Reputation Effects in Two-Sided Incomplete-Information Games

Ayça Özdoğan *
University of Minnesota

April 15, 2010

Abstract

This paper studies reputation effects in a class of games with imperfect public monitoring and two long-lived players, both of whom have private information about their own type and uncertainty over the types of the other player. Players may be either a *strategic* type who maximizes expected utility or a (simple) *commitment* type who plays a prespecified action every period. As in standard models, the reputation of a strategic type of player for being the commitment type is established by mimicking the behavior of the commitment type. The distinct feature of our model is that *both* strategic players aim to establish a (false) reputation for being the commitment type. The class of games we consider encompasses a wide range of economic interactions between two parties that involve hidden-information (e.g. between a regulator and a regulatee) or hidden-action (e.g between an employer and an employee), where the reputation concerns of both parties are apparent. In both games, one party (principal) prefers that the other party (agent) play in a specific way and use costly auditing to enforce this behavior. The principal aims to establish a reputation for being *diligent*; whereas the agent want to build a reputation for being *virtuous*. Extending the techniques of Cripps, Mailath, and Samuelson (2004), we find that neither strategic player can sustain a reputation for playing a noncredible behavior, i.e. a behavior which is not optimal given that the opponent is best responding in the stage game. Hence, in this class, the true types of both players will be revealed eventually in all Nash equilibria and the asymmetric information does not affect equilibrium analysis in the long-run. In fact, we show that this is the only class of two-sided incomplete information games (with simple commitment types) where reputations disappear in the long-run, in all equilibria. To do so, we provide an example where reputations for noncredible behavior are sustained in a Nash equilibrium.

*I am grateful to my advisor Aldo Rustichini for his invaluable guidance and encouragement throughout this project. I also thank Chris Phelan and Jan Werner for their very helpful comments and support. In addition, I thank Rajesh Aggarwal, Beth Allen, Ioanna Grypari, David Rahman, Gina Pieters, Itai Sher, Yuichiro Waki, Randall Wright and participants of Mathematical Economics Workshop at the University of Minnesota, seminars at Bogazici University, Middle East Technical University, Bilkent University and TOBB-ETU for their comments and suggestions. Financial support from Graduate Research Partnership Program Fellowship is gratefully acknowledged. All remaining errors are mine. Email ayca@econ.umn.edu for comments and questions.

1 Introduction

This paper studies long-run effects of reputations in a class of games with imperfect public monitoring[1] and two long-lived players, both of whom have private information about their own type and uncertainty over the types of the other player. Players may be either a *strategic* type who maximizes expected utility or a *commitment (behavioral)* type who plays a prespecified action every period. As in standard models, the reputation of a strategic type of player for being commitment type is established by mimicking the behavior of the commitment type.[2] The distinct feature of our model is that, since there is uncertainty over the types of both players, *both* players aim to establish a false reputation for being the commitment type in order to induce the opponent player behave in a specific way. We believe that wanting to establish a reputation is a key concern for *all* parties involved in several economic interactions. Specifically, in the economic applications that can be described by the class of games we study (that is to be discussed in the subsequent sections), the reputation concerns of both parties are apparent. Our paper examines the sustainability of reputations in these games when both players have concerns about their own reputation and the reputation of the opponent player. We find that neither (strategic type of) player can sustain a false reputation for playing a noncredible behavior, i.e. a behavior which is not optimal given that the opponent player is best responding to this in the stage game where the true types (strategic) were to be known. Hence, in this class, the true types of both players will be revealed eventually and the asymmetric information does not affect equilibrium analysis in the long-run.

This class of games encompass a wide range of economic applications between two parties that involve hidden-information (e.g. between a regulator and regulatee) or hidden-action (e.g. between an employer and employee). The common feature of these economic interactions is that one party (the principal) prefers that the other party (the agent) play in a certain way (e.g. to be truthful in a hidden-information setting, and to exert high effort in a hidden-action setting) and use costly auditing to enforce this behavior. The principal can choose either to be lazy or diligent in auditing the agent,[3] which results in different probabilities of detecting an undesirable behavior of the agent. The agent prefers the principal to be lazy. Suppose that the payoffs of the stage game is defined such that if the agent thinks that the principal is diligent, she is better off by behaving properly, i.e. the way the principal wants her to play. Otherwise, if she thinks that the principal is lazy, she has an incentive to play some other action, which is not preferred by the principal. On the other hand, when the principal thinks that the agent chooses the proper action, he would like to be lazy in auditing. Otherwise, he is better off by being diligent. Suppose that the agent believes that with some probability the principal is a *tough* type who always chooses to be diligent in

[1]The actions of players are not observable; however, players observe a public signal probability of which depends on the action profile chosen.

[2]In most studies, a player aims to establish reputation of being the Stackelberg type, by mimicking this type's commitment action, i.e. the Stackelberg action (an action one would like to commit given that such a commitment induces a best response from the opponent player).

[3]This can be interpreted as principal allocating time and/or resources for auditing the agent among its other tasks

auditing. Then, the principal aims to establish a reputation for being the tough type, by choosing diligence every period, in order to enforce the agent to choose the proper behavior. Suppose also that the agent could be a *virtuous* type who acts properly every period. Then, the agent aims to establish a reputation for being the virtuous type to induce the principal to be lazy. The fact that the actions of the two parties are not observable to each other prevents them learning each other's true type.

These games belong to a class, which we call *one-sided binding moral hazard at the commitment profile*, as only one of the players has an incentive to deviate from the commitment action given that her opponent is of the commitment type or plays like the commitment type. [4] The properties of these games will be discussed in detail in subsequent sections. However, we would like to point out that this class captures games with the following property: The interest of one of the players is aligned with the behavior of the commitment type of the opponent player; whereas the interest of the other player conflicts with the behavior of the commitment type of opponent. For instance, the interest of the strategic principal is aligned with the behavior of the virtuous type of the agent since he prefers the agent act properly. However, the strategic agent's interest conflicts with the behavior of the tough principal since she prefers the principal to be lazy. We study long-run behavior of reputations in games with one-sided binding moral hazard at commitment profile when the actions of the players are not perfectly observable (imperfect public monitoring).

Long-run equilibrium analysis requires to examine the evolution of reputations. The theory of reputations in standard infinitely-repeated games focuses on only one of the players having uncertainty over her types (call this player the *informed* player) and wanting to build a false reputation. The main results about how reputations evolve are: (1) There is a reputation-building phase in which the uninformed player is eventually convinced to see the commitment action if the informed player has been playing it. (2) There is a reputation-destruction phase, when the actions of the informed player is not observable, if the reputation is for a noncredible behavior, i.e. the commitment action that the informed player is mimicking is not optimal given that the uninformed player is best responding. The reputation is destructed optimally by the informed player as a trade off between the loss of reputation and the payoff gain by deviating to an action other that the commitment action, knowing that such a deviation cannot be detected unambiguously when her actions are not observable. We show that the central results about the evolution of reputations are robust introducing uncertainty over the types of the second player in games with one-sided binding moral hazard at the commitment profile. Extending the techniques of Cripps, Mailath, and Samuelson (2004), we find that neither player can sustain reputation for playing a strategy that is not equilibrium of the complete-information stage game - the game without uncertainty over the types. Hence, in this class, the true types of the players will be revealed eventually and the asymmetric information about the types

[4]The commitment action profile in the principal-agent game is (proper behavior, diligence). The only player who has an incentive to deviate from this profile is the principal, i.e. the principal wants to deviate to be lazy given that the agent chooses to behave properly. For the agent, the best reply against a diligent principal is to behave properly.

of players does not affect equilibrium analysis in the long-run.

In order to illustrate some important features of these games, we provide an example that describes a hidden-information framework between a regulator and a regulatee in section 1.1. The reader who wants to continue with a discussion of the general set up and our results may jump to section 1.2.

1.1 Example

Consider the repeated interaction between a regulator and a regulatee [5] where the possible actions for the regulatee are to be truthful or untruthful about some noisy information she gets regarding the state of nature (that is realized at the end of the period); and those for the regulator are to be lazy or diligent in auditing the regulatee. For instance, the regulatee could be a bank which gets a noisy information about its own financial health and the regulator could be a government official. [6,7] The actions of the regulatee to be truthful or untruthful are not observable to the regulator. However, the regulator observes if the message sent by the regulatee is correct or not (i.e. if it matches the realized state of nature or not). Since the regulatee's information about the state of nature is noisy; an incorrect message can come from a truthful behavior, as well as a correct message can come from an untruthful behavior. Similarly, the actions of the regulator (being lazy or diligent), which induce different probability of audit, are not observable to the regulatee. [8] However, she observes if there is an audit or not at the end of the period, which may result after a lazy or a diligent behavior. The probability of correct message is higher if the regulatee is truthful; similarly, the probability of an audit is higher if the regulator is diligent. The regulator prefers the regulatee to be truthful and the regulatee prefers the regulator to be lazy.

Table 1: Expected Payoff Matrix

	L	D
T	x, y	$x - l_1, y - c$
U	$x + g, z$	$x - l_2, z - c + d$

where $y, z, g, c, d > 0$, $l_2 > l_1 \geq 0$ and $y > y - c > z - c + d > z$.

[5] This game can be considered as a variant of *inspection games* extensively studied in the literature. We refer the reader Rudolf, Bernhard, and Zamir (2002) for a discussion on inspection games.

[6] The other possible applications for which these games can be used to model include analyzing tax evasion through the interaction between a tax payer and tax collecting agency or the asset market manipulation via strategic announcements of an insider in the presence of a regulator.

[7] Benabou and Laroque (1992) provide a model of repeated strategic communication that analyzes manipulation in asset markets, where they extend Sobel (1985)'s model to the case in which the sender (insider) has noisy private information about the value of an asset. The sender can deceive public and distort the asset price through strategic announcements. However, their model is missing a strategic receiver who can audit the sender. Ozdogan (2009) provides a model with a strategic receiver.

[8] One can interpret this as the regulator chooses between two mixed strategies; or allocates some resources or time to auditing among its other tasks.

