Sustainable Reputations with Biased Review Platforms

Gökmen Öz[*,†]

Abstract

We analyze how much information rent a long-lived agent can extract against an infinite sequence of short-lived players in a moral hazard game played repeatedly. Given the existence of a committed type, we show that the long-lived agent can attain even higher payoffs than the mixed Stackelberg payoff. We assume that principals do not observe past play. Instead, they observe a review score (an aggregate review score as in Yelp, Foursquare ...) announced by a central mechanism called review platform. By censoring the past actions of the long-lived player, the platform enables her to attain the highest payoff that can ever be achieved.

Keywords: Reputations, Review Platform, Repeated Game, Mixed Stackelberg payoff

[*] Ph.D. Candidate at The Pennsylvania State University. Address: 1003 W Aaron Dr, Apt 9A, State College, PA, 16803. Email: guo108@psu.edu.

[†] I would like to thank Kalyan Chatterjee for making this paper possible and for being a great advisor. I would like to also thank Mehmet Ekmekci, Henrique Roscoe De Oliveria and Edward Green for their both insightful and technical contributions.

1 Introduction

Some goods do not reveal their quality at the time of the purchase. Therefore, a certain degree of trust is required for the trade to happen. However, trust is susceptible to exploitation. As an example, imagine a souvenir shop in a tourist destination so the probability of a customer comeback is low. With the quality of an item not observable by buyers at the time of a purchase, a seller is tempted to reduce the quality. As a result, rational consumers hesitate to shop at this store even when the quality is at its highest. The inefficiency described here has two sources. One is the seller's lack of commitment and the other is the inability of buyers to monitor the quality.

Online review platforms tackle the latter by providing a buyer with reviews from the past experiences, and an aggregate summary such as an average score. However, the reviews carry partial information. As a result, sellers can still build a reputation for being trustworthy.

In addition to the imperfectness of information on review platforms, there is also a question of trustworthiness. Mayzlin, Dover and Chevalier [17] show that when posting a review has no cost as in TripAdvisor, hotels feed made-up reviews in order to inflate their rating or malign a competitor. In light of this fact, we will analyze in this paper how much information rent a patient long-run player can gain against myopic short-run players in a repeated game with a review platform.

To the best of our knowledge, Fudenberg & Levine [8] presents the first model in which a long-run player with an unknown type repeatedly plays a simultaneous-move game against a sequence of short-lived opponents. Their result states that if there is a positive probability that the long-run player will always play the pure Stackelberg strategy, then her payoff in any Nash equilibrium exceeds a bound that converges to the pure Stackelberg payoff as his discount factor approaches one. Ekmekci [6] improves the highest attainable payoff to the long-run player by constructing a finite rating system. The rating system ensures that the long-run player attains almost her mixed Stackelberg payoff. To put it in a nutshell, the earlier literature lacks the characterization of an equilibrium that yields payoffs to the long-run player greater than or equal to the mixed Stackelberg payoff. Our paper demonstrates that it is possible that the mixed Stackelberg payoff can be attained. We construct a review platform which is similar to Ekmekci's [6] rating system and show that if there is a positive probability that the long-run player is a commitment type who plays the pure Stackelberg strategy every period, then the strategic type can achieve strictly higher payoffs than the mixed Stackelberg payoff.

The existence of a crazy role-model who takes a fixed action that other sellers are motivated to imitate is crucial in our model. The idea of a role-model approach was introduced by Kreps, Milgrom, Roberts & Milson [13].

Most of the studies where the seller's type is unknown take the complete information case as the benchmark. Fudenberg, Kreps & Maskin [7] show that the highest payoff to a long-run player cannot exceed the pure Stackelberg payoff in a repeated complete information game played against a sequence of myopic buyers.

Even though the early literature improves upon the complete information benchmark by adding crazy committed types, it lacks a characterization of equilibrium that yields payoffs close to the (mixed) Stackelberg payoff to the long-run player.

Later studies provide characterizations of such equilibria that attain almost the Stackelberg payoff to the long-run player. Pei [5] defines different types for the long-run player over the cost of producing a high-quality product. As a result, he characterizes an equilibrium in which the long-run player achieves a payoff arbitrarily close to her Stackelberg payoff. Liu [15] makes similar improvements to the long-run player's payoff by imposing a cost to short-run players for information acquisition.

Cripps, Mailath and Samuelson [4] (CMS hereafter) presents a negative result by showing that in games with imperfect monitoring, it is impossible to maintain a permanent reputation. However, there are various ways to sustain reputation effects for a long time. One of them is imposing a stochastic process governing the type of the long-run player through time. Holmstrom [11], Cole, Dow and English [3], Mailath and Samuelson [16], Phelan [2], Wiseman [18] and [19] are some examples of stochastic type. In our model, the type of the long-run player is determined once and for all before the game begins.

In contrast to CMS, reputation dynamics do not necessarily degenerate in review platforms despite the abundance of noisy but informative feedbacks. The reason for the persistent reputation dynamics is that information acquired by buyers may not be accumulated. An explanation for why the information may not be accumulated is that most of the time buyers only look at an aggregate summary rather than the complete history of reviews. Yelp provides a 5-star rating system based on the reviews that the restaurants receive. By looking only at an aggregate measure, buyers miss most of the past data and this behavior allows the seller to build reputation even in the distant future. Another possible explanation is driven by the fact that some review platforms such as eBay provide censored information. Doing so hinders buyers from stripping out the noise in the reviews.

Best and Quigley [1] considers a piece of similar machinery to ours to get around the CMS result. A review aggregator keeps myopic buyers uncertain about the past actions of a seller by garbling the history. Since they examine a persuasion game similar to Kamenica & Gentzkow [12], their technique does not apply to our model.

As we mentioned, we follow Ekmekci [6]'s lead in proposing a mechanism that determines the information revealed to the short-run player each period. We call that mechanism our review platform. Hence, one can interpret the information at a given period as an aggregate measure observed by the short-run player. A lot of past data is censored in our model. The reason for the censoring may be either due to how the review platform is designed or to the fact that buyers find it too costly to gather all the information available in the platform. An example for the former is eBay showing feedbacks given to a seller only in the most recent month.

In this paper, we focus on a specific form of information censoring. A central authority of a review platform announces a score from a finite set based on the past performance of the seller. The first score of the game is determined randomly. At the end of each period, unless the score is at its lowest, it may remain the same, increase or decrease by one depending on the past actions of the long-run player. We deviate from Ekmekci [6] in two ways. The first is by allowing for a jump of two scores only if the current score is the lowest. Even though this jump may seem inconsequential when the number of total scores is large, it significantly increases the highest payoff attainable to the long-run player. The second is by giving the review platform the knowledge of the type of the long-run player. We are not pleased to make this assumption, yet it can be justified in some cases. Consider a journalist who has an inside information on a politician but does not have enough hard evidence. For the sake of his credibility, he may rather base his reports on the politician's actions. Similarly, the review platform we design will condition the scores only on the actions of the long-run player except the initial period when there is no action observed.

Our review platform is a reminiscent of the machinery in Kamenica & Gentzkow [12]. Every score delivered to the buyers is associated with a posterior belief. Yet in our model, the posterior beliefs never hit the boundaries of the unit simplex. Posterior belief at a boundary refers to a complete information game where the payoff to the long-run player is dramatically lower than what we find.

It is fair to take the Stackelberg payoff as an upper bound as to how much payoff the long-run player can achieve. However, in the presence of a crazy type who is committed to a fixed action, the opportunistic long-run player can enjoy higher payoffs than the Stackelberg payoff. To fix this idea, suppose that buyers will purchase a product only if it meets a certain quality criteria and they mistakenly believe that the seller is likely to be committed to producing the high quality of a product, which is actually her dominated action. In this case, without any information revelation, the opportunistic type can constantly cheat buyers by producing the low quality. The Stackelberg action in this example would be mixing the low quality with adequate high quality in order to make buyers indifferent. Therefore, the opportunistic type enjoys strictly higher than the Stackelberg payoffs.

The new upper bound that we set on the payoff to the long-run player depends on the prior belief of buyers about her type. As is usual in the literature, we would like to consider small probabilities of the long-run player being a crazy type. We show that even when the prior belief is very small, the long-run player achieves strictly higher than the Stackelberg payoff, in contrast to the previous literature.

Besides the highest attainable payoff to the long-run player, another focus of attention of the literature is efficiency. According to Fudenberg & Levine [9], the inefficiency may be severe if the imperfection of the monitoring is large. Yet the review platform we propose offers equilibria with efficient allocations even though the monitoring is quite limited.

In our model, the long-run player is either a commitment type who plays the pure Stackelberg action or a normal type who has a moral hazard problem. The review platform observes the past play and the past scores and announces a new score. The rule governing the transition from the past scores to the new one is called a transition rule.

We construct the transition rule as a function of a parameter and show the existence of a value of the parameter for which our results are satisfied. We cannot write down its value explicitly except asymptotically as discount factor goes to one and the number of possible scores goes to infinity.

In the next section, we describe the model in detail. In Section 3, we present the first step of our result. We present our result in two steps for the sake of clarity. In Section 4, we go over an example. Section 5 illustrates numerically the nature of the equilibrium and concludes. Section 6 concludes our result with the second step. Section 7 analyzes the payoffs to the short-run players and Section 8 is the appendix.

2 Model

We study a product choice game played repeatedly between a long-run player (Player 1, Seller, she) and an infinite sequence of short-run players (Player 2, Buyer, he). At every period, a new short-run player enters the game and lives for one period.

2.1 The stage game

	B	N
H	$p-c$,v_H-p	$-c$,0
L	p,v_L-p	0,0

The pure actions available to player 1 are high and low effort: $A_1 = \{H, L\}$. Let $\Delta(A_1)$ denote the set of all probability distributions over A_1 and α_1 be a generic element of $\Delta(A_1)$.

Player 2's pure actions are buy or not: $A_2 = \{B, N\}$. Similarly, the set of all actions for player 2 is $\Delta(A_2)$ and α_2 denotes a generic element of $\Delta(A_2)$.

