcal{H}_1 \to \Delta\big(A_1 \times \{0,c\}\big)$, where $\{0,c\}$ denotes his choice of whether to erase his action. Player 2's strategy is $\sigma_2 : \mathcal{H}_2 \to \Delta(A_2)$. For every Nash equilibrium (σ_1, σ_2), we construct a stationary strategy for the opportunistic type $\sigma_1^* : \mathcal{H}_2 \to \Delta\big(A_1 \times \{0,c\}\big)$ such that for every h_2 that occurs with positive probability under (σ_1,σ_2), the opportunistic type's mixed action at h_2, denoted by $\sigma_1^*(h_2)$, equals player 2's expectation of the opportunistic type's action at h_2 under strategy σ_1. By construction, we have $\mathcal{H}(\sigma_1,\sigma_2) = \mathcal{H}(\sigma_1^*,\sigma_2)$.

First, we verify that (σ_1^*,σ_2) is a stationary equilibrium. Since σ_1 best replies to σ_2, σ_1^* also best replies to σ_2. This is because player 1's best reply depends only on player 2's strategy σ_2 and σ_2 is measurable with respect to $\mathcal{H}_2$. Since σ_2 best replies to σ_1, σ_2 also best replies to σ_1^* given that player 2's best reply depends only on her expectation of player 1's action, which remains the same under σ_1^*.

Next, we verify that (σ_1,σ_2) and (σ_1^*,σ_2) are payoff equivalent. Since both σ_1 and σ_1^* best reply to σ_2, player 1's discounted average payoffs under (σ_1,σ_2) and (σ_1^*,σ_2) are the same, which establishes the equivalence on player 1's payoff. In order to establish the equivalence on player 2's payoffs, we make use of two observations. First, for every $h_2 \in \mathcal{H}_2$, conditional on observing history h_2, player 2's payoffs are the same under (σ_1,σ_2) and (σ_1^*,σ_2), which we denote by $v(h_2)$. This is because her expectations of player 1's actions are the same. Second, by construction, the unconditional distributions over player 2's histories are the same under (σ_1,σ_2) and (σ_1^*,σ_2), which we denote by $\mu \in \Delta(\mathcal{H}_2)$. We use 2_k to denote the short-run player who arrives in period k. Let $\mu_k(h_2)$ denote the probability that player 2_k observes history h_2 under (σ_1,σ_2). Let $\mu_k^*(h_2)$ denote the probability that player 2_k observes history h_2 under (σ_1^*,σ_2). Since player 2's prior belief assigns probability $(1-\overline{\delta})\overline{\delta}^k$ to the calendar time being k, we have

$$\mu(h_2) = \sum_{k=0}^{+\infty} (1-\overline{\delta})\overline{\delta}^k \mu_k(h_2) = \sum_{k=0}^{+\infty} (1-\overline{\delta})\overline{\delta}^k \mu_k^*(h_2).$$

This implies that

$$\mathbb{E}^{(\sigma_1,\sigma_2)}\Big[\sum_{k=0}^{+\infty}(1-\overline{\delta})\overline{\delta}^k u_2(a_{1,k},a_{2,k})\Big] = \sum_{k=0}^{+\infty}\Big\{(1-\overline{\delta})\overline{\delta}^k \sum_{h_2\in\mathcal{H}_2} v(h_2)\mu_k(h_2)\Big\} = \sum_{h_2\in\mathcal{H}_2} v(h_2)\Big(\sum_{k=0}^{+\infty}(1-\overline{\delta})\overline{\delta}^k \mu_k(h_2)\Big)$$

$$= \sum_{h_2\in\mathcal{H}_2} v(h_2)\mu(h_2) = \sum_{h_2\in\mathcal{H}_2} v(h_2)\Big(\sum_{k=0}^{+\infty}(1-\overline{\delta})\overline{\delta}^k \mu_k^*(h_2)\Big) = \mathbb{E}^{(\sigma_1^*,\sigma_2)}\Big[\sum_{k=0}^{+\infty}(1-\overline{\delta})\overline{\delta}^k u_2(a_{1,k},a_{2,k})\Big].$$

B Proof of Proposition 1

Fix any equilibrium (σ_1, σ_2). Let $V(h_2)$ denote player 1's continuation value at h_2. Let

$$\overline{V} \equiv \sup_{h_2 \in \mathcal{H}(\sigma_1,\sigma_2)} V(h_2), \tag{B.1}$$

which is player 1's *highest* continuation value at histories that occur with positive probability under (σ_1, σ_2). Suppose by way of contradiction that $\overline{V} > 0$. The definition of $\overline{V}$ implies that for every

$$0 < \varepsilon < \min\Big\{\frac{\overline{V}}{2}, \frac{(1-\delta)(\overline{c}-c)}{\delta}\Big\}, \tag{B.2}$$

there exists $h_2 \in \mathcal{H}(\sigma_1, \sigma_2)$ such that $V(h_2) > \overline{V} - \varepsilon$. We examine player 1's incentive at h_2. Player 1's payoff from playing $\underline{a}_1$ and then erasing it is at least $(1-\delta)(u_1(\underline{a}_1, \beta(h_2)) - c) + \delta(\overline{V} - \varepsilon)$, where $\beta(h_2) \in \Delta(A_2)$ is player 2's action at h_2. Recall the definition of action a_1' in Section 2. Player 1's payoff for taking any action a_1 with $a_1 \succsim a_1'$ is at most $(1-\delta)u_1(a_1, \beta(h_2)) + \delta\overline{V}$, which is strictly less than $(1-\delta)(u_1(\underline{a}_1, \beta(h_2)) - c) + \delta(\overline{V} - \varepsilon)$ when $c < \overline{c}$ and ε satisfies (B.2). This implies player 1 has no incentive to play any action that is weakly greater than a_1' at history h_2. Therefore, player 2 has a strict incentive to take action $\underline{a}_2$ at h_2. Therefore,