Row player is the regulatee who chooses to be truthful (T) or untruthful (U) and the column player is the regulator who chooses to be diligent (D) or lazy (L). The regulatee's best response against the choice of being lazy is to be untruthful, since she has an expected gain of g. However, the regulatee's best response when the regulator is diligent in auditing is to be truthful, since the expected loss from untruthfulness when the regulator is diligent l_2 is higher than l_1. [9] For the regulator, the best response when the regulatee is truthful is to be lazy, since the diligence in auditing has a cost of c. On the other hand, the regulator's best response is to be diligent when the regulatee is untruthful, since there is a expected gain d from possible detection of the untruthful behavior. The regulator gets his highest payoff when he is lazy and the regulatee is truthful; whereas the regulatee gets her highest payoff when the regulator is lazy and she is untruthful. Thus, the regulator wants to convince the regulatee that he is diligent to enforce truthfulness. However, the best response of the regulator if the regulatee is truthful is to be lazy. On the other hand, the regulatee prefers the regulator to be lazy, and thus the regulatee wants to convince the regulator that she is truthful to enforce laziness in auditing. However, the regulatee's best response once she thinks that the regulator is lazy is to be untruthful. [10]

Let α_1 be the strategy of the regulatee and α_2 be the strategy of the regulator (in the stage game). There is unique mixed strategy Nash equilibrium of the stage game, which is $\alpha_1(T) = \frac{d-c}{d}$ and $\alpha_2(D) = \frac{g}{g+l_2-l_1}$ that provides a payoff of $x - \frac{g.l_1}{g+l_2-l_1}$ to the regulatee and $y - \frac{c(y-z)}{d}$ to the regulator. We want to point out that the unique equilibrium profile is Pareto dominated by the profile $(Truthful, Lazy)$ which gives x and y to the regulatee and regulator, respectively. Also, note that as the expected gain d from detecting untruthfulness for the regulator approaches to the cost of monitoring c, then the regulatee, in equilibrium, chooses to be truthful with a smaller probability. On the other hand, as the expected gain g from untruthfulness increases, the regulator would choose to be diligent with a higher probability.

Suppose that there are private types for both of the players and players have uncertainty over the types of the other player. More specifically, the regulator believes that the regulatee is a *virtuous* type, who is truthful in every period, with some probability. The regulatee wants to use regulator's uncertainty over her types and pretend to be the virtuous type (by acting like the virtuous type) to enforce the regulator to be lazy in the continuation play. On the other hand, the regulatee believes that the regulator is a *tough* type, who is diligent in every period, with some probability. Then the regulator may find it worthwhile to exploit regulatee's uncertainty over his type by pretending to be the tough type to induce truthfulness. Since the actions of the players are not observable to each other, they can't learn each other's true types for sure. We would like to point out that the incentives of the regulator is aligned with the virtuous type of

[9]We allow l_1, the expected loss for the regulatee when the regulator is diligent and the regulatee is truthful, to be positive as well as zero because of the noise in signals regarding her behavior.

[10]We can define the repeated interaction between an employee and employer via a game with the same stage game expected payoff matrix, where the available actions for the employee are high effort and low effort, and for the employer are to be diligent and lazy in auditing. The actions of the players are not observable. The employer observes the output produced by the employee and the employee observes if there has been an audit or not by the employer.

the regulatee in the sense that the regulator prefers the regulatee to be truthful and a virtuous regulatee is always truthful. On the other hand, the incentives of the regulatee conflicts with that of a tough regulator as the regulatee prefers the regulator to be lazy, but a tough regulator is always diligent. We characterize the equilibria of the one-shot game with the incomplete-information on both sides and how those depend on the parameters of the stage-game in the Appendix.

We analyze what happens to the reputations (for being tough and virtuous) of the regulator and the regulatee in the limit in order to understand long-run equilibrium behavior and payoffs. For instance, if there were to be a Nash equilibrium where both of the reputations are sustainable, i.e. the types of the players are not revealed, this means that each player should be seeing similar public signals on average from both types of the opponent player, since they cannot distinguish between the types (given that the signals are statistically informative about both players' actions). But, this is only possible if both types of the players act the same on average in the limit. Then the regulator should be diligent and the regulatee should be truthful on average indefinitely, which wouldn't be efficient since this profile is Pareto inferior to always playing the profile $(Truthful, Lazy)$. We show this is not the case: neither of the players' reputation is sustainable, i.e. reputations of being tough and virtuous disappear in the limit, when players are not indeed those types. Hence, the asymmetric information over the types of players does not affect the equilibrium analysis, in the sense that any Nash equilibrium of the game converges to the Nash equilibrium of the game without uncertainty over the types.

1.2 Approach and Result

We present a class of games with imperfect public monitoring [11] and incomplete-information over the types of both players, which captures the repeated interaction between a regulator and regulatee described in section 1.1.

Each player, in our model, can be either a *commitment* type who plays an exogenously specified stage game action every period independently of the history of the play [12] or a *strategic* type who maximizes expected payoffs. The stage games and the stage game commitment action profile we allow are restricted. We say that a player is *subject to binding moral hazard at the commitment action profile* if the player finds it suboptimal to play the commitment action given that the other player is playing her commitment action or is the commitment type. The key condition of our model is that the stage games we consider have "one-sided binding moral hazard at the commitment profile," which requires exactly one player being subject to binding moral at the commitment action profile. Hence, for only one of the players (say player 2 - the principal), the best response to the commitment action of the opponent is not his commitment

[11] Actions of the players are not observable to each other; instead, players observe a public signal, distribution of which depends on the action profile chosen.

[12] A repeated game (behavior) strategy is called a *simple* strategy if it is a constant function of histories, i.e. for every period independent of history of the play, it assigns the same (possibly mixed) stage game action.

action, while for the other player (player 1 - the agent), her best response to the commitment action of the opponent player is her commitment action. So, there is only one player who has incentives to deviate to another action at the commitment action profile. Moreover, we require a monitoring structure that ensures that the public signals are statistically informative about each player's actions; and also, that allows players to infer the opponent's beliefs about their own type (thus makes the beliefs of each player *public*). The assumptions we make on monitoring structure enable players to identify any fixed stage game action of the other player from frequencies of signals after sufficiently many observations and compute the other player's posterior belief about their own types.

In this setting, we show that reputations of both of the players disappear eventually if the commitment actions are not part of an equilibrium for the strategic types in the stage game. More precisely, if for both of the players, their best response to the best response of the opponent to their commitment strategy is not their commitment strategy, then in any Nash equilibrium of the incomplete-information game, the true types will be revealed (almost surely).

The techniques of the proofs are borrowed from Cripps, Mailath, and Samuelson (2004), who study games with imperfect public monitoring in which only one of the players has uncertainty about the types of the other player. The key condition we require on the stage games, i.e. the *one-sided binding moral hazard at the commitment action profile*, makes the long-run behavior of the reputation of the player who is subject to binding moral hazard (the principal) independent of the reputation of the player who is not (the agent); whereas, this condition necessitates the behavior of the reputation of the player who is *not* subject to binding moral hazard (the agent) depend on the behavior of the reputation of the player who is subject to binding moral hazard (the principal). Let player 2 (the principal) be the one who is subject to binding moral hazard at the commitment action profile. For instance, if player 2 were to play the commitment action all the time, there is no incentive for player 1 (the agent) to play something other than her commitment strategy since player 1's best response to the commitment action of player 2 is her commitment action. Thus, player 1's type won't be revealed unless player 2's type is revealed. In other words, there won't be any set of histories with positive probability on which player 1's type is revealed but not player 2's. After establishing that player 2's type is (almost) revealed in the long-run, we can show that player 1's true type should be revealed as well. Hence, the one-sided binding moral hazard condition allows us to break the analysis of the long-run behavior of the reputations of the two players into two stages:

1. We first show that the reputation of player 2 is not sustainable in the long-run: Suppose on the contrary that there is a positive probability set of events in which the type of player 2 is not revealed, i.e. if both types of player 2 are given strictly positive probability in the long-run. It means that the play of the strategic type of player 2 is not distinguishable from the play of the commitment type in the limit (given that the public signals are statistically informative about a player's actions). Hence, in the event in which the type is not revealed, the strategic and commitment type must be playing

the same way on average. The strategic type of player 1 is going to best reply to the commitment action of player 2, which is the same as her commitment action. Thus, player 1 must be playing the same strategy on average, too. Hence, the posterior beliefs about player 1's type will not change in these histories. For any (fixed) belief about player 1's type, since the commitment strategy of player 2 is not his best response to player 1's commitment action, he can gain by playing the best response to player 1's commitment action (which is different than his commitment action), knowing that the deviations from equilibrium play cannot be unambiguously detected by player 1 (because of imperfect monitoring). Then, the strategic type of player 2 has a trade off between the loss of reputation and the current payoff gain. The convergence of beliefs in the limit ensures that any signal has an arbitrarily small effect on players' beliefs. And thus, when reputations are public, player 2 eventually knows that player 1's beliefs have nearly converged and be sure that deviations from the commitment strategy have arbitrarily small effect on the payoffs from the continuation play (due to discounting). But then the strategic and the commitment type of player 2 are expected to play differently, which contradicts to the belief of player 1 about both types of player 1 playing the same strategy in the limit (on a positive set of histories). This establishes the fact that there cannot be positive set of events in which player 1's type is not revealed eventually.

Moreover, the disappearance of player 2's reputation is uniform across all Nash equilibria, i.e. there is some period T after which reputation of Player 2 of being the commitment type converges to zero (almost surely) in all Nash equilibria (according to the probability measure induced by the strategic type of player 2's play).

2. Given that player 2's type is revealed (almost surely) after some T in any Nash equilibria, the game "approaches" to the one where there is uncertainty over the types of player 1 only. We show that player 1's reputation for being the commitment type converges to zero, in the same fashion. The only problem could be that even though player 2's type is revealed, he could keep playing the commitment strategy on average. In this case, the strategic player 1 would give a best response which coincide with her commitment type and the type wouldn't be revealed. But this can't happen since in the events that type of player 1 is not revealed and the strategic and commitment type of player 1 play approximately the same way, so the strategic type of player 2 finds it optimal to play the best response against player 1's commitment action, which will decrease the reputation of player 2 even more after T. [13] So, player 1 expects to see a best reply to her commitment strategy from the strategic type of player 2 with a high probability. Thus, player 1 best responds to the strategic type of player 2 who gives a best reply to the commitment strategy of player 1 with high probability. However, as the strategic type player 1's best response to player 2's strategy is different than her commitment strategy, the strategic and the commitment type of player 1 are expected to

[13]More precisely, the histories for which player 2's reputation can be rebuilt would have measure 0.

play differently, which reveals her true type eventually.

