Sustainable Reputations with Rating Systems*

Mehmet Ekmekci†
Princeton University

September 2006

Abstract

We study a class of moral hazard games between a long lived agent and an infinite sequence of short lived principals. We assume that principals cannot observe past signals. To facilitate the analysis of applications such as online auctions (e.g. eBay), online shopping search engines (e.g. BizRate.com) and consumer reports, we assume that a central mechanism observes all past signals, and makes public announcements every period. The set of announcements and the mapping from observed signals to the set of announcements is called a rating system. We show that absent reputation effects, information censoring cannot improve attainable payoffs. However, if there is an initial probability that the agent is a commitment type that plays a particular strategy every period, then there exists a rating system and an equilibrium of the resulting game such that, the expected present discounted payoff of the agent is almost his Stackelberg payoff after every history. This is in contrast to Cripps, Mailath and Samuelson (2004), where it is shown that reputation effects do not last long in such games if principals can observe all past signals. We also construct rating systems that increase payoffs of almost all principals, while decreasing the agent's payoff.

*I am indebted to Faruk Gul for guidance at every stage of this work. I am also grateful for suggestions and comments to Wioletta Dziuda, Aytek Erdil, Gabor Virag and Andrea Wilson. All remaining errors are my own.
†Department of Economics, Fisher Hall, Princeton, NJ 08544. E-mail: mekmekci@princeton.edu.

INTRODUCTION

Most economic activities leave room for abuse by one of the parties. When shopping from an online store, buyers make the payment before they receive the product, and the store "promises" to deliver the product on time as advertised. Some goods have characteristics that reveal themselves only after they are used for some period of time, such as resistance, or endurance. At the time of the payment a user cannot observe these characteristics. Thus, there needs to be some kind of "trust" for the activity to be undertaken.

Due to the recent development of internet and emergence of online marketplaces, every day thousands of transactions are executed between people who do not know each other. In these markets, buyers can see at most the photos of the product they are buying. They cannot be sure at the time of the payment whether the seller will deliver the product as advertised. There is even no guarantee that the seller will indeed send the product after receiving the payment. The most obvious examples of such an uncertain environment are internet auction sites such as eBay.

In economics, these problems fall under the category of "moral hazard problems." In a typical moral hazard problem, one of the parties (agent) to an agreement has an incentive, after the agreement is made, to act in a manner that brings additional benefits to himself at the expense of the other party (principal). Foreseeing this, principal does not want to make the agreement in the very first place. In the examples above, the seller is the agent and the buyer is the principal. The seller has an incentive to not deliver the product, or deliver a low quality product once he gets the payment. Foreseeing this, the buyer does not engage in the trade activity.

Economists have been trying to understand moral hazard problems and find ways to regulate the agent. Legal punishment systems constitute one way of achieving trust. The principal can go to the court, and claim his rights if the agent cheats. Hence, the agent can commit himself to following the requirements of the agreement because cheating is costly to him. However these punishment systems are in general too costly and in some situations it is difficult to prove that the agent has cheated because of incomplete contracts or imperfect monitoring. Besides, legal sanctions are still not strong enough to prevent fraud in online auctions.

Moral hazard may be mitigated if the agent is going to face with similar situations in the future. In such repeated settings, how or if the activity will be performed in the future might depend on whether the parties are happy about the outcomes of the activities in the past. This kind of punishment through the play of the game in the future will enable the agent to commit not to cheat.

On eBay people can sell and buy objects through auctions. After every transaction the buyer and seller may leave feedbacks about the transaction, and these feedbacks become part of their identity which is publicly available to all users. In most cases, a seller does not encounter with the same buyer again, so a potential buyer cannot use his own past experience to evaluate the seller. However, the feedbacks are publicly observable, so the potential buyer learns past performance of the seller and decides how much to bid accordingly. If the agent cheats he receives a bad feedback, and future buyers will not be willing to trade with him (or will bid low). So the agent's incentive to cheat is weakened or even may be eliminated by the loss of potential future profits that cheating will cause.

However, feedbacks carry only partial information about the action of the seller. Therefore these settings carry various potential inefficiencies. There are various examples of situations, apart from online marketplaces, where actions are not observed perfectly. In such situations, due to imperfection there will be times when the agent will be punished although he didn't cheat. This in turn will cause inefficiencies that can be as severe as the activity not being undertaken at all in the very beginning. Fudenberg and Levine (1994) and Dellarocas (2004) study examples of these situations and analyze possible inefficiencies.

Reputation has been proposed as a way to escape the "curse" of imperfect monitoring. Kreps and Wilson (1982), and Milgrom and Roberts (1982) are the first to discuss the idea that even a small amount of uncertainty about the type of agent might be enough to sustain cooperation and induce trust, at least in the early stages of a long-term relationship. In these papers reputation is modeled as a situation where there is a long run player (agent) facing a sequence of short run opponents (principals). The long run player has private information about his type and discounts the future with a factor $\delta < 1$. Later, Fudenberg and Levine (1992) show that in these games the long run player can achieve payoffs arbitrarily close to what he would get if he could commit himself publicly to playing a particular strategy in the stage game. Their result says that reputation can bring in the commitment power of the agent in repeated games with imperfect monitoring. Although their finding seems to be very powerful, it has only short-term implications: i) The payoff is calculated at the beginning of the game (first period). ii) How the game is going to be played in the distant rounds of the game is not explored.

The long run features of the equilibrium play of these games have been explored only recently by Cripps, Mailath and Samuelson (2004) (CMS hereafter). They show that reputation effects do not last long in repeated games with one sided private information and imperfect monitoring: i) The payoff of the long run player will eventually be close to an equilibrium payoff of the repeated game without any uncertainty on the type of the long run player. ii) The play of the game on the equilibrium path will be eventually very similar to that of the repeated game without any uncertainty on the type of the long run player.

CMS further suggest that one should incorporate some other mechanism into the model in order to prevent reputation effects from disappearing. The purpose of this paper is to propose such a mechanism in a model of repeated principal-agent problems with incomplete information.

The examples in the literature that sustain non-disappearing reputation effects assume that the type of the player is governed by a stochastic process through time, rather than being determined once and for all at the beginning of the game. This is called replacement. Holmstrom (1999), Cole, Dow and English (1995), Mailath and Samuelson (2001), and Phelan (2001) maintain permanent reputations by assuming particular types of replacement in their models.

We believe that replacement is an exogenous feature of a model, and if one wants to *design* a mechanism that enhances efficiency by using reputation effects, replacement cannot be a choice variable. In this paper we take a mechanism design approach to address the question: Is it possible to have reputation effects at work forever in a moral hazard environment where the type of the agent is determined at the beginning of the game?

CMS gives a negative answer to this question *if the short run players observe all of the past signals in the game.* However, in this paper we propose a new mechanism that determines the information each short run player observes about the past play of the game. In other words, we allow for the choice of a structure on the way information about the past signals is transmitted to the short run players. This changes the negative result of CMS. We construct information transmission mechanisms that enable the long run player to build reputation at all times.

In contrast to the literature, we assume that an institution (e.g. a central computer) observes the play of the game each period (that is either all actions, or the realizations of the signals). Short run players do not have any information about the past play of the game other than what the institution provides. In other words, an institution censors the information observed by the short-run players.

Information censoring could be done in many ways. Showing only summary statistics about past performance data, like the average performance or time weighted average of past data, showing only the most recent data, refining the performance data into a binary form and showing the sum of the past n performances are some examples. There are various examples of information censoring in practice. Institutions that make consumer reports collect data about a product, or a firm over time, and these data are "processed" before the customer sees them. Every new piece of information is not reflected in the report in its most transparent form. Another example is "shopping search engines". These engines give scores to online stores based on the information they collect about them either by customer reviews or by shopping from the stores themselves. The scores are updated as new information arrives, and usually the customer does not have access to all pieces of information separately. Online marketplaces such as eBay also provide censored information about the seller and the buyer. Sometimes they show only feedbacks given during the most recent month. In other cases each user is given a score, an integer between 1 and 100, that is the difference between the number of good and bad feedbacks in the last 6 months.

In this paper we focus on one class of censoring methods, which we call "Rating Systems". Rating systems describe past performance with a number from a finite rating set $S = \{0, 1, ..., n\}$ in a particular way. At the initial period the party to be rated is assigned a random number from the rating set. At the end of each period, depending on the observed performance of the party at that period, the rating may decrease or increase by one, or stay the same. If the current rating is already the lowest (highest) possible rating, then at the end of the period the rating either stays the same or increases (decreases) by one. The rule governing the transition from one rating to another after an observed performance level (e.g. signal, feedback, etc.) is called a "transition rule".

In many economic settings of interest regulating institutions are not allowed to make private talk with the players. We define a public institution as one that provides precisely the same information to a period's players about its future behavior. In this paper we will allow for only public rating systems.

Dellarocas (2004), Bakos and Dellarocas (2003) study a repeated model of bilateral exchange environment with a moral hazard problem on the long run player's side. In particular, they analyze a stylized model of repeated auctions. In this model the seller is the long run player and the buyers are short run players. Their model is a repeated moral hazard game with imperfect monitoring without uncertainty on the type of the long run player. Short run players do not observe the outcomes of past play. Instead, an institution called a reputation mechanism observes past outcomes with some imperfection, and may disclose some or all of this information. They use the techniques developed in Fudenberg and Levine (1994) to show that maximum efficiency can be attained by a two state randomization device, and that any attainable payoff vector is bounded away from the Pareto frontier of the stage game. They also show that if the monitoring imperfection is sufficiently large, then the resulting inefficiency may be as severe as the destruction of all trade possibilities.

In our model there is incomplete information about the type of the long-lived player. One of these types is a "normal" type that has the usual payoff structure and has the moral hazard problem. The other types are commitment types that play the same strategy every period, either for morality reasons or because they are boundedly rational, or simply because their payoff structure is different. Reputation is the belief of others about the type of the long-lived player after observing his past behavior. The future play is predicted using this belief and the strategic considerations of the players. This is the approach taken by Kreps, Milgrom, Roberts and Wilson.

Information disclosure by institutions, in particular rating systems, serves a dual role in our model. The first one is learning the type of the long run player and transmitting this information to the short run players. This will enable the short run players know the true type of principal they are facing against with a high probability. Unless the principal fully mimics a commitment type the institution learns his true type after observing sufficient number of signals. However if this information is fully disclosed to the agents, and if it turns out that long run player is not a commitment type with high probability, then the play is almost like the one of perfect information and inefficiency is inevitable. For this reason, a rating system should be "forgetting" some of the past data, and allow normal type of the long run player to build a reputation even in the distant future. This constitutes the second role of rating systems, committing to discard some information in order to maintain the agent's incentives for not cheating at all times.

In theorem 0, we show that if there is no commitment type, the equilibrium payoff set of the game with any public institution is bounded away from the Pareto frontier of the stage game. This is the analog of the inefficiency results in repeated moral hazard games with imperfect monitoring

(Fudenberg and Levine (1994)). We use this theorem as a benchmark for our results on how addition of a small uncertainty on the type of the long run player can facilitate indefinite efficient play.

In theorem 1, we construct a particular rating system. We show that under a mild assumption on the commitment types, our rating system allows the long run player to get almost his Stackelberg payoff after every history in an equilibrium of the game. This is our main technical contribution to the literature, to provide a mechanism by which reputation brings in the commitment power of the agent even in distant future.

Theorems 2 and 3 are about the payoffs of the short run players. This is a crucial point ignored in previous studies on reputation. First, we make an assumption that the more the effort level the long run player commits to, the better off the short run players are. This is easily satisfied if the underlying mechanism is a Dutch or English auction. Theorem 2 says that we are able to regulate the agent to exert any effort level smaller than the most hard working commitment type in the type space, if commitment types arise with small enough probabilities. If commitment types have non-negligible probabilities, then in theorem 3 we regulate the agent to exert effort levels that are less than the most hard working commitment type, unless he prefers to mimic a commitment type that exerts less effort than these effort levels. In theorem 3, we also show that distant short run players know almost surely the type of the long run player. Hence, these players enjoy informational rents from the experimentation held during the early rounds of the game.

In theorem 2 we also show that for each point on the Pareto frontier of the underlying game, there exists a rating system that implements that point in the long run. This enables us to interpret rating systems as emerging from a competitive environment.

Next we discuss two applications that fit our model, and then introduce our model with and without incomplete information. Later we present our main result on permanent reputations and welfare of buyers after which we conclude.

Applications

In this section we will discuss two different models of moral hazard that have been studied in the literature. Our framework will be general enough to include these models. The first one is online auctions, such as eBay. A long run monopolist sells an item to a group of buyers that live only one period. This model is very similar to the model that Dellarocas (2004) analyzes.

The second application is a quality game where a monopolist is selling an object to a buyer. The buyer may pay a fixed price p and get the object or may choose not to buy the object. Signals that are informative about seller's actions occur only after the buyer buys the object.

The eBay Model

A long run monopolist is selling a product to n short run buyers. The valuations of each buyer for the marginal quality of the product, v_i are drawn independently from a continuous distribution F whose support is $[0, 1]$. Each buyer makes a bid from a finite set, and the product is given to the highest bidder. The winner pays the second highest bid to the seller.

The quality of the product is either high or low. A buyer with a marginal valuation v has a utility v if the product's quality is high and he gets the product, 0 if the product's quality is low.

The seller may choose to exert effort that costs him $c > 0$, or may choose not to exert any effort. The quality of the product depends on the effort choice of the seller. Product's quality is high with probability q_H if the seller exerts effort, and it is $q_L < q_H$ if he does not exert any effort.

We allow each buyer who wins the object to leave a feedback from a finite set $R = \{\oslash, 1, 2, ..., r\}$ where $\oslash$ represents not leaving any feedback. The feedback decisions of buyers are exogenous and depend on the quality of the product which is privately observed by the buyer who gets the product. More precisely, there are two probability distributions over R, for each quality level.

The timing of the actions at any period is as below:

- Buyers observe feedbacks from previous transactions.
- Buyers place their bids and seller chooses effort level simultaneously.
- The buyer who wins the object observes the quality level of the product and leaves a feedback.

Not Exerting effort is a dominant strategy for the seller in a one stage game. Therefore, in the unique Nash equilibrium of the one period game, the seller does not exert effort, and buyers bid low.

Let $p(\alpha)$ be the equilibrium expected payment made to the seller if he could commit to exerting effort with probability α. If $p(1) - p(0) > c$, he would like to commit to exerting effort, rather than playing the Nash equilibrium.