$$\overline{V} - \varepsilon < V(h_2) \leq (1-\delta)u_1(\underline{a}_1, \underline{a}_2) + \delta\overline{V}$$

which implies that $\overline{V} - \frac{\varepsilon}{1-\delta} < u_1(\underline{a}_1, \underline{a}_2)$ for every ε that satisfies (B.2). Since ε can be arbitrarily close to 0, we know that $\overline{V} \leq u_1(\underline{a}_1, \underline{a}_2) \equiv 0$ and the opportunistic type will take action $\underline{a}_1$ at every on-path history.

Suppose by way of contradiction that player 2 takes action $a_2 \succ \underline{a}_2$ with strictly positive probability at some history $h_2 \in \mathcal{H}(\sigma_1, \sigma_2)$. Assumption 1 and the definition of $\underline{a}_2$ imply that player 2 will never take actions lower than $\underline{a}_2$. This implies that player 2's action at h_2 strictly FOSDs $\underline{a}_2$. Since $u_1(a_1, a_2)$ is strictly increasing in a_2, we know that $V(h_2) > 0$. This contradicts our conclusion that $\overline{V} = 0$. Since player 2 takes action $\underline{a}_2$ at every on-path history, player 1 has no incentive to take any action that is strictly greater than $\underline{a}_1$.

C Proof of Lemma 2

First, we show that for every $\lambda \in [0, 1]$, player 2 has at most two pure-strategy best replies to $\lambda a_1^* + (1-\lambda)\underline{a}_1$. Suppose by way of contradiction that there exist a_2, a_2', a_2'' with $a_2 \succ a_2' \succ a_2''$ such that all three actions best reply to $\lambda^* a_1^* + (1-\lambda^*)\underline{a}_1$ for some $\lambda^* \in [0, 1]$. Then the last part of Assumption 2 implies that there exist $\lambda, \lambda', \lambda'' \in [0,1] \backslash \{\lambda^*\}$ such that a_2 best replies to $\lambda a_1^* + (1-\lambda)\underline{a}_1$, a_2' best replies to $\lambda' a_1^* + (1-\lambda')\underline{a}_1$, and a_2'' best replies to $\lambda'' a_1^* + (1-\lambda'')\underline{a}_1$. Notice that $\lambda, \lambda', \lambda''$ can be the same. Therefore, *either* at least two of $\{\lambda, \lambda', \lambda''\}$ are strictly more than λ^*, *or* at least two of $\{\lambda, \lambda', \lambda''\}$ are strictly less than λ^*. In the first case, a_2 best replies to $\lambda^* a_1^* + (1-\lambda^*)\underline{a}_1$ and there exists an action that is strictly lower than a_2 that best replies to an action that FOSDs $\lambda^* a_1^* + (1-\lambda^*)\underline{a}_1$. This contradicts Assumption 1 that $u_2(a_1, a_2)$ has strictly increasing differences. A similar contradiction arises when at least two of $\{\lambda, \lambda', \lambda''\}$ are strictly less than λ^*. Hence, player 2 has at most two pure-strategy best replies to any $\lambda a_1^* + (1-\lambda)\underline{a}_1$.

Next, let

$$A_2^* \equiv \{a_2 \in A_2 | \text{there exists } \lambda \in [0,1] \text{ s.t. } a_2 \text{ best replies to } \lambda a_1^* + (1-\lambda)\underline{a}_1\}, \tag{C.1}$$

which is the set of player 2's pure best replies against player 1's actions that are mixtures between a_1^* and $\underline{a}_1$. We show that there exists $\lambda \in [0,1]$ such that it is optimal for player 2 to mix between $a_2 \in A_2^*$ and $a_2' \in A_2^*$ with $a_2 \succ a_2'$ against $\lambda a_1^* + (1-\lambda)\underline{a}_1$ *if and only if* there exists no $a_2'' \in A_2^*$ such that $a_2 \succ a_2'' \succ a_2'$. This is because when $u_2(a_1, a_2)$ has strictly increasing differences, our earlier conclusion implies that there exist $0 \equiv \lambda_0 < \lambda_1 < ... < \lambda_n \equiv 1$ such that for every $a_2 \in A_2^*$, there exists $j \in \{1, 2, ..., n\}$ such that a_2 is a strict best reply to $\lambda a_1^* + (1-\lambda)\underline{a}_1$ for every $\lambda \in (\lambda_{j-1}, \lambda_j)$. Since $u_2(a_1, a_2)$ has strictly increasing differences, player 2's best reply is increasing in λ. The upper-hemi-continuity of best reply correspondences implies that for every $j \in \{1, 2, ..., n-1\}$, player 2 has 2 pure best replies to $\lambda_j a_1^* + (1-\lambda_j)\underline{a}_1$ which are her strict best replies when $\lambda \in (\lambda_{j-1}, \lambda_j)$ and when $\lambda \in (\lambda_j, \lambda_{j+1})$. Therefore, every pair of mixed actions in $\mathcal{B}$ can be ranked according to FOSD. Since $u_2(a_1, a_2)$ has strictly increasing differences and $a_1^* \succ \underline{a}_1$, a_2^* is the highest action in A_2^* and $\underline{a}_2$ is the lowest action in A_2^*. Since $u_1(a_1, a_2)$ is strictly increasing in a_2, for every $a_1 \in A_1$ and $v \in [u_1(a_1, \underline{a}_2), u_1(a_1, a_2^*)]$, there exists a unique $\beta \in \mathcal{B}$ such that $u_1(a_1, \beta) = v$.