Our results imply that the asymmetric information about the types of players does not interfere in the long-run equilibrium behavior. We show that continuation play in every Nash equilibrium of the incomplete-information game is converges to an equilibrium of the complete-information game.

The implication of our results for the games between a regulator and a regulatee, presented in Section 1.1, is that the reputation of being tough for the regulator disappears in the long-run (regardless of the long-run behavior of the regulatee's reputation of being virtuous) since the regulator is the player who is subject to binding moral hazard at the commitment profile. After his true type is almost known, the regulatee starts to take advantage of regulator's uncertainty over her type and regulatee's reputation of being virtuous disappears eventually as well. So, the regulatee waits for the revelation of the regulator's type, before revealing her type.

These results suggest that if there is a possibility that the regulator is replaced every period so that the uncertainty about the regulator's type renewed every period, then the regulatee is never fully convinced that the regulator is tough. In this situation, the regulator keeps playing *diligent* and will not have incentive to deviate since the regulatee is not convinced. [14] So, one way to sustain the reputations is to introduce the possibility that keeps the uncertainty over the type of regulator every period. However, this will create an equilibrium which is not efficient. That is why we believe that the replacements should be strategically scheduled.

We would like to point out that our model satisfy all the assumptions of Cripps, Mailath, and Samuelson (2004), and thus if there is uncertainty about the types of one player only (either of the players), their result hold and the true type of the informed player is revealed eventually. We are able to show that Cripps, Mailath, and Samuelson (2004)'s result is robust to introducing uncertainty over the type of the second player in this class of games. Also, we believe that the existence of a commitment type of player 1 (the player who is not subject to binding moral hazard) speeds up the revelation of the true type of player 2; whereas, existence of a commitment type of player 2 postpones the revelation of true type of player 1, compared to the corresponding one-sided incomplete-information games.

1.3 Related Literature

Most of the early literature on games with reputation concerns focus on settings in which a long-lived player faces a sequence of short-lived players, each of whom plays only once but observes the previous play and believes that the long-lived player might be committed to some exogenously specified strategy. In this environment, Fudenberg and Levine (1989) and Fudenberg and Levine (1992) provide a lower bound

[14] Mailath and Samuelson (2001) and Phelan (2006) provide models where the long-lived informed player's type is governed by a stochastic process that has long-run implications.

on the long-lived player's average payoff, namely the Stackelberg payoff, [15] given that she is sufficiently patient. Following the tradition of long-lived player gains from short-lived players' uncertainty over her types, Schmidt (1993) and Celentani, Fudenberg, Levine, and Pesendorfer (1996) show reputation effects arise in settings that involve two long-lived players and can be stronger than when the uninformed player is short-lived.

However, Cripps, Mailath, and Samuelson (2004) show that a long-lived player can maintain a permanent reputation for playing a commitment strategy in a game with imperfect monitoring only if that strategy plays an equilibrium of the corresponding complete-information stage game. They also prove that the continuation play in every Nash equilibrium of the incomplete-information game is Nash equilibrium of the complete-information game. Thus, the powerful results about the lower bounds on the long-lived informed player's average payoff are short-run reputation effects, where the long-lived informed player's payoff is calculated at the beginning of the game. Cripps, Mailath, and Samuelson (2007) extend their earlier result to games with two long-lived players where the uninformed long-lived player has private beliefs over the types of informed long-lived player, so that the reputation of the informed player is *private*. Also, an earlier result on the long-run properties of reputations is by Benabou and Laroque (1992), who study a game with a long-lived player who can be one of two types, honest or opportunist, and a continuum of myopic players in asset markets. They focus on the Markov perfect equilibrium of this game where the actions of the long-lived player is not observable. They show that the long-lived player reveals her type in any Markov perfect equilibrium. All these results are for games with imperfect public monitoring and uncertainty over the types of only one of the players. We show that this result extends to games with one-sided binding moral hazard at the commitment action profile.

The paper is organized as follows: Section 2 describes the model, Section 3 states the main results of the paper and Sections 4 and 5 are devoted to the proofs of Proposition 1 and 2 respectively, which lead to the proof of our main result Theorem 1.

2 Model

In this section, we first define the complete-information game, the game without uncertainty over the types of players (i.e. the game when both players are *strategic* types). Then we present the incomplete-information game by adding commitment types of players to the model.

[15] Stackelberg payoff is the payoff players receive in the stage game when they play their Stackelberg action (i.e. the action players would like to commit given that such a commitment induces a best response from the opponent player) and the opponent best responds to it. Stackelberg action is the action players would like to choose in an extensive-form game when they move the first.

2.1 Complete-information game

We define an infinitely repeated game with imperfect public monitoring. The stage game is a two-player finite strategic form game. Player 1 ("she") chooses an action $i \in I \equiv \{1, ..., \bar{I}\}$ and player 2 ("he") chooses an action $j \in J \equiv \{1, ..., \bar{J}\}$. The public signal y is drawn from the finite set Y. The probability that the public signal is realized under the action profile (i, j) is given by ρ_{ij}^y. The ex post (realized) stage game payoff to player 1 (resp., 2) from action i (resp., j) and signal y is given by $u_1(i, y)$ (resp. $u_2(j, y)$). The ex ante (expected) stage game payoffs are $\pi_1(i, j) = \sum_y u_1(i, y)\rho_{ij}^y$ and $\pi_2(i, j) = \sum_y u_2(j, y)\rho_{ij}^y$.

Both players are long-lived with discount factor $\delta_1 < 1$ for player 1 and $\delta_2 < 1$ for player 2. The set of histories is denoted by $h_t^f \equiv ((i_0, j_0, y_0), ..., (i_{t-1}, j_{t-1}, y_{t-1})) \in H_t^f \equiv (I \times J \times Y)^t$. Each player observes the realization of the public signal and his or her own past actions. Player 1's private history is denoted by $h_{1t} \equiv ((i_0, y_0), ..., (i_{t-1}, y_{t-1})) \in H_{1t} \equiv (I \times Y)^t$. Similarly, player 2's private history is denoted by $h_{2t} \equiv ((j_0, y_0), ..., (j_{t-1}, y_{t-1})) \in H_{2t} \equiv (J \times Y)^t$. And, the public history observed by both players is $h_t \equiv (y_0, ..., y_{t-1}) \in H_t \equiv Y^t$. The filtration on $(I \times J \times Y)^\infty$ induced by the private histories of player $m = 1, 2$ is denoted by $\{\mathcal{H}_{mt}\}_{t=0}^\infty$, while the filtration induced by the public histories is denoted by $\{\mathcal{H}_t\}_{t=0}^\infty$. Player 1's strategy, $\sigma \equiv \{\sigma_t\}_{t=0}^\infty$, is a sequence of maps $\sigma_t : H_{1t} \to \Delta(I)$. Similarly, Player 2's strategy, $\tau \equiv \{\tau_t\}_{t=0}^\infty$, is a sequence of maps $\tau_t : H_{2t} \to \Delta(J)$.

The payoffs in the infinitely repeated game are normalized discounted sum of stage game payoffs, $(1 - \delta_m) \sum_{s=0}^\infty \delta_m^s \pi_m(i_s, j_s)$ for player $m = 1, 2$. The average discounted payoffs in period t is denoted by $\pi_{mt} \equiv (1 - \delta_m) \sum_{s=t}^\infty \delta_m^{s-t} \pi_m(i_s, j_s)$.

We require some conditions on the monitoring technology. We assume that the public signals have full support (Assumption 1). So, every public signal is possible after any action profile and players can not infer the actions chosen by the other player perfectly after a signal.

Assumption 1 *(Full support) For all $(i, j) \in I \times J$ and $y \in Y$, $\rho_{ij}^y > 0$.*

Full support assumption prevents perfect inference of actions after any signal. We also assume "individual full rank" conditions, so that after sufficiently many observations, any fixed stage game action of either player can be identified from the frequencies of the signals (Assumptions 2 and 3).

Assumption 2 *(Individual 1 full rank) For all $j \in J$, the I columns in the matrix $(\rho_{ij}^y)_{y \in Y, i \in I}$ are linearly independent.*

Assumption 3 *(Individual 2 full rank) For all $i \in I$, the J columns in the matrix $(\rho_{ij}^y)_{y \in Y, j \in J}$ are linearly independent.*

Assumption 2 and 3 ensure that, for each player, the distribution of signals generated by any (possibly mixed) action is statistically distinguishable from any other for any given action of the other player. Note that these conditions require that $|Y| \geq \max\{|I|, |J|\}$.

A strategy profile (σ, τ) induces a probability distribution $P_{(\sigma,\tau)}$ over $H^f_\infty \equiv (I \times J \times Y)^\infty$. We denote the expectation with respect to this distribution by $E_{(\sigma,\tau)}$.

Definition 1 *A Nash equilibrium of the complete information game is a strategy profile (σ^*, τ^*) such that $E_{(\sigma^*,\tau^*)}[\pi_{10}] \geq E_{(\sigma',\tau^*)}[\pi_{10}]$ for all σ' and $E_{(\sigma^*,\tau^*)}[\pi_{20}] \geq E_{(\sigma^*,\tau')}[\pi_{20}]$ for all τ'.*

This definition implies that under the equilibrium strategy profile, player m's strategy maximizes continuation expected utility after any history that occurs with positive probability. Note that, with the full support assumption, all public histories occur with positive probability. Hence, any Nash equilibrium outcome is also the outcome of a perfect Bayesian equilibrium.

2.2 Incomplete-information game

We model uncertainty about players' preferences with Harsanyi (1967)'s notion of games with incomplete information by introducing a *commitment* type for each player. At time $t = -1$, before the game starts, nature selects a type for both players and tells each player her or his own type privately. With probability $1 - \mu_0 > 0$, player 1 is a "strategic" type, denoted by n, with the preferences described above and with probability $\mu_0 > 0$, she is a "commitment" type, denoted by c, who plays the action $s_1 \in \Delta(I)$ in each period regardless of history. Similarly, with probability $1 - \gamma_0 > 0$, player 2 is a "strategic" type, denoted by n, whose preferences are described above and with probability $\gamma_0 > 0$, he is a "commitment" type, denoted by c, who plays the action $s_2 \in \Delta(J)$ in each period independent of history. [16]

The stage games we consider are restricted. We denote the stage game commitment actions by (s_1, s_2) and define (s_1, s_2) to be *commitment profile*. Let $r_1 \equiv BR_1(s_2)$ and $r_2 \equiv BR_2(s_1)$ be the best responses of strategic type of player 1 and 2 against the commitment action of their opponent, respectively.