In the infinite horizon repeated game, for some parameter values, there is no trade at all even when the seller is very patient. For most parameter values his payoffs will be far away from the payoff he would get by committing to a particular strategy.

Quality Game

A monopolist is selling a product. First, the buyer decides whether to buy the product at a fixed price p. If he does not buy the product no trade occurs and both players get utility 0. If the buyer decides to buy the product the seller may either exert effort or not. If he exerts effort, the quality of the product is high with probability $q_H < 1$, the expected utility of the buyer is $vq_H > p$, and the utility of the seller is $p - c > 0$. However, if he does not exert effort, the quality of the product is high with probability q_L, $0 < q_L < q_H$, the expected utility of the buyer is $vq_L < p$, and the utility of the seller is $p > 0$. The quality of the product is publicly observed only if the buyer pays the price. The effort decision of the seller is not observable.

The Stackelberg strategy of the seller is to exert effort with a probability just enough to induce the buyer to pay the price, that is to exert effort with a probability α such that $v(\alpha q_H + (1-\alpha)q_L) = p$. But in the unique subgame perfect Nash equilibrium of the game the buyer does not purchase the product, and the seller exerts no effort if the buyer buys.

If the game is repeated with the same monopolist every period and a different buyer each period, the seller's payoff is always strictly less than his Stackelberg payoff for any discount factor. For some parameter values, there is no trade at all even if the seller is very patient.

MODEL

We study an infinitely repeated moral hazard game between one long run player (player 1) and an infinite sequence of short run players. Each short run player lives for one period and serves as player 2 in the following stage game.

The Stage Game

Let $A_1 = \{L, H\}$ be the set of actions available to player 1in the stage game. Let $A_2 = \{0, 1, 2, ..., m\}$ be the set of actions for the short run player in each stage game. An action $a_1 \in A_1$ describes player 1's level of effort. Here the long run player can either exert high effort H or low effort L. The action $a_2 \in A_2$ specifies how much a short run player pays (or bids) in the stage game. The action 0 corresponds to exit (or non-participation). Actions are taken simultaneously.

Let $A = A_1 \times A_2$ be the set of all action profiles. For any finite set X, let $\Delta(X)$ denote the set of all probability distributions over X. In particular, $\Delta(A_k)$ is the set of all mixed strategies in the stage game for k. Let s_k denote a generic element of $\Delta(A_k)$. We will refer to $s_1(H)$ as player 1's effort level. Without risk of confusion we write $a_k \in A_k$ to denote the mixed strategy s_k such that $s_k(a_k) = 1$.

Player k receives payoff $U_k(s)$ when the stage game strategy profile is $s = (s_1, s_2)$, where $U_k(s) = \sum_{(a_1,a_2)\in A} u_k(a_1, a_2)s_1(a_1)s_2(a_2)$ and $u_k : A \to \mathbb{R}$. Let B_2 denote the best response correspondence of player 2. That is,

$B_2(s_1) = \{s_2 \in \Delta(A_2)|U_2(s_1, s_2) \geq U_2(s_1, s_2') \text{ for all } s_2' \in \Delta(A_2)\}$.

Let $V_1(s_1)$ denote the best commitment payoff for player 1 given s_1. That is;

$$V_1(s_1) = \max_{s_2 \in B_2(s_1)} U_1(s_1, s_2)$$

Similarly, we define follower's payoff as $V_2(s_1) = \max_{s_2\in\Delta(A_2)} U_2(s_1, s_2)$. Let V_1^s denote the Stackelberg payoff for player 1: $V_1^s = \max_{s_1\in\Delta(A_1)} V_1(s_1)$.

The following conditions on the payoffs characterize a moral hazard game:

Condition 1 *(constant effort cost)* $u_1(L, a_2) - u_1(H, a_2) = c > 0$ *for* $a_2 \in \{1, 2, ..., m\}$ *and* $u_1(L, 0) > u_1(H, 0)$.

Condition 1 says that the exerting high effort is costly and the cost is constant across all non-zero actions of player 2. This condition ensures that exerting low effort is a dominant strategy for player 1.

Condition 2 *(more money is better) If* $a_2, a_2' \in \{1, 2, ..., m\}$*, and* $a_2 > a_2'$*,* $u_1(a_1, a_2) - u_1(a_1, a_2') > 0$ *for each* $a_1 \in A_1$.

Condition 2 says that for any fixed action of player 1, he prefers that player 2 plays a higher action. In our applications, a_2 is interpreted as the amount of money player 2 bids, so this condition says that player 1 prefers that buyers bid (accordingly pay) more.

Condition 3 *(increasing payments with increasing effort)* $s_1(H) > s_1'(H)$, $a_2 \in B_2(s_1)$, $a_2' \in B_2(s_2')$ *implies* $a_2 \geq a_2'$. $B_2(L) = \{0\}$, $B_2(H) = \{m\}$.

This condition says that best response correspondence of player 2 is weakly increasing with the effort level of player 1. Also the low effort level induces player 2 not to participate and high effort level induces him to bid the highest amount possible. In our applications, this means buyers bid more if they anticipate that the seller will exert more effort. This condition is satisfied if the product is sold using a Dutch or English auction.

Condition 4 *(Stackelberg strategy is not part of a Nash equilibrium):* $V_1^s > u_1(L, 0)$.

Conditions 1-3 ensure that in the unique Nash Equilibrium of a moral hazard game, player 1 exerts low effort and player 2 does not participate (or bids 0). Condition 4 says that player 1's Stackelberg payoff is bigger than his Nash equilibrium payoff (otherwise he does not want to commit, so the problem is not interesting).

Condition 5 *(Stackelberg strategy is unique)* $\Psi^s = \{s_1 | V_1(s_1) = V_1^s$ *for* $s_1 \in \Delta(A_1)\}$ *is a singleton.*

This condition says that the Stackelberg strategy is unique. We use α^s to refer to the effort level in the Stackelberg strategy.

Signal Structure

Short run players do not observe the action of player 1, but only observe a public signal that is correlated with the action. The public signal is denoted y, and is drawn from a finite set of realizations, Y. Let $\rho_a \in \Delta(Y)$ denote the probability distribution of the signal given the action profile a. Hence the probability of observing signal y given a is $\rho_a(y)$. We impose the following condition on the signal structure.

Assumption 1 *(Identification) If* $a = (a_1, a_2)$, $a' = (a'_1, a_2)$, $a_1 \neq a'_1$, *and* $a_2 > 0$, *then* $\rho_a \neq \rho_{a'}$.

This assumption ensures that if every short run player chooses a non-zero action and if the long run player chooses the same action in each stage game, then an outside observer who observes player 2's action would eventually be able to learn player 1's action. We allow the signals to be uninformative when player 2 plays action 0.

Repeated Game with Institution

The stage game described above is infinitely repeated. Player 1's discount factor is $\delta < 1$. Each short run player (player 2) can only observe the outcome in his own period.

An institution observes the action of player 2, and the signal y every period. The institution delivers a message $r \in R$ to the current period players before they play the stage game. We depart from the literature by assuming that a short run player cannot observe past play of the game. However, before the stage game is played, he learns the public announcement made by the institution (r) that may provide some information about the past play. Player 1 knows his own actions, public announcements, and the actions of the short run players.

Player 1, before playing period t action has a private history, consisting of the messages, actions, and signals. Let $H_1 = R \times A \times Y$, then a private history for player 1 is $h_{1t} \in H_{1t} \equiv R \times H_1^{t-1}$. Player 2 at period t has a private history $h_{2t} \in H_2 \equiv R$. The timing of the flow of information and actions at period t is summarized below:

1- Institution makes a public announcement $r \in R$.

2- Player 1 and player 2 choose their actions simultaneously.

3- A signal $y \in Y$ is realized.

Signal y can be interpreted in various ways. It can be a payoff relevant variable (e.g. quality of a product, how satisfactory the service was) or a payoff irrelevant variable (e.g. feedbacks from player 2, outcomes of auditing reports). In our model, it is crucial that the signal is not observed by the population. Therefore we focus on applications in which this assumption is more easily satisfied, such as online marketplaces or consumer reports.

In this paper we view institutions as mechanisms that transmit information to a period's players about past play outcomes. An institution maps histories that consist of past actions of player 2, signals, and its own past messages into a probability distribution over a set of messages. Among several such mechanisms, we will look for an information transmission mechanism that facilitates efficient trade. For our results it will be sufficient to focus on those that have a Markovian structure. In particular, we will use institutions whose message at period t (r_t) evolves according to a Markov process called a Markov transition rule. In general, the transition probabilities will depend on the action of the short run player and the observed signal, however, for our results it will suffice to consider transition probabilities that only depend on the observed signal. Below we give the formal definition of a Markov transition rule.

Definition 1 *Let R be a countable set. A map $\tau : R \times A_2 \times Y \to \Delta(R)$ is called a Markov transition rule.*

We are now ready to give a formal definition of a special class of institutions called rating systems. A rating system consists of a set of messages called state space, an initial distribution on the state space, and a Markov transition rule on the state space. At time 0, the state is determined by the initial distribution, and every period the state is publicly announced before actions are taken. After the actions are taken, the new state is determined by the previous state and the outcome of the play.

Definition 2 *A triplet $\phi = (R, \tau, p_0)$ is called a rating system where R is a countable set, τ is a Markov transition rule, and $p_0 \in \Delta(R)$.*

At time 0, a rating system announces a message r_0 drawn according to an initial distribution p_0. At the end of each period t, the random variable $r_{t+1} \in R$ is realized by the transition rule τ, the signal at period t, and $r_t \in R$. At the beginning of period $t+1$, r_{t+1} is announced to period $t+1$ players.

The Incomplete Information Game

A type space $W = \{0, 1, ..., l\}$ is a finite set. Prior to time $t = 0$, nature chooses a type for player 1 according to a probability distribution $\eta \in \Delta(W)$. We will use η_j for $\eta(j)$, and assume without loss of generality that $\eta_j > 0$ for $j = 0, 1, ..., l$. The normal type of player 1 has the payoff structure as described in the stage game, and type 0 represents the normal type. Each type $j > 0$ represents a commitment type that plays H with a probability α^j every period. Let the index set be ordered such that $\alpha^j > \alpha^{j-1}$ for $j = 2, 3...l$. We will call the strategies of the commitment types s^j, that is $s^j \in \Delta(A_1)$ where $s^j(H) = \alpha^j$. We call $\Gamma = (W, \eta)$ a type model, and assume that no commitment type forces the short run player to play 0, the uninformative action. More precisely, $B_2(s^1) \neq \{0\}$. Also, when we use the term player 1 without specifying a type, we mean type 0 of player 1.

Let $H = \bigcup_{t=0}^{\infty} H_{1t}$ be the set of all finite histories of player 1. Player 1's strategy is a map $\sigma_1 : H \to \Delta(A_1)$. Player 2's strategy is a collection of maps $\sigma_2 = \{\sigma_{2t}\}_{t=0,1,...}$ where $\sigma_{2t} : R \to \Delta(A_2)$. The strategy spaces of players 1 and 2 are denoted by $\sum_1$ and $\sum_2$ respectively.

For a given rating system ϕ, each strategy profile (σ_1, σ_2) induces a probability distribution P over all action profiles, signals and messages. The payoff to player 1 and period-t player 2 of a strategy profile (σ_1, σ_2) are:

$$U_1(\sigma_1, \sigma_2) = E^P(\sum_{t=0}^{\infty} \delta^t u_1(a_1, a_2) | \text{player 1's type is 0})$$

$$U_{2t}(\sigma_1, \sigma_2|r_t) = E^P(u_2(a_1, a_2)|r_t)$$

We also define the payoff of player 1 after a history $h_{1t} \in H$ as:

$$U_1(\sigma_1, \sigma_2|h_{1t}) = E^P(\sum_{t'=t}^{\infty} \delta^{t'-t} u_1(a_1, a_2)|h_{1t}, \text{player 1's type is 0})$$

We will call a game $G(\delta, W, \eta, \phi)$, that has the payoffs and strategy spaces as described above, a repeated incomplete information game with institution ϕ. We will use $G(\delta, \Gamma, \phi)$ interchangeably with $G(\delta, W, \eta, \phi)$, where $\Gamma = (W, \eta)$.

Having described the payoffs and the strategy spaces, we can now define the equilibrium concept we will be using. For a history $h \in H_1$ of player 1, we denote the latest message in h by $r(h)$.

Definition 3 *Strategies (σ_1, σ_2) are Perfect Bayesian equilibrium of $G(\delta, \Gamma, \phi)$ if*
i) $U_1(\sigma_1, \sigma_2|h_{1t}) \geq U_1(\sigma_1', \sigma_2|h_{1t}) \; \forall \; \sigma_1' \in \sum_1, \forall h_{1t} \in H$.
ii) $U_{2t}(\sigma_1, \sigma_2|r_t) \geq U_{2t}(\sigma_1, \sigma_2'|r_t) \; \forall \sigma_2' \in \sum_2$, $t = 0, 1, ..., \forall r_t \in R$ with $P(r_t) > 0$.
iii) Let $h_{r_t} = \{h_{1t}|r(h_{1t}) = r_t \text{ for } h_{1t} \in H_{1t}\}$. If $P(r_t) = 0$, $\sigma_{2t}(r_t)$ is a best response to a belief κ where $\kappa \in \Delta(h_{r_t})$, and player 1's type is 0.

This is a game of incomplete information therefore the proper solution concept is Perfect Bayesian equilibrium. Players are required to update their beliefs using Bayesian rule whenever possible. Player 1 knows his type, and there is no incomplete information on his side, thus he does not perform any Bayesian updating. However, short run players learn some information about the type of player 1 through the messages delivered by the institution. They use the message to form their beliefs about player 1's type. But Bayesian updating is not possible after histories with probability zero. In our game, the institution is not a player, so any deviation from the equilibrium is attributed to players' past behavior. We require the short run players to play a best response to some arbitrary belief on player 1's histories that are consistent with the message they observed, and a belief that player 1's type is the normal type.

If the support of the signal distribution ρ_a is constant across all action profiles $a \in A$, then the set of all histories that can occur with positive probability given a strategy profile is constant across all strategy profiles. So the set of all Perfect Bayesian equilibria is the same as the set of all Nash equilibria.