Remark: We explain why the second part of Assumption 2 is generically satisfied. If there exist $a_2 \in A_2$ and $a_1, a_1' \in A_1$ such that there is a unique $\lambda \in [0,1]$ where a_2 best replies to $\lambda a_1 + (1-\lambda)a_1'$, then the proof of Lemma 2 suggests that there exists $\lambda^* \in [0,1]$ such that player 2 has at least 3 pure best replies to $\lambda^* a_1 + (1-\lambda^*)a_1'$. Once we depict player 2's payoff from each of her pure actions as a function of the probability with which player 1 plays a_1' (as opposed to a_1), having three best replies to $\lambda^* a_1 + (1-\lambda^*)a_1'$ implies that three of these linear functions intersect at λ^*, which can only occur under knife-edge (u_1, u_2).

D Proof of Proposition 2

Recall that $\pi(h_2)$ is the probability with which player 2's belief assigns to the honest type after observing her history h_2. Since $\pi(h_2) = 0$ for every $h_2 \notin \mathcal{H}_*$, Proposition 1 implies that $\beta(h_2) = \underline{a}_2$ for every $h_2^k \notin \mathcal{H}_*$. Therefore, at any history that contains any unerased action that is not a_1^*, the opportunistic-type of player 1 will play $\underline{a}_1$ and will not erase his action. Since $u_1(a_1, a_2)$ is strictly decreasing in a_1, taking any action other than $\underline{a}_1$ and a_1^* is strictly dominated by taking action $\underline{a}_1$. This implies that player 1 only takes actions a_1^* and $\underline{a}_1$ at any on-path history. Hence, player 2's action at every on-path history belongs to $\mathcal{B}$.

Let V_k be player 1's continuation value at h_*^k and let $\overline{V} \equiv \sup_{k \in \mathbb{N}} V_k$. Suppose by way of contradiction that $\overline{V} = 0$, then player 2 takes action $\underline{a}_2$ at every on-path history. This implies that the opportunistic type takes action $\underline{a}_1$ at every on-path history. According to Bayes rule, player 2 will assign probability 1 to the honest type at history h_*^1, which implies that she will have a strict incentive to take action a_2^* at h_*^1. As a result, $V_1 > (1-\delta)u_1(\underline{a}_1, a_2^*) > 0$, which contradicts the hypothesis that $\overline{V} = 0$. This implies that $\overline{V} > 0$.

Fix any ε that satisfies (B.2). According to Proposition 1, for every h_2 that satisfies $V(h_2) > \overline{V} - \varepsilon$, the opportunistic type has a strict incentive to take action $\underline{a}_1$ at history h_2. Since the opportunistic type's continuation value is 0 at any history that does not belong to $\mathcal{H}_*$, there exists $t \in \mathbb{N}$ such that the opportunistic type's continuation value reaches $\overline{V} - \varepsilon$ at h_*^{t-1}, at which point he will have no incentive to take action a_1^*. Therefore, player 2 will assign probability 1 to the commitment type at every h_*^k with $k \geq t$, at which point the opportunistic type's continuation value is $u_1(\underline{a}_1, a_2^*) - c$ and has no incentive to take action a_1^*. Let p_k be the probability that the opportunistic type takes action a_1^* at history h_*^k. The definition of t implies that $p_k^* = 0$, $\beta_k = a_2^*$, and $\pi_k = 1$ for every $k \geq t$, and $p_k^* > 0$, $\beta_k \neq a_2^*$, and $\pi_k < 1$ for every $k < t-1$.

Next, we show that $p_k^* < 1$ for every $k < t$. Suppose by way of contradiction that $p_k^* = 1$ for some k. Then player 2 strictly prefers to play a_2^* at h_*^k. Player 1's incentive to play a_1^* instead of $\underline{a}_1$ implies that

$$V_k = (1-\delta)u_1(a_1^*, a_2^*) + \delta V_{k+1} \geq \max\Big\{u_1(\underline{a}_1, a_2^*) - c, (1-\delta)u_1(\underline{a}_1, a_2^*)\Big\}. \tag{D.1}$$

If $c < u_1(\underline{a}_1, a_2^*) - u_1(a_1^*, a_2^*)$, then $V_{k+1} > u_1(\underline{a}_1, a_2^*) - c$. For every $t-1 > s \geq k+1$, taking action a_1^* is optimal at h_*^s, which implies that $V_{s+1} > V_s > u_1(\underline{a}_1, a_2^*) - c$. This implies that $V_{t-1} \geq V_{t+1} > V_k$. At history h_*^{t-1}, playing a_1^* is not optimal, which implies that

$$V_{t-1} = \max\{(1-\delta)u_1(\underline{a}_1, \beta_{t-1}), u_1(\underline{a}_1, \beta_{t-1}) - c\} \leq \max\{u_1(\underline{a}_1, a_2^*) - c, (1-\delta)u_1(\underline{a}_1, a_2^*)\}, \tag{D.2}$$

where the last inequality comes from Assumption 1 that $u_1(a_1, a_2)$ is strictly increasing in a_2. Inequalities

(D.1) and (D.2) together imply that $V_{t-1} \leq V_k$. This contradicts $V_{t-1} > V_k$, which implies that taking action $\underline{a}_1$ is weakly optimal for the opportunistic type at every history. Hence, at every h_*^k with $k < t$, *either* player 1 has an incentive to take $\underline{a}_1$ and then erase it, *or* he has an incentive to take $\underline{a}_1$ and then not erase it.