Definition 2 *Player $m = 1, 2$ is subject to binding moral hazard at the commitment profile (s_m, s_{-m}) if $r_m \neq s_m$.*

Since Assumption 1 ensures that deviations by players are not unambiguously detectable, if player m is subject to binding moral hazard at strategy profile (s_m, s_{-m}), then he has strict incentive to deviate from the profile (s_m, s_{-m}).

Definition 3 *A game has one-sided binding moral hazard at the commitment profile if there is only one player who is subject to binding moral hazard at the commitment strategy profile (s_1, s_2).*

[16] Instead of modelling the incomplete-information by behavioral types, we could have modeled the commitment types as agents whose payoffs are different from those of the strategic ones, as in Koren (1992). Then players would know their own payoffs and have uncertainty over the payoffs of the other player.

From thereafter, when we use the term one-sided binding moral hazard, it refers to the one at the commitment profile. An example of a stage game with one-sided binding moral hazard is given in Table 2.2. It is a specific example of the games we discussed in section 1.1 between a regulator and regulatee with numerical values for the parameters. The commitment profile is (T, D) (which corresponds to $Truthful$ and $Diligent$). This game has one-sided binding moral hazard, since only player 2 is subject to binding moral hazard at (T, D). However, for a prisoners' dilemma game, if the commitment profile is given by (C, C), both players are subject to binding moral hazard at the commitment profile.

The stage games we allow have one-sided binding moral hazard at the commitment profile. The importance of this condition will be discussed in the following sections. We make the following assumption on expected stage game payoffs.

Assumption 4 *The stage game satisfies the following:*

1. *The stage game has one-sided binding moral hazard at the commitment profile* (s_1, s_2).
2. *Each player has a unique best reply* r_m *to* s_{-m} *and* (s_1, r_2) *and* (r_1, s_2) *are not stage game Nash equilibria.*

Table 2: Regulatee-Regulator game: One-sided binding moral hazard at (T, D)

	L	D
T	2, 3	0, 2
U	3, 0	-1, 1

Table 3: Prisoners' dilemma: Two-sided binding moral hazard at (C, C)

	C	D
C	2, 2	-1, 3
D	3, -1	0, 0

Note that the first requirement already implies that either (s_1, r_2) or (r_1, s_2) is not Nash equilibrium, depending on who is subject to binding moral hazard at (s_1, s_2). For instance, if player 2 is the one who is subject to binding moral hazard, then $r_1 = s_1$ and (r_1, s_2) is not a Nash equilibrium. Moreover, the second condition necessitates s_1 to be a pure action, whereas the commitment action of player 2 could be a mixed-action.

Let $\hat{\sigma}$ denote the repeated game strategy of playing s_1 in each period independent of history and $\hat{\tau}$ denote the repeated game strategy of playing s_2 in each period independent of history. Since r_1 is unique best response to s_2, by Assumption 4, the best response of player 1 in the repeated game, i.e. $BR_1(\hat{\tau})$, is a singleton and prescribes playing r_1 in each period for every history. Similarly, the best response of player 2 against the commitment strategy of player 1, $BR_2(\hat{\sigma})$, is a singleton and prescribes playing r_2 in each period for every history. Since (s_1, r_2) and (r_1, s_2) are not stage game Nash equilibrium, $(\hat{\sigma}, BR_2(\hat{\sigma}))$ and $(BR_1(\hat{\tau}), \hat{\tau})$ are not Nash equilibrium of the complete-information infinitely repeated game. The unique stage game best responses guarantee that there are no multiple best responses to the commitment strategies in the infinitely repeated game. Since $(\hat{\sigma}, BR_2(\hat{\sigma}))$ and $(BR_1(\hat{\tau}), \hat{\tau})$ are not Nash equilibrium in the complete-information infinitely repeated game, each player has an incentive to deviate to a strategy other than the repeated game commitment strategy, given that opponent is best responding to the commitment strategy.

The first requirement of Assumption 4 implies that the commitment profile (s_1, s_2) is not a Nash equilibrium of the stage game, thus repetition of this commitment profile in every period independent of history can not be a Nash equilibrium of the complete-information infinitely repeated game (when the types are known and indeed strategic). So, at least one of the players has a profitable deviation from the repeated commitment profile. Since only one player is subject to binding moral hazard, for only one player, there is a profitable deviation in the complete-information repeated game. This condition requires that one of the players best response against the commitment action of the opponent player should be the same as her commitment action, i.e. $r_m = s_m$ for one of the players. Hence, for one player the best response in the complete-information infinitely repeated game is the same as her commitment strategy in the repeated game. For notational clarity, we let player 2 be the player who is subject to binding moral hazard at (s_1, s_2) and player 1 be the one whose best response to $\hat{\tau}$ in repeated game is the same as her commitment strategy $\hat{\sigma}$.

Example given by Table 2.2 illustrates the stage games we allow. Note that it is a numerical example of the games we discussed in section 1.1 between a regulator and regulatee. The stage game has a unique Nash equilibrium in mixed strategies. Let $\alpha_1 \in \Delta(I)$ be the strategy of player 1 (regulatee) and $\alpha_2 \in \Delta(J)$ be the strategy of player 2 (regulator). The unique Nash equilibrium of the stage game is $\alpha_1(T) = \frac{1}{2}$ and $\alpha_2(D) = \frac{1}{2}$, providing a payoff vector of (1, 1.5) to player 1 and player 2, respectively. The minmax value for player 1 is 0 (the action that minmaxes player 1 is D) and the minmax value for player 2 is 1 (the action that minmaxes player 2 is U). We introduce incomplete information about the types of both players by assuming there is a probability $\mu_0 > 0$ that player 1 is the Stackelberg type who plays her Stackelberg action [17] T in every period and there is a probability $\gamma_0 > 0$ that player 2 is the Stackelberg type who plays his Stackelberg action D in every period. [18] This game satisfies the conditions of Assumption 4.

[17] This is an action to which a player would like to commit given that such a commitment induces a best response from the opponent player, i.e. this is the action a player would like to choose in an extensive-form game when she/he moves the first.

[18] Note that the commitment types in the game between a regulatee and regulator, i.e. *virtuous* and *tough*, are indeed their

The first requirement of Assumption 4 is met since the unique best response for player 1 against D, the commitment action of player 2, is T and (T, D) is not a Nash equilibrium of the stage game; and for player 2, the unique best response against the commitment action of player 1 is L and (T, L) is not a Nash equilibrium of the stage game. It is easy to check that the commitment profile (T, D) is not Nash equilibrium of the stage game and only player 2 is subject to binding moral hazard at (T, D). Thus, player 1's best response to D is the same as her commitment action (Stackelberg action) T. Hence, only player 2 has an incentive to deviate from the commitment profile in the infinitely repeated complete-information game.

Let $K = \{c, n\}$ and $L = \{c, n\}$ be the type spaces for player 1 and player 2, respectively. The repeated game strategy for player 1, σ, is a sequence of maps $\sigma_t : H_{1t} \times K \to \Delta(I)$. For player 2, the repeated game strategy is denoted by τ is a sequence of maps $\tau_t : H_{2t} \times L \to \Delta(J)$. Let σ be denoted as $\sigma \equiv (\hat{\sigma}, \tilde{\sigma})$ where $\hat{\sigma}$ is the strategy of the c type of player 1 that prescribes playing s_1 in each period independent of history and $\tilde{\sigma}$ is the repeated game strategy of n type of player 1. Similarly, we denote the repeated game strategy of c type of player 2 by $\hat{\tau}$ which prescribes playing s_2 every period regardless of history and $\tilde{\tau}$ is the strategy of n type of player 2, and $\tau \equiv (\hat{\tau}, \tilde{\tau})$. A state of the world in the incomplete information game, ω, is a type for player 1, a type for player 2, and a sequence of actions and public signals. The set of states is $\Omega \equiv K \times L \times H^f_\infty$, where $H^f_\infty = (I \times J \times Y)^\infty$. The priors (μ_0, γ_0), the strategies $\sigma \equiv (\hat{\sigma}, \tilde{\sigma})$ of player 1, and the strategies $\tau \equiv (\hat{\tau}, \tilde{\tau})$ of player 2 jointly induce a probability measure $Q_{(\sigma,\tau;\mu,\gamma)}$ on $(\Omega, \mathcal{F}) \equiv (K \times L \times H^f_\infty, 2^K \otimes 2^L \otimes \mathcal{H}^f_\infty)$. The probability measure $Q_{(\sigma,\tau;\mu,\gamma)}$ describes how an uninformed observer of the game expects the play to evolve. We denote the expectation with respect to $Q_{(\sigma,\tau;\mu,\gamma)}$ by $E_{(\sigma,\tau)}$. Let $E_{(\sigma,\tau)}[.|\mathcal{H}_{1t}]$ and $E_{(\sigma,\tau)}[.|\mathcal{H}_{2t}]$ denote players expectations with respect to $Q_{(\sigma,\tau)}$ conditional on the filtration induced by the private histories, $\mathcal{H}_{1t}$ and $\mathcal{H}_{2t}$, respectively.

The strategy profile $(\hat{\sigma}, \tau)$ and $(\tilde{\sigma}, \tau)$, where $\tau = (\hat{\tau}, \tilde{\tau})$, induce probability measure $Q^{c\cdot}$ and $Q^{n\cdot}$, which describes how the play evolves when player 1 is the commitment and strategic type, respectively. The probability measure $Q^{k\cdot} \equiv Q_{(\sigma_k,\tau)}$, where σ_k is the strategy of the k type of player 1, describes how the game evolves if player 1 is of type k. The associated expectation is denoted by $E^{k\cdot} \equiv E_{(\sigma_k,\tau)}$. Similarly, the strategy profile $(\sigma, \hat{\tau})$ and $(\sigma, \tilde{\tau})$, where $\sigma = (\hat{\sigma}, \tilde{\sigma})$, induce probability measure $Q^{\cdot c}$ and $Q^{\cdot n}$, which describe how the play evolves when player 2 is the commitment and strategic type, respectively. So, the probability measure $Q^{\cdot l} \equiv Q_{(\sigma,\tau_l)}$, where τ_l is the strategy of the l type of player 2, describes how the game evolves if player 2 is of type l, and the associated expectation is $E^{\cdot l} \equiv E_{(\sigma,\tau_l)}$.