If the support of the signal distribution ρ_a is not constant across all action profiles $a \in A$, then the set of all histories that can occur with positive probability will depend on the strategy profile. The equilibrium restriction (iii) might seem weak because it gives too much freedom in how we can choose beliefs of the short run players. Therefore the equilibrium set might be larger than a more strict equilibrium concept such as sequential equilibrium. However, we will design our rating systems in such a way that all histories for player 2 occur with positive probability given any strategy profile. So, even when constant support assumption is not satisfied, we will not make use of out of equilibrium beliefs to support a particular equilibrium.

Complete Information Case

In this section we analyze the complete information case of our model. The purpose of this exercise is to understand the effects of information censoring on the equilibrium payoff set when there is no room for reputation formation. Our result will be negative, in a sense that information censoring can only decrease the equilibrium payoff set compared to a model in which information is not censored at all.

Most of the literature on repeated games assumed that short run players could observe all of the past signals. However, we drop this assumption and instead assume that there is an institution that observes signals every period and conveys some information to the short run players every period. In fact these models are a special case of our model, where the institution delivers all information about the past play to every short run player. These institutions are called transparent institutions and are denoted ϕ^*. The formal definition of transparent institutions (that is, message space, transition rule, and initial state) is given in the appendix.

The messages of a transparent institution include all information about past play of the game. We allow for public randomizations and each short run player observes the outcomes of these randomizations as well. When institutions are transparent, the informational assumptions of our model coincide with those in the standard reputation literature where all past signals and actions of player 2 are observable by the current period players.

If there is no incomplete information about player 1's type, Perfect Bayesian equilibrium puts the same restrictions as Perfect equilibrium.

Theorem 0 *Let ϕ be a Markov institution, and $W = \{w_0\}$. The payoff to the long run player in any Perfect Bayesian equilibrium of $G(\delta, W, \eta, \phi)$ is no more than the highest payoff he can get in some Perfect Bayesian equilibrium of $G(\delta, W, \eta, \phi^*)$ for some transparent institution ϕ^*.*

This result says that Markov institutions can do no better than transparent institutions if there is no incomplete information about the type of player 1. The power of this theorem is most emphasized when used with theorem 6.1 of Fudenberg, Levine (1994), which says that the payoff to the long run player is generically bounded away from his most preferred commitment payoff. We use this result as a benchmark for discussing reputation effects in our model.

Although theorem 0 focuses on Markov institutions, the result is still valid if we allowed for more complicated information transmission mechanisms, in particular all public institutions. A public institution is one that provides precisely the same information about its future behavior to a period's players. Every Markovian institution is public, because the future information transmission rule is common knowledge among player 1 and period-t player 2 at period t. If the transition rule depends on some information that is not in the history of player 2 but is in the history of player 1, then the corresponding institution fails to be a public institution.

Permanent Reputations with Rating Systems

Having discussed the inability of rating systems to increase efficiency under complete information, we will now assume that the type space W has commitment types. Each of these types plays a particular strategy every period, independent of the history. When there is incomplete information about the type of player 1, he will be able to build a reputation by imitating a certain type, or trying not to look like a particular type. Our result is that rating systems are able to bring in commitment power to the long run player forever. First a definition of a subset of type space is given.

Definition 4 *A type space W is called a Stackelberg space if $\alpha^l > \alpha^s$ and $\alpha^1 \leq \alpha^s$ implies $u_1(L, a_2) < V_1^s$ for each $a_2 \in B_2(s^1)$.*

Stackelberg spaces have a commitment type that exerts more effort than the Stackelberg strategy. Moreover if there is a lazy type that exerts less effort than the Stackelberg strategy, then Stackelberg payoff is more than the payoff player 1 can get if player 2 played a best response to the laziest type. Our main result is below.

Theorem 1 *For any Stackelberg space W, η, and $u < V_1^s$, there exists $\overline{\delta} < 1$ such that for $\delta \geq \overline{\delta}$ there is a rating system ϕ, and a Perfect Bayesian equilibrium of $G(\delta, W, \eta, \phi)$ where the payoff to the normal type of the long run player is at least u after every history.*

The result says that with a suitable choice of a rating system, there is a Perfect Bayesian equilibrium of the game in which the long run player can get almost his most preferred commitment payoff after every history. This does not mean that his stage game payoffs are always almost his commitment payoff after every history. On the equilibrium path, the long run player may get period payoffs that are less than his Stackelberg payoff. However these periods do not follow each other frequently enough, so his discounted payoff calculated at the beginning of every history becomes almost his Stackelberg payoff.

Although Theorem 1 is stated at the payoff level it has corresponding behavioral implications as well. On the equilibrium path, the frequency with which the long run player exerts his most preferred commitment effort level is almost 1 after every history. So reputation never ceases to give the long run player commitment power when we intervene information transmition. This turns over the results of CMS.

Our restriction on Stackelberg spaces stems from 2 facts. The first one is that we need a commitment type that normal type of player 1 can imitate and build a reputation upon. This is standard in the literature. The second one is a restriction our rating systems require. The lowest states are where the probability of the laziest type is the highest. In the lowest state the payoff to player 1 is at least the payoff he can get by not exerting any effort and enjoying the payments by player 2. We need this payoff to be less than the Stackelberg payoff to make these states less desirable to player 1.

Previous studies approximated the payoff of the long run player by a payoff he could get by mimicking a particular commitment type in the support of the prior distribution of commitment types. However in our model, we only require the existence of a commitment type that exerts more effort than the Stackelberg strategy.

The intuition for Theorem 1 is as follows. The long run player is regulated to play a particular action through offering him continuation payoffs depending on the signal realization. The incentives that should be provided are at the order of $1-\delta$, so with a finite number of states, long run player's payoffs at every state get close to each other as $\delta \to 1$.

However, so far we haven't mentioned reputation, and all the lines up to here could be said in a repeated setting with complete information. But we know that efficiency result fails in an imperfect monitoring world with complete information (theorem 0). The missing part is that, when there is complete information, if at a state where the long run player's continuation payoff is the highest, he exerts high effort with some probability, his payoff at this state will be strictly below $u_1(H, a_2)$ for any positive a_2. But $u_1(H, a_2)$ is strictly below Stackelberg payoff by definition.

We get around this by using the institutions for separating the types. The commitment types and the normal type behave differently at most of the states. Thus, the process by which the state moves depends on the type. Rating systems enable the concentration of each commitment type at a different subset of the state space. Thus, even though the normal type exerts low effort when at states where a commitment type plays H more than his Stackelberg strategy, short run players continue to play their best response to the commitment type. Because observing those states, the type of the long run player is most likely to be the commitment types. A sketch of the proof is illustrated in the following example.

Example and sketch of proof

There are 2 actions available for each player, $A_1 = \{H, L\}$ and $A_2 = \{B, N\}$. The signal space is $Y = \{g, n\}$ and y represents the realization of the signal, where the probability distribution of y conditional on player 1's actions is given by:

$$pr(y = g|a_1 = H) = 2/3 = \ pr(y = n|a_1 = L)$$

Player 1 is the row player, player 2 is the column player in the stage game with the payoff matrix given below:

			s=1	1<s=103	103<s<143	s=143
b	0	signal y=g	(0,0.125)	(0.375,0.5)	(0.375,0.5)	(0.125,0)
		signal y=n	(0,0.125)	(0.375,0.25)	(0.375,0.25)	(0.125,0)
	1	signal y=g	-	-	(0.25,0.25)	(0.125,0)
		signal y=n	-	-	(0.25,0.25)	(0.125,0)

Figure 1: The numbers in each cell correspond to $(\pi(s,b,y)(s^-), \pi(s,b,y)(s^+))$. In some cells there are no numbers to emphasize that these states are never visited. For each δ we choose a different π whose values differ from those of this table by an amount at the order of $1-\delta$.

	B	N
H	$1,2$	$-1,0$
L	$2,-2$	$0,0$

In this game, the Stackelberg strategy of player 1 is to play H with a probability 0.5, and his Stackelberg payoff is 1.5. The type space is $W = \{0,1\}$ where type 0 is the normal type of player 1 and has the payoff structure above. Type 1 always plays action H at every period of the game. The prior probability distribution of the type of player 1 is $\eta \in \Delta(W)$, where $\eta(0) = 2/3$, and $\eta(1) = 1/3$.

Let $\alpha^* = 0.5$, $W^* = \{0.5, 1\}$ where 0.5 (1) is the probability with which the normal type (type 1) plays H mostly (always). We define $W_2^* = \{(0.5,1)\}$ as the set of transition regions in our state space, in this example there is only one transition region. Let $S = \{1, 2, ..., 143\}$ be the rating set and $B = \{0,1\}$ be a binary set whose value is determined by a public randomization at a rating. The state space is $\Omega = S \times B$, and we divide the state space into regions in the following way: $\Omega_I = \{1\} \times \{0,1\}$, $\Omega_{0.5} = \{2, ..., 101\} \times \{0,1\}$, $\Omega_{(0.5,1)} = \{102, 103, 104\} \times \{0,1\}$, $\Omega_1 = \{105, ..., 142\} \times \{0,1\}$ and $\Omega_L = \{143\} \times \{0,1\}$. The sensitivity of rating changes to observed signals is going to be different in each region. Therefore, in equilibrium normal type of player 1 spends most time in the region $\Omega_{0.5}$, and the commitment type spends most time in the region Ω_1. Transition region is where the reputation of player 1 for being a commitment type increases gradually from some value below 0.5 to some value above 0.5.

Let the transition rule be a map $\tau : \Omega \times A_2 \times Y \to \Delta(\Omega)$, indicating the rule by which the rating system determines the announcement at time $t+1$. We will construct τ by two other maps, π and v. The map $\pi : \Omega \times Y \to \Delta(\{s^+, s^0, s^-\})$ will determine the probability of a rating increase, decrease or stay the same at any state. For instance, $\pi(132, 1, g)(s^+)$ reads as the probability of a rating increase to 133 if the current state is $(132, 1) \in \Omega$ and the signal y is g. The map $v : S \to [0,1]$ indicates the probability that the binary variable is 0 at any rating. We set $v(s) = 1$ for $s \leq 104$ and for $s = 143$ and $v(s) = 1/2$ for all other ratings. We summarize the values of π in Figure 1.

We define the transition rule as follows. Suppose at time t the state is $w = (s, b)$ and the signal is y. Then first the rating at period $t+1$ is determined by π. When the outcome of π is s^+, s^0 or

	s=1	1<s=103	103<s<143	s=143
b=0	(0,N)	(0.5,B)	(0.5,B)	(0,B)
b=1	-	-	(0,B)	-
v(s)	1	1	0.5	1

Figure 2: The entries in the first two rows indicate the probability of high effort for player 1, and the pure strategy action of player 2 at any state. Third row is the probability that the outcome of the public randomization is 0 at any rating.

s^-, then the rating at period $t+1$ (s') is $s+1$, s or $s-1$ respectively. Once the rating at $t+1$ is determined, the public randomization v decides the value of the binary variable b' depending only on the period $t+1$ rating s', and the new state (s', b') is announced at the beginning of period $t+1$.

Consider the strategies $\sigma_1 : \Omega \to [0,1]$ indicating the probability of H for player 1 and $\sigma_2 : \Omega \to [0,1]$ indicating the probability of B for player 2:
$\sigma_1(s,1) = 0$ for each $s \in S$, $\sigma_1(1,0) = \sigma_1(143,0) = 0$, $\sigma_1(s,0) = 0.5$ for $s \in \{2,...,142\}$. $\sigma_2(s,1) = 0$ for $s \leq 103$, $\sigma_2(s,1) = 1$ for $s > 103$, $\sigma_2(s,0) = 1$ for $s > 1$. We summarize the strategies and the values of the map v in Figure 2 .

We will first show that the strategy σ_1 is optimal. The strategy profile (σ_1, σ_2) is Markovian, so the probability distribution of the stream of future payoffs depends on the current state. Let $V(s,b)$ denote the expected present discounted value of the strategy profile to player 1 when the state is $(s,b) \in \Omega$. We also define $V(s) = v(s)V(s,0) + (1-v(s))V(s,1)$. These values are determined by the following recursive equations:

$$V(s,b) = \left[(1-\delta)u_1(\sigma_1(s,b), \sigma_2(s,b)) + \delta E(V(s')|(s,b))\right] \ where$$

$$\begin{aligned} E(V(s')|s,b) \quad = \quad & \\ & [\sigma_1(s,b)(2/3) + (1-\sigma_1(s,b))(1/3)] \left[\sum_{w' \in \Omega} \tau(s,b,a_2,g)(w')V(w')\right] + \\ & [\sigma_1(s,b)(1/3) + (1-\sigma_1(s,b))(2/3)] \left[\sum_{w' \in \Omega} \tau(s,b,a_2,n)(w')V(w')\right] \end{aligned}$$

We use the method of guess and verify to find $V(s)$. Our guess is $V(s) = 1.5 - \frac{12(1-\delta)}{\delta}(143-s)$ and this can be verified by putting these values in the above system of equations. Once we know the continuation values, it is straightforward to check that σ_1 is indeed optimal.

The second task is to show that σ_2 is optimal. Actually we will show that σ_2 is optimal in the long run. The strategy profile (σ_1, σ_2) induces a Markov transition matrix over ratings for each type of player 1. Let P^0 and P^1 be 143×143 transition matrices where

$$
\begin{aligned}
P^0[i,j] &= \sum_{b\in B}[v(i)(1-b)+(1-v(i))b]\times \\
&\quad \{[\sigma_1(s,b)(2/3)+(1-\sigma_1(s,b))(1/3)]\,(\sum_{w'\in\{j\}\times B}\tau(i,b,a_2,g)(w'))+ \\
&\quad [\sigma_1(s,b)(1/3)+(1-\sigma_1(s,b))(2/3)]\,(\sum_{w'\in\{j\}\times B}\tau(i,b,a_2,n)(w'))\}
\end{aligned}
$$

$P^1[i,j]$ is as above where $\sigma_1(s,b)$ is replaced by 1. These matrices are positive recurrent and irreducible, so a unique stationary distribution π^0 and π^1 exists for P^0 and P^1. In the long run, player 2 calculates the probability $p(s,b)$ that player 1 plays action H at state $w=(s,b)$ as below:

$$
p(s,b)=\frac{\pi^1(s)\eta(1)}{\pi^0(s)\eta(0)+\pi^1(s)\eta(1)}1+\frac{\pi^0(s)\eta(0)}{\pi^0(s)\eta(0)+\pi^1(s)\eta(1)}\sigma_1(s,b)
$$

where $pr(k,s)=\frac{\pi^k(s)\eta(k)}{\pi^0(s)\eta(0)+\pi^1(s)\eta(1)}$ is the probability a short run player observing a rating s assigns to player 1 being a type k. The reputation of player 1 for being a commitment type is $pr(1,s)$, and is increasing in the rating s. In our example, $pr(1,s)>(<)0.5$ for $s>(<)103$. It is now straightforward to check that σ_2 is optimal for player 2 in the long run. The strategy of player 2 is a pure strategy, and is a strict best response. So when the distribution of types across ratings is close to the stationary distributions, the strategy σ_2 is still a best response. Therefore, after a finite time T, σ_2 is optimal. We will define the transition rule in the initial periods in a way that player 1's strategy is still a best response at all times, and player 2 plays different than σ_2 for some finite time T, but plays σ_2 after time T. We explain a precise definition of the transition rule and equilibrium strategies in the appendix.