The players move simultaneously. The normal form of the stage game is depicted above. It is a standard product choice game where a buyer pays a fixed price p in order to buy the good and receives a value v_H if the seller exerted high effort and v_L if she exerted low effort. Buyers are interested in trade only if a certain amount of high effort is put by the seller: $v_H > p, v_L < p$. The seller has to incur a constant cost of c to exert high effort. Therefore, she has a dominant action of exerting low effort. She receives a payment of p which is higher than her cost if the product is sold. Hence, trade always makes her better off. In the unique Nash equilibrium of the stage game, player 1 exerts low effort and player 2 does not buy and the unique Nash equilibrium payoff profile is normalized to be $(0, 0)$.

Let us highlight the crucial conditions implied by the assumptions we made on the normal form game. Note that these conditions also exist in Ekmekci [6].

Condition 1: The incentive to exert low effort $c > 0$ is independent of whether the good is sold or not. This assumption is made for the sake of simplicity. We will discuss in more detail how it makes the construction easier later.

Condition 2: Trade is always profitable to the seller. This condition is one of those which allow us to have a punishment and reward phase for the seller depending on her past actions. The review platform punishes her with a higher probability of no-trade state in future.

Condition 3: There exists a non-degenerate α_1 such that for any α_1' with $\alpha_1'(H) \geq \alpha_1(H)$, the best response of player 2 to α_1' includes buying: $\alpha_1'(H)v_H + \alpha_1'(L)v_L \geq p$.

Condition 4: Player 1 prefers committing to exerting high effort to her Nash equilibrium payoff: $p - c > 0$.

2.2 Incomplete Information

For the sake of simplicity, we assume that there are only two types, commitment and normal as we call them. Prior to time 0, nature chooses player 1 to be either the commitment type with probability λ or the normal type with probability $1 - \lambda$. Let us denote the type space as $\Omega = \{n, c\}$.

The payoff structure of the normal type of player 1 is as shown in the stage game. The commitment type, on the other hand, exerts high effort with probability 1 at every period.

2.3 Repeated game with review platform

The stage game is played repeatedly between a seller and an infinite sequence of buyers with a different buyer each period. Player 1 discounts the future with $\delta < 1$. Player 2 lives for one period and only observes the outcome in his own period.

A review platform observes the past play and announces a public score $s \in S = \{1, 2, ..., K\}$ at the beginning of each period. The set of scores, S, is finite. The initial score is chosen randomly.

The review platform R is denoted as $\{S, \pi_0, P\}$ where S is a finite set of scores, $\pi_0 \in \Delta(S)$ is the initial distribution that the initial score is drawn from and P is a transition rule. A transition rule maps histories that consist of past actions and scores to a distribution of probabilities over S. Theorem 1 assumes for simplicity that the platform has the information of player 1's type ahead of the game and as a result, the initial score may condition on the type. Theorem 2 gets rid of this assumption.

For our purpose, it is sufficient to focus on Markovian transition rules. We define a Markovian transition rule as a map $P : S \times A \to \Delta(S)$ because even when the platform knows the type of player 1, it does not condition the transition probabilities on that information.

At time 0, s_0 is announced according to π_0. Then, players move simultaneously to play the stage game. The review platform observes the actions taken by the players and updates the score for time 1. This process is repeated infinitely.

2.4 Equilibrium definition:

At time t before the stage game is played, player 1 has a private history of h_1^t which consists of all past actions $\{(a_{1,0}, a_{2,0}), ..., (a_{1,t-1}, a_{2,t-1})\}$, and review scores $\{s_0, ...s_t\}$. Player 2 does not observe anything other than the current public score of s_t, so h_2^t only consists of s_t. Let H_1^t and H_2^t denote the set of all possible histories for player 1 and 2. Then, we can write $H_1^0 = S$, $H_1^t = (\Pi_{i=0}^{i=t} S \times \Pi_{i=1}^{i=t} A)$ for $t > 1$ and $H_2^t = S$ where $A = A_1 \times A_2$.

Let $H_1 = \cup_{t=0}^{\infty} H_1^t$ be the set of all possible histories for player 1. We use σ_1 to denote the strategy of player 1 which maps player 1's histories to the set of probability distributions over her action space, $\sigma_1 : H_1 \to \Delta(A_1)$. Since there are only two pure actions, we can identify a mixed strategy by the probability of playing one of the actions, say H: $\sigma_1 := \sigma_1(H)$. The strategy of a player 2 that lives at time t is a map $\sigma_{2,t} : S \to \Delta(A_2)$. Similarly, we will use $\sigma_{2,t}$ to denote $\sigma_{2,t}(B)$ from now on. Let σ_2 be the collection of all period-t strategies. The strategy spaces of player 1 and player 2 are Σ_1 and Σ_2 respectively.

Let $u_1(a_1, a_2)$ and $u_2(a_1, a_2)$ denote the stage game payoffs of player 1 and 2 when the actions a_1 and a_2 are taken.

A review platform R, a strategy profile (σ_1, σ_2) and the type model (Ω, λ) altogether induce a probability distribution Z over action profiles and scores. The discounted average payoffs to the normal type of player 1 and period-t player 2 at time t are:

$$U_1(\sigma_1, \sigma_2 | h_1^t) = E^Z\Big((1-\delta)\sum_{t=0}^{\infty} \delta^t u_1(a_1, a_2) | h_1^t\Big)$$
$$U_{2t}(\sigma_1, \sigma_2 | h_2^t) = E^Z\big(u_2(a_1, a_2 | h_2^t)\big)$$

Definition 1: A strategy profile (σ_1, σ_2) along with the beliefs $\lambda_t(s)_{t\in\mathbb{N}, s\in S}$ is a Perfect Bayesian Equilibrium if for all t,

$$(i)\ U_1(\sigma_1, \sigma_2|h_1^t) \geq U_1(\sigma_1', \sigma_2|h_1^t)\ \ \forall\sigma_1' \in \Sigma_1,\ \ \forall h_1^t \in H_1$$
$$(ii)\ U_{2t}(\sigma_1, \sigma_2|h_2^t) \geq U_{2t}(\sigma_1, \sigma_2'|h_2^t)\ \ \forall\sigma_2' \in \Sigma_2,\ \ \forall h_2^t \in S$$
$$(iii)\ \text{The beliefs of player 2 are updated according to Bayes' rule.}$$

The review platform we are going to construct does not allow for any out of equilibrium path. Therefore, $Z(s)$ is strictly positive for all $s \in S$. Player 1 knows her type but player 2 will need to update his belief using Bayes' rule when he observes a score.

3 Main Result

Let B_2 denote the best response correspondence of player 2. It can be written as

$$B_2(\alpha_1) = \{\alpha_2 \in \Delta(A_2)|U_2(\alpha_1, \alpha_2) \geq U_2(\alpha_1, \alpha_2')\ \forall\alpha_2' \in \Delta(A_2)\}$$

Let $\bar{v}(\sigma_1)$ denote the Stackelberg payoff of player 1. That is,

$$\bar{v} = \max_{\alpha_1\in\Delta(A_1)} \max_{\alpha_2\in B_2(\alpha_1)} U_1(\alpha_1, \alpha_2)$$

and let α_s denote the Stackelberg strategy. Then, we can write

$$\bar{v} = \alpha_s(p - c) + (1 - \alpha_s)p$$

The key here is that the Stackelberg payoff is a concept which has nothing to do with the different types of player 1. With a commitment type in the game, who is fine with receiving a less payoff than $\bar{v}$, the normal type can extract the rest of the remaining surplus. Then, let us define the adjusted Stackelberg payoff to the normal player 1. That is,

$$\bar{V} = \frac{\bar{v} - \lambda(p - c)}{1 - \lambda}$$

which comes from the condition that average payoff to player 1 is $\bar{v}$ in its maximum: $\lambda(p-c) + (1 - \lambda)\bar{V} = \bar{v}$. It is easy to check that $\bar{V}$ is strictly greater than $\bar{v}$. Finally, we assume that the prior belief λ is strictly less than α_s. Otherwise, hiding all the information from the buyers would be an easy solution. Our main result follows as below.

Theorem 1: For any $\lambda \in (0, \alpha_s)$ and $u < \bar{V}$, there exists a $\bar{K}$ such that for all $K > \bar{K}$, there exists a $\bar{\delta}$ such that for all $\delta > \bar{\delta}$ there exists a review platform $R^* = \{S, \pi^0, P^*\}$ and a stationary Perfect Bayesian Equilibrium $\{\sigma_1^*(s), \sigma_2^*(s), \lambda^*(s)\}_{s\in\{1,\ldots,K\}}$ where the payoff to the normal type of the long-run player is at least u after every history.

Proof Sketch: We present here the sketch of the proof and leave the formal proof to the appendix.

(i): The transition rule of the review platform induces a steady state distribution over the scores for each of the types. We first show that for the commitment type, each steady-state probability of the scores from 1 to $K/2$ is arbitrarily close to 0 for some large values of δ and K.

(ii): Next, by using (i), we show in the appendix that the probability of visiting score 1 for the normal player 1 is arbitrarily close to 0 for some large values of K and δ.

(iii): According to our equilibrium strategy, score 1 is the only state in which trade does not occur and (ii) states that the probability of visiting such a state even for the normal type is almost 0. The next step is to show that there is no unnecessary reputation built-up. In other words, we show that there exists a transition rule such that the posterior belief upon observing score K is equal to α_s.

(iv): Knowing the type of player 1, the platform chooses the posteriors to be equal to the steady-state probabilities. To make this idea more concrete, consider the case when the platform did not know the type of player 1. If the review platform did not know the true type of player 1, the early posteriors would be close to the prior belief λ and it would take an infinite amount of updates for the posteriors to reach the steady-states. Therefore at any point in time, the posterior at the highest score K would be less than the minimal belief for trade to take place. Therefore no buyer would buy at the highest score when player 1 does not exert any effort and our equilibrium unravels. In the equilibrium that we construct, player 2 is indifferent between buying and not buying at all the scores in which trade exists. As such, any perturbation of the transition rule or of the equilibrium strategy of player 1 will cause inefficiencies. Therefore, we assumed that the review platform knows player 1's type so it can choose the initial distribution to be equal to the steady-state distribution over type space and scores. □

Note that $\bar{V}$ is strictly greater than the Stackelberg payoff $\bar{v}$ for all $\lambda > 0$. Therefore, Theorem 1 states that even for small probabilities of a presence of the commitment type, the payoff the normal type of player 1 is higher than the Stackelberg payoff.