We will denote $Q_{(\sigma,\tau)}$, $E_{(\sigma,\tau)}$, $Q_{(\sigma_k,\tau)}$, $E_{(\sigma_k,\tau)}$, $Q_{(\sigma,\tau_l)}$ and $E_{(\sigma,\tau_l)}$ by Q, E, $Q^{k\cdot}$, $E^{k\cdot}$, $Q^{\cdot l}$ and $E^{\cdot l}$, respec-

Stackelberg types and the commitment profile $(Truthful, Diligent)$ are the Stackelberg actions of regulatee and regulator, respectively.

tively. Players' payoffs in the repeated game is then

$$
\begin{aligned}
E^{k\cdot}[\pi_{10}] &= E^{k\cdot}\big[(1-\delta_1)\sum_{t=0}^{\infty}\delta_1^t \pi_1(i_t, j_t)\big] \\
E^{\cdot l}[\pi_{20}] &= E^{\cdot l}\big[(1-\delta_2)\sum_{t=0}^{\infty}\delta_2^s \pi_2(i_t, j_t)\big]
\end{aligned}
$$

We analyze the game assuming that players are indeed "strategic."

Definition 4 *A Nash equilibrium of the incomplete information game is a strategy profile* $(\tilde{\sigma}, \tilde{\tau})$ *such that*

$$
\begin{aligned}
E^{n\cdot} \equiv E_{\tilde{\sigma},\tau}[\pi_{10}] \geq E_{\sigma',\tau}[\pi_{10}], \quad \forall \sigma' \\
E^{\cdot n} \equiv E_{\sigma,\tilde{\tau}}[\pi_{20}] \geq E_{\sigma,\tau'}[\pi_{20}], \quad \forall \tau'
\end{aligned}
$$

Player 1's posterior belief in period t about player 2's type is given by $\mathcal{H}_{1t}$ - measurable random variable

$$\gamma_t \equiv Q^{k\cdot}(c \mid \mathcal{H}_{1t}) : \Omega \to [0,1],$$

and player 2's posterior belief in period t about player 1's type is given by $\mathcal{H}_{2t}$ - measurable random variable

$$\mu_t \equiv Q^{\cdot l}(c \mid \mathcal{H}_{2t}) : \Omega \to [0,1].$$

Our main theorem will establish that the reputations for being the commitment types cannot be sustainable indefinitely, i.e. $\mu_t \to 0$ and $\gamma_t \to 0$ almost surely (with respect to the probability distribution induced by the strategies of the strategic type of the players). We should point out that players' beliefs about each other's type is private. This means players do not know the beliefs of the other player about their own types perfectly. We impose a condition on the public monitoring structure that rules out the dependence of beliefs on player's own past actions, and thus enables players to infer opponent's beliefs about their own types. We assume that the monitoring structure is such that the informativeness of the public signal about any player's action is independent of the other player's action (Assumption 5), and as a consequence, the reputations become *public*. Let Prob$(i \mid y, j, \alpha_1)$ be the posterior probability of "player 1 having chosen pure action i", given mixed α_1 and given that player 2 observed signal y after playing action j, and Prob$(j \mid y, i, \alpha_2)$ is the corresponding posterior probability of player 2's action.

Assumption 5 *(Independence) For any* $\alpha_1 \in \Delta(I)$ *and* $\alpha_2 \in \Delta(J)$*, and any signal* $y \in Y$*,*

$$
\begin{aligned}
\textit{Prob}(i \mid y, j, \alpha_1) &= \textit{Prob}(i \mid y, j', \alpha_1), \quad \textit{for all} \quad j, j' \\
\textit{Prob}(j \mid y, i, \alpha_2) &= \textit{Prob}(j \mid y, i', \alpha_2), \quad \textit{for all} \quad i, i'.
\end{aligned}
$$

Assumption 5 implies that for all $\alpha_1 \in \Delta(I)$ and $j, j' \in J$,

$$\frac{\text{Prob}(y \mid i, j)\alpha_1(i)}{\sum_{i\in I} \alpha_1(i)\text{Prob}(y \mid i, j)} = \frac{\text{Prob}(y \mid i, j')\alpha_1(i)}{\sum_{i\in I} \alpha_1(i)\text{Prob}(y \mid i, j')} \tag{1}$$

Similarly, for all $\alpha_2 \in \Delta(J)$ and $i, i' \in I$,

$$\frac{\text{Prob}(y \mid i, j)\alpha_2(j)}{\sum_{j\in J} \alpha_2(j)\text{Prob}(y \mid i, j)} = \frac{\text{Prob}(y \mid i', j)\alpha_2(j)}{\sum_{j\in J} \alpha_2(j)\text{Prob}(y \mid i', j)} \tag{2}$$

Assumption 5 ensures that public signals allow players to infer the other player's beliefs about their type since the information that public signal provides about the player's action is independent of the opponent's behavior. In other words, this monitoring structure allows players to calculate opponent's inference about their reputation without knowing the opponent's action, thus reputations of both players becomes *public* and beliefs are common knowledge.

Assumption 5 holds if each player has individual specific signals, i.e. the public signal y is such that $y = (y_1, y_2) \in Y = Y_1 \times Y_2$ where y_1 is a signal of player 1's action and y_2 is a signal of player 2's, with

$$\rho_{ij}^{y} = \rho_i^{y_1}\rho_j^{y_2}, \quad \forall i, j, y.$$

Such monitoring structure is called *product structure*. In the motivating example game presented in section 1.1, each player had separate public signal, distribution of each depended only on player's own action, independent of the other players action. [19,20]

For games with product structure, every sequential equilibrium payoff in the complete-information infinitely repeated game (equilibrium when private histories are used as beliefs) is also a public perfect equilibrium payoff. [21] We can show that Assumption 5 implies every sequential equilibrium payoff in the complete-information game (equilibrium when private histories are used as beliefs) is also a public perfect equilibrium payoff, in the same spirit of Fudenberg and Levine (1994). Thus, we believe there is no loss of generality if we restrict attention to public strategies. A public strategy $\sigma \equiv \{\sigma_t\}_{t=0}^{\infty}$ for player 1 is a sequence of maps $\sigma_t : H_t \to \Delta(I)$ and that for player 2, $\tau \equiv \{\tau_t\}_{t=0}^{\infty}$, is a sequence of maps $\tau_t : H_t \to \Delta(J)$. The strategy profile (σ, τ) induces a probability distribution $Q_{(\sigma,\tau)}$ over $H_{\infty}^{f} \equiv (I \times J \times Y)^{\infty}$. Let $E_{(\sigma,\tau)}[. \mid \mathcal{H}_t]$ denote players expectations with respect to $Q_{(\sigma,\tau)}$ conditional on the filtration induced by the public history, $\mathcal{H}_t$.

[19] For regulatee, the public signals are correct and incorrect message; whereas, for regulator, the public signals are there has been an audit or no audit; probability of which depends on their own actions only.

[20] In games with product structure, pure-action profiles satisfy pairwise identifiability condition of Fudenberg, Levine, and Maskin (1994) Folk theorem result for games with imperfect public monitoring. Hence, any feasible and Pareto efficient payoff dominating a Nash equilibrium payoff of the stage game can be attained as an equilibrium payoff of the repeated (complete-information) game if players are sufficiently patient.

[21] See Mailath and Samuelson (2006) (p.330) and Fudenberg and Levine (1994) (Theorem 5.2) for further discussion.

Note that due to Assumption 5 (i.e. under a monitoring structure such as the product structure), γ_t and μ_t can be viewed as $\mathcal{H}_t$ - measurable random variable $Q(c \mid \mathcal{H}_t)$ on Ω. This property enables both players to calculate the opponent player's beliefs about themselves. So, in period t, strategic type of player 1 is maximizing $E_{\tilde{\sigma},\tau}[\pi_{1t} \mid \mathcal{H}_t]$, and a strategic player 2 is maximizing $E_{\sigma,\tilde{\tau}}[\pi_{2t} \mid \mathcal{H}_t]$, that depend on the information sets generated by public histories.

At any Nash equilibrium of the incomplete information game, γ_t is a bounded martingale with respect to the measure Q and filtration $\{\mathcal{H}_{1t}\}_t$ (and also with respect to filtration $\{\mathcal{H}_t\}_t$ by Assumption 5). Therefore, γ_t converges Q-almost surely (and also $Q^{\cdot n}$ - almost surely and Q^{nn} - almost surely, since $Q^{\cdot n}$ and Q^{nn} are absolutely continuous with respect to Q) to a random variable γ_∞ on Ω. Similarly, at any Nash equilibrium of the incomplete information game, μ_t is a bounded martingale with respect to the measure Q and filtration $\{\mathcal{H}_{2t}\}_t$ (and also with respect to filtration $\{\mathcal{H}_t\}_t$), and thus converges Q-almost surely (and hence $Q^{n\cdot}$ - almost surely and Q^{nn} - almost surely, since $Q^{n\cdot}$ and Q^{nn} are absolutely continuous with respect to Q) to a random variable μ_∞ on Ω. [22]

Our main theorem states that the reputations disappears eventually (cannot be sustained indefinitely), that is $\mu_\infty = 0$ and $\gamma_\infty = 0$.

3 Main Results

Our main result is that neither player can sustain a reputation for playing a strategy that is not part of a Nash equilibrium of the complete-information stage game, for games with one-sided binding moral hazard at the commitment profile, under imperfect public monitoring.