The choice of the number of states is determined as follows. $P^1(i,i+1)=P^1(i+1,i)$ for states $(i,b)\in\Omega_1$, and $P^1(i,i+1)>P^1(i+1,i)$ for states $(i,b)\in\Omega_{0.5}$. Therefore, as the number of states in these two regions increase, $\sum_{(i,b)\in\Omega_1}\pi^1(i)$ approaches 1. Similarly, $P^0(i,i+1)<P^1(i+1,i)$ for states $(i,b)\in\Omega_1$, and $P^1(i,i+1)=P^1(i+1,i)$ for states $(i,b)\in\Omega_{0.5}$. Therefore, $\sum_{(i,b)\in\Omega_{0.5}}\pi^0(i)$ approaches 1 as the number of states in both regions increase. Then we choose the number of states big enough, and divide the state space into regions such that $pr(1,s)>0.5$ for $(s,b)\in\Omega_1$ and $pr(1,s)<0.5$ for $(s,b)\in\Omega_{0.5}$.

The Fate of Short run Players

So far studies concentrated on when and how reputation brings in the commitment power of the long run player. Interpreting their results, CMS say that in some circumstances we care about the fate of the short run players, even those in the distant future. We agree with them, and further argue that we might want to regulate the long run player to exert more effort than his most preferred commitment effort level. In our online shopping example the rater might be interested in regulating the store to be trustable not only as much as is enough to induce the customers to buy, but more. The interests of a consumer report might be more aligned with those of the customers than those of firms. The objective of the designer of an online auction site might be to minimize the number of cheatings while still enabling trade.

The regulating power of rating systems depend on the type space. In general, the higher α^l is, the larger is the set of implementable payoffs. We first present our results when the probability of commitment types is very small.

Assumption 2 *For $s_1, s_1' \in \Delta(A_1)$, $s_1(H) > s_1'(H)$ implies $V_2(s_1) \geq V_2(s_1')$.*

This assumption says that short run players are better off if long run player commits to exerting higher levels of effort.

Definition 5 *For a type space W, the set $CP(W) = \{(V_1(s_1), V_2(s_1)) : s_1(H) < \alpha^l\}$ is called the commitment payoff set of W. The set $IRP(W)$ is the convex hull of $CP(W)$, and is a subset of the set of individually rational payoffs of the stage game.*

$IRP(W)$ is the set of payoff vectors obtained when player 1 plays H with a probability $\alpha < \alpha^l$, and player 2 plays a best response to α.

Theorem 2 *For every $\epsilon > 0$, $(U, V) \in IRP(W)$, there exists $\overline{\eta}_0 < 1$, a natural number T, $\overline{\delta} < 1$, such that for $\delta \geq \overline{\delta}$, $1 > \eta_0 > \overline{\eta}_0$ there is a rating system ϕ, and a Perfect Bayesian equilibrium of $G(\delta, W, \eta, \phi)$ satisfying the following: i) the payoff to the normal type of the long run player is at least $U - \epsilon$ after every history. ii) unconditional expected payoff of every short run player after period T is at least $V - \epsilon$.*

The very idea behind reputation is that a small amount of uncertainty on the type of the long run player can give him his commitment power. Our theorem says that if the uncertainty is indeed small enough, then we can regulate the long run player to exert any effort level α that is less than the effort level of the most hard working commitment type, α^l. Moreover, short run players distant enough in the future get an expected payoff of what they can get at best when the long run player's effort level is α. The short run players in the early rounds of the play do not get as much payoff as the ones in the distant future, because they are informationally inferior to them. The type of the long run player is almost revealed through the signals during these rounds, and in the distant future the effort level becomes close to α.

Although the large increase in the attainable payoff set when the complete information repeated game is perturbed slightly is important, there might be situations where the commitment types are likely to exist. In these environments, we would like to separate the types, and enable the short run player play his best response to each of them.

Definition 6 *Effort level $\alpha \in [0,1)$ is called enforcable (with respect to W) if the following conditions hold: i) $\alpha < \alpha^l$, ii) if $s_1(H) = \alpha$, $V_1(s_1) \geq V_1(s^k)$ for each k such that $\alpha^k < \alpha$ iii) $\alpha^1 < \alpha$ implies $u_1(a_2, L) < V_1(s_1)$ for each $a_2 \in B_2(s^1)$.*

Enforcable effort levels are those that are lower than the effort level of the most hard working commitment type and are not less preferred by the long run player to any smaller effort level a commitment type exerts. condition iii) is a restriction on the laziest commitment type. Player 1's commitment payoff from an enforcable effort level is more than the payoff he gets when player 2 believes he is the laziest type.

Definition 7 *For a type model Γ, let the set $P(\Gamma) = \{(V_1(s_1), v_2)|\ s_1(H)$ is enforcable and $u_2 = \sum_{k\geq 1} \eta_k V_2(s^k) + \eta_0 V_2(s_1)\}$. $CHP(\Gamma)$ is the convex hull of $P(\Gamma)$.*

The set $P(\Gamma)$ is the set of payoff vectors that are obtained when player 1 commits to playing H with an enforcable probability, and the short run player plays a best response to the type of player 1 with complete information.

Theorem 3 *For every $\epsilon > 0$, Γ, $(U, V) \in CHP(\Gamma)$ there exist a natural number T,$\overline{\delta} < 1$ such that for $\delta \geq \overline{\delta}$ there is a rating system ϕ, and a Perfect Bayesian equilibrium of $G(\delta, \Gamma, \phi)$ satisfying the following: i) the payoff to the normal type of the long run player is at least $U - \epsilon$ after every history. ii) the unconditional expected payoff of every short run player after period T is at least $V - \epsilon$.*

This theorem says that for any enforcable effort level α, we can give player 1 a payoff of more than almost what he gets by committing to exerting effort level α. At the same time, short run players distant in the future get an expected payoff of more than almost what they would get by acting optimally if they knew the true type of player 1, where the normal type is committed to play H with probability α. This theorem is of interest to us, because short run players in the distant future enjoy the benefits from learning the type of player 1, and still the reputation effects never disappear and regulate player 1.

CONCLUSION

This paper demonstrated that reputations can be sustained permanently if information transmission is regulated by a central mechanism. Recent results in reputation literature pointed out that reputation effects do not last long if past outcomes of a long term play is observed by all players. These results made it impossible to explain how markets carrying too much uncertainty can operate for long periods of time.

In this paper we introduce a rating system that collects information about past outcomes in a repeated game, processes the information, and announces the refined information to the players. Similar practices are employed by online markets, such as eBay. We construct a rating system that enables a long lived player (e.g. sellers on internet) to build a reputation for being a commitment type. This brings in the commitment power of this player after any history of the game.

Our rating system observes the same information that short run players observe in the traditional models. However, our rating system throws away some of this information in a systematic way to enable reputation effects in our model. One could view our rating system as an asymptoticaly optimal way of forgetting information to enable reputation effects at all times while regulating the agent.

The welfare of buyers have been ignored in past studies of reputation. However, our setting makes it easier to regulate the long lived player to undertake actions that make short lived players better off. We construct rating systems that implement the Pareto frontier of the stage game when commitment types arise with small probability. The regulating power of rating systems depend on the most hard working commitment type that can arise with positive probability.

When commitment types arise with high probabilities, we could regulate the long run player to follow a strategy that gives him a payoff that is more than the payoff he could get by imitating a type in the set of commitment types. However, in this case we managed to separate all types, thus buyers in distant future know almost the true type of the long run player.

References

[1] Abreu, D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041-1063.

[2] Bakos, Y., and C. Dellarocas (2003): "Cooperation Without Enforcement? A Comparative Analysis of Litigation and Online Reputation as Quality Assurance Mechanisms," *MIT Sloan Working Paper*, No. 4295-03.

[3] Cole, H. L., J. Dow, and W.B. English (1995): "Default, Settlement, and Signalling: lending Resumption in a Reputational Model of Sovereign Debt," *International Economic Review*, 36, 365-385.

[4] Cripps, M. W., G.J. Mailath, and L. Samuelson (2004): "Imperfect Monitoring and Impermanent Reputations," *Econometrica*, 72, 407-432.

[5] Dellarocas, C. (2003): "The Digitization of Word-of-Mouth: Promise and Challenges of Online Feedback Mechanisms," *Management Science*, 49, 1407-1424.

[6] Dellarocas, C. (2004): "Efficiency and Robustness of Binary Feedback Mechanisms in Trading Environments with Moral Hazard," *MIT Sloan Working Paper*, No. 4297-03.

[7] Fudenberg, D., and D. K. Levine (1989): "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57, 759-778.

[8] Fudenberg, D., and D. K. Levine (1992): "Maintaining a Reputation When Strategies are Imperfectly Observed," *Review of Economic Studies*, 59, 561-579.

[9] Fudenberg, D., and D. K. Levine (1994): "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, 62, 103-135.

[10] Fudenberg, D., and D. K. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997-1040.

[11] Holmstrom, B. (1999): "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66, 169-182.

[12] Kreps, D., P. R. Milgrom, D. J. Roberts, and R. Wilson (1982): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 27, 245-252.

[13] Kreps, D., and R. Wilson (1982): "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253-279.

[14] Mailath, G. J., and L. Samuelson (2001): "Who Wants a Good reputation?" *Review of Economic Studies*, 68, 415-441.

[15] Milgrom, P. R., and D.J. Roberts (1982): "Predation, Reputation and Entry Deterrence," *Journal of Economic Theory*, 27, 280-312.

[16] Phelan, C. (2001): "Public Trust and Government Betrayal," *Research Department Staff Report 283, Federal Reserve Bank of Minneapolis*.

[17] Resnick, P., R. Zeckhauser, E. Friedman, and K. Kuwabara (2000): "Reputation Systems," *Communications of the ACM*, 43, 45-48.

APPENDIX

DEFINITIONS

Definition 8 *Let D be a countable set. A set of all finite histories is* $H_I(D) = \cup_{t=0,1,\ldots} D \times (D \times Y \times A_2)^t$.

Definition 9 *Let D be a countable set. A transition rule $\tau : H_I(D) \times Y \times A_2 \to \Delta(H_I(D))$ is called a transparent transition rule where $\sum_{d \in D} \tau(h, y, a_2)(h \cup \{d, y, a_2\}) = 1$ for each $h \in H_I$.*

Definition 10 *An institution $\phi = (H_I(D), \tau, p_0)$ where τ is a transparent transition rule, and support of p_0 is D is called a transparent institution, and is denoted ϕ^*.*

Proof of Theorem 0

We will show that for each equilibrium strategy profile (σ_1, σ_2) of the game $G(\delta, W, \eta, \phi)$, there exists a transparent institution ϕ^* whose public randomization outcomes are the messages of the institution ϕ, and there exists a strategy profile (σ'_1, σ'_2) that yields the same payoff to players 1 and 2, and is an equilibrium of the game $G(\delta, W, \eta, \phi^*)$.

Let (σ_1, σ_2) be a perfect equilibrium of $G(\delta, W, \eta, \phi)$. For $\phi = (R, \tau, p_0)$, let $D = R$, $\tau^* : H_I(D) \times Y \times A_2 \to \Delta(H_I(D))$ such that, $\tau^*(h, y, a_2)(h \cup \{r', y, a_2\}) = \tau(r(h), y, a_2)(r')$ where $r(h) \in R$ represents the last message $r \in R$ in h.

τ^* is a transparent transition rule, and ϕ^* is a transparent institution where $\phi^* = (H_I(D), \tau^*, p'_0)$. The support of p'_0 is the same as the support of p_0, and $p'_0(d) = p_0(d)$ for each $d \in D$. Let $H_{k,t}$ and $H'_{k,t}$ be the set of histories of length t, H_k and H'_k be the set of all finite histories for player k in the game $G(\delta, W, \eta, \phi)$ and $G(\delta, W, \eta, \phi^*)$ respectively. The histories in the latter game include the histories in the former game, let $\varkappa : H'_1 \to H_1$ be a map where $\varkappa(h'_1)$ is the actions, signals and messages of ϕ observed by player 1 in a history in H'_1. We will construct an equilibrium in the latter game that yields the same payoffs to both players as in the former game.

Define $\sigma'_2(h'_2) = \sigma_2(r(h'_2))$. Define σ'_1 as follows:

$\sigma'_1(h_{1,0}) = \sigma_1(h_{1,0})$ for $h_{10} \in H_{1,0} = H'_{1,0}$

Let $\psi_1(r) \in \Delta(\varkappa(h'_{1,1}))$ be the belief of player 2 at period 1 over the histories of length 1 for player 1 observing a message r in the former game.

$\sigma'_1(h'_{1,1}) = \sum_{h_{1,1} \in \varkappa(h'_{1,1})} \psi_1(r(h'_{1,1}))(h_{1,1})\sigma_1(h_{1,1})$

At any period t,

$\sigma'_1(h'_{1,t}) = \sum_{h_{1,t} \in \varkappa(h'_{1,t})} \psi_t(r(h'_{1,t}))(h_{1,t})\sigma_1(h_{1,t})$

It is straightforward to check that σ'_1, σ'_2 is a perfect equilibrium of $G(\delta, W, \eta, \phi^*)$, and yields the same payoffs to players 1 and 2 as in the former game.

Proofs of Theorems 1-3

We will prove the theorems for 2 cases.

Case 1: $B_2(\alpha^k)$ is a singleton for each $k > 0$.

Case 2: There exists a $k > 0$ such that $B_2(\alpha^k)$ is not a singleton.

We will first prove our theorems 1-3 for case 1 for expositional reasons, and then explain how we extend the proofs for case 2.