In the first case, $V_k = u_1(\underline{a}_1, \beta_k) - c$ and in the second case, $V_k = (1-\delta)u_1(\underline{a}_1, \beta_k)$. Not erasing $\underline{a}_1$ is preferred to erasing $\underline{a}_1$ if and only if $(1-\delta)u_1(\underline{a}_1, \beta_k) \geq u_1(\underline{a}_1, \beta_k) - c$, or equivalently,

$$u_1(\underline{a}_1, \beta_k) \leq c/\delta. \tag{D.3}$$

Since player 2's action at every on-path history belongs to $\mathcal{B}$, which by Lemma 2 can be completely ranked via FOSD, inequality (D.3) is equivalent to β_k being lower than some cutoff in the sense of FOSD.

In the next step, we show that there exists no $k < t$ such that player 1 prefers to erase $\underline{a}_1$ at h_*^k and prefers not to erase $\underline{a}_1$ at h_*^{k+1}. Suppose by way of contradiction that there exists such a k, then it must be the case that $\beta_k \succeq \beta_{k+1}$. Player 1 weakly prefers *playing* a_1^* *at* h_*^k *and then playing* $\underline{a}_1$ *and not erasing at* h_*^{k+1} to the following two strategies (i) *playing* $\underline{a}_1$ *and erasing in every subsequent period after reaching* h_*^k as well as (ii) *playing* $\underline{a}_1$ *and not erasing in every subsequent period after reaching* h_*^k. This implies that

$$(1-\delta)u_1(a_1^*, \beta_k) + \delta(1-\delta)u_1(\underline{a}_1, \beta_{k+1}) \geq u_1(\underline{a}_1, \beta_k) - c \tag{D.4}$$

and

$$(1-\delta)u_1(a_1^*, \beta_k) + \delta(1-\delta)u_1(\underline{a}_1, \beta_{k+1}) \geq (1-\delta)u_1(\underline{a}_1, \beta_k). \tag{D.5}$$

When $c < \bar{c}$, we have $u_1(\underline{a}_1, \beta_k) - c > u_1(a_1^*, \beta_k)$. Therefore, (D.4) together with $\beta_k \succeq \beta_{k+1}$ implies that $(1-\delta)u_1(\underline{a}_1, \beta_{k+1}) > u_1(a_1^*, \beta_k) \geq u_1(a_1^*, \beta_{k+1})$. Inequality (D.5) implies that

$$u_1(a_1^*, \beta_k) \geq u_1(\underline{a}_1, \beta_k) - \delta u_1(\underline{a}_1, \beta_{k+1}) \geq (1-\delta)u_1(\underline{a}_1, \beta_k).$$

This leads to a contradiction. Hence, there exists $t_0 \leq t$ such that at history h_*^k, player 1 erases $\underline{a}_1$ with probability 1 if $k > t_0$, and erases $\underline{a}_1$ with zero probability if $k < t_0$.

In the last step, we show that the length of the reputation-building phase t is bounded below a linear function of $(1-\delta)^{-1}$, which is uniform across all equilibria. Let $a_2' \in A_2 \backslash \{a_2^*\}$ denote player 2's action such that player 2 is indifferent between a_2' and a_2^* when player 1 takes a mixed action $x^* a_1^* + (1-x^*)\underline{a}_1$ for some $x^* \in [0,1]$. This action a_2' is uniquely defined under Assumption 1. The opportunistic type has an incentive to take action $\underline{a}_1$ and then erase it at every h_*^k with $t_0 < k < t$, which implies the following upper bound on the extent to which player 1's continuation value increases with the length of his record:

$$V_{k+1}-V_k = u_1(\underline{a}_1, \beta_{k+1})-u_1(\underline{a}_1, \beta_k) \leq (1-\delta)\Big(u_1(\underline{a}_1, \beta_{k+1})-c-u_1(a_1^*, \beta_k)\Big) \leq (1-\delta)\Big(u_1(\underline{a}_1, a_2^*)-c-u_1(a_1^*, \underline{a}_2)\Big).$$

Let $\Delta \equiv u_1(\underline{a}_1, a_2^*) - c - u_1(a_1^*, \underline{a}_2)$. The rest of the proof considers three classes of equilibria separately.

First, consider any equilibrium in which player 1 has an incentive not to erase $\underline{a}_1$ at h_*^0. Player 1's continuation value at $h_*^{t_0}$ is at most $\frac{(1-\delta)c}{\delta}$ and his continuation value at h_*^t is $u_1(\underline{a}_1, a_2^*) - c$, which is strictly greater than $u_1(a_1^*, a_2^*)$ when $c < \overline{c}$. This implies that

$$t \geq t - t_0 \geq \frac{u_1(\underline{a}_1, a_2^*) - c - \frac{(1-\delta)c}{\delta}}{\Delta}(1-\delta)^{-1}.$$

Second, consider any equilibrium in which player 2 takes action a_2^* with zero probability at h_*^0. Player 1's continuation value at h_*^0 is at most $\max\{\frac{(1-\delta)c}{\delta}, u_1(\underline{a}_1, a_2') - c\}$. This implies that

$$t \geq \frac{u_1(\underline{a}_1, a_2^*) - c - \max\{\frac{(1-\delta)c}{\delta}, u_1(\underline{a}_1, a_2') - c\}}{\Delta}(1-\delta)^{-1}.$$