Theorem 1 *Suppose Assumptions 1-5 hold. In any Nash equilibrium of the incomplete-information game, reputations of players cannot be sustained indefinitely:*

$$\mu_t \to 0, \quad Q^{n\cdot} - \textit{almost surely},$$
$$\gamma_t \to 0, \quad Q^{\cdot n} - \textit{almost surely}.$$

Moreover, the convergence is uniform. [23]

We prove Theorem 1 with the help of two propositions. The first proposition argues that reputation of player 2, who is subject to binding moral hazard at the commitment profile, disappears uniformly in any Nash equilibrium of the incomplete-information game, i.e. $\gamma_t \to 0, \quad Q^{\cdot n}-$almost surely and convergence is uniform across all Nash equilibria. The second proposition argues that if player 2's reputation disappears uniformly, then player 1's reputation disappears uniformly in any Nash equilibrium as well.

[22] The proof is presented in the Appendix, Lemma 7.

[23] Note that $\mu_t \to 0 \quad Q^{nn}$ – almost surely, and $\gamma_t \to 0 \quad Q^{nn}$ – almost surely, since Q^{nn} is absolutely continuous with respect to $Q^{n\cdot}$ and $Q^{\cdot n}$.

One-sided binding moral hazard condition implies that the player whose reputation disappears in the long-run, independent of the asymptotic behavior of the other player's reputation, is the one who is subject to binding moral hazard at the commitment profile. After establishing that player's type is (almost) revealed in the long-run, we can show that other player's true type should be revealed as well. Hence, the one-sided binding moral hazard condition allows us to break the analysis of the long-run behavior of the reputations of the two players into two stages.

The proof of Theorem 1 is immediate by Proposition 1 and 2. Sections 4 and 5 are devoted to the proofs of Proposition 1 and 2.

Proposition 1 *Suppose monitoring technology satisfies Assumptions 1, 3 and 5, and the stage game satisfies one-sided binding moral hazard at the commitment profile. Then, in any Nash equilibrium of the incomplete-information game, reputation of player 2, who is subject to binding moral hazard at the commitment profile, cannot be sustained indefinitely:*

$$\gamma_t \rightarrow 0, \quad Q^{\cdot n} - \textit{almost surely}.$$

Moreover, the disappearance of player 2's reputation is uniform. That is for all $\varepsilon > 0$, there exists T, such that for all Nash equilibria $(\tilde{\sigma}, \tilde{\tau})$ of the incomplete-information game,

$$Q^{\cdot n}_{\sigma,\tilde{\tau}}(\gamma_t(\sigma, \tilde{\tau}) < \varepsilon, \forall t > T) > 1 - \varepsilon,$$

where $Q^{\cdot n}_{\sigma=(\hat{\sigma},\tilde{\sigma}),\tilde{\tau}}$ is the probability measure induced on Ω by $(\sigma, \tilde{\tau})$ and the strategic type of player 2 and $\gamma_t(\sigma, \tilde{\tau})$ is the associated reputation of player 2.

We should point out that the disappearance of player 2's reputation is independent of the asymptotic behavior of player 1's reputation, whether player 1's reputation of being commitment type disappears or not. So, player 2's reputation of being the commitment type converges to zero $Q^{\cdot n}$-almost surely, whether the uncertainty over the types of player 1 is resolved or not. This leads to the following corollary.

Corollary 1 *Suppose monitoring technology satisfies Assumptions 1, 3 and 5, and the stage game satisfies one-sided moral hazard at the commitment profile. Suppose there exists a Nash equilibrium $(\tilde{\sigma}, \tilde{\tau})$ that induces a set of histories on which the reputation of player 2 (who is subject to binding moral hazard) does not disappear, i.e. there exists $A \in \Omega$ such that $\gamma_t(\omega) \rightarrow \gamma_\infty > \eta$ for some $\eta > 0$, given that the reputation of player 1 is sustained on these histories, $\mu_t(\omega) \rightarrow \mu_\infty > \epsilon$ for some $\epsilon > 0$ for all $\omega \in A$. Then $Q^{\cdot n}(A) = 0$, where $Q^{\cdot n}$ denotes $Q_{(\sigma,\tilde{\tau})}$.*

Suppose that there exists a Nash equilibrium profile that induces histories (with positive measure) on which the reputation of player 1 does not disappear. This means that player 2 believes that player 1 plays the commitment action s_1 on average in the long-run, which induces him to deviate from (s_1, s_2) eventually. Also, the other immediate corollary of Proposition 1 is that there is no histories with positive measure where player 1's reputation disappears in the long-run, but player 2's not.

Corollary 2 *Suppose monitoring technology satisfies Assumptions 1, 3 and 5, and the stage game satisfies one-sided moral hazard at the commitment profile. Suppose there exists a Nash equilibrium $(\tilde{\sigma}, \tilde{\tau})$ that induces a set of histories on which player 1's reputation disappears but not player 2's (who is subject to binding moral hazard), i.e. there exists $A \in \Omega$ such that $\mu_t(\omega) \to 0$, but $\gamma_t(\omega) \to \gamma_\infty > \eta$ for some $\eta > 0$ and for all $\omega \in A$. Then $Q^{\cdot n}(A) = 0$, where $Q^{\cdot n}$ denotes $Q_{(\sigma, \tilde{\tau})}$.*

Having established that player 2 reveals his true type eventually regardless of the asymptotic behavior of player 1's reputation, the game can be considered to be the one with "almost" one-sided incomplete-information where the uncertainty is about the types of player 1 only. The next proposition gives the sufficient conditions for the disappearance of player 1's reputation.

Proposition 2 *Suppose the monitoring technology satisfies Assumptions 1, 2 and 5, and player 2's reputation γ_t converges uniformly to zero $Q^{\cdot n}$-almost surely in any equilibrium. Suppose also (s_1, r_2) is not a Nash equilibrium of the stage game. Then, in any Nash equilibrium of the incomplete-information game, player 1's reputation disappears eventually:*

$$\mu_t \to 0, \quad Q^{n\cdot} - \textit{almost surely}$$

where the convergence is uniform, i.e. for all $\varepsilon > 0$, there exists T, such that for all Nash equilibria $(\tilde{\sigma}, \tilde{\tau})$,

$$Q^{n\cdot}_{\tilde{\sigma},\tau}(\mu_t(\tilde{\sigma}, \tau) < \varepsilon, \forall t > T) > 1 - \varepsilon,$$

where $Q^{n\cdot}_{\tilde{\sigma},\tau=(\hat{\tau},\tilde{\tau})}$ is the probability measure induced on Ω by $(\tilde{\sigma}, \tau)$ and the strategic type of player 1 and $\mu_t(\tilde{\sigma}, \tau)$ is the associated reputation of player 1.

Proposition 2 implies that the sustainability of player 1's reputation depends on that of player 2's reputation. If player 2's reputation disappears, player 1's reputation disappears eventually, given that (s_1, r_2) is not a Nash equilibrium of the stage game (and thus $(\hat{\sigma}, BR_2(\hat{\sigma}))$ is not a Nash equilibrium of the repeated complete-information game).

Corollary 3 *Suppose the monitoring technology satisfies Assumptions 1, 2 and 5. Suppose there exists a Nash equilibrium that induces a set of histories $A \in \Omega$ with $Q(A) > 0$ on which the reputation of player 2 does not disappear, i.e. $\gamma_t(\omega) \to \gamma_\infty > \eta$ for some $\eta > 0$ and for all $\omega \in A$. Suppose also the stage game best reply of player 1 against s_2 is the same as her commitment action, i.e. $r_1 = s_1$. Then, $\mu_t(\omega) \to \mu_\infty > \epsilon$ for some $\epsilon > 0$ and $Q^{n\cdot}$-almost surely in A.*

Corollary 3 says that if the uncertainty over player 2's type persists, the uncertainty over player 1's type persists as well, since then player 1 expects to see s_2 on average in the long-run and gives a best response to it ($r_1 = s_1$). However, by Proposition 1, the uncertainty over player 1's type can persist only if either (s_1, s_2) a Nash equilibrium of the stage game or there is a mechanism that replenish the uncertainty over player 2's type. One such mechanism can be introducing a possibility for replacing the type of player 2 every period. With such a mechanism, player 2 need to mimic the commitment type always to convince player 1, since player 1 is never fully convinced because of the replacement possibility. Hence, player 2's type will not be revealed. As player 2's type is not revealed, player 1's type will not be revealed as well.

The implication of these results for the regulatee-regulator game presented in Section 1.1 is that the reputation of being tough for the regulator disappears in the long-run since regulator is the player who is subject to binding moral hazard at the commitment profile (by Proposition 1). After his true type is almost known, the regulatee starts to take advantage of regulator's uncertainty over her type and regulatee's reputation of being virtuous disappears eventually as well (by Proposition 2). Furthermore, the set of histories where the regulatee's true type is almost known, but regulator's true type is not revealed has measure zero. We can interpret this as the regulatee waits for the revelation of the true type of the regulator to exploit regulator's uncertainty over her types. One way to make both reputations sustainable is to introduce the possibility that the type of the regulator changes every period (with some probability). Then, the regulator can never convince the regulatee perfectly that he is tough, so he needs to be diligent every period. Both reputations can be made permanent this way. However, from the welfare point of view, this is inefficient. In order to get the efficient stage game outcome played (frequently), we need a mechanism that allows for some deterioration for the reputation of the regulator up to a lower bound, so that the regulatee will not start exploiting this deterioration. By strategically scheduled replacement periods, there are periods of $(Truthful, Lazy)$ which Pareto dominates $(Truhful, Diligent)$. There is an optimal schedule for replacement periods that depends on the parameter values of the payoffs, as well as the values for the prior beliefs and discount factors.

3.1 One-sided binding moral hazard at the commitment profile

The condition of one-sided binding moral hazard at the commitment profile (s_1, s_2) is crucial for our results. If none of the players has an incentive to deviate at (s_1, s_2) in the complete-information stage game, it means (s_1, s_2) is a Nash equilibrium of the stage game and thus repetition of (s_1, s_2) every period independent of history is a Nash equilibrium of the repeated complete-information game. Hence, the reputations can be sustained in that situation as the commitment strategies $(\hat{\sigma}, \hat{\tau})$ is a Nash equilibrium for strategic types. If, on the other hand, the stage game has two-sided binding moral hazard at the commitment profile, i.e. both players have an incentive to deviate at (s_1, s_2) in the stage game, then the results are not clear. We can construct a Nash equilibrium where the reputations do not necessarily disappear.

Consider the following stage game and suppose that there are $cooperative$ types for both players and both strategic players have an incentive to deviate from the commitment action profile (C, C). Let $\mu_1 = \mu_2 = \mu_0$ be the prior belief that the players are $cooperative$ type.