Proof of Theorem 1

Theorem 1:

For any Stackelberg space W, η, and $u < V_1^s$, there exists $\overline{\delta} < 1$ such that for $\delta \geq \overline{\delta}$ there is a rating system ϕ, and a Perfect Bayesian equilibrium of $G(\delta, W, \eta, \phi)$ where the payoff to the normal type of the long run player is at least u after every history.

Proof: (Case 1)

We will use $V_1(\alpha)$ to refer to $V_1(s_1)$ where $s_1(H) = \alpha$. Fix W, η and $u < V_1^s$. Choose $\alpha^* \in (0,1)$ such that:

i) $\alpha^* > \alpha^s$, $\nexists j : \alpha^s < \alpha^j < \alpha^*$, ii) $V_1(\alpha^*) > u$, iii) $V_1(\alpha^*) > V_1(\alpha^j)$ for each $j > 0$.

Our assumptions ensure that such an α^* exists. In the equilibrium we will construct, player 1 is going to play H with probability α^* most of the time.

Let $W^* = \{\alpha^j | \; l \geq j > 0\} \cup \{\alpha^*\} = \{\beta^1, ..., \beta^{l+1}\}$ such that $\beta^j < \beta^{j+1}$ and let $\beta^k = \alpha^*$. Let $B(\beta^j) = B_2(s_1)$ where $s_1(H) = \beta^j$. Let $W_2^* = \{(\beta, \overline{\beta}) \in W^* \times W^* | \overline{\beta} > \beta$, and $\overline{\beta} > \alpha > \beta$ implies $\alpha \notin W^*\}$.

Let $S = \{1, 2, ..., N\}$ be a set of ratings, and let $B = \{0, 1\}$ be a binary set that is the set of possible outcomes of public randomizations. The state space is $\Omega = S \times B = \{(s, b) | s \in S, b \in B\}$. The number of ratings N and the transition rule on the state space Ω is going to be the rating system. Let $\Omega_I \subset \Omega$ be $\{(1, 0), (1, 1)\}$, $\Omega_L \subset \Omega$ be $\{(N, 0), (N, 1)\}$. For all $\beta \in W^* \cup W_2^*$, we will define regions Ω_β that are disjoint subsets of Ω and $\bigcup_{\beta \in W^* \cup W_2^* \cup \{I, L\}} \Omega_\beta = \Omega$.

For any state $w \in \Omega$, we will define the probability of a rating increase and decrease conditional on the action of player 1. Later we will show that these probabilities can be generated by the signal structure. We will choose a number m big enough to ensure that this is doable, and we will calculate how big this number should be later.

Let $\pi : \Omega \times A_1 \rightarrow \Delta(\{s^+, s^0, s^-\})$ be a map indicating the probability of a rating increase, decrease or not change at any state for each action of player 1.

For $w = (1, 0) \in \Omega_I$,

if $\beta^1 < \alpha^*$, let

$$\pi(w, H)(s^+) = \pi(w, L)(s^+) = \frac{V_1(\alpha^*) - u_1(L, B(\beta^1))}{m}$$

$$\pi(w, H)(s^-) = \pi(w, L)(s^-) = 0$$

if $\beta^1 = \alpha^*$, let

$$\pi(w, H)(s^+) = \pi(w, L)(s^+) = \frac{V_1(\alpha^*) - u_1(L, 0)}{m}$$

$$\pi(w, H)(s^-) = \pi(w, L)(s^-) = 0$$

For $s = 1$, $w = (1, 1) \in \Omega_I$

$$\pi(w, H)(s^+) = \pi(w, L)(s^+) = 1/2$$

$\pi(w,H)(s^-) = \pi(w,L)(s^-) = 0$

For $s = N$, $w = (N,b) \in \Omega_L$, let

$\pi(w,H)(s^+) = \pi(w,L)(s^+) = 0$

$\pi(w,H)(s^-) = \pi(w,L)(s^-) = \frac{u_1(L,B(\beta^{l+1}))-V_1(\alpha^*)}{m}$

For $w = (s,0) \in \Omega_\beta$,

if $\beta < \alpha^*$, let

$\pi(w,H)(s^+) = \frac{V_1(\alpha^*)-u_1(L,B(\beta))+c}{m(1-\beta)} + 1/4$

$\pi(w,H)(s^-) = \pi(w,L)(s^-) = \frac{V_1(\alpha^*)-u_1(L,B(\beta))+c}{m(1-\beta)} \times \beta + 1/4$

$\pi(w,L)(s^+) = 1/4$

if $\beta = \alpha^*$, let

$\pi(w,H)(s^+) = \frac{c}{m} + 1/4$

$\pi(w,H)(s^-) = \pi(w,L)(s^-) = \frac{c}{m} \times \beta + 1/4$

$\pi(w,L)(s^+) = 1/4$

if $\beta > \alpha^*$, let

$\pi(w,H)(s^+) = \frac{c}{m} + 1/4$

$\pi(w,H)(s^-) = \pi(w,L)(s^-) = \frac{c}{m} \times \beta + 1/4$

$\pi(w,L)(s^+) = 1/4$

For $w = (s,1) \in \Omega_\beta$,

$\pi(w,H)(s^+) = \pi(w,H)(s^-) = 1/4$

$\pi(w,L)(s^+) = \pi(w,L)(s^-) = 1/4$

Let $\theta : W_2^* \to \mathbb{N}$ be a function that determines the number of states in every transition region W_2^*. This number will be determined by an algorithm that stops in finite iterations. The map π is going to be determined by the same algorithm. Let $\phi : W_2^* \times \mathbb{N} \to \mathbb{R}$ be a map that determines the relative likelihood of the types $(\beta,\overline{\beta})$ being at any state in a transition region Ω_{w^*} where $w^* = (\beta,\overline{\beta})$. Let $s = \max\{s' | (s',b) \in \Omega_\beta\}$ be the highest rating before the region Ω_{w^*}.

Algorithm for $w^* = (\beta,\overline{\beta})$ where $\overline{\beta} = \alpha^*$:

Step 1: Let $\phi > 0 : B(\beta), B(\beta)+1 \in B_2(\phi/(1+\phi)\beta + 1/(1+\phi)1)$. Set $\phi_1 = \phi + \varepsilon$ such that $B(\beta)+1 \notin B_2(\phi_1/(1+\phi_1)\beta + 1/(1+\phi_1)1)$, $i = 1$.

Step 2:

Let $b = \frac{c}{m}$, $a = 1/4 + \zeta$, $k = \frac{u_1(B_2(\alpha^*),L)-u_1(B_2(\phi_i/(1+\phi_i)\beta+1/(1+\phi_i)\alpha^*),L)}{m}$

$\pi((s+i),b,H)(s^+) = b + a$

$\pi((s+i),b,H)(s^-) = \pi((s+i),b,L)(s^-) = b\overline{\beta} + a - k$

$\pi((s+i),b,L)(s^+) = a$

Where $\zeta > 0$ is a small number such that for $\phi = \phi_i \frac{a+b\beta}{a+b\overline{\beta}} < \phi_i$, $B_2(\phi/(1+\phi)\beta + 1/(1+\phi)1)$ is a singleton. Set $\phi(w^*,i) = \phi_i$

Set $\phi_{i+1} = \phi_i \frac{a+b\beta}{a+b\overline{\beta}} < \phi_i$. If $B(\overline{\beta}) = B_2(\phi_{i+1}/(1+\phi_{i+1})\beta + 1/(1+\phi_{i+1})\overline{\beta})$, then set $\theta(w^*) = i+1$ and go to Step 3. Otherwise increase i by 1 and go to Step 2.

Step 3: Set $\phi(w^*, \theta(w^*)) = \phi_{\theta(w^*)}$

Algorithm for $w^* = (\beta, \overline{\beta})$ **where** $\overline{\beta} \neq \alpha^*$:

Step 1: Let $\phi > 0 : B(\beta), B(\beta)+1 \in B_2(\phi/(1+\phi)\beta + 1/(1+\phi)\overline{\beta})$, if there is no such ϕ, set $\phi = 1$. Set $\phi_1 = \phi + \varepsilon$ such that $B(\beta)+1 \notin B_2(\phi_1/(1+\phi_1)\beta + 1/(1+\phi_1)\overline{\beta})$, $i = 1$.

Step 2:

Let $b = \frac{c}{m}$, $a = 1/4 + \zeta$, $k = \frac{u_1(B_2(\alpha^*),L) - u_1(B_2(\phi_i/(1+\phi_i)\beta + 1/(1+\phi_i)\overline{\beta}),L)}{m}$

$\pi((s+i), b, H)(s^+) = b + a$

$\pi((s+i), b, H)(s^-) = \pi((s+i), b, H)(s^-) = b\alpha^* + a - k$

$\pi((s+i), b, H)(s^+) = a$

Where $\zeta > 0$ is a small number such that for $\phi = \phi_i \frac{a+b\beta}{a+b\overline{\beta}} < \phi_i$, $B_2(\phi/(1+\phi)\beta + 1/(1+\phi)\overline{\beta})$ is a singleton. Set $\phi(w^*, i) = \phi_i$

Set $\phi_{i+1} = \phi_i \frac{a+b\beta}{a+b\overline{\beta}} < \phi_i$. For $\beta > \alpha^*$, if $B(\overline{\beta}) = B_2(\phi_{i+1}/(1+\phi_{i+1})\beta + 1/(1+\phi_{i+1})\overline{\beta})$, then set $\theta(w^*) = i+1$ and go to Step 3. For $\beta = \alpha^*$, if $B(\overline{\beta}) = B_2(\phi_{i+1}/(1+\phi_{i+1})0 + 1/(1+\phi_{i+1})\overline{\beta})$, then set $\theta(w^*) = i+1$ and go to Step 3. Otherwise increase i by 1 and go to Step 2.

Step 3: Set $\phi(w^*, \theta(w^*)) = \phi_{\theta(w^*)}$

Seperation and Construction of Ω

Fix the prior probability distribution of types, η. With a minor abuse of notation, let $\eta_\beta = \eta(j)$ where $\beta = \alpha^j$, $\eta_{\alpha^*} = \eta(0)$.

Let $v : W^* \cup W_2^* \cup \{I, L\} \to [0,1]$ be a map indicating the probability that the value of the binary variable is 0 in any region. We will set $v(w) = 1$ for $w \in W_2^* \cup \{I, L\}$. For $w = \beta \in W^*$, if $\beta \leq \alpha^*$, $v(w) = 1$. If $\beta > \alpha^*$, $v(w) = \frac{u_1(L, B_2(\beta)) - V_1(\alpha^*)}{u_1(L, B_2(\beta)) - V_1(\beta)} \in (0,1)$.

State Space

We will construct the state space by putting the regions in an order first. The number of states in a region is going to be determined by a map $\Upsilon : W^* \times \mathbb{N} \to \mathbb{N}$ where the first argument is the number of states in the first region after region I. We will show that as this increases, each type is going to spend more time in a region designed for it. Note that the number of states in a region $w \in W_2^*$ is determined by an algorithm, and is a constant number θ_w. Let $\Omega(T_1) = \{\Omega_I, \Omega_{\beta^1}, \Omega_{\beta^1,\beta^2}, ..., \Omega_{\beta^{l+1}}, \Omega_L\}$ as follows:

$\{1\} \times \{0,1\} = \Omega_I$,

$\{2, ..., T_1 + 1\} \times \{0,1\} = \Omega_{\beta^1}$,

$\{T_1 + 2, ..., T_1 + 1 + \theta_{\beta^1,\beta^2}\} \times \{0,1\} = \Omega_{\beta^1,\beta^2}$,

$\{T_1 + 2 + \theta_{\beta^1,\beta^2}, ..., T_1 + 1 + \theta_{\beta^1,\beta^2} + \Upsilon(\beta^2, T_1)\} \times \{0,1\} = \Omega_{\beta^2}$,

....

$\Omega_L = \{\max \Omega_{\beta^{l+1}} + 1\} \times \{0,1\}$

For $(s,b) \in \Omega(T_1)$, let $g(s) \in W^* \cup W_2^* \cup \{I, L\}$ where $(s,b) \in \Omega_{g(s)}$.

Equilibrium Strategies

For the normal type, consider the strategy $\sigma : \Omega(T_1) \to [0,1]$ that determines the effort level at any state. If $g(s) \in W^*$, define $\sigma(s,0) = 1$ if $g(s) < \alpha^*$. If $g(s) \geq \alpha^*$, $\sigma(s,0) = \alpha^*, \sigma(s,1) = 0$ for each $s \in S$. If $s = 1$ or $s = N$, $\sigma(s,0) = 0$. If $s = (\beta^1, \beta^2) \in W_2^*$, and $\beta^2 \leq \alpha^*$, $\sigma(s,0) = 1$, $\sigma(s,1) = 0$. If $\beta^2 > \alpha^*$, $\sigma(s,0) = \alpha^*$, $\sigma(s,1) = 0$.

Consider a pure strategy σ_2 for player 2, where $\sigma_2 : \Omega(T_1) \to A_2$. For $w \in \Omega_I$, if $\beta^1 = \alpha^*$, $\sigma_2(w) = 0$, or if $\beta^1 < \alpha^*$, $\sigma_2(w) = B(\beta^1)$. For $w \in \Omega_\beta$ where $\beta \in W^* \backslash \{\alpha^*\}$, $\sigma_2(w) = B(\beta)$. For $w = (s,b) \in \Omega_{\beta^k}$, $\sigma_2(s,0) = B(\alpha^*)$. For $w \in \Omega_L$, $\sigma_2(w) = B(\beta^{l+1})$. For $w = (s,b) \in \Omega_{(\beta_1,\beta_2)}$ where $(\beta^1,\beta^2) \in W_2^*$, let $i = s + 1 - \min \Omega_{(\beta^1,\beta^2)}$, $\phi = \phi((\beta^1,\beta^2), i)$. If $\beta^2 < \alpha^*$, $\sigma_2(s,0) = B_2(\phi/(1+\phi)\beta^1 + 1/(1+\phi)\beta^2)$. If $\beta^2 = \alpha^*$, $\sigma_2(s,0) = B_2(\phi/(1+\phi)\beta^2 + 1/(1+\phi)1)$. If $\beta^2 > \alpha^*$, $\sigma_2(s,0) = B_2(\phi/(1+\phi)\beta^1 + 1/(1+\phi)\beta^2)$.