4 Example

Consider the stage game below. The payoff matrix satisfies all the conditions we stated before. The unique Nash equilibrium outcome of the stage game is $(0, 0)$. The Stackelberg strategy of player 1 is to play H with probability 0.5, and hence her Stackelberg payoff is 1.5.

	B	N
H	1,2	−1,0
L	2,−2	0,0

Figure 1 visualizes all payoff profiles of the stage game. $(1,2)$ is the outcome when player 1 is committed to her pure Stackelberg action, and $(1.5,0)$ is when she is committed to her mixed Stackelberg strategy in a complete information game. Overall, 1.5 is the highest feasible payoff to player 1. We do not claim to obtain any non-feasible payoff. However, we do make the distinction of the payoff to the commitment and the payoff to the normal type of player 1. The equilibrium we propose yields an average payoff profile arbitrarily close to $(1.5,0)$. In other words, player 1 extracts all the information rent that is shared by the commitment and the normal type of player 1. Given that player 2 buys almost surely, the commitment type's payoff is arbitrarily close to 1. Hence, given $\lambda < 0.5$, the profile of the normal type of player 1 and player 2 will be arbitrarily close to a point on the line between $(1.5,0)$ and $(2,-2)$, marked by the blue point. This gives player 1 a strictly higher payoff than 1.5.

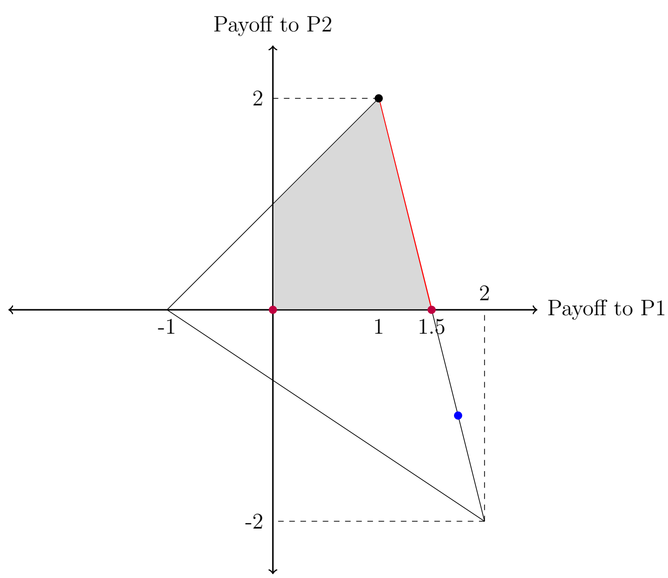

Figure 1: Payoff profile space of the stage game

The prior belief λ on player 1 being the commitment type is 1/3. $\lambda = 0$ represents the complete information case. In the complete information game, the highest payoff player 1 can achieve is 1. This is because if there is no commitment type, player 1 needs to exert high effort with positive probability whenever player 2 is buying. As a result, exerting high effort whenever her product is purchased is her best reply, from which her stage game payoff cannot exceed the payoff of 1, so is her average discounted payoff. This proof does not apply to the incomplete information game because with the existence of a commitment type, there will be some scores under which the best response of player 1 when her product is purchased is cheating player 2

$$
a_1 = H : \begin{bmatrix} 5/12 & 0 & 7/12 & 0 & \dots & & & \\ 1/6 & 1/6 & 2/3 & 0 & \dots & & & \\ & \ddots & \ddots & \ddots & & & & \\ \dots & 0 & 1/6 & 1/6 & 2/3 & 0 & \dots & \\ & & & \ddots & \ddots & \ddots & & \\ & & \dots & 0 & 1/6 & 1/6 & 2/3 & \\ & & & \dots & 0 & 1/6 & 5/6 & \end{bmatrix}
$$

$$
a_1 = L : \begin{bmatrix} 5/12 & 0 & 7/12 & 0 & \dots & & & \\ 1/6 & 5/6 & 0 & \dots & & & & \\ & \ddots & \ddots & \ddots & & & & \\ \dots & 0 & 1/6 & 5/6 & 0 & \dots & & \\ & & & \ddots & \ddots & \ddots & & \\ & & \dots & 0 & 1/6 & 5/6 & 0 & \\ & & & \dots & 0 & 1/6 & 5/6 & \end{bmatrix}
$$

Figure 2: The transition probabilities based on the action of player 1

by exerting low effort.

In the incomplete information game, Ekmekci [6] attains almost the Stackelberg payoff of 1.5 for player 1.

Let $S = \{1, 2, ..., K\}$ be the set of scores. Our theorem says that for any $u < 1.75$, there exists a $\bar{K}$ such that for all $K > \bar{K}$, there exists a $\bar{\delta}$ such that for all $\delta > \bar{\delta}$, there exists a review platform and a Perfect Bayesian Equilibrium where the payoff to the normal type of the long-run player is at least u after every history.

Step 1: Constructing the transition rule P. [1] Recall that the review platform updates the score based on whether H or L is observed. The commitment type always plays H and the normal type is free to choose between the two. We construct two transition matrices for high and low effort. Each matrix consists of transition probabilities from the most recent score to the new score given the action played by player 1. The rows represent the most recent score and the columns represent the new score. Consider the 2nd row. The punishment for playing L is losing the probability of upgrade which is 2/3 and having it added to the probability of staying at the same score. For all the scores except the lowest one, the score can be downgraded or upgraded by one, or remain the same. When the score is 1 at its lowest, it either jumps to 3 or remains at 1. This jump that we allow for, unlike Ekmekci [6], makes the computations harder but improves the highest payoff to player 1 significantly.

The first matrix is ergodic but the second is not. If L was to be played in all the states,

[1] Referring to the appendix for the general form of the transition probabilities, we note that the transition rule depends on a parameter η. We show in the appendix that a value of η with the appropriate properties can always be found.

there would be no way of reaching the high states. However, in the equilibrium we construct, there is always a positive probability of H being played except the highest score. Therefore, averaging the transition matrix for H, and that for L; we obtain an ergodic transition matrix. Hence there is a unique stationary distribution for each of them. Let $\pi_c^* \in \Delta(S)$ and $\pi_n^* \in \Delta(S)$ denote the stationary distributions of the commitment and the normal type respectively.

Step 2: Equilibrium strategies. Consider the strategies $\sigma_1^* : S \to [0,1]$ and $\sigma_2^* : S \to [0,1]$:

$$\sigma_1^*(s) = \begin{cases} 0 & \text{if } s \in \{1, K\} \\ \frac{0.5-\lambda^*(s)}{1-\lambda^*(s)} & \text{otherwise} \end{cases}$$

where $\lambda^*(s)$ is the posterior belief of player 2 on player 1 being the commitment type upon observing the score s. It is formed according to Bayes' rule:

$$\lambda^*(s) = \frac{\lambda \pi_c(s)}{\lambda \pi_c(s) + (1-\lambda)\pi_n(s)}$$

and

$$\sigma_2^*(s) = \begin{cases} 0 & \text{if } s = 1 \\ 1 & \text{otherwise} \end{cases}$$

The equilibrium we are proposing has 3 phases:

(i): Score 1 serves a punishment phase. Player 2 does not buy the product. Condition 2 states that player 1 strictly prefers trade so no-trade is a punishment. The transition rule at this score is independent of player 1's action. Hence she plays L.

(ii): Middle scores from 2 to K-1 serve as reputation building phase. Player 1 puts the least amount of effort necessary for player 2 to buy the product. Hence, buying is one of the best responses for player 2.

(iii): Score K serve as the reputation milking phase. Since player 2 thinks that player 1 is likely to be the commitment type, he buys the product. The transition rule at this score is independent of actions again, therefore player 1 plays L.

Step 3: Perturbing P to make σ_1^* optimal. Given σ_2^* and $\delta < 1$, the proposed strategy σ_1^* for the normal type of player 1 is not optimal. We need to perturb P so that the difference in the continuation values of any two adjacent scores must be $1.5(1-\delta)/\delta$. In order to understand this argument, let us calculate the difference between expected continuation values to player 1 from playing H and L. At any middle score $s \in \{2, 3, ..., K-1\}$:

$$V(s|H) = (1-\delta)u_1(H, B) + \delta[2/3V(s+1) + 1/6V(s) + 1/6V(s-1)]$$
$$V(s|L) = (1-\delta)u_1(L, B) + \delta[5/6V(s) + 1/6V(s-1)]$$

Since player 1 builds reputation at middle scores by mixing H and L, we must have $V(s) = V(s|H) = V(s|L)$ which implies $V(s+1) - V(s) = 1.5(1-\delta)/\delta$.
Let $P(i, j, a_1)$ denote the probability of moving from score i to score j when action a_1 is played. We perturb P in the following way:

$$P^*(s, s-1, a_1) = 1/6 + (K-s)(1-\delta)/\delta$$
$$P^*(s, s, a_1) = \begin{cases} 1/6 - (K-s)(1-\delta)/\delta & \text{if } s > 1 \text{ and } a_1 = 1 \\ 5/6 - (K-s)(1-\delta)/\delta & \text{if } s > 1 \text{ and } a_1 = 0 \\ 5/12 + (K-s)(1-\delta)/(2\delta) & \text{if } s = 1 \end{cases}$$
$$P^*(1, 3, a_1) = 7/12 - (K-s)(1-\delta)/(2\delta)$$

For any fixed K, the perturbed transition probabilities are well defined for δ sufficiently close to 1. The perturbed transition rule ensures that given player 2's strategy σ_2^*, σ_1^* is optimal.