Third, consider any equilibrium such that in period 0, player 1 has a strict incentive to erase $\underline{a}_1$ *and* player 2 takes action a_2^* with positive probability. The definition of t implies that $p_t^* = 0$. Inequality (3.8) implies that there exists $\phi > 0$ such that $p_{k-1} - p_k \leq \phi(1-\overline{\delta})$. Player 2's incentive to take action a_2^* in period 0 implies that she expects player 1 to take action a_1^* in period 0 with probability at least x^*. According to (3.5) and (3.6), we know that there exists $x_0 \geq x^*$ such that

$$\frac{\pi}{1-\pi} = \frac{1}{1-\overline{\delta}+\overline{\delta}p_0^*} \cdot \frac{x_0 - p_0^*}{1-x_0}. \tag{D.6}$$

Therefore, for any fixed $\pi \in (0,1)$, there exists $\underline{p}_0 \in (0,1)$ such that (D.6) holds only if $p_0^* > \underline{p}_0$. Hence,

$$t \geq \frac{\underline{p}_0}{\phi(1-\overline{\delta})} \geq \frac{\underline{p}_0}{\phi(1-\delta)}.$$

The three cases together imply that for every $\pi \in (0,1)$, there exists a constant $\lambda > 0$ such that $t > \frac{\lambda}{1-\delta}$ in every equilibrium of the reputation game where player 1's effective discount factor is δ.

E Proofs of Corollaries 1 and 2

Proof of Corollary 1: Recall that p_0^* is strictly decreasing in $l \equiv \frac{\pi}{1-\pi}$ and equals x_0 when $l = 0$. The right-hand-side of inequality (3.8) implies that when l is close to 0, p_0^* is close to x_0, and $p_1^* - p_0^*$ is close to 0. Therefore, for every $\widehat{\delta}$ and $\overline{\delta}$, there exists l close enough to 0 such that the number of periods for p_t^* to reach 0 is strictly greater than the upper bound on t driven by the opportunistic type's incentive to play a_1^*. This implies that player 1's payoff is no more than $(1-\delta)c/\delta$ in every equilibrium when π is low.

Proof of Corollary 2: We know from the proof of Theorem 1 that the opportunistic type has an incentive to play $\underline{a}_1$ and then not erase it at history h_*^0. Therefore, in order to show the first part, we only need to show that he has an incentive to play $\underline{a}_1$ and then erase it at h_*^0. According to (3.2), the opportunistic type's indifference condition at h_*^0 will then imply that his continuation value at h_*^0 is exactly $(1-\delta)c/\delta$.

Suppose by way of contradiction that when $\pi > \pi^*$ and there exists an equilibrium in which player 1 erases $\underline{a}_1$ at h_*^0 with zero probability. We consider two cases. First, suppose player 2 does not play a_2^* for sure at h_0^*, then the probability with which player 1 plays a_1^* at h_*^0 cannot exceed π^*. This cannot happen when $\pi > \pi^*$ and player 1 does not erase his action at h_*^0. Second, suppose player 2 plays a_2^* for sure at h_*^0, then the opportunistic type of player 1 reaches his highest continuation value at h_*^0, which implies that he has a strict incentive to play $\underline{a}_1$ at h_0^*. The hypothesis that he will not erase $\underline{a}_1$ at h_*^0 implies that $u_1(\underline{a}_1, a_2^*) < \frac{c}{\delta}$. Recall that $\underline{\delta} \equiv \frac{c}{u_1(\underline{a}_1, a_2^*)}$. Inequality $u_1(\underline{a}_1, a_2^*) < \frac{c}{\delta}$ contradicts our requirement that $\delta \equiv \overline{\delta} \cdot \widehat{\delta} > \underline{\delta}$.

For the second part, notice that when player 1's payoff is $\frac{(1-\delta)c}{\delta}$ in every equilibrium, player 2's action at h_*^0, denoted by $\beta_0 \in \Delta(A_2)$, satisfies

$$V_0 = u_1(\underline{a}_1, \beta_0) - c = \frac{(1-\delta)c}{\delta}. \tag{E.1}$$

Recall the definition of $\mathcal{B}$ in (3.1), that player 2's (potentially mixed) action at every on-path history belongs to $\mathcal{B}$, and that each pair of elements in $\mathcal{B}$ can be ranked according to FOSD. Since $\beta_0 \in \mathcal{B}$ and $u_1(a_1, a_2)$ is strictly increasing in a_2, equation (E.1) uniquely pins down player 2's action at h_*^0, denoted by β_0. Similarly, the values of V_1, V_2,... are pinned down by V_0 via equation (3.3), which also pin down player 2's actions at h_*^1, h_*^2,... Under generic parameter values, $\beta_0, \beta_1, ..., \beta_{t-1}$ are non-trivially mixed actions, which pin down player 2's belief about player 1's action at histories $h_*^0, h_*^1, ..., h_*^{t-1}$. This pins down the opportunistic type of player 1's actions at all on-path histories.