Seperation

For $w^* = (\beta_1, \beta_2) \in W_2^*$, $\beta \neq \alpha^*$, $s = \max\{s' | (s',b) \in \Omega_{\beta_1}\}$, $s_2 = \min\{s' | (s',b) \in \Omega_{\beta_2}\}$, let $\rho(w^*, \beta) =$

$\Pi_{i=2,\ldots\theta(w^*)} \frac{\pi((s+i),0,H)(s^-)}{\beta\pi((s+i-1),0,H)(s^+)+(1-\beta)\pi((s+i-1),0,L)(s^+)}$

$\underline{\rho}(w^*, \beta) =$

$\frac{\pi((s+1),0,H)(s^-)}{v(\beta_1)(\beta\pi(s,0,H)(s^+)+(1-\beta)\pi(s,0,L)(s^+))+(1-v(\beta_1))(\beta\pi(s,1,H)(s^+)+(1-\beta)\pi(s,1,L)(s^+))}$

$\overline{\rho}(w^*, \beta) =$

$\frac{v(\beta_2)(\beta\pi(s_2,0,H)(s^-)+(1-\beta)\pi(s_2,0,L)(s^-))+(1-v(\beta_2))(\beta\pi(s_2,1,H)(s^-)+(1-\beta)\pi(s_2,1,L)(s^-))}{\beta\pi((s+\theta(w^*)),0,H)(s^+)+(1-\beta)\pi((s+\theta(w^*)),0,L)(s^+)}$

For $\beta = \alpha^*$, replace β by the equilibrium effort level (σ) at that state.

Determination of Number of states

We will determine the function Υ now. $\Upsilon(\beta^1, T) = T$. This is an identity function for the first region after the initial region. $\Upsilon(\beta^i, T)$ is defined recursively as follows:

$\{\frac{\eta_i}{\Upsilon(\beta^i,T)} \frac{1}{\underline{\rho}(w^*,\beta^i)}\} / \{\frac{\eta_{i+1}}{\overline{\Upsilon}(\beta^{i+1},T)} \overline{\rho}(w^*, \beta^{i+1}) \rho(w^*, \beta^{i+1})\} = \phi(w^*, 1)$ for $w^* = (\beta^i, \beta^{i+1}) \in W_2^*$.

$\Upsilon(\beta^{i+1}, T) = \lfloor \overline{\Upsilon}(\beta^{i+1}, T) \rfloor$.

Markov processes τ_β

Having described the state space and transition probabilities, now we will describe the Markov processes with which each type wanders around the states. We call the map $\tau : \Omega \to \Delta(\Omega)$ a transition rule. Every transition rule determines a stochastic process for each type given an initial

probability distribution $p_0 \in \Delta(\Omega)$. The processes for the commitment types is easy, because they play the same strategy at every period. For $\beta \neq \alpha^*$, $g(s) = w$, $g(s+1) = w'$, $g(s-1) = w''$.
$\tau_\beta(s,b)(s+1,0) = v(w')(\pi(s,b,H)(s^+)\beta + \pi(s,b,L)(s^+)(1-\beta))$
$\tau_\beta(s,b)(s+1,1) = (1-v(w'))(\pi(s,b,H)(s^+)\beta + \pi(s,b,L)(s^+)(1-\beta))$
$\tau_\beta(s,b)(s-1,0) = v(w'')(\pi(s,b,H)(s^-)\beta + \pi(s,b,L)(s^-)(1-\beta))$
$\tau_\beta(s,b)(s-1,1) = (1-v(w''))(\pi(s,b,H)(s^-)\beta + \pi(s,b,L)(s^-)(1-\beta))$
$\tau_\beta(s,b)(s,0) = v(w)(1 - \sum_{b'\in B}\tau_\beta(s,b)(s+1,b') - \sum_{b'\in B}\tau_\beta(s,b)(s-1,b'))$
$\tau_\beta(s,b)(s,1) = (1-v(w))(1 - \sum_{b'\in B}\tau_\beta(s,b)(s+1,b') - \sum_{b'\in B}\tau_\beta(s,b)(s-1,b'))$
For $\beta = \alpha^*$, replace β by $\sigma(s,b)$ in the equations above.

Let π_{τ_β} be the stationary distribution of τ_β. Also denote by $\pi^t_{\tau_\beta}$ the time t distribution of the process that starts at state $(1,0)$, and evolves according to τ_β.

Lemma 1 *There is a T_1, Markov processes τ_β on $\Omega(T_1)$ defined as above, and a T such that for $t \geq T$, $B_2(\frac{\sum_{\beta\in W^*\setminus\{\alpha^*\}}\pi^t_{\tau_\beta}(w,b)\eta_\beta\beta + \pi^t_{\tau_{\alpha^*}}(w,b)\eta_{\alpha^*}\sigma(w,b)}{\sum_{\beta\in W^*\setminus\{\alpha^*\}}\pi^t_{\tau_\beta}(w,b)\eta_\beta + \pi^t_{\tau_{\alpha^*}}(w,b)\eta_{\alpha^*}}) = \sigma_2(w,b)$ for each state $(w,b) \in \Omega(T_1)$*

Proof. Satisfied by construction. Fix a type β. The probability of a state increase at a state is the same as the probability of a state decrease at a state for type β for states in Ω_β. At states in $\Omega_{\beta'}$, $\beta' < \beta$ ($\beta' > \beta$), probability of a state increase is strictly bigger (smaller) than probability of a state decrease for type β. As the number of these states gets large (remember number of states in other regions $\Omega_{w'}$ for $w' \in W_2^* \cup \{I, L\}$ is constant), $\sum_{(w,b)\in\Omega_\beta}\pi_{\tau_\beta}(w,b) \times \Upsilon(\beta,T) \to 1$ as T gets large, and $\pi_{\tau_\beta}(w,b) = \pi_{\tau_\beta}(w',b)$ if $w, w' \in \Omega_\beta$. For $(\beta,\beta') \in W_2^*$, $\pi_{\tau_\beta}(w,b) = \underline{\rho}(w^*,\beta)\pi_{\tau_\beta}(w',b)$ where $w \in \Omega_\beta$, $w^* = (\beta,\beta')$, w' is the first state in Ω_{w^*}. $\pi_{\tau_{\beta'}}(w',b) = \pi_{\tau_{\beta'}}(w'',b)\overline{\rho}(w^*,\beta')\rho(w^*,\beta')$ for $w'' \in \Omega_{\beta'}$. Moreover, $\frac{\pi_{\tau_\beta}(w,b)}{\pi_{\tau_{\beta'}}(w,b)} \to_{T\to\infty} \infty$ if $w \in \Omega_\beta$, $(\beta,\beta') \notin W_2^*$. By our choice of Υ, we made sure that $\frac{\pi_{\tau_\beta}(w',b)\eta_\beta}{\pi_{\tau_{\beta'}}(w',b)\eta_{\beta'}} \to \phi((\beta,\beta'),1)$ if w' is the first state in $\Omega_{(\beta,\beta')}$. $\frac{\pi_{\tau_\beta}(w',b)\eta_\beta}{\pi_{\tau_{\beta'}}(w',b)\eta_{\beta'}} \to \phi((\beta,\beta'),i)$ if w' is the i^{th} state in $\Omega_{(\beta,\beta')}$ by our algorithm.

As t gets large, $\pi^t_{\tau_\beta}$ gets close enough to π_{τ_β} for each type β. In our construction, we chose the relative likelihoods such that player 2's best response is strict, so for t large enough, our claim becomes true. ■

Lemma 2 *For each $\delta > 0$, there is a T, and $\epsilon > 0$ such that for transition rules $\{\tau'_{\beta,t}\}_{t=1,...,T}$ where $|\tau_\beta - \tau'_{\beta,t}| < \epsilon$ for each type β and t $\leq$T implies $|p_0(\Pi_{t=1}^T\tau'_{\beta,t})(\tau'_\beta)^s - \pi^{T+s}_{\tau_\beta}| < \delta$.*

Proof. This follows starightforward from the continuity of matrix multiplication, and convergence of markov processes. ■

State Space

We will construct the state state space and the transition rule. Let's first index the signal space such that $Y = \{y_1, y_2, ..., y_z\}$. Define a $z \times 1$ column vector $M(a_2)$ for each $a_2 \in A_2$ as:

$M(a_2)(j) = 1$ if $\rho^{y_j}_{Ha_2} > \rho^{y_j}_{La_2}$,
$M(a_2)(j) = 0$ if $\rho^{y_j}_{Ha_2} < \rho^{y_j}_{La_2}$,
$M(a_2)(j) = 1/2$ if $\rho^{y_j}_{Ha_2} = \rho^{y_j}_{La_2}$.
Let $\sum_{y_j \in Y} \rho^{y_j}_{Ha_2} M(a_2)(j) = M(a_2, H)$, and $\sum_{y_j \in Y} \rho^{y_j}_{La_2} M(a_2)(j) = M(a_2, L)$.

Before defining the transition rule depending on the state, action of player 2, and the signal, We will define the building blocks of transition rule. Let $\pi^* : \Omega \times A_2 \times Y \to \Delta(\{s^-, s^0, s^+\})$ be a map that determines the probability of a rating increase, decrease or stay the same at any state depending on the action of player 2 and the signal. For $a_2 \neq 0$, $\pi^*(s, b, a_2, j)(s^+) = \frac{(M(a_2)(j) - M(a_2,L)) \times (\pi(s,b,H)(s^+) - \pi(s,b,L)(s^+))}{M(a_2,H) - M(a_2,L)} + \pi(s, b, L)(s^+)$, $\pi^*(s, b, a_2, j)(s^-) = \pi(s, b, H)(s^-)$. $\pi^*(s, b, 0, j)(s^+) = \pi(s, b, L)(s^+)$, $\pi^*(s, b, 0, j)(s^-) = \pi(s, b, L)(s^-)$.

Note that, $\pi^*(s, b, a_2, j)(s^+) + \pi^*(s, b, a_2, j)(s^-) < 1$,

and $\pi^*(s, b, a_2, j)(s^+), \pi^*(s, b, a_2, j)(s^-) \in [0, 1]$ if m in the definition of π is big enough.
$\pi^*_{\varepsilon,\delta} : \Omega \times A_2 \times Y \to \Delta(\{s^-, s^0, s^+\})$ is defined as:
$\pi^*_{\varepsilon,\delta}(s, b, a_2, j)(s^+) = \pi^*(s, b, a_2, j)(s^+)$ for $s \neq 1$,
$\pi^*_{\varepsilon,\delta}(1, b, a_2, j)(s^+) = \pi^*(1, b, a_2, j)(s^+) - (\max \Omega - 1)\frac{(1-\delta)}{\delta} + \frac{\varepsilon}{m}$
$\pi^*_{\varepsilon,\delta}(1, b, a_2, j)(s^-) = \pi^*(1, b, a_2, j)(s^-) = 0$,
$\pi^*_{\varepsilon,\delta}(s, b, a_2, j)(s^-) = \pi^*(s, b, a_2, j)(s^-) + (\max \Omega - s)\frac{(1-\delta)}{\delta} - \frac{\varepsilon}{m}$ for $(s, b) \in \Omega_w \neq \Omega_I$.

Transition Rules

For $t \geq 1$, let $z_t = \{0, 1\}^t$, $z_0 = \oslash$, and $Z = \bigcup_{t=0,1\ldots,\infty} z_t$, $R = Z \times \Omega$. $v(s, 0) = v(g(s)), v(s, 1) = 1 - v(g(s))$, $n(z_t)$ is the number of 1's in the string z_t.

Let $\phi_{\delta,\varepsilon} = (p_0, R, \tau^*)$ be a rating system where $p_0(\oslash, 1, 0) = 1$, and τ^* is defined recursively with the equilibrium strategies as follows:
$\sigma_1(z_t, s, b) = \sigma(g(s), b)$, $\sigma_{2t}(z_T, s, b) = \sigma_2(g(s), b)$ for $t \geq T$.
For $t < T$, $\sigma_{2t}(z_t, s, b)$ is defined recursively as follows:
$u_1(s, b) = u_1(L, \sigma_2(g(s), b))$
$\sigma_{20}(\oslash, s, b) \in B_2(\sum_{i \neq k} \eta_i \beta^i + \eta_k \sigma_1(\oslash, s, b))$
$\tau^*(\oslash, s, b, j, a_2)(\oslash \cup \{0\}, s + 1, b') =$
$1/2 v(s + 1, b')(1 - \frac{(1-\delta)(u_1(s,b) - u_1(L, \sigma_{20}(\oslash, s,b)))}{\varepsilon\delta^T})\pi^*(s, b, j, a_2)(s^+)$
$\tau^*(\oslash, s, b, j, a_2)(\oslash \cup \{1\}, s + 1, b') =$
$1/2 v(s + 1, b')(1 + \frac{(1-\delta)(u_1(s,b) - u_1(\sigma_{20}(\oslash, s,b)))}{\varepsilon\delta^T})\pi^*(s, b, j, a_2)(s^+)$
$\tau^*(\oslash, s, b, j, a_2)(\oslash \cup \{0\}, s - 1, b') =$
$1/2 v(s - 1, b')(1 - \frac{(1-\delta)(u_1(s,b) - u_1(L, \sigma_{20}(\oslash, s,b)))}{\varepsilon\delta^T})\pi^*(s, b, j, a_2)(s^-)$
$\tau^*(\oslash, s, b, j, a_2)(\oslash \cup \{1\}, s - 1, b') =$
$1/2 v(s - 1, b')(1 + \frac{(1-\delta)(u_1(s,b) - u_1(\sigma_{20}(\oslash, s,b)))}{\varepsilon\delta^T})\pi^*(s, b, j, a_2)(s^-)$
$\sigma_{2t}(z_t, s, b) \in B_2(\sum_{i \neq k} pr(\beta = \beta^i | \tau^*, s, b)\beta^i + pr(\beta = \alpha^* | \tau, s, b)\sigma_1(z_t, s, b))$
$\tau^*(z_t, s, b, j, a_2)(z_t \cup \{0\}, s + 1, b') =$