Step 4: Fixing K and δ. We have shown that given the proposed σ_2^*, σ_1^* is optimal. Now we will show that for some (K, δ), σ_2^* is optimal given the beliefs held by player 2. For any middle score and the posterior upon observing that score, σ_1^* is high enough for player 2 to buy the product. In order for player 2 to buy at the score of K as well, we need the reputation of player 1 at the highest score, $\lambda^*(K)$, to be at least 0.5. The perturbed transition rule together with σ_1^* induces a posterior belief upon observing each score. Suppose that K is 50 and δ is 0.9996. With given parameters, $\lambda^*(K)$ becomes 0.501. Hence, player 2 buys the product even though the normal type of player 1 exerts low effort. Note that, as K increases, $\lambda^*(K)$ may drop below 0.5. However, as we show in the appendix, there is always a transition rule which makes $\lambda^*(K)$ at least 0.5. We could have constructed a transition rule to obtain a much higher reputation than 0.5 at the highest score as Ekmekci [6] did. However, while constructing the transition rule and the equilibrium, we economize on the reputations that player 1 would have because higher reputation implies a higher amount of effort having been spent. Therefore, we would like the belief at the score of K as close as possible to 0.5. We show in the appendix that there is a review platform that makes $\lambda^*(K)$ equal to 0.5

Step 5: Early periods. If the review platform did not observe the type of player 1, the early scores announced by the platform would not be as informative. In order to understand this argument better, consider the very first period at time 0. Since there would be nothing to observe yet, the initial score would have no information and thus the prior belief would not be updated. Yet the posterior beliefs would converge to the steady state beliefs with time. We cannot overcome this issue by any perturbation because player 2 is indifferent in all scores bigger than 1. Hence, we assume that the review platform can choose an initial distribution depending on player 1's type. That is,

$$\pi^0(w,s) = \begin{cases} \pi_c^*(s) & \text{if } \omega = c \\ \pi_n^*(s) & \text{if } \omega = n \end{cases}$$

4.1 No-jump Case

In this section, we are going to analyze the example in more detail to show the improvement in the payoff to the long-run player relative to Ekmekci [6]. To start with, as we mentioned earlier, we first change his setup by allowing for jumps between the review scores. However, there is another essential step forward to our results. That is economizing the reputations. We present two equilibria. The first one is the most economizing equilibrium given a review platform with no jump as in as the rating system of Ekmekci [6]. In the second one, player 1 plays H and L with equal probabilities whenever he is indifferent as she does in Ekmekci [6] but under a review platform with a jump as in ours. To make this part more clear, we are going to find v_1 and v_2 in Table 1 where the payoff row represents the payoff in the limit to player 1. According to

	Example 1	Example 2	Ekmekci	Us
Economized Reputations	Yes	No	No	Yes
Jump	No	Yes	No	Yes
Payoff	v_1	v_2	1.5	1.75

Table 1: Payoffs to the long-run player under different settings and equilibria

the example, our review platform can achieve payoffs to the long-run player arbitrarily close to 1.75 whereas Ekmekci [6] obtains almost 1.5. Suppose the transition probabilities without the perturbation are as follows in Figure 3: Furthermore, we propose the same strategies whenever player 1 is indifferent for $s < K$:

$$\sigma_1^*(s) = \begin{cases} 0 & \text{if } s \in \{1, K\} \\ \frac{0.5-\lambda^*(s)}{1-\lambda^*(s)} & \text{otherwise} \end{cases}$$

Then, as long as K is greater than 3, the posterior belief at K is going to be higher than 0.5, hence player 1 is able to cheat player 2 by exerting no effort. In this equilibrium, there is no reputation wasted. In other words, player 1 is never exerting more effort than necessary. However, if we wrote down the recursive formula and computed the value functions, each

$$a_1 = H : \begin{bmatrix} 5/8 & 3/8 & 0 & \dots & & & & \\ 1/8 & 5/8 & 1/4 & 0 & \dots & & & \\ & \ddots & \ddots & \ddots & & & & \\ \dots & 0 & 1/8 & 5/8 & 1/4 & 0 & \dots & \\ & & & \ddots & \ddots & \ddots & & \\ & & \dots & 0 & 1/8 & 5/8 & 1/4 & \\ & & & \dots & 0 & 1/8 & 7/8 & \end{bmatrix}$$

and

$$a_1 = L : \begin{bmatrix} 5/8 & 3/8 & 0 & \dots & & & & \\ 1/8 & 7/8 & 0 & \dots & & & & \\ & \ddots & \ddots & & & & & \\ \dots & 0 & 1/8 & 7/8 & 0 & \dots & & \\ & & & \ddots & \ddots & \ddots & & \\ & & \dots & 0 & 1/8 & 7/8 & 0 & \\ & & & \dots & 0 & 1/8 & 7/8 & \end{bmatrix}$$

.

Figure 3: The transition probabilities without jump

payoff would be less than 1.5. Finding the same value functions as in Ekmekci [6] as a result of the recursive equations makes perfect sense analytically because the only thing we changed is the randomization probabilities which should not affect the outcome. The intuition though for the values not increasing is that as player 1 exerts less effort, with the platform Ekmekci [6] designed, the probability of visiting the lowest score in which no trade occurs significantly increases because his rating system does not have the upward shift. The upward shift is; between any pair of adjacent scores, the steady-state probability of the higher score is greater than the lower one. Recall that in our platform, that probability is almost 0. Hence, our platform enables the long-run player to economize his reputation without the worry over no-trade possibility.

Next, we are going to look at another equilibrium when the platform has the same attributes as ours but player 1 is not smart enough to economize on his reputations. Suppose the transition probabilities are as shown at the beginning of this section in Figure 2, and player 1 plays H and L with probability 0.5 whenever he is indifferent. Recall that in order to achieve the properties of our equilibrium, the number of total scores K must be a large number. Moreover, the posterior belief at the highest score is 0.5 so that player 2 is indifferent. With player 1 exerting more effort at the middle scores, the probability of her visiting the highest score will be higher than before, decreasing the posterior belief below 0.5 In order to bring it back up to 0.5, we would need to update the transition probabilities by the parameter η (see Appendix for more information on η). In our example, when the review platform is designed as we proposed and player 1 economizes perfectly all her reputations, η goes to 1.5 as the total number of scores K increases for δ close enough to 1. Therefore, for a value of η slightly bigger than 1.5, there are values of K and δ such that the payoff to the long-run player is almost 1.75. When player 1 is playing H and L with probability 0.5 whenever she is indifferent, we would need to update η to be slightly bigger than a value of around 1.66. Remember that our set of recursive equations in the Appendix gives us the following formula for the value functions:

$$V(s) = 2 - \frac{\eta - 1}{2} - \frac{(K-s)(1-\delta)\eta}{\delta}$$

Hence, the values converge to 1.67 for every score. The updated version of Table 1 is shown below. As we see, the jump between the review scores is essential in passing the threshold

	Example 1	Example 2	Ekmekci	Us
Economized Reputations	Yes	No	No	Yes
Jump	No	Yes	No	Yes
Payoff	1.5	1.67	1.5	1.75

Table 2: Payoffs to the long-run player under different settings and equilibria

of mixed Stackelberg payoff. However, it is not enough to attain the maximum information rent available. In order for that to happen, we need player 1 to be smart enough to distribute her reputations over the scores wisely. Our equilibrium represents such a situation and the resulting payoff is the highest that she could ever attain in any setting. It is worth to note that the platform in Example 2 and the one we proposed are not identical. Nonetheless, the functional forms are. The review platform has been defined on only one input argument which is η. The updated η in Example 2 gives us the highest attainable payoff to the long-run player given the restriction that she exerts high effort with probability half whenever she is indifferent.

5 A Numerical Illustration

As in the example, let us fix K at 50 and δ at 0.9996. With η equal to 1.505, the resulting posterior beliefs look like in Figure 2. Note that Lemma 3 provides an analytical solution for each belief. That is, $\lambda^*(s) = 0.5^{K-s+1}$. As a result, the beliefs float right above zero for all except the very top scores. This is intuitive because, at the steady state, the probability of being at one of the top scores is higher than one of the lower scores. At all the low scores where the posterior belief is small, the normal type of player 1 plays H with a probability close to 0.5.

Figure 2 implies that reputation building is slower at low scores, which is a realistic finding because buyers hesitate to trust when the review scores are low.

The recent literature has attempted to explore the limits of attainable payoffs to the long-run player and we believe that this paper sets a new bar by reaching payoffs that were considered

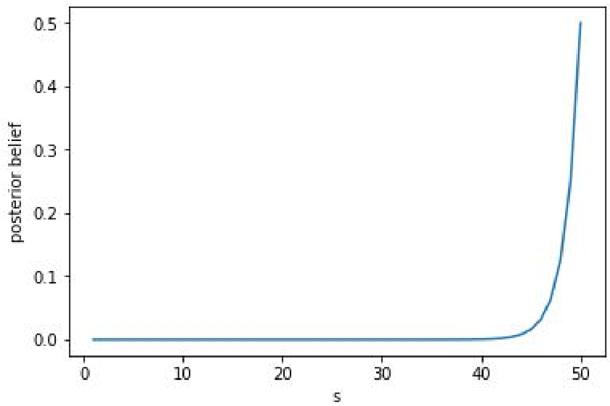

Figure 4: Steady state posterior belief distribution over the scores

to be non-feasible.

6 Review Platform without the information of the type

We have assumed that the review platform is informed of the type of the long-run player. While we are aware that this assumption might be challenged, we have made it for the sake of simplicity and our results do not rely on it. In this section, we demonstrate that attaining as high payoffs to the long-run player as we have is still possible with the review platform who is unaware of the type of the long-run player. However, our new equilibrium will depart from the previous one in a way. That is, we would need time-dependent strategies for the long-run player for some initial finite number of periods.