F Proof of Lemma 4

The payoff lower bound is straightforward. The rest of this proof establishes the payoff upper bound. When $\widetilde{q}$ is the Dirac measure on 0, the opportunistic type never takes a_1^* since it is strictly dominated by playing $\underline{a}_1$ and then erasing it. We consider two cases. If player 2's action at the null history β_0 satisfies $u_1(\underline{a}_1, \beta_0) \leq \frac{c}{\delta}$, then not erasing $\underline{a}_1$ is optimal for the opportunistic type, in which case his payoff is no more than $(1-\delta)u_1(\underline{a}_1, \beta_0) \leq \frac{(1-\delta)c}{\delta}$. If β_0 is such that $u_1(\underline{a}_1, \beta_0) > \frac{c}{\delta}$, then player 1 has a strict incentive to erase $\underline{a}_1$ in period 0, in which case player 2's belief assigns probability π to player 1's action being a_1^*. Therefore, player 2's action is at most $\overline{a}_2^{\pi}$, in which case player 1's payoff is no more than $u_1(\underline{a}_1, \overline{a}_2^{\pi}) - c$.

G Proof of Theorem 2

We use the same notation as in the proof of Theorem 1. Similar to the proof of Proposition 2, we know that in every equilibrium, there exists $t \in \mathbb{N}$ such that the opportunistic type takes action a_1^* with positive probability until his record length reaches $t-1$ and moreover, the opportunistic type cannot have a strict incentive to play a_1^* at h_*^k for every $k \leq t-1$. Therefore, the opportunistic type's equilibrium payoff is bounded above by $\max\{\frac{(1-\delta)c}{\delta}, u_1(\underline{a}_1, \beta_0) - c\}$. This payoff upper bound is no more than (4.1) unless (i) $u_1(\underline{a}_1, \beta_0) > \frac{c}{\delta}$ and (ii) there exists an action strictly greater than $\overline{a}_2^{\pi}$ that belongs to the support of β_0.

The first condition implies that at every h_*^k with $k \leq t-1$, the opportunistic type strictly prefers to erase $\underline{a}_1$ after taking it at h_*^k. Given our definition of $\overline{a}_2^{\pi}$, the second condition implies that the expected probability with which the opportunistic type of player 1 takes action a_1^* must be strictly bounded above 0 in order to provide player 2 an incentive to take actions that are strictly greater than $\overline{a}_2^{\pi}$ starting from period 0.

Fix any $\widehat{\delta} \in (0,1)$. Suppose by way of contradiction that for every $\overline{\delta}$ close enough to 1, there exist a disclosure policy with unconditional distribution $\widetilde{q} \in \Delta(\mathbb{N} \cup \{s^*\})$ and an equilibrium under $\widetilde{q}$ such that the opportunistic type's payoff is strictly greater than (4.1). Since at every history h_*^k with $k < t$, the opportunistic type is indifferent between playing a_1^* and playing $\underline{a}_1$ and then erasing it, we have

$$V_{k+1} - V_k = (1-\delta)\Big(u_1(\underline{a}_1, \beta_{k+1}) - c - u_1(a_1^*, \beta_k)\Big). \tag{G.1}$$

This implies that t is bounded from above by some linear function of $(1-\delta)^{-1}$.

Since player 2 has an incentive to take some action that is strictly greater than $\overline{a}_2^{\pi}$ at h_*^0 and player 2's action increases in the length of player 1's good record in the sense of FOSD, there exists $x > \pi$ such that player 2's belief assigns probability at least x to a_1^* at every h_*^k with $k \leq t-1$. Importantly, this x depends only on (u_1, u_2) and does not depend on $\widehat{\delta}$, $\overline{\delta}$, and the honest type's disclosure policy. Recall that μ_k^* is the probability that the history is h_*^k conditional on player 1 being the opportunistic type and that p_k^* is the probability with which the opportunistic type plays a_1^* at h_*^k. Player 2's incentive constraint at history h_*^k implies that

$$\frac{\pi\widetilde{q}(k) + (1-\pi)\mu_k^* p_k^*}{(1-\pi)\mu_k^*(1-p_k^*)} \geq \frac{x}{1-x} \text{ for every } k \in \{1, ..., t-1\}. \tag{G.2}$$

Since $\pi < x$, (G.2) is true for every $k \leq t-1$ only when $\sum_{j=0}^{t-1} \mu_j^* p_j^*$ is bounded above 0. Lemma 3 implies that $\mu_j^* p_j^* \leq 1 - \overline{\delta}$ for every $j \leq t-1$. Therefore, $\sum_{j=0}^{t-1} \mu_j^* p_j^*$ is bounded above 0 if and only if t is bounded below by a linear function of $(1-\overline{\delta})^{-1}$. For any fixed $\widehat{\delta} \in (0,1)$, there exists $\overline{\delta}$ close to 1 such that the lower bound on t is strictly greater than the upper bound on t implied by the opportunistic type's incentive constraints. This rules out equilibria in which the opportunistic type's payoff being strictly greater than (4.1). Similarly, fix any equilibrium and the resulting distribution over player 1's actions, player 2's payoff cannot be greater than their payoff when they can observe player 1's realized pure action before choosing their action. This upper bound cannot be greater than $\pi u_2(a_1^*, a_2^*) + (1-\pi)u_2(\underline{a}_1, \underline{a}_2)$ as $\overline{\delta} \to 1$ since $\Pr(\mathcal{E}_k) \leq 1 - \overline{\delta}$ for every $k \leq t$ and t is at most proportional to $(1-\delta)^{-1}$.

H Proof of Proposition 3

Suppose by way of contradiction that there exists an equilibrium in which the honest type erases a_1^* with positive probability at some on-path history. Let $t \in \mathbb{N}$ denote the smallest integer k such that the honest type erases his record with positive probability at h_*^k. Hence, it is optimal for the honest type not to erase any action until period t, after which he erases the record in every subsequent period. We call this strategy σ_1^*. We consider two cases separately.