$1/2v(s+1,b')(1-\frac{(1-\delta)(u_1(s,b)-u_1(L,\sigma_{2t}(z_t,s,b)))}{\varepsilon\delta^{T-t}})\pi^*(s,b,j,a_2)(s^+)$
$\tau^*(z_t,s,b,j,a_2)(z_t\cup\{1\},s+1,b')=$
$1/2v(s+1,b')(1+\frac{(1-\delta)(u_1(s,b)-u_1(\sigma_{2t}(z_t,s,b)))}{\varepsilon\delta^{T-t}})\pi^*(s,b,j,a_2)(s^+)$
$\tau^*(z_t,s,b,j,a_2)(z_t\cup\{0\},s-1,b')=$
$1/2v(s-1,b')(1-\frac{(1-\delta)(u_1(s,b)-u_1(L,\sigma_{2t}(z_t,s,b)))}{\varepsilon\delta^{T-t}})\pi^*(s,b,j,a_2)(s^-)$
$\tau^*(z_t,s,b,j,a_2)(z_t\cup\{1\},s-1,b')=$
$1/2v(s-1,b')(1+\frac{(1-\delta)(u_1(s,b)-u_1(\sigma_{2t}(z_t,s,b)))}{\varepsilon\delta^{T-t}})\pi^*(s,b,j,a_2)(s^-)$
$\tau^*(z_T,s,b,j,a_2)(b_T,s+1,b')=v(s+1,b')\pi^*_{\varepsilon n(z_T),\delta}(s,b,j,a_2)(s^+)$
$\tau^*(z_T,s,b,j,a_2)(b_T,s-1,b')=v(s-1,b')\pi^*_{\varepsilon n(z_T),\delta}(s,b,j,a_2)(s^-)$
$\tau^*(z_T,s,b,j,a_2)(b_T,s,b')=$
$v(s,b')(1-\pi^*_{\varepsilon n(z_T),\delta}(s,b,j,a_2)(s^-)-\pi^*_{\varepsilon n(z_T),\delta}(s,b,j,a_2)(s^+))$

Theorem: *There exists $\varepsilon>0$, $\overline{\delta}<1$ such that for $\delta<\overline{\delta}$, above strategies are perfect Bayesian Nash equilibrium of $G(\phi_{\varepsilon,\delta},\delta,W,\eta)$.*

Note that the strategies are stationary for $t\geq T$. Let the average discounted payoff of the repeated game to player 1 be $V(z_t,s)$. It is easy to check that $V(z_T,s)=u_1(\alpha^*)-(\max\Omega-s)\frac{1-\delta}{\delta}m+\varepsilon n(z_T)$. Moreover, one can easily verify that $V(z_t,s)=u_1(\alpha^*)-(\max\Omega-s)\frac{1-\delta}{\delta}m+\delta^{T-t}(\varepsilon n(z_t)+\varepsilon(T-t)1/2)$. Once we know the continuation values, we can check that there is no profitable deviation for player 1 at any history, so the following becomes true: $\sigma(g(s),b)\in\arg\max_{\sigma_1\in\Sigma_1}(1-\delta)u_1(\sigma_1,\sigma_{2t}(z_t,s,b))+\delta\sum_{z',s'}(pr(z',s'|s,b,\sigma,\tau^*)V(z',s'))$. To check that strategy of player 2 is optimal, we prove the following claim:

Claim: There exists $\varepsilon>0,\overline{\delta}<1$, such that for $\delta>\overline{\delta}$, $\sigma_2(g(s),b)\in\arg\max_{\sigma\in\overline{\sigma}_2}u_2((\sum_{i=1,..,l}pr(w=w_k|\tau^t_{\varepsilon,\delta},z_T,s,b)\alpha_k+pr(w=w_0|\tau^t_{\varepsilon,\delta},s,b)\sigma_{1t}(z_T,s,b)),\sigma)$

Proof. Follows from lemmas 1 and 2. Note that for each $\zeta_1>0$, there exists $\varepsilon_1>0$, and δ_1 such that $|\pi^t_{\tau_{\varepsilon,\delta}}-\pi^t_{\tau^*}|<\zeta_1\ \forall t\geq T$, for $\varepsilon<\varepsilon_1$, $\delta>\delta_1$. Also There exists a $\delta_2(\varepsilon_1)$ such that for $\delta>\delta_2(\varepsilon_1)$, $pr(w=w_k|\tau^t_{\varepsilon,\delta},b_T,s,b)-pr(w=w_k|\tau^*_{\varepsilon,\delta},b_T,s,b)<\zeta_1$. Choosing ε_1 small enough yields us the result (we use lemma 2 here). ■

Proof of Theorem 2

Theorem 2:

For every $\epsilon>0$, $(U,V)\in IRP(W)$, there exists $\overline{\eta}_0<1$, a natural number T, $\overline{\delta}<1$, such that for $\delta\geq\overline{\delta}$, $1>\eta_0>\overline{\eta}_0$ there is a rating system ϕ, and a Perfect Bayesian equilibrium of $G(\delta,W,\eta,\phi)$ satisfying the following: i) the payoff to the normal type of the long run player is at least $U-\epsilon$ after every history. ii) unconditional expected payoff of every short run player after period T is at least $V-\epsilon$.

Proof: (Case 1)

We will implement payoffs in $CP(W)$. By an initial randomization over the payoffs in $CP(W)$, we can generate any payoff in the convex hull of $IRP(W)$. Fix W, and $(U,V)\in CP(W)$. Then there is an $\alpha^* < \alpha^l$, $\epsilon > 0$ such that $U - \varepsilon < u_1(\alpha^*) - \epsilon$, $V - \varepsilon < V_2(s_1)$ where $s_1(H) = \alpha^*$, $\nexists \alpha^k > \alpha^*$ with $u_1(\alpha^k) \geq u_1(\alpha^*) - \epsilon$. Such an α^* exists because of assumption 2.

In the equilibrium we will construct, player 1 is going to play H with probability α^* most of the time, and will have a payoff of $u_1(\alpha^*) - \epsilon$. Let $W^* = \{\alpha^j | \, l \geq j, \alpha^j > \alpha^*\} \cup \{\alpha^*\} = \{\beta^1, ..., \beta^f\}$ such that $\beta^j < \beta^{j+1}$ and $\beta^1 = \alpha^*$. Let $B(\beta^j) = B_2(s_1)$ where $s_1(H) = \beta^j$. Let $W_2^* = \{(\beta, \overline{\beta}) \in W^* \times W^* | \overline{\beta} > \beta$, and $\overline{\beta} > \alpha > \beta$ implies $\alpha \notin W^*\}$.

Let $S = \{1, 2, ..., N\}$ be a set of ratings, and let $B = \{0, 1\}$ be a binary set that is the set of possible outcomes of public randomizations. Our state space is $\Omega = S \times B = \{(s, b) | s \in S, b \in B\}$. The number of ratings N and the transition rule on the states Ω is going to be the rating system. Let $\Omega_I \subset \Omega$ be $\{(1, 0), (1, 1)\}$, $\Omega_L \subset \Omega$ be $\{(N, 0), (N, 1)\}$. For all $\beta \in W^* \cup W_2^*$, we will define regions Ω_β that are disjoint subsets of Ω and $\bigcup_{\beta \in W^* \cup W_2^* \cup \{I, L\}} \Omega_\beta = \Omega$.

For any state $w \in \Omega$, we will define the probability of a rating increase and decrease conditional on the action of player 1. Later we will show that these probabilities can be generated by the signal structure. We will choose a number m big enough to ensure that this is doable, and we will calculate how big this number should be later.

Let $\pi : \Omega \times A_1 \rightarrow \Delta(\{s^+, s^0, s^-\})$ be a map indicating the probability of a rating increase, decrease or not change at any state for each action of player 1.

For $s = 1$, $w = (1, b) \in \Omega_I$,

$$\pi(w, H)(s^+) = \pi(w, L)(s^+) = \tfrac{u_1(\alpha^*) - \epsilon - u_1(L, 0)}{m}$$
$$\pi(w, H)(s^-) = \pi(w, L)(s^-) = 0$$

For $s = N$, $w = (N, b) \in \Omega_L$, let

$$\pi(w, H)(s^+) = \pi(w, L)(s^+) = 0$$
$$\pi(w, H)(s^-) = \pi(w, L)(s^-) = \tfrac{u_1(L, B(\beta^r)) - u_1(\alpha^*) + \epsilon}{m}$$

For $w = (s, 0) \in \Omega_\beta$,

if $\beta = \alpha^*$, let

$$\pi(w, H)(s^+) = \tfrac{c}{m} + 1/4$$
$$\pi(w, H)(s^-) = \pi(w, L)(s^-) = \tfrac{c}{m} \times \beta + 1/4 + \tfrac{\epsilon}{m}$$
$$\pi(w, L)(s^+) = 1/4$$

if $\beta > \alpha^*$, let

$$\pi(w, H)(s^+) = \tfrac{c}{m} + 1/4$$
$$\pi(w, H)(s^-) = \pi(w, L)(s^-) = \tfrac{c}{m} \times \beta + 1/4$$
$$\pi(w, L)(s^+) = 1/4$$

For $w = (s, 1) \in \Omega_\beta$,

$$\pi(w, H)(s^+) = \pi(w, H)(s^-) = 1/4$$
$$\pi(w, L)(s^+) = \pi(w, L)(s^-) = 1/4$$

Let $\theta : W_2^* \rightarrow \mathbb{N}$ be a function that determines the number of states in every transition region W_2^*. This number will be determined by an algorithm that stops in finite iterations. The map π is going to be determined by the same algorithm. Let $\phi : W_2^* \times \mathbb{N} \to \mathbb{R}$ be a map that determines the relative likelihood of the types $(\beta, \overline{\beta})$ being at any state in a transition region $w^* = (\beta, \overline{\beta})$. Let $s = \max\{s' | (s', b) \in \Omega_\beta\}$ be the highest rating before the region Ω_{w^*}.

Algorithm for $w^* = (\beta, \overline{\beta})$:

Step 1: Let $\phi > 0 : B(\beta), B(\beta)+1 \in B_2(\phi/(1+\phi)\beta + 1/(1+\phi)\overline{\beta})$, if there is no such ϕ, set $\phi = 1$. Set $\phi_1 = \phi + \varepsilon_2$ such that $B(\beta)+1 \notin B_2(\phi_1/(1+\phi_1)\beta + 1/(1+\phi_1)\overline{\beta})$, $i = 1$.

Step 2:

Let $b = \frac{c}{m}$, $a = 1/4 + \zeta$, $k = \frac{u_1(B_2(\alpha^*),L) - u_1(B_2(\phi_i/(1+\phi_i)\beta + 1/(1+\phi_i)\alpha^*),L)}{m}$

$\pi(s+i, H)(s^+) = b + a$

$\pi(s+i, H)(s^-) = \pi(s+i, L)(s^-) = b\alpha^* + a - k + \frac{\epsilon}{m}$

$\pi(s+i, L)(s^+) = a$

Where $\zeta > 0$ is a small number such that for $\phi = \phi_i \frac{a+b\beta}{a+b\overline{\beta}} < \phi_i$, $B_2(\phi/(1+\phi)\beta + 1/(1+\phi)\overline{\beta})$ is a singleton. Set $\phi(w^*, i) = \phi_i$

Set $\phi_{i+1} = \phi_i \frac{a+b\beta}{a+b\overline{\beta}} < \phi_i$. For $\beta > \alpha^*$, if $B(\overline{\beta}) = B_2(\phi_{i+1}/(1+\phi_{i+1})\beta + 1/(1+\phi_{i+1})\overline{\beta})$, then set $\theta(w^*) = i+1$ and go to Step 3. For $\beta = \alpha^*$, if $B(\overline{\beta}) = B_2(\phi_{i+1}/(1+\phi_{i+1})0 + 1/(1+\phi_{i+1})\overline{\beta})$, then set $\theta(w^*) = i+1$ and go to Step 3. Otherwise increase i by 1 and go to Step 2.

Step 3: Set $\phi(w^*, \theta(w^*)) = \phi_{\theta(w^*)}$

Seperation and Construction of Ω

Fix the prior probability distribution of types, η. With a minor abuse of notation, let $\eta_i = \eta(j)$ where $\alpha^j = \beta^i$, $\eta_1 = \eta(0)$.

Let $v : W^* \cup W_2^* \cup \{I, L\} \to \Delta([0,1])$ be a map indicating the probability distribution of the value of the binary variable in any region. We will set $v(w) = 1$ for $w \in W_2^* \cup \{I, L\}$. For $w = \beta \in W^*$, if $\beta = \alpha^*$, $v(w) = 1$. If $\beta > \alpha^*$, $v(w) = \frac{u_1(L, B_2(\beta)) - u_1(\alpha^*) + \epsilon}{u_1(L, B_2(\beta)) - V_1(\beta)} \in (0,1)$.

State Space

We will construct the state space by putting the regions in an order first. The number of states in a region is going to be determined by a map $\Upsilon : W^* \times \mathbb{N} \to \mathbb{N}$ where the first argument is the number of states in the first region after region I. We will show that as this increases, each type is going to spend more time in a region designed for it. Note that the number of states in a region $w \in W_2^*$ is determined by an algorithm, and is a constant number θ_w. Let $\Omega(T_1) = \{\Omega_I, \Omega_{\beta^1}, \Omega_{\beta^1,\beta^2}, ..., \Omega_{\beta^f}, \Omega_L\}$ as follows:

$\{1\} \times \{0,1\} = \Omega_I$,

$\{2, ..., T_1 + 1\} \times \{0,1\} = \Omega_{\beta^1}$,

$\{T_1 + 2, ..., T_1 + 1 + \theta_{\beta^1,\beta^2}\} \times \{0,1\} = \Omega_{\beta^1,\beta^2}$,

$\{T_1 + 2 + \theta_{\beta^1,\beta^2}, ..., T_1 + 1 + \theta_{\beta^1,\beta^2} + \Upsilon(T_1, \beta^2)\} \times \{0,1\} = \Omega_{\beta^2}$,

....
$\Omega_L = \{\max \Omega_{\beta^f} + 1\} \times \{0,1\}$
For $(s,b) \in \Omega(T_1)$, let $g(s) \in W^* \cup W_2^* \cup \{I, L\}$ where $(s,b) \in \Omega_{g(s)}$.

Equilibrium Strategies

For the normal type, consider the strategy $\sigma : \Omega(T_1) \to [0,1]$ that determines the effort level at any state. If $g(s) \in W^*$, define $\sigma(s,0) = 1$ if $g(s) < \alpha^*$. If $g(s) \geq \alpha^*$, $\sigma(s,0) = \alpha^*, \sigma(s,1) = 0$ for each $s \in S$. If $s = 1$ or $s = N$, $\sigma(s,b) = 0$. If $s = (\beta^1, \beta^2) \in W_2^*$, and $\beta^2 \leq \alpha^*$, $\sigma(s,0) = 1$, $\sigma(s,1) = 0$. If $\beta^2 > \alpha^*$, $\sigma(s,0) = \alpha^*$, $\sigma(s,1) = 0$.