Table 4: Two-sided binding moral hazard at (C, C)

	C	D_1	D_2
C	4, 4	0, 0	0, 5
D_1	5, 0	1, 3	0, 0
D_2	0, 0	0, 0	1, 3

Let $Y_1 = \{H, L\}$ and $Y_2 = \{h, l\}$ be the public signal spaces for player 1 and 2, $y \in Y_1 \times Y_2$, and the probability distributions over signals are given as below:

$$prob(H|C) = p, \quad prob(H|D_1) = r, \quad prob(H|D_2) = q$$
$$prob(h|C) = p, \quad prob(h|D_1) = q, \quad prob(h|D_2) = r$$

where $p > 1/2 > q > r$.

The strategy profile represented by the following automaton is a public perfect equilibrium for the complete-information infinitely repeated game. The states are $W = \{w_{CC}, w_1, w_2\}$ and the initial state is w_{CC}, and players choose the following action profiles corresponding to each state:

$$\begin{aligned} f(w_{CC}) &= CC, \\ f(w_1) &= D_1 D_1, \\ f(w_2) &= D_2 D_2. \end{aligned}$$

and the transition is

$$t(y) = \begin{cases} w_{CC} & \text{if} \quad y = Hh \quad \text{or} \quad Ll, \\ w_1 & \text{if} \quad y = Lh, \\ w_2 & \text{if} \quad y = Hl, \end{cases}$$

Note that w_1 is the punishment state for player 1 and w_2 is the punishment state for player 2. The equilibrium path can be described by an ergodic Markov chain on the state space with stationary distribution putting more weight on w_{CC}. This strategy profile is also an equilibrium profile for low enough μ_0, and the types will not be revealed because of frequent play of CC.

However, one can also construct Nash equilibria on which the uncertainty over the types of both players is going to be revealed eventually. In fact, if one of the player's type is revealed, the other's type is going to be revealed as well.

We want to point out that Cripps, Mailath, and Samuelson (2004) show that if there is uncertainty over the types of only one of the players, the true type (which is strategic) of this player will be revealed eventually in any Nash equilibria. Adding uncertainty over the types of the other players may change this result in some Nash equilibria.

3.2 Equilibrium behavior

After establishing that the true types will be (almost) known in the long-run and the information structure of the game approaches to that of the complete-information game, we expect to see such a convergence result holds for the equilibrium behavior. We show that any Nash equilibrium of the incomplete-information game converges to a public perfect equilibrium of the complete-information game, following the definitions and the methods provided by Cripps, Mailath, and Samuelson (2004) for one long-lived and a sequence of short-lived players.

We will introduce some notation before stating the result about long-run equilibrium behavior. Let $t' = 0, 1, ...$ denote the time periods of the continuation play of the game that starts at some period t. A pure public (continuation game) strategy ς_1 for player 1 is a sequence of maps $\varsigma_{1t'} : H_{t'} \rightarrow I$ for $t' = 0, 1, ...$ Similarly, a pure public strategy for player 2 in the continuation game t' is ς_2, a sequence of maps $\varsigma_{2t'} : H_{t'} \rightarrow J$ for $t' = 0, 1, ...$ Let $S_1 = I^{\bigcup_{t'=0}^{\infty} Y^{t'}}$ and $S_2 = J^{\bigcup_{t'=0}^{\infty} Y^{t'}}$ be the set of pure strategies for player 1 and player 2, respectively. [24] Note that S_1 and S_2 include the pure strategies in the original game as well. Player m's payoff is given by, [25]

$$U_m(\varsigma_1, \varsigma_2) = E_{(\varsigma_1, \varsigma_2)}\left[(1 - \delta_m) \sum_{t'=0}^{\infty} \delta_m^{t'} \pi_m(i_{t'}, j_{t'})\right]$$

We now define the mixed strategies (of the repeated continuation game) as the probability distributions over the set of pure strategies, i.e. ϑ_m be probability measures on $(S_m, \mathcal{S}_m)$. [26] Let Θ_m denote the set of all probability measures ϑ_m, $m = 1, 2$. Note that Θ_m is sequentially compact with respect to the product topology. Since players' payoffs are discounted, the utility function $U_m : \Theta_1 \times \Theta_2 \rightarrow \Re$ is continuous for each $m = 1, 2$ with this topology. We refer Fudenberg and Tirole (1991) for a detailed discussion and proof of the continuity of the utility function (due to discounting $\delta_1, \delta_2 < 1$).

[24] The sets S_1 and S_2 are countable products of finite sets I and J. Define σ-algebras for each set that are generated by cylinder sets and denote by $\mathcal{S}_m$, $m = 1, 2$. $(S_m, \mathcal{S}_m)$ is equipped with the product topology.

[25] Even though the strategies are pure, the payoffs are random because of imperfect public monitoring.

[26] Note that by Kuhn's theorem, we can replace mixed strategies by behavior strategies for games with perfect recall.

A sequence of measures ϑ_1^n converges to $\hat{\vartheta}_1$ if the following holds: For every $T \geq 0$,

$$\vartheta_1^n|_{I^{Y^T}} \rightarrow \hat{\vartheta}_1|_{I^{Y^T}}$$

and, similarly, ϑ_2^n converges to $\hat{\vartheta}_2$ if for every $T \geq 0$,

$$\vartheta_2^n|_{J^{Y^T}} \rightarrow \hat{\vartheta}_2|_{J^{Y^T}}$$

Pick an equilibrium $(\tilde{\sigma}, \tilde{\tau})$ of the incomplete-information game and a public history h_t. These strategies specifies behavior strategies in the continuation game, $\tilde{\sigma}_{h_t}$ and $\tilde{\tau}_{h_t}$, which are realization equivalent to the mixed strategies $\tilde{\vartheta}_1^{h_t}$ and $\tilde{\vartheta}_2^{h_t}$ (for the continuation game), by Kuhn's Theorem. The following theorem states that the limit of every convergent subsequence of $(\tilde{\vartheta}_1^{h_t}, \tilde{\vartheta}_2^{h_t})$ is a Nash equilibrium of the complete-information game. Similar convergence result about the asymptotic equilibrium behavior can be found Cripps, Mailath, and Samuelson (2004) for one long-lived and a sequence of short-lived players. We present the appropriate modifications of their proofs for our model in the Appendix.

Theorem 2 *Suppose Assumptions 1-5 are satisfied. For any Nash equilibrium of the incomplete-information game and for almost all sequences of public histories $\{h_t\}_t$ (with respect to measure Q^{nn}), the limit of every convergent subsequence of continuation equilibrium profiles $(\tilde{\vartheta}_1^{h_t}, \tilde{\vartheta}_2^{h_t})$ is a public perfect equilibrium of the complete-information game (game with strategic types of players).*

Theorem 2 implies that the equilibrium of the incomplete-information regulatee-regulator game converges to a public perfect equilibrium of the complete-information regulatee-regulator game, and thus the the equilibrium payoff in the complete-information game converges to an equilibrium payoff of the complete-information game. We would like to point out that by our assumptions on the monitoring technology (Assumptions 2, 3 and 5), the imperfect public monitoring Folk theorem holds.[27] Hence, any feasible and individually rational payoff vector of stage game in the complete-information game can be attained as an equilibrium payoff of the complete-information repeated game, if players are sufficiently patient. However, Theorem 2 neither constrain the possible set of equilibrium payoffs, nor answers if any particular Nash equilibrium strategy profile (or payoff) of the complete-information game can be achieved as a limit of a Nash equilibrium of the incomplete-information game.

However, we are able to provide a lower bound for the equilibrium payoff vector if we allow the discount factor of player 2 (who is subject to one-sided binding moral hazard) to vary, i.e. for any fixed prior beliefs (μ_0, γ_0), if player 2's discount factor is above some $\bar{\delta}_2(\mu_0, \gamma_0) \rightarrow 1$, the lower bound on his equilibrium payoff converges to $\pi_2(s_1, s_2) = \pi_2(r_1, s_2)$, whereas player 1's payoff converges to $\pi_1(r_1, s_2)$.

[27] We refer the reader to Fudenberg, Levine, and Maskin (1994).

Theorem 3 *Suppose Assumption 3 holds. Let the game satisfies one-sided binding moral-hazard, with player 2 being subject to binding moral hazard at the commitment action $s_2 \in \Delta(J)$. Then there exists a constant $k(\mu_0, \gamma_0)$, independent of player 2's discount factor, such that in any Nash equilibrium of the incomplete-information game, strategic type of player 2's payoff in any is at least*

$$\delta_2^k \min_{\alpha_1 \in BR_1(s_2)} \pi_2(\alpha_1, s_2) + (1 - \delta_2^k) \min_{(i,j)} \pi_2(i, j)$$

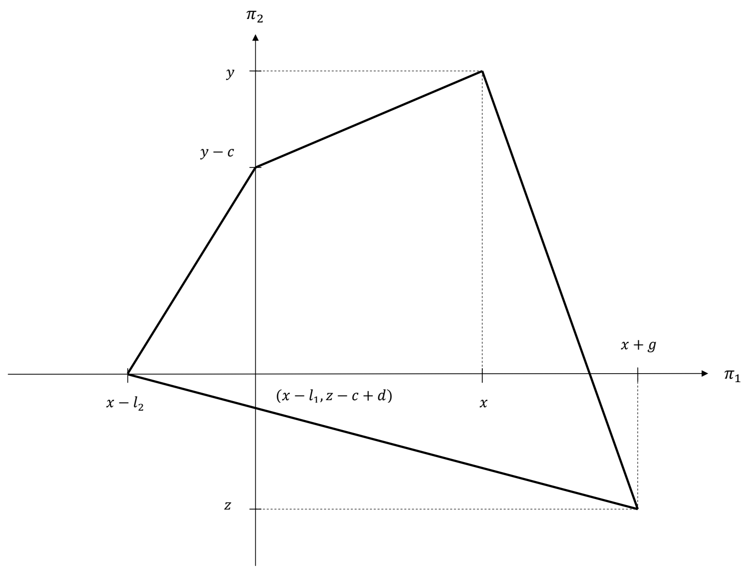
Suppose, in addition, s_2 minmaxes player 1. Then player 1's temporary payoff is close to her minmax payoff, $\pi_1(r_1, s_2)$.