First, suppose β_{t+1} weakly FOSDs β_t. Since $c > 0$ and the honest type chooses H in every period, his payoff from σ_1^* is strictly less than his payoff from the following strategy: Do not erase any action until period $t+1$ and erase every action taken after period $t+1$. This leads to a contradiction.

Next, suppose β_t strictly FOSDs β_{t+1}. Player 2 does not play a_2^* for sure at h_*^{t+1} and therefore, the opportunistic type plays $\underline{a}_1$ with positive probability at h_*^{t+1}. If the opportunistic type reaches h_*^{t+1} with positive probability, then it is optimal for him to play a_1^* at h_*^t and then play $\underline{a}_1$ at h_*^{t+1}. However, this gives the opportunistic type a strictly lower payoff compared to playing $\underline{a}_1$ at h_*^t. This leads to a contradiction and implies that the opportunistic type does not reach h_*^{t+1} with positive probability in equilibrium. If this is the case, then player 2 assigns probability 1 to the honest type at history h_*^{t+1}, in which case she will have a strict incentive to play a_2^* at h_*^{t+1}. This contradicts our earlier hypothesis that β_t is strictly greater than β_{t+1}.

I Player 2's Posterior Belief about Calendar Time

This appendix explains how to compute the short-run player's posterior belief about calendar time after they observe their history. Our calculation is based on (i) their prior belief about calendar time and about the long-run player's type and (ii) players' strategy profile. Such a calculation is *not* required for the proofs of our results, since we directly work with the short-run player's posterior belief about the long-run player's *action* after observing their history, which is what matters for the short-run player's incentives.

Recall that the short-run player's prior belief assigns probability π to the honest type and probability $1-\pi$ to the opportunistic type. The honest type's strategy is playing a_1^* in every period and never erasing any action. Fix the opportunistic type's strategy σ_1, the short-run player's strategy σ_2, and any history h of player 2's that occurs with positive probability under the above strategy profile. According to Bayes rule,

$$\Pr\Big(\text{calendar time is } t \Big| \text{P2's history is } h\Big) \quad = \quad \frac{\Pr(\text{P2's history is } h|\text{calendar time is } t)\cdot \Pr(\text{calendar time is } t)}{\Pr(\text{P2's history is } h)}.$$

As long as h occurs with positive probability, the denominator $\Pr(\text{P2's history is } h)$ is strictly positive.

We explain how to compute each of the three terms on the right-hand-side of the above equation. The term $\Pr(\text{calendar time is } t)$ is given by the short-run player's prior belief about calendar time:

$$\Pr(\text{calendar time is } t) = (1-\overline{\delta})\overline{\delta}^t. \tag{I.1}$$

As for the term $\Pr(\text{P2's history is } h|\text{calendar time is } t)$, the law of total probability implies that

$$\Pr\Big(\text{P2's history is } h\Big|\text{calendar time is } t\Big) = \pi\cdot \Pr\Big(\text{P2's history is } h\Big|\text{calendar time is } t \text{ and P1 is honest}\Big)$$
$$+(1-\pi)\cdot \Pr\Big(\text{P2's history is } h\Big|\text{calendar time is } t \text{ and P1 is opportunistic}\Big). \tag{I.2}$$

The honest type's strategy implies that the red term equals 1 if and only if h consists only of t actions, all of which are a_1^*, and equals 0 otherwise. The blue term is pinned down by (σ_1, σ_2).

As for the term $\Pr(\text{P2's history is } h)$, we apply the law of total probability and obtain that

$$\Pr(\text{P2's history is } h) = \sum_{k=0}^{+\infty} \Pr(\text{calendar time is } k)\cdot \Pr\Big(\text{P2's history is } h\Big|\text{calendar time is } k\Big), \tag{I.3}$$

where $\Pr(\text{calendar time is } k) = (1-\overline{\delta})\overline{\delta}^k$ and $\Pr(\text{P2's history is } h|\text{calendar time is } k)$ can be computed via (I.2). One can then compute the short-run player's posterior belief about calendar time after she observes history h by plugging (I.1), (I.2), and (I.3) into the expression for $\Pr(\text{calendar time is } t|\text{P2's history is } h)$.

Next, we use an example to explain how to apply the above formula to compute the short-run player's posterior belief about calendar time. Suppose that the stage-game is the product choice game in the introduction. The honest type plays G in every period and never erases any action. The opportunistic type can take any action and he can also erase his action in period t by the end of period t.

For future reference, we use h_*^k to denote player 2's history where she observes k actions, all of which are G, and we use h_-^k to denote player 2's history where she observes k actions, all of which are B.

As an example, suppose that the opportunistic type uses the following strategy, which is measurable with respect to player 2's history:

- At h_*^0, he mixes between G and B with probability $1/2$ each and does not erase his action.
- At h_*^k for every $k \geq 1$, he plays B and then erases his action.
- At any other history, he plays B and does not erase his action.

One can verify that only $\{h_*^k, h_-^k\}_{k \in \mathbb{N}}$ occur with positive probability under such a strategy.

First, we compute $\Pr(\text{P2's history is } h|\text{calendar time is } t \text{ and P1 is opportunistic})$, which is the distribution over player 2's history conditional on calendar time being t and player 1 being the opportunistic type.

- If $t = 0$, player 2's history is h_*^0 for sure.
- If $t = 1$, player 2's history is h_*^1 with probability $1/2$ and is h_-^1 with probability $1/2$, depending on which action the opportunistic type took in period 0.
- If $t = k$ for every $k \geq 2$, player 2's history is h_*^1 with probability $1/2$ and is h_-^k with probability $1/2$.