When the designer of the review platform does not know the type of the long-run player, the current transition rule does not work. In order to understand why it does not, let us consider the very first period. In the first period, there is no past action played and the platform's only option is to condition the score on the type of player 1. If the designer does not have access to that information, the first score will be uninformative. Hence, the belief will not be updated. As the game progresses and the platform observes more actions, the beliefs will get closer to their steady-state levels. However, they will never be equal to those. This fact makes it hard to economize the reputations. The reason is that what we mean by economizing the reputations is exerting such an amount of effort to make player 2 indifferent given the belief he holds. Since a posterior belief upon a given score is never equal to the posterior beliefs when the same score was observed before, we cannot maintain a stationary equilibrium which yields as high payoff as we have found. Therefore, we need the long-run player to play non-stationary strategies. Nevertheless, this is not enough either. Even though we can take care of the middle scores by imposing non-stationary strategies, there is no way to fix the issue at score

K. That is, we cannot enable the normal player 1 to cheat because, in the cheating state, she strictly prefers action L so cannot randomize between the two actions. With the uninformative initial score, the posterior belief at score K will always be less than what is needed in order for player 2 to buy the product unconditionally. Therefore, our proposed equilibrium will not work in this case. In addition to non-stationary strategies, we also need to tweak the platform to make the posterior belief at score K slightly higher than what it was before. As such, after a finite amount of periods, the posterior belief will exceed the minimum belief player 2 ought to have to buy unconditionally. Let us call that finite period N. Up until N, player 2 does not buy if the normal player 1 plays according to σ_1^* because as we said there will be some scores under which the beliefs will never be high enough to sustain σ_1^* and score K is one of them. Without a reward period, we can not incentivize the normal player 1 to put any effort at the middle scores. For this reason, we need to update the strategies and the review platform. We let stage-game Nash equilibrium be played until period N and adjust the platform such that the unconditional probability of announcing every score stays unchanged even though the normal player 1 is playing a completely different strategy. The way we do it is as the following. Remember that the normal player 1 had a transition rule P_n^* equal to a combination of $P^*(H)$ and $P^*(L)$, that is $\sigma_1^* P^*(H) + (1-\sigma_1^*)P^*(L)$ with some abuse of notation. Let us denote the new adjusted transition rule function by P^{**}. P_c^{**} will be the same as $P^*(H)$ so we set $P^{**}(H) = P^*(H)$. On the other hand, P_n^{**} is $P^{**}(L)$ until period N now, therefore we set $P^{**}(L) = P_n^* = \sigma_1^* P^*(H) + (1-\sigma_1^*)P^*(L)$. To put it in simple terms, the platform announces the scores until period N as if the normal player 1 is playing according to σ_1^* even though she is not. From N on, she will be playing time-dependent strategies due to the argument we made previously. As a result, we can assert the following theorem.

Theorem 2: For any $\lambda \in (0, \alpha_s)$ and $u < \bar{V}$, there exists a $\bar{K}$ such that for all $K > \bar{K}$, there exists a $\bar{\delta}$ such that for all $\delta > \bar{\delta}$ there exists a review platform $R^{**} = \{S, \pi_0^{**}, P^{**}\}$ and a Perfect Bayesian Equilibrium $\{\sigma_{1,t}^{**}(s), \sigma_{2,t}^{**}(s), \lambda_t^{**}(s)\}_{t=0,1,2,\ldots;s\in\{1,\ldots,K\}}$ where the payoff to the normal type of the long-run player is at least u after every history.

Proof: We provide here a sketch of the proof. The formal proof can be seen in the appendix.

i) Constructing the platform: In order to make player 2 strictly prefer buying at the score of K, we tweak the transition rule by slightly increasing the probability of staying at K once the score is K when the action. Since the normal type does not play H at the score K, this tweak does not affect the transition rule regarding the normal type. No matter how small the tweak is, there will be a time period before infinity that the belief upon observing K will exceed the α_s and stay above it forever.

ii) Equilibrium strategies: Let us call N the period when the belief upon observing the score of K would exceed α_s the first time. Note that K is not necessarily observed in the period N. Since N is finite, we do not worry about what happens before N by considering the values of δ approaching 1. As a result, we propose no-trade strategy profile until N. From N onward, we propose again the minimal effort level for each middle score, that makes player 2 indifferent. For scores 1 and K, the normal player 1 still exerts no effort.

iii) The intuition behind the optimality of the strategies: As we discussed, the platform now knowing the type of player 1 brought an issue to our equilibrium where player 2 was indifferent at all the scores expect the lowest one when there is no trade. We solved the issue for the score of K by tweaking the platform to make the option of buying strictly preferable. However, there may always be some other scores that need to be taken care of. By having introduced the time-dependent strategies, we do that. However, the transition rule regarding the normal type of player 1 is distorted with the time-dependent strategies. Therefore, we need to make sure that the belief upon observing K remains above α_s. We do that in the appendix.

iv) The periods until N: Up until N, we proposed no-trade strategies to be played. Then, with the given transition rule, the posterior belief at the score of K would be nowhere near α_s. Actually, it would be much higher than that because having exerted no effort thus far, the probability of visiting the highest score for the normal type would be quite low compared to what it would be if she played the minimal effort to make player 2 indifferent at every middle score. In order to solve this problem, we update the transition rule as follows. Before the time N, the platform announces a score as if the normal player 1 is following a strategy that prescribes a positive amount of effort for every middle score. This strategy is nothing but what the time-dependent strategies of the periods after N are converging to. In other words; even though the normal type is not putting any effort, the platform is covering her by keeping the unconditional probability of observation of the higher scores up.

7 Regulation

As most of the previous literature, we focus on how a long-run player benefits from reputation effects. However, the purpose of a review platform is supposed to make short-run buyers better off. The reason the literature has focused on the payoffs to the seller is that the non-trivial results lie in the search of those payoffs. Nevertheless, some earlier literature analyzes if it serves the buyers to reveal full information or conceal some of it in order to avoid the cold start problem pointed out by Lillethun [14]. The cold start problem is when buyers are certain that they are facing a strategic seller and therefore afraid to buy. This implies that it may be better for short-run players to be uninformed of some information known by the platform. In our model, depending on our assumption, the platform may observe first the type of the seller and then certainly observes her action each period. We find that a cold start problem exists if the platform knows and reveals the type of seller to the buyers. We need to elaborate on the cold start problem in our context because its definition is slightly different here. We define the cold start problem as the existence of the equilibrium of no-trade. As such, if the platform knows and reveals the type of the buyer, the game turns into a complete information game where no-trade equilibrium is possible. However, if the type of the seller is not revealed, we can use the Fudenberg / Levine [8] that the probability of no-trade even in the worst equilibrium for the seller is arbitrarily small. We propose a simple review platform which is fully informative on the most recent action of the seller in order to obtain a grim-trigger equilibrium. The platform has two possibles scores, the lower of which is announced if the seller plays L and the higher of which is announced when the seller plays H; and the punishment is no-trade. Under such a review platform, the only stationary equilibrium is the seller exerting high effort and the buyers buying every period.

Theorem 3: There exists a review platform that survives the cold start problem. It is defined for values of δ close to 1 as R:

i) R does not know or reveal w,

ii) $K = 2$,

iii) $P(1, 2, H) = P(2, 2, H) = 1$ and $P(1, 1, L) = P(2, 1, L) = 1$.

Proof: See Appendix.

8 Appendix

Proof of Theorem 1

i) Constructing the transition rule: The transition rule we propose has the following structure:

$$a_1 = H : \begin{bmatrix} 1-\gamma & 0 & \gamma & 0 & \dots & & \\ P^*(2,1,H) & P^*(2,2,H) & P^*(2,3,H) & 0 & \dots & & \\ & \ddots & \ddots & \ddots & & & \\ \dots & 0 & P^*(s,s-1,H) & P^*(s,s,H) & P^*(s,s+1,H) & 0 & \dots \\ & & & \ddots & \ddots & \ddots & \\ & & \dots & 0 & P^*(K-1,K-2,H) & P^*(K-1,K-1,H) & P^*(K-1,K,H) \\ & & & \dots & 0 & P^*(K,K-1,H) & P^*(K,K,H) \end{bmatrix}$$

and

$$a_1 = L : \begin{bmatrix} 1-\gamma & 0 & \gamma & 0 & \dots & & \\ P^*(2,1,L) & P^*(2,2,L) & 0 & \dots & & & \\ & \ddots & \ddots & \ddots & & & \\ \dots & 0 & P^*(s,s-1,L) & P^*(s,s,L) & 0 & \dots & \\ & & & \ddots & \ddots & \ddots & \\ & & \dots & 0 & P^*(K-1,K-2,L) & P^*(K-1,K-1,L) & 0 \\ & & & \dots & 0 & P^*(K,K-1,L) & P^*(K,K,L) \end{bmatrix}$$

where

$$
\begin{aligned}
\gamma &= \frac{2-\alpha_s(\eta-1)}{2\eta} - \frac{(K-1)(1-\delta)}{2\delta} \\
P^*(s,s-1,a_1) &= \alpha_s\frac{\eta-1}{\eta} + \frac{(K-s)(1-\delta)}{\delta} \quad a_1 \in \{H,L\} \\
P^*(s,s,1) &= \alpha_s\frac{\eta-1}{\eta} - \frac{(K-s)(1-\delta)}{\delta} \quad s \in \{2,3,...,K-1\} \\
P^*(s,s,0) &= P^*(s,s,1) + \frac{1}{\eta} \\
P^*(s,s+1,1) &= \frac{1}{\eta}
\end{aligned}
$$

The transition rule has a parameter η for now. Hence, we denote it by $P^*(\eta)$.

The transition rule we constructed for the example has the same form where $\eta = 1.5$.

ii) The conditions implied by the transition rule: The vectors of steady state probabilities for the normal and the commitment type, π_n^* and π_c^* induced by $P^*(\eta)$ and σ_1^* can be found as shown below:

$$\pi_\omega^* = \pi_\omega^* P_\omega^* \quad \omega \in \Omega \tag{1}$$

where P_n^* and P_c^* are the K x K transition matrices for the normal and the commitment types of player 1. Their each element can be found as shown below:

$$
\begin{aligned}
&P_n^*(i,j) = \sigma_1^*(i)P^*(i,j,H) + (1-\sigma_1^*(i))P^*(i,j,L) \quad i,j \in S \\
&P_c^*(i,j) = P^*(i,j,H) \quad i,j \in S
\end{aligned}
$$