Consider a pure strategy σ_2 for player 2, where $\sigma_2 : \Omega(T_1) \to A_2$. For $w \in \Omega_I$, if $\beta^1 = \alpha^*$, $\sigma_2(w,b) = 0$, or if $\beta^1 < \alpha^*$, $\sigma_2(w,b) = B(\beta^1)$. For $w \in \Omega_\beta$ where $\beta \in W^* \backslash \{\alpha^*\}$, $\sigma_2(w,b) = B(\beta)$. For $w \in \Omega_{\beta^k}$, $\sigma_2(w,0) = B(\alpha^*)$. For $w \in \Omega_L$, $\sigma_2(w,b) = B(\beta^f)$. For $w = (s,b) \in \Omega_{(\beta_1,\beta_2)}$ where $(\beta_1,\beta_2) \in W_2^*$, let $i = s + 1 - \min \Omega_{(\beta_1,\beta_2)}$, $\phi = \phi((\beta_1,\beta_2), i)$. If $\beta_2 < \alpha^*$, $\sigma_2(w,0) = B_2(\phi/(1+\phi)\beta_1 + 1/(1+\phi)\beta_2)$. If $\beta_2 = \alpha^*$, $\sigma_2(w,0) = B_2(\phi/(1+\phi)\beta_1 + 1/(1+\phi)1)$. If $\beta_2 > \alpha^*$, $\sigma_2(w,0) = B_2(\phi/(1+\phi)\beta_1 + 1/(1+\phi)\beta_2)$.

Seperation

For $w^* = (\beta_1, \beta_2) \in W_2^*$, $\beta \neq \alpha^*$, $s = \max\{s' | (s',b) \in \Omega_{\beta_1}\}$, $s_2 = \min\{s' | (s',b) \in \Omega_{\beta_2}\}$, let $\rho(w^*, \beta) =$
$\Pi_{i=2,\ldots\theta(w^*)} \frac{\pi((s+i),0,H)(s^-)}{\beta\pi((s+i-1),0,H)(s^+)+(1-\beta)\pi((s+i-1),0,L)(s^+)}$
$\underline{\rho}(w^*, \beta) =$
$\frac{\pi((s+1),0,H)(s^-)}{v(\beta_1)(\beta\pi(s,0,H)(s^+)+(1-\beta)\pi(s,0,L)(s^+))+(1-v(\beta_1))(\beta\pi(s,1,H)(s^+)+(1-\beta)\pi(s,1,L)(s^+))}$
$\overline{\rho}(w^*, \beta) =$
$\frac{v(\beta_2)(\beta\pi(s_2,0,H)(s^-)+(1-\beta)\pi(s_2,0,L)(s^-))+(1-v(\beta_2))(\beta\pi(s_2,1,H)(s^-)+(1-\beta)\pi(s_2,1,L)(s^-))}{\beta\pi((s+\theta(w^*)),0,H)(s^+)+(1-\beta)\pi((s+\theta(w^*)),0,L)(s^+)}$
For $\beta = \alpha^*$, replace β by the equilibrium effort level (σ) at that state.

Determination of Number of states We will determine the function Υ now. $\Upsilon(\beta^1, T) = T$. This is an identity function for the first region after the initial region. $\Upsilon(\beta^i, T)$ for $i > 1$ is defined recursively as follows:

$\{\frac{\eta_i}{\Upsilon(\beta^i,T)} \frac{1}{\underline{\rho}(w^*,\beta^i)}\} / \{\frac{\eta_{i+1}}{\overline{\Upsilon}(\beta^{i+1},T)} \overline{\rho}(w^*, \beta^{i+1}) \rho(w^*, \beta^{i+1})\} = \phi(w^*, 1)$ for $w^* = (\beta^i, \beta^{i+1}) \in W_2^*$.

$\Upsilon(\beta^{i+1}, T) = \lfloor \overline{\Upsilon}(\beta^{i+1}, T) \rfloor$.

Let $r_1 = \frac{\pi(1,0,L)(s^+)}{\pi(2,0,L)(s^-)}$, $r_2 = \frac{(1-\alpha^*)\pi(2,0,L)(s^+)+\alpha^*\pi(2,0,H)(s^+)}{\pi(2,0,L)(s^-)} < 1$,

Let $\{\frac{\eta_1}{r_1 + \frac{1}{1-r_2}} r_2^{T-1} \frac{1}{\underline{\rho}(w^*,\beta^i)}\} / \{\frac{\eta_2}{\overline{\Upsilon}(\beta^2,T)} \overline{\rho}(w^*, \beta^2) \rho(w^*, \beta^2)\} = \phi(w^*, 1)$ for $w^* = (\beta^1, \beta^2) \in W_2^*$.

Markov processes τ_β

Having described the state space and transition probabilities, now we will describe the Markov processes with which each type wanders around the states. We call the map $\tau : \Omega \to \Delta(\Omega)$ a transition rule. Every transition rule determines a stochastic process given an initial probability distribution $p_0 \in \Delta(\Omega)$. The processes for the commitment types is easy, because they play the same strategy at every period. For $\beta \neq \alpha^*$, $g(s) = w$, $g(s+1) = w'$, $g(s-1) = w''$.
$\tau_\beta(s,b)(s+1,0) = v(w')(\pi(s,b,H)(s^+)\beta + \pi(s,b,L)(s^+)(1-\beta))$
$\tau_\beta(s,b)(s+1,1) = (1-v(w'))(\pi(s,b,H)(s^+)\beta + \pi(s,b,L)(s^+)(1-\beta))$
$\tau_\beta(s,b)(s-1,0) = v(w'')(\pi(s,b,H)(s^-)\beta + \pi(s,b,L)(s^-)(1-\beta))$
$\tau_\beta(s,b)(s-1,1) = (1-v(w''))(\pi(s,b,H)(s^-)\beta + \pi(s,b,L)(s^-)(1-\beta))$
$\tau_\beta(s,b)(s,0) = v(w)(1 - \sum_{b' \in B} \tau_\beta(s,b)(s+1,b') - \sum_{b' \in B} \tau_\beta(s,b)(s-1,b'))$
$\tau_\beta(s,b)(s,1) = (1-v(w))(1 - \sum_{b' \in B} \tau_\beta(s,b)(s+1,b') - \sum_{b' \in B} \tau_\beta(s,b)(s-1,b'))$
For $\beta = \alpha^*$, replace β by $\sigma(s,b)$ in the equations above.

Let π_{τ_β} be the stationary distribution of τ_β. Also denote by $\pi^t_{\tau_\beta}$ the time t distribution of the process that starts at state $(1,0)$, and evolves according to τ_β.

Lemma 3 *There is a $\overline{\eta} < 1$ such that for $\eta_1 > \overline{\eta}$, there is a T_1, Markov processes τ_β on $\Omega(T_1)$ defined as above, and a T such that for $t \geq T$, $B_2(\frac{\sum_{\beta \in W^* \setminus \{\alpha^*\}} \pi^t_{\tau_\beta}(w,b)\eta_\beta\beta + \pi^t_{\tau_{\beta^k}}(w,b)\eta_k\sigma(w,b)}{\sum_{\beta \in W^* \setminus \{\alpha^*\}} \pi^t_{\tau_\beta}\eta_\beta + \pi^t_{\tau_{\beta^k}}\eta_k}) = \sigma_2(w,b)$*

Proof. Satisfied by construction. Fix a type β. The probability of a state increase at a state is the same as the probability of a state decrease at a state for type β for states in Ω_β. At states in $\Omega_{\beta'}$, $\beta' < \beta$ ($\beta' > \beta$), probability of a state increase is strictly bigger (smaller) than probability of a state decrease for type β. As the number of these states get large (remember number of states in other regions $\Omega_{w'}$ for $w' \in W_2^* \cup \{I, L\}$ is constant), $\sum_{(w,b) \in \Omega_\beta} \pi_{\tau_\beta}(w,b) \times \Upsilon(\beta, T) \to 1$ as T gets large, and $\pi_{\tau_\beta}(w,b) = \pi_{\tau_\beta}(w',b)$ if $w, w' \in \Omega_\beta$. For $(\beta,\beta') \in W_2^*$, $\pi_{\tau_\beta}(w,b) = \underline{\rho}(w^*,\beta)\pi_{\tau_\beta}(w',b)$ where $w \in \Omega_\beta$, $w^* = (\beta,\beta')$, w' is the first state in Ω_{w^*}. $\pi_{\tau_{\beta'}}(w',b) = \pi_{\tau_{\beta'}}(w'',b)\overline{\rho}(w^*,\beta')\rho(w^*,\beta')$ for $w'' \in \Omega_{\beta'}$. Moreover, $\frac{\pi_{\tau_\beta}(w,b)}{\pi_{\tau_{\beta'}}(w,b)} \to_{T \to \infty} \infty$ if $w \in \Omega_\beta$, $(\beta,\beta') \notin W_2^*$. By our choice of Υ, we made sure that $\frac{\pi_{\tau_\beta}(w',b)\eta_\beta}{\pi_{\tau_{\beta'}}(w',b)\eta_{\beta'}} \to \phi((\beta,\beta'),1)$ if w' is the first state in $\Omega_{(\beta,\beta')}$. $\frac{\pi_{\tau_\beta}(w',b)\eta_\beta}{\pi_{\tau_{\beta'}}(w',b)\eta_{\beta'}} \to \phi((\beta,\beta'),i)$ if w' is the i^{th} state in $\Omega_{(\beta,\beta')}$ by our algorithm.

For types $\alpha^j \notin W^*$, the probability of a decrease is strictly less than the probability of a state increase in Ω_{β^1}, and this is also true for the normal type β^1. So for η_k large enough, the lemma becomes true.

As t gets large, $\pi^t_{\tau_\beta}$ gets close enough to π_{τ_β} for each type β. In our construction, we chose the relative likelihoods such that player 2's best response is strict, so for t large enough, our claim becomes true. ■

State Space and Transition Rules

Please look at the same section in the proof of Theorem 1.

Theorem 2:

There exists $\varepsilon > 0$, $\overline{\delta} < 1$ such that for $\delta < \overline{\delta}$, above strategies are perfect Bayesian Nash equilibrium of $G(\Phi, \delta, W, \eta)$.

Proof: The proof is very similar to the proof of theorem 1. The remaining part is to show that short run players' payoffs are larger than $V - \varepsilon$ after time T. This follows by observing that the probability that the normal type is at a state in region β^1 at some period $t > T$ can be made close enough to 1 by making ϵ close enough to 0.

Proof of Theorem 3

Theorem 3:

For every $\epsilon > 0$, Γ, $(U, V) \in CHP(\Gamma)$ there exist a natural number T,$\overline{\delta} < 1$ such that for $\delta \geq \overline{\delta}$ there is a rating system ϕ, and a Perfect Bayesian equilibrium of $G(\delta, \Gamma, \phi)$ satisfying the following: i) the payoff to the normal type of the long run player is at least $U - \epsilon$ after every history. ii) the unconditional expected payoff of every short run player after period T is at least $V - \epsilon$.

Proof: (Case 1)

We will use $V_1(\alpha)$ to refer to $V_1(s_1)$ where $s_1(H) = \alpha$. First we will implement payoffs in $P(\Gamma)$. By an initial randomization over the payoffs in $P(\Gamma)$, we can generate any payoff in the convex hull of $P(\Gamma)$, $CHP(\Gamma)$. Fix W, η and $(U, V) \in P(\Gamma)$. Choose $\alpha^* \in (0, 1)$ such that:
i) $\alpha^* > \alpha^s$, $\nexists j : \alpha^s < \alpha^j < \alpha^*$, ii) $V_1(\alpha^*) > U$, iii) $V_1(\alpha^*) > V_1(\alpha^j)$ for each $j > 0$, iv) $V < \sum_{k \geq 1} \eta_k V_2(s^k) + \eta_0 V_2(\alpha^*)$

Our assumptions ensure that such an α^* exists. In the equilibrium we will construct, player 1 is going to play H with probability α^* most of the time. The rest of the construction is the same as in the proof of Theorem 1. The payoff of short run players after some time T is close to V because the probability that a type is in the region designed for him is close to 1.

Proofs of Theorems 1-3 - Case 2

Let $B^-(\alpha^i) = \min B(\alpha^i)$, $B^+(\alpha^i) = \max B(\alpha^i)$. In this case, we will divide each region Ω_w for $w \in W^* \backslash \{\beta^{k+1}, \beta^1\}$ into 2 sub-regions, Ω_{w-} and Ω_{w+}. If $k \neq l$ and $B^-(\beta^{k+1}) \neq B^+(\beta^{k+1})$, then we divide the region Ω_w for $w = \beta^{k+1}$ into 3 regions Ω_{w-}, Ω_{w-+} and Ω_{w+}. The map π and v will be changed as follows:

For states $(s, 0) \in \Omega_{\beta-}$ where $\beta < \alpha^s$, $\pi(w, H)(s^+) = \frac{V_1(\alpha^*) - u_1(L, B^-(\beta)) + c}{m(1-\beta)} + 1/4$ and $\pi(w, H)(s^-) = \pi(w, L)(s^-) = \frac{V_1(\alpha^*) - u_1(L, B^-(\beta)) + c}{m(1-\beta)} \times \beta + 1/4$. For states $(s, 0) \in \Omega_{\beta+}$ we make the change by putting $B^+(\beta)$ in the previous equations instead of $B^-(\beta)$. Also change $B(\beta^{l+1})$ with $B^-(\beta^{l+1})$ in the definition of π for region L. For the initial region and for the region Ω_{β^1}, use $B^+(\beta^1)$ instead of $B(\beta^1)$. We will change v for regions Ω_{w-} and Ω_{w+} for regions $w > \alpha^s$ in a similar way by changing $B(\beta)$ by $B^-(\beta)$ and $B^+(\beta)$. For Ω_{w-+}, v will be the solution of $v(u_1(L, B^+(\beta)) - \beta c) + (1-v)(u_1(L, B^-(\beta))) = V_1(\alpha^*)$

The function that determines the number of states in each region, Υ will change as well. In this case, the ratio of $\Upsilon(w-, T)$ and $\Upsilon(w+, T)$ will be determined by $\phi(w, w', 1)$ and $\phi(w'', w, \theta_{w'',w})$ jointly.

The equilibrium strategy of player 1 is unchanged. Player 2 plays $B^-(\beta)$ in regions $\Omega_{\beta-}$ and $B^+(\beta)$ in regions $\Omega_{\beta+}$. Also in the region $\Omega_{\beta-+}$, player 2 plays $B^+(\beta)$ if the public randomization b is 0 and $B^-(\beta)$ otherwise.