Our formula implies that at h_*^0, player 2's posterior belief assigns probability 1 to the calendar time being 0. At any h_*^k with $k \geq 2$, player 2's posterior belief assigns probability 1 to the calendar time being k. At any h_-^j with $j \geq 1$, player 2's posterior belief assigns probability 1 to the calendar time being j. At history h_*^1, player 2's posterior belief assigns:

- Probability 0 to calendar time being 0.
- Probability $\frac{(1+\pi)(1-\overline{\delta})}{1+\pi-2\pi\overline{\delta}}$ to calendar time being 1.
- Probability $\frac{(1-\pi)\overline{\delta}^{k-1}(1-\overline{\delta})}{1+\pi-2\pi\overline{\delta}}$ to calendar time being k for every $k \geq 2$.

One can verify that the sum of these probabilities is 1.

References

[1] Besley, Timothy and Andrea Prat (2006) "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability," *American Economic Review*, 96(3), 720-736.

[2] Bhaskar, V and Caroline Thomas (2019) "Community Enforcement of Trust with Bounded Memory," *Review of Economic Studies*, 86(3), 1010-1032.

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

[4] Bolton, Gary, Ben Greiner and Axel Ockenfels (2013) "Engineering Trust: Reciprocity in the Production of Reputation Information," *Management Science*, 59(2), 265-285.

[5] Cai, Hongbin, Ginger Zhe Jin, Chong Liu and Li-an Zhou (2014) "Seller Reputation: From Word-of-Mouth to Centralized Feedback," *International Journal of Industrial Organization*, 34, 51-65.

[6] Clark, Daniel, Drew Fudenberg and Alexander Wolitzky (2021) "Record-Keeping and Cooperation in Large Societies," *Review of Economic Studies*, 88, 2179-2209.

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

[8] Ekmekci, Mehmet, Leandro Gorno, Lucas Maestri, Jian Sun and Dong Wei (2022) "Learning from Manipulable Signals," *American Economic Review*, 112(12), 3995-4040.

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

[10] Ely, Jeffrey, Drew Fudenberg and David Levine (2008) "When is Reputation Bad?" *Games and Economic Behavior*, 63(2), 498-526.

[11] Fudenberg, Drew and Kevin He (2018) "Learning and Type Compatibility in Signaling Games," *Econometrica*, 86(4), 1215-1255.

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

[13] Fudenberg, Drew and Eric Maskin (1986) "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54(3), 533-554.

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

[15] Fudenberg, Drew and David Levine (1993) "Steady State Learning and Nash Equilibrium," *Econometrica*, 61(3), 547-573.

[16] Hauser, Daniel (2023) "Censorship and Reputation," *American Economic Journal: Microeconomics*, 15(1), 497-528.

[17] Heller, Yuval and Erik Mohlin (2018) "Observations on Cooperation," *Review of Economic Studies*, 88, 1892-1935.

[18] Hu, Ju (2020) "On the Existence of the Ex Post Symmetric Random Entry Model," *Journal of Mathematical Economics*, 90, 42-47.

[19] Kamenica, Emir and Matthew Gentzkow (2011) "Bayesian Persuasion," *American Economic Review*, 101(6), 2590-2615.

[20] Kaya, Ayca (2009) "Repeated Signaling Games," *Games and Economic Behavior*, 66(2), 841-854.

[21] Kovbasyuk, Sergey and Giancarlo Spagnolo (2023) "Memory and Markets," *Review of Economic Studies*, forthcoming.

[22] Levine, David (2021) "The Reputation Trap," *Econometrica*, 89(6), 2659-2678.

[23] Li, Yingkai and Harry Pei (2021) "Equilibrium Behaviors in Repeated Games," *Journal of Economic Theory*, 193, 105222.

[24] Lipnowski, Elliot, Doron Ravid, and Denis Shishkin (2022) "Persuasion via Weak Institutions," *Journal of Political Economy*, 130(10), 2705-2730.

[25] Livingston, Jeffrey (2005) "How Valuable Is a Good Reputation? A Sample Selection Model of Internet Auctions," *Review of Economics and Statistics*, 87(3),453-465.

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

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

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

[29] Nosko, Chris and Steven Tadelis (2015) "The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment," NBER Working Paper, 20830.

[30] Pei, Harry (2023) "Reputation Effects under Short Memories," Working Paper.

[31] Phelan, Christopher (2006) "Public Trust and Governmental Betrayals," *Journal of Economic Theory*, 130(1), 27-43.

[32] Smirnov, Aleksei, and Egor Starkov (2022) "Bad News Turned Good: Reversal Under Censorship," *American Economic Journal: Microeconomics*, 14(2), 506-560.

[33] Sugaya, Takuo and Alexander Wolitzky (2020) "Do a Few Bad Apples Spoil the Barrel?: An Anti-Folk Theorem for Anonymous Repeated Games with Incomplete Information," *American Economic Review*, 110, 3817-3835.

[34] Sun, Yiman (2023) "A Dynamic Model of Censorship," *Theoretical Economics*, forthcoming.

[35] Tadelis, Steven (2016) "Reputation and Feedback Systems in Online Platform Markets," *Annual Review of Economics*, 8, 321-340.

[36] Takahashi, Satoru (2010) "Community Enforcement When Players Observe Partners' Past Play," *Journal of Economic Theory*, 145, 42-62.

[37] Vong, Allen (2022) "Certification for Consistent Quality Provision," Working Paper.

[38] Wong, Yu Fu (2023) "Dynamic Monitoring Design," Working Paper.