The set (1) of equations implies the following conditions:

$$
\begin{aligned}
\pi_c^*(1)\gamma &= \pi_c^*(2)P_c^*(2,1) \\
\pi_n^*(1)\gamma &= \pi_n^*(2)P_n^*(2,1) \\
\Rightarrow \lim_{K(1-\delta)\to 0} \frac{\pi_c^*(2)}{\pi_c^*(1)} &= \lim_{K(1-\delta)\to 0} \frac{\pi_n^*(2)}{\pi_n^*(1)} = \frac{2-\alpha_s(\eta-1)}{2\alpha_s(\eta-1)} > 1.5 \qquad (2) \\
\pi_c^*(3) &= \pi_c^*(2)\frac{\frac{\eta-(\eta-1)(1-\alpha_s)}{\eta}+\epsilon_2}{\frac{\eta-1}{\eta}\alpha_s+\epsilon_3} \qquad (3) \\
\pi_n^*(3) &= \pi_n^*(2)\frac{\frac{\eta-(\eta-1)(1-\alpha_s)-(1-\sigma_1^*(2))}{\eta}+\epsilon_2}{\frac{\eta-1}{\eta}\alpha_s+\epsilon_3} \qquad (4) \\
\Rightarrow \lim_{K(1-\delta)\to 0} \frac{\pi_c^*(3)}{\pi_c^*(2)} &= \frac{1+\alpha_s(\eta-1)}{\alpha_s(\eta-1)} > 1 \quad \text{and} \\
\lim_{K(1-\delta)\to 0} \frac{\pi_n^*(3)}{\pi_n^*(2)} &= \frac{1+\alpha_s(\eta-1)-(1-\sigma_1^*(2))}{\alpha_s(\eta-1)} > 1 \\
\pi_c^*(4)\Big(\alpha_s\frac{\eta-1}{\eta}+\epsilon_4\Big) &= \pi_c^*(3)\Big(1-(1-\alpha_s)\frac{\eta-1}{\eta}+\epsilon_3\Big) \qquad (5) \\
&\quad - \pi_c^*(2)\frac{1}{\eta} - \pi_c^*(1)\Big(\frac{2-\alpha_s(\eta-1)}{2\eta}-\epsilon_1/2\Big) \\
\pi_n^*(4)\Big(\alpha_s\frac{\eta-1}{\eta}+\epsilon_4\Big) &= \pi_n^*(3)\Big(1-(1-\alpha_s)\frac{\eta-1}{\eta}-(1-\sigma_1^*(3))\frac{1}{\eta}+\epsilon_3\Big) \qquad (6) \\
&\quad - \pi_n^*(2)\frac{\sigma_1^*(2)}{\eta} - \pi_n^*(1)\Big(\frac{2-\alpha_s(\eta-1)}{2\eta}-\epsilon_1/2\Big) \\
\pi_c^*(s+1)\Big(\alpha_s\frac{\eta-1}{\eta}+\epsilon_{s+1}\Big) &= \pi_c^*(s)\Big(1-(1-\alpha_s)\frac{\eta-1}{\eta}+\epsilon_s\Big) \qquad (7) \\
&\quad - \pi_c^*(s-1)\frac{1}{\eta} \\
\pi_n^*(s+1)\Big(\alpha_s\frac{\eta-1}{\eta}+\epsilon_{s+1}\Big) &= \pi_n^*(s)\Big(1-(1-\alpha_s)\frac{\eta-1}{\eta}-(1-\sigma_1^*(s))\frac{1}{\eta}+\epsilon_s\Big) \qquad (8) \\
&\quad - \pi_n^*(s-1)\frac{\sigma_1^*(s-1)}{\eta} \qquad s\in\{4,...,K-1\} \\
\pi_n^*(K)\alpha_s(\eta-1) &= \pi_n^*(K-1)\sigma_1^*(K-1) \qquad (9) \\
\pi_c^*(K)\alpha_s(\eta-1) &= \pi_c^*(K-1) \qquad (10)
\end{aligned}
$$

where $\epsilon_s = \frac{(K-s)(1-\delta)}{\delta}$.

iii) Equilibrium strategies: The strategies are as follows:

$$\sigma_1^*(s) = \begin{cases} 0 & \text{if } s \in \{1, K\} \\ \frac{0.5-\lambda^*(s)}{1-\lambda^*(s)} & \text{otherwise} \end{cases}$$

$$\sigma_2^*(s) = \begin{cases} 0 & \text{if } s = 1 \\ 1 & \text{otherwise} \end{cases}$$

iv) Existence of optimal η: We pin down η by the condition $\lambda^*(K) = \alpha_s$. Hence, we show that there exists such an $\eta \in (1, 2)$.

Consider when η is 1. Equations (11), (12) and (13) imply that $\pi_c^*(s+1)/\pi_c^*(s)$ converges to infinity as ϵ_s goes to 0. Therefore, $\pi_c^*(K)$ converges to 1.

Consider a positive number smaller than α_s, say $\alpha_s/2$. Let us define N^0 to be $N^0 = \{s \in S \setminus K : \sigma_1^*(s) > \alpha_s/2\}$ and $N^1 = \{s \in S \setminus K : \sigma_1^*(s) \le \alpha_s/2\}$. If s is in N^0, we will show that $\pi_n^*(s+1)/\pi_n^*(s)$ grows infinitely for some large values of K and δ. Combining (8) and (9), we can write the same relation between any adjacent scores s and $s+1$ for $s > 2$. Also referring to (2) and (4) we can say that if s is in N^0, $\pi_n^*(s+1)/\pi_n^*(s)$ goes to infinity as η approaches 1. Therefore, $\sum_{s \in N^0} \pi_n^*(s)$ converges to 0.

If s is in N^1, then there exists a $b > 0$ such that $\pi_n^*(s) < \pi_c^*(s)/b$. Therefore, $\sum_{s \in N^1} \pi_n^*(s) < \sum_{s \in N^1} \pi_c^*(s)/b$.
Since for every $s < K$, s is either in N^0 or N^1, we conclude $\pi_n^*(K)$ goes to 1.

When η is 2, the probability of downgrade and the probability of upgrade at any $s > 1$ would be equal to each other in the limit of ϵ_s going to 0 for every s if $\sigma_1^*(s)$ was α_s for every middle score. Given that $\sigma_1^*(s)$ is always less than α_s, the upward drift for the normal type fades away. However, there is always a strict upward drift for the commitment type. Therefore, the posterior belief at score K approaches to 1 as η goes to 2.

We have shown that for small values of η, the prior is not updated at score K, and for high values of K, the posterior at score K converges to infinity. Since the posterior at score K is continuous with respect to η, by intermediate value theorem, there has to be an η such that $\lambda^*(K)$ is exactly α_s.

v) $\lim_{K(1-\delta)\to 0, \delta \to 1} \pi_n^*(1)$=0: First, we prove that the probability of being at one of a lower portion of the scores for the commitment type gets arbitrarily close to 0 for some large values of K and δ. By using this result, we then show that the probability of visiting the lowest score even for the normal type is arbitrarily close to 0.

We can state the first result more formally as follows.

Lemma 1: For every $\eta \in (1, 2)$ and $\rho > 0$, there exist a $\bar{K}$ such that for all $K > \bar{K}$, there exists a $\bar{\delta}$ such that for all $\delta > \bar{\delta}$ and $s < K/2$, $\pi_c^*(s) < \rho$.

Proof of Lemma 1: We prove the lemma by showing that there is always an upward drift ($\pi_c^*(s+1) > \pi_c^*(s)$) for the committed type. This leads to steady-states of lower scores approaching to 0 as the total number of scores and the discount factor increase. The equations (7) and (10) imply that for some large values of δ,

$$\frac{\pi_c^*(s+1)}{\pi_c^*(s)} \approx \frac{1}{\alpha_s(\eta-1)} > 1 \tag{11}$$

for $s > 2$. For s equal to 1 and 2, the equations (2) and (3) imply

$$\frac{\pi_c^*(3)}{\pi_c^*(2)} \approx \frac{1+\alpha_s(\eta-1)}{\alpha_s(\eta-1)} > 1 \tag{12}$$

and

$$\frac{\pi_c^*(2)}{\pi_c^*(1)} \approx \frac{2-\alpha_s(\eta-1)}{2\alpha_s(\eta-1)} > 0.5 \tag{13}$$

Let us define $\epsilon = [1 + 1/(\alpha_s(\eta-1))]/2$ and $s^o = \max\{s \in S | s < K/2\}$. Then, there exist a K and a δ such that we can write $\pi_c^*(s^o) < \pi_c^*(K)/(1+\epsilon)^{K-s^o}$. Hence, $\pi_c^*(s^o)$ and $\pi_c^*(s)$ for all $s < s^o$ approaches to 0. $\square$

Now, using Lemma 1, we show that the probability of visiting score 1 for the normal type of player 1 is arbitrarily close to 0 for some large values of K and δ.

Lemma 1.1: For every $\eta \in (1,2)$ and $\rho > 0$, if $\pi_n^*(s) > \rho$ for some $s \in \{s \in S | 1 < s < K/2\}$; then there exists a $\bar{K}$ such that for all $K > \bar{K}$, there exists a $\bar{\delta}$ and a $\rho' > 0$ such that for all $\delta > \bar{\delta}$, $\sigma_1^*(s) > \alpha_s - \rho'$.

Proof of Lemma 1.1: For any $\eta \in (1,2)$ and $\rho > 0$, consider the $\bar{K}$ and the correspondence $\bar{\delta}(K)$, satisfying Lemma 1. For each $K > \bar{K}$ and $\delta > \bar{\delta}(K)$, $\pi_c^*(s)$ for $s < K/2$ is arbitrarily close to 0. Since $\pi_n^*(s) > \rho$ for a $\rho > 0$, the Bayesian update implies that the posterior probability is arbitrarily close to 0, which in turn implies that $\sigma_1^*(s)$ is arbitrarily close to α_s. $\square$

Lemma 1.2: For every $\eta \in (1,2)$ and K, there exist a $\rho > 0$ and $\bar{\delta}$ such that for all $\delta > \bar{\delta}$, $(\alpha_s - \rho)P^*(s, s+1, H) + (1 - \alpha_s + \rho)P^*(s, s+1, L) \geq P^*(s, s-1, a_1)$ for $s \in \{2, 3, ..., K-1\}$.

Proof of Lemma 1.2: Since K is finite, it is enough to show

$$\alpha_s P^*(s, s+1, H) + (1-\alpha_s)P^*(s, s+1, L) > P^*(s, s-1, a_1)$$

which can be simplified as

$$\alpha_s \frac{1}{\eta} > \left(\alpha_s \frac{\eta - 1}{\eta} + \epsilon_s\right)$$

Since η is fixed at a value less than 2, there exist a small enough ϵ_s that makes the inequality satisfied. Such small ϵ_s can be obtained by high enough δ. $\square$

Lemma 2: For every $\eta \in (1,2)$ and $\rho > 0$, there exist a $\bar{K}$ such that for every $K > \bar{K}$, there exists a $\bar{\delta}$ such that for all $\delta > \bar{\delta}$; $\pi_n^*(1) < \rho$.

Proof of Lemma 2: For a fixed $\eta \in (1,2)$, suppose, by contradiction that there exists a $p > 0$ such that for every K and δ, $\pi_n^*(1) > p$.

Followed by (2), $\pi_n^*(2) > p$ as well. Then, by Lemma 1.1, for every $p' > 0$ there exist some large values of K and δ such that $\sigma_1^*(2) > \alpha_s - p'$. Lemma 1.2 in turn implies that the probability of upgrade is higher than the probability of downgrade for score s. Also supported by equation (4), we can say that $\pi_n^*(3)$ is greater than p for some large values of K and δ, followed by Lemma 1.1 implying that $\sigma_1^*(3)$ being arbitrarily close to α_s, and Lemma 1.2 and equation (6) together implying $\pi_n^*(4) > \pi_n^*(3)$.

The low of motion with $\pi_n^*(s)$ together with the initial probability of $\pi_n^*(1)$ implies a bound, typically less than K, on the first state where the upward drift becomes downward. Let us call that score s^o. It is important to note that s^o is independent of K as long as K and δ are large enough.

In order for the upward drift to end, a necessary condition is that there exists a $\beta > 0$ such that $\sigma_1^*(s^o) < \alpha_s - \beta$. However, this is a contradiction. For any fixed s^0 and $\beta > 0$, if we continue increasing K, $\pi_c^*(s)$ for $s < s^0$ gets even closer to 0 and thus $\sigma_1^*(s)$ for $s < s^0$ gets even closer to α_s. Therefore, there exists a large enough K after which there exists high enough values of δ such that $\sigma_1^*(s^o) > \alpha_s - \beta$. $\square$

vi) Early periods: We take care of the early rounds by choosing the initial distribution to reflect the steady states:

$$\pi_0^*(w,s) = \begin{cases} \pi_c^*(s) & \text{if } w = c \\ \pi_n^*(s) & \text{if } w = n \end{cases}$$

In Theorem 2, we get rid of the assumption that the review platform knows the type of player 1. We chose to present our main result in two theorems in the hope of clarity. The goal of Theorem 1 is to give the reader the essence of our method.

vii) Calculating the payoff: In order to achieve the highest payoff to the normal player 1, we do not only need trade happening, but we also need her to exert as low effort as possible. In other words, the normal type of player 1 should not build any reputation more than she needs. At any middle score, player 2 is indifferent, so there is no unnecessary reputation at middle scores. The transition rule implies that at score K, the normal player 1 plays L and player 2 plays B. In order for the equilibrium to work, the posterior belief at K must be no less than α_s. Since we do not want any unnecessary reputation, it should not be greater than

α_s either. Therefore, we pick an η such that $\lambda^*(K)$ is exactly ${\alpha_s}^2$. So far, we have proved the existence of such η. Also, we have shown that the probability of visiting the only state in which trade does not occur, which is score 1, is almost 0 even for the normal type of player 1.

The rest of the work is only showing the closed form expression of the payoff for each score. The value of a score of s can be determined by the following recursive equation:

$$V(s) = (1-\delta)U_1(\sigma_1^*(s), \sigma_2^*(s)) + \delta[\sigma_1^*(s)\sum_{s' \in S} \tau(s,s',H)V(s') \\ + (1-\sigma_1^*(s))\sum_{s'\in S}\tau(s,s',L)V(s')]$$

and we find

$$V(s) = 2 - \alpha_s(\eta - 1) - \frac{(K-s)(1-\delta)\eta}{\delta}$$

As $K(1-\delta)$ goes to 0, the values approach $2-\alpha_s(\eta-1)$. In the limit, the values are also equal to the adjusted Stackelberg paoyff, that is $\frac{2-\alpha_s-\lambda}{1-\lambda}$. Therefore, η must be asymptotically equal to $1 + \frac{\alpha_s - \lambda}{(1-\lambda)\alpha_s}$. Note that the functional form of η confirms that it is always between 1 and 2. □

Before we close up, we show how the beliefs are distributed because we need the following result in the proof of Theorem 2.

Lemma 3: For every $s > 2$; the posterior belief $\lambda^*(s)$ is equal to α_s^{K-s+1}.

Proof of Lemma 3: Remember that we pinned down η by the condition $\lambda^*(K) = \alpha_s$. Within a period, we have the following condition

$$\sigma_1^*(K-1) = \frac{\alpha_s - \lambda^*(K-1)}{1-\lambda^*(K-1)} \tag{14}$$

$$= \alpha_s - (1-\alpha_s)\frac{\lambda\pi_c^*(K-1)}{(1-\lambda)\pi_n^*(K-1)} \tag{15}$$

By combining (9), (10) and (15); we find $\sigma_1^*(K-1) = \frac{\alpha_s}{1+\alpha_s}$. Plugging $\sigma_1^*(K-1)$ back in (14), we find $\lambda^*(K-1) = \alpha_s^2$. By iterating towards lower scores until $s = 2$, we find $\lambda^*(s) = \alpha_s\lambda^*(s+1)$. □

Lemma 3 highlights a crucial insight of our machinery. What we do in order to prove the lemma is to find a fixed point, that is σ_1^*. Once a vector of strategies is plugged in, the transition rule P^* yields a K dimensional vector of beliefs which, in turn, yields a vector of new strategies. We have shown that σ_1^* is a fixed point of this process. We will use this argument in more detail in the next section.

[2] It is important to emphasize that all we need is for the transition probabilities such that the buyer is indifferent at K because the other states will be taken care with randomization.

Proof of Theorem 2: i) Constructing the platform. We denote the transition rule function that we proposed in the previous proof as P^*. Let us define P^{**} as the following:

$$P^{**}(K, K-1, H) = P^*(K, K-1, H) - \epsilon$$
$$P^{**}(K, K, H) = P^*(K, K, H) + \epsilon$$
$$P^{**}(i, j, a_1) = P^*(i, j, a_1)$$

Since the normal type will still not be exerting any effort at score K, this update does not change her transition rule. However, it increases the stationary probability of visiting K for the commitment type while decreasing those of the other scores. Having said that, the recursive equation at every score for the normal type stays the same. Despite the fact that there is going to be a small amount of loss; by choosing a small ϵ and large N, K and δ; we can still stay arbitrarily close to the Pareto frontier. Before we finish up this section, it is worthwhile to note that the platform does not violate the time-independence. It is still defined on a finite space.

ii) Equilibrium strategies. We propose $\sigma_{1,t}^{**}$ and $\sigma_{2,t}^{**}$ as

$$\sigma_{1,t}^{**} = \begin{cases} 0 & \text{if } t < N \\ \frac{\alpha_s - \lambda_t^{**}(s)}{1-\lambda_t^{**}(s)} & \text{if } t \geq N, 1 < s < K \\ 0 & \text{if } t \geq N, s \in \{1, K\} \end{cases}$$

and

$$\sigma_{2,t}^{**} = \begin{cases} 0 & \text{if } t < N \\ 1 & \text{if } t \geq N, s > 1 \\ 0 & \text{if } t \geq N, s = 1 \end{cases}$$

We consider the first N periods as the experimentation stage. Since we focus on the payoff as δ goes to 1, we will ignore the payoff loss in this stage. During experimentation, the normal type and the short-run buyers play the no-trade strategies. We will update the platform to take care of the experimentation stage in the last section of the proof.

From N onward, player 1 plays time-dependent strategies. The reason that they are required is the following. Suppose that we have a stationary equilibrium. For some of the scores, the belief at any time would be lower than its steady-state level. Hence, the strategy of player 1 would not be optimal. This argument concludes that there cannot be a stationary equilibrium that achieves arbitrarily close payoff to $\bar{V}$. As a result, we impose the time-dependent strategies for player 1 from N onward. The drawback of time-dependent strategies is that now we have a heterogeneous transition matrix for the normal type because her transition depends on her strategy which is now distorted with the time-dependence. Therefore, we need to show that the distortion caused by the time-dependent strategies does not cause the belief at the highest score to move away from where it was at N.

iii) Checking if the proposed strategies are optimal: We need to show that even though the transition rule is distorted after period N with time-dependent strategies, the distorted distribution can be approximated by the stationary distribution which would be reached if there was no distortion.

Now, let us return the argument we pointed out at the end of the previous section. We have shown that σ_1^* is a fixed point of a function that we are about to call f^*. Let us explain f^* step by step. The strategy of the normal type is defined on $[0, \alpha_s]^{K-2} \times \{0\}^2$ which is a compact and convex set. The function f^* takes a K dimensional vector of strategies as its input and by combining the transition rule P^*, Bayesian update and our equilibrium definition (in order); returns a new vector of strategies. Since f^* is continuous, by Brouwer's fixed point theorem, it has a fixed point. In Theorem 1, we showed that σ_1^* is a fixed point of f^* and σ_1^* is such that $\lambda^*(s) = \alpha_s$. Let us call f^{**} the composite function associated with P^{**}. f^{**} carries the same properties with f^*. Therefore, it has a fixed point as well. Let us call σ_1^{**} a fixed point of f^{**}. By Fudenberg & Tirole [10], we know that the equilibrium set is upper hemicontinuous. Moreover, f^* and f^{**} each return only a vector of singletons, not a correspondence. Therefore; for any sequence of f^n with $f^n \to f^*$ and $\sigma_1^n = f^n(\sigma_1^n)$, it is true that $\sigma_1^n \to \sigma_1^*$. Hence, there is a small tweak ϵ such that $||\sigma_1^* - \sigma_1^{**}|| < \epsilon'$ for any $\epsilon' > 0$. Note that σ_1^{**} is only a vector of stationary strategies, not our equilibrium strategies. However, we are going to show that our equilibrium vector is arbitrarily close to σ_1^{**}.

In addition to the stationary parameters that we have defined, let us call $(\pi_c^{**}(s), \pi_n^{**}(s))$ the stationary distributions associated with P^{**} and $\sigma_1 **$. Notice that we denote our equilibrium parameters with time subscript and the stationary levels without one. Due to how we tweaked the transition rule, we are going to have a large enough N such that $\lambda_N^{**}(K) > \alpha_s$. We can write

$$\pi_n^{**}(K) = \pi_n^{**}(K-1)P_n^{**}(K-1, K) + \pi_n^{**}(K)P_n^{**}(K, K) \tag{16}$$

where $P_n^{**}(i,j) = P^{**}(i,j,H)\sigma_1^{**}(i) + P^{**}(i,j,L)(1 - \sigma_1^{**}(i))$. In addition, the transition between two consecutive periods implies

$$\pi_{n,N+1}^{**}(K) = \pi_{n,N}^{**}(K-1)P_{n,N}^{**}(K-1, K) + \pi_{n,N}^{**}(K)P_{n,N}^{**}(K, K) \tag{17}$$

where $P_{n,N}^{**}(i,j) = P^{**}(i,j,H)\sigma_{1,N}^{**}(i) + P^{**}(i,j,L)(1 - \sigma_{1,N}^{**}(i))$. Combining (16) and (17), we obtain

$$
\begin{aligned}
\pi^{**}_{n,N+1}(K) - \pi^{**}_n(K) = {} & [\pi^{**}_{n,N}(K-1) - \pi^{**}_n(K-1)]P^{**}_n(K-1,K) + [\pi^{**}_{n,N}(K) - \pi^{**}_n(K)]P^{**}_n(K,K) \\
& + \frac{\sigma^{**}_{1,N}(K-1) - \sigma^{**}_1(K-1)}{\eta}\pi^{**}_{n,N}(K-1)
\end{aligned}
$$

Let us define $\Delta\sigma^t_1(s) = \sigma^{**}_{1,t}(s) - \sigma^{**}_1(s)$. If we rewrite the above equation, it becomes

$$
\begin{aligned}
\pi^{**}_{n,N+1}(K) - \pi^{**}_n(K) = {} & [\pi^{**}_{n,N}(K-1) - \pi^{**}_n(K-1)]P^{**}_n(K-1,K) + [\pi^{**}_{n,N}(K) - \pi^{**}_n(K)]P^{**}_n(K,K) \\
& + \frac{\Delta\sigma^N_1(K-1)}{\eta}\pi^{**}_{n,N}(K-1) \qquad (18)
\end{aligned}
$$

We know that we can choose a large enough N such that $\pi^{**}_{n,N}$ will be close enough to $\pi^{**}_n(K)$ so that λ^{**}_N will still be greater than α_s.

Now, we show that $\Delta\sigma^t_1(K-1) < \Delta\sigma^N_1(K-1)$ for every $t > N$. This means that the strategies and so the other variables such as the transition probabilities and the beliefs stay around the steady-state levels.

Now, consider $\Delta\sigma^{N+1}_1(K-1)$. It can be expressed as

$$
\begin{aligned}
& \frac{\alpha_s - \lambda^{**}_{N+1}(K-1)}{1 - \lambda^{**}_{N+1}(K-1)} - \frac{\alpha_s - \lambda^{**}(K-1)}{1 - \lambda^{**}(K-1)} \\
= & (1-\alpha_s)\frac{\lambda}{1-\lambda}\Big[\frac{\pi^{**}_c(K-1)}{\pi^{**}_n(K-1)} - \frac{\pi^{**}_{c,N+1}(K-1)}{\pi^{**}_{n,N+1}(K-1)}\Big] \qquad (19)
\end{aligned}
$$

where

$$
\begin{aligned}
\pi^{**}_{n,N+1}(K-1) = {} & \pi^{**}_{n,N}(K-2)P^{**}_{n,N}(K-2,K-1) + \pi^{**}_{n,N}(K-1)P^{**}_{n,N}(K-1,K-1) \\
& + \pi^{**}_{n,N}(K)P^{**}_{n,N}(K,K-1) \qquad (20) \\
= {} & \pi^{**}_{n,N}(K-2)P^{**}_n(K-2,K-1) + \pi^{**}_{n,N}(K-1)P^{**}_n(K-1,K-1) \\
& + \pi^{**}_{n,N}(K)P^{**}_n(K,K-1) \\
& + \frac{\Delta\sigma^N_1(K-2)\pi^{**}_{n,N}(K-2)}{\eta} - \frac{\Delta\sigma^N_1(K-1)\pi^{**}_{n,N}(K-1)}{\eta} \qquad (21)
\end{aligned}
$$

As a result, we can consider $\Delta\sigma^{N+1}_1(K-1)$ as a function of $\Delta\sigma^N_1(K-2)$ and $\Delta\sigma^N_1(K-1)$. Let us call that function g^{N+1}. Note that $g^{N+1}(0,0) = 0$ for a small ϵ because when the strategies at time N are at their steady-state levels, so will be the strategies at $N+1$. Moreover, for a small ϵ and large N, we can write:

$$\begin{aligned}\Delta\sigma_1^{N+1}(K-1) &\approx \frac{\partial g(\Delta\sigma_1^{N+1}(K-1), \Delta\sigma_1^{N+1}(K-2))}{\partial \Delta\sigma_1^{N}(K-2)}\Delta\sigma_1^{N}(K-2) \\ &+ \frac{\partial g(\Delta\sigma_1^{N+1}(K-1), \Delta\sigma_1^{N+1}(K-2))}{\partial \Delta\sigma_1^{N}(K-1)}\Delta\sigma_1^{N}(K-1) \\ &\approx \frac{\alpha_s^2(1+\alpha_s+\alpha_s^2)(\eta-1)}{(1+\alpha_s)^2\eta}\Delta\sigma_1^{N}(K-2) - \frac{\alpha_s}{(1+\alpha_s)\eta}\Delta\sigma_1^{N}(K-1) \qquad (22) \\ &< \frac{\alpha_s}{1+\alpha_s}\max\{\Delta\sigma_1^{N}(K-1), \Delta\sigma_1^{N}(K-2)\}\end{aligned}$$

where (22) can be written by Lemma 3. Let $\gamma_N = \max_{s\in S}\{|\Delta\sigma_1^N(s)|\}$. Then, the condition above implies that there exists a small enough tweak such that $|\Delta\sigma_1^{N+1}(K-1)| < \gamma_N/2$ for a large N. Moreover, we can write the same condition for every score $s < K$. Consequently, we can say that the deviation from the stationary levels is contracting. This process can be iterated over infinitely many time periods. Hence, the strategies σ_2^{**} of player 2 are still optimal.

Let us explain what is happening intuitively. Suppose that there are only 3 scores: 1, 2 and 3. By tweaking the review platform, we have slightly increased the posterior belief at the score of 3 and decreased the beliefs at 1 and 2. As a result of this, the effort level at 2 now has to be higher. However, the effect on the tweak cannot perpetuate because since the effort level at 2 increased, the probability of being at the score of 2 (for the normal type) will decrease, which results in an increase in the belief. All in all, the parameters stay around their steady-state levels which are arbitrarily close to those in the equilibrium proposed by Theorem 1. Note that a (K, δ) tuple that works for Theorem 1 may not work to sustain this equilibrium. We would need even higher values of those in this one.

iv) Experimentation Stage: We take care of the experimentation stage as the following so that by the time N, the beliefs will be close to their steady-state levels.

$$P^{**}(i,j,a_1) := \begin{cases} P^{**}(i,j,a_1) & \text{if } a_1 = H, t < N \\ \sigma_1^{**}(i)P^{**}(i,j,H) + (1-\sigma_1^{**}(i))P^{**}(i,j,L) & \text{if } a_1 = L, t < N \\ P^{**}(i,j,a_1) & \text{if } t \geq N \end{cases}$$

Given the above transition rule, even though the normal type of player 1 plays L until period N, the transition rule announces a score as if she follows σ_1^{**}.

Proof of Theorem 3: Suppose, by contradiction, that there is an equilibrium with $(\sigma_1(s), \sigma_2(s) = (0,0)$ for every $s \in \{1,2\}$. Then, $\lambda(2)$ becomes 1 and player 2 strictly prefers buying at score 2. This is contradiction to no-trade at every score. □

References

[1] J. Best and D. Quigley. "Persuasion for the long run". 2018.

[2] P. C. "Public trust and government betrayal". *Journal of Economic Theory*, (127), 2006.

[3] H. L. Cole, D. J., and W. B. English. "Default, settlement, and signalling: Lending resumption in a reputational model of sovereign debt". *International Economic Review*, (36), 1995.

[4] M. Cripps, G. J. Mailath, and L. Samuelson. "Imperfect monitoring and impermanent reputations". *Econometrica*, (72), 2004.

[5] H. Di Pei. "Trust and betrayals reputation building and milking without commitment". 2018.

[6] M. Ekmekci. "Sustainable reputations with rating systems.". *Journal of Economic Theory*, (146), 2011.

[7] D. Fudenberg, D. Kreps, and E. Maskin. "Repeated games with long-run and short-run players.". *Review of Economic Studies,*, (57(4)), 1990.

[8] D. Fudenberg and D. K. Levine. "Reputation and equilibrium selection in games with a patient player.". *Econometrica*, (57(4)), 1989.

[9] D. Fudenberg and D. K. Levine. "Efficiency and observability with long-run and short-run players.". *Journal of Economic Theory*, (62), 1994.

[10] D. Fudenberg and J. Tirole. "Game theory.". *Cambridge, Massachusetts*, (p. 31), 1991.

[11] B. Holmstrom. "Managerial incentive problems: A dynamic perspective". *The Review of Economic Studies*, (66), 1999.

[12] E. Kamenica and M. Gentzkow. "Bayesian persuasion.". *The American Economic Review*, (101.6), 2011.

[13] M. P. R. J. . W. R. Kreps, D. M. "Rational cooperation in the finitely repeated prisoners' dilemma.". *Journal of Economic Theory,*, (27(2)), 1982.

[14] E. Lillethun. "Optimal information design for reputation building". 2017.

[15] Q. Liu. "Information acquisition and reputation dynamics". *The Review of Economic Studies*, (78.4), 2011.

[16] G. J. Mailath and L. Samuelson. "Who wants a good reputation?". *The Review of Economic Studies*, (68), 2001.

[17] Y. D. Mayzlin, Dina and J. Chevalier. "Promotional reviews: An empirical investigation of online review manipulation". *American Economic Review*, (104(8)), 2014.

[18] T. Wiseman. "Reputation and impermanent types". *Games and Economic Behavior*, (62), 2008.

[19] T. Wiseman. "Reputation and exogenous private learning". *Journal of Economic Theory*, (144), 2009.