Proof. To be completed. ■

Theorem 3 says that if player 2 is sufficiently patient, he can get a payoff of at least $\min_{\alpha_1 \in BR_1(s_2)} \pi_2(\alpha_1, s_2)$, i.e. for any $\epsilon > 0$, there exists a $\delta_2^*(\mu_0, \gamma_0) \in (0, 1)$ such that for all $\delta_2 \in (\delta_2^*, 1)$, strategic type of player 2's payoff is at least $\min_{\alpha_1 \in BR_1(s_2)} \pi_2(\alpha_1, s_2) - \epsilon$. However, we should point out that the lower bound on equilibrium payoffs given by Theorem 1 is a "temporary" lower bound on player 2's equilibrium payoff, in the sense that for any fixed discount factor δ_2 of player 2, γ_t will eventually be small enough so that $\delta_2 < \delta_2^*(\mu_t, \gamma_t)$. Hence, at the beginning of the game, there will be a period of (s_1, s_2) where players collect their payoffs $\pi_1(s_1, s_2)$ and $\pi_2(s_1, s_2)$), but eventually player 2 cannot resist to deviate which will reveal his true type.

We would like to make a couple observations. First, note that player 2's commitment action is allowed to be a mixed-action. In fact, if player 2 can commit to a mixed stage game action, he can get a higher lower bound for his equilibrium payoffs. Consider the regulator-regulatee game given in Table 2.2. Suppose the commitment action of the regulator is $s_2 = D$. Then $\min_{\alpha_1 \in BR_1(s_2)} \pi_2(\alpha_1, s_2) = \pi_2(r_1, s_2) = 2$. However, instead, if the regulator's commitment action randomizes between L and D, giving a weight on D a little more than $\frac{1}{2}$, then the best reply of the regulatee is still T and the regulator can achieve $\frac{5}{2}$. The regulatee's payoff, on the other hand, is bound by her minmax payoff (if $s_2 = D$).

The following graph is the set of feasible payoff vectors of the regulatee-regulator game, depending on the parameters of the stage game, where the origin $(x - l_1, z - c + d)$ is the minmax payoff of each player. Suppose the commitment action of the regulator is D. Then, the regulator collects a payoff of $y - c$ and the regulatee collects her minmax payoff $x - l_1$ at the beginning of the game. If the discount factor of the regulator is very high, the equilibrium payoffs they get in the beginning of the game is inefficient.



4 Player 2's reputation disappears uniformly

We first show that either the true type of player 2 is revealed or player 1's expectation of the strategy played by the strategic type of player 2 is in the limit the same as the strategy played by the commitment type, given that the public signals are statistically informative about player 2's behavior (Lemma 1). In other words, if player 1 is not eventually convinced that player 2 is strategic, she must be convinced that player 2 is mimicking the commitment type on average in the long-run. We show this with the extensively used merging of beliefs argument modified for imperfect public monitoring games.[28] Then, we show that if there is a set of histories with positive measure in which player 2's reputation does not disappear, in those histories, player 1 must be convinced that player 2 will play the commitment strategy in the continuation play; and moreover, reputations being public implies that player 2 knows about player 1's beliefs about his behavior (Lemma 2). Hence, player 2 believes that the strategic type of player 1 should be best responding to the commitment strategy of player 2 (Lemma 3), which also coincide with the strategy of commitment type of player 1 (since player 1 is *not* subject to binding moral hazard at the commitment type, and thus $r_1 = s_1$). Then player 2 has an incentive to deviate from his commitment strategy (as player 2 is subject to binding moral hazard at the commitment profile), knowing that these deviations will not be detected due to imperfect monitoring [29] and player 1's beliefs have nearly converged, and thus the effect of deviations on player 1's beliefs will be arbitrarily small. However, the long-run effect of many such deviations,

[28]See Sorin (1999) and Cripps, Mailath, and Samuelson (2004).

[29]Player 2's incentive to deviate from the commitment strategy is stronger in two-sided incomplete-information game compared to one-sided incomplete information game where there is uncertainty only over the types of player 2.

which generate different distributions over the public signals (Assumption 3), reveals that player 2 plays a strategy different than the commitment strategy. This provides the ground for the desired contradiction to the hypothesis of having a positive measure set of histories in which player 1 is convinced that player 2 is playing the commitment strategy on average in the long run.

4.1 Player 1's posterior beliefs about player 2

The following Lemma and Corollary establish that either player 1's expectation of the strategy played by the strategic type of player 2 is in the limit the same as the strategy played by the commitment type of player 2, or player 1's posterior probability that player 2 is the commitment type converges to zero (given that player 2 is indeed strategic). The proof is as the one provided by Cripps, Mailath, and Samuelson (2004).

Lemma 1 *Suppose Assumptions 1, 3 and 5 are satisfied. In any Nash equilibrium of the incomplete-information game,*

$$\lim_{t\to\infty} \gamma_t(1-\gamma_t)\|\hat{\tau}_t - E^{\cdot n}[\tilde{\tau}_t|\mathcal{H}_t]\| = 0, \quad Q-a.s. \tag{3}$$

Note that since $\hat{\tau}_t$ is a simple strategy, it can be replaced by s_2.

Proof. Let $\gamma_{t+1}(h_t; i_t, y_t)$ denote player 1's belief in period $t+1$ after playing i_t and observing public signal y_t in period t, given public history h_t. By Bayes' rule,

$$\gamma_{t+1}(h_t; i_t, y_t) = \frac{\gamma_t \text{prob}[y_t \mid h_t, i_t, c]}{\gamma_t \text{prob}[y_t \mid h_t, i_t, c] + (1-\gamma_t)\text{prob}[y_t \mid h_t, i_t, n]}$$

Since the probability of observing the signal y_t from the commitment type of player 2 is $\sum_{j\in J} s_2^j \rho_{i_t j}^{y_t}$ and from the strategic type is $E^{\cdot n}[\sum_{j\in J} \tilde{\tau}_t^j \rho_{i_t j}^{y_t} \mid h_t]$, we can rewrite the above expression as,

$$\begin{aligned}
\gamma_{t+1}(h_t; i_t, y_t) &= \frac{\gamma_t \sum_{j\in J} s_2^j \rho_{i_t j}^{y_t}}{\gamma_t \sum_{j\in J} s_2^j \rho_{i_t j}^{y_t} + (1-\gamma_t)E^{\cdot n}[\sum_{j\in J} \tilde{\tau}_t^j \rho_{i_t j}^{y_t} \mid h_t]} \\
&= \frac{\gamma_t \sum_{j\in J} s_2^j \rho_{i_t j}^{y_t}}{\sum_{j\in J} \rho_{i_t j}^{y_t}\Big(\gamma_t s_2^j + (1-\gamma_t)E^{\cdot n}[\tilde{\tau}_t^j \mid h_t]\Big)}
\end{aligned}$$

The difference between $\gamma_{t+1}(h_t; i_t, y_t)$ and $\gamma_t(h_t)$ gives,

$$\begin{aligned}
|\gamma_{t+1}(h_t;i_t,y_t) - \gamma_t(h_t)| &= \left|\frac{\gamma_t \sum_{j\in J} s_2^j \rho_{i_t j}^{y_t}}{\sum_{j\in J} \rho_{i_t j}^{y_t}(\gamma_t s_2^j + (1-\gamma_t)E^{\cdot n}[\tilde{\tau}_t^j \mid h_t])} - \gamma_t\right| \\
&= \left|\frac{\gamma_t(1-\gamma_t)\sum_{j\in J} s_2^j \rho_{i_t j}^{y_t} - \gamma_t(1-\gamma_t)\sum_{j\in J}\rho_{i_t j}^{y_t} E^{\cdot n}[\tilde{\tau}_t^j \mid h_t])}{\sum_{j\in J} \rho_{i_t j}^{y_t}(\gamma_t s_2^j + (1-\gamma_t)E^{\cdot n}[\tilde{\tau}_t^j \mid h_t])}\right| \\
&= \frac{\gamma_t(1-\gamma_t)\left|\sum_{j\in J}\rho_{i_t j}^{y_t}(s_2^j - E^{\cdot n}[\tilde{\tau}_t^j \mid h_t])\right|}{\sum_{j\in J} \rho_{i_t j}^{y_t}(\gamma_t s_2^j + (1-\gamma_t)E^{\cdot n}[\tilde{\tau}_t^j \mid h_t])}
\end{aligned}$$

Note that the denominator $\sum_{j\in J} \rho_{i_t j}^{y_t}(\gamma_t s_2^j + (1-\gamma_t)E^{\cdot n}[\tilde{\tau}_t^j \mid h_t]) < \max_{j\in J}\rho_{i_t j}^{y_t} < 1$. Thus,

$$|\gamma_{t+1}(h_t;i_t,y_t) - \gamma_t(h_t)| \geq \gamma_t(1-\gamma_t)\left|\sum_{j\in J}\rho_{i_t j}^{y_t}(s_2^j - E^{\cdot n}[\tilde{\tau}_t^j \mid h_t])\right|$$

which implies

$$\max_{y\in Y}|\gamma_{t+1}(h_t;i_t,y_t) - \gamma_t(h_t)| \geq \gamma_t(1-\gamma_t)\left|\sum_{j\in J}\rho_{i_t j}^{y_t}(s_2^j - E^{\cdot n}[\tilde{\tau}_t^j \mid h_t])\right|$$

Since the stochastic process γ_t is a martingale on $[0,1]$ (with respect to Q and filtration $\{\mathcal{H}_t\}_t$ by Lemma 7) and bounded martingales converge almost surely, thus $|\gamma_{t+1} - \gamma_t| \to 0$ Q-almost surely. This implies, for any $y \in Y$,

$$\gamma_t(1-\gamma_t)\left|\sum_{j\in J}\rho_{i_t j}^{y_t}(s_2^j - E^{\cdot n}[\tilde{\tau}_t^j \mid h_t])\right| \to 0, \qquad Q - \text{a.s.} \tag{4}$$

Since (4) holds for all y, it can be restated as

$$\gamma_t(1-\gamma_t)\left\|\Pi_{i_t}(s_2 - E^{\cdot n}[\tilde{\tau}_t \mid h_t])\right\| \to 0, \qquad Q - \text{a.s.}$$

where Π_{i_t} is a $|Y| \times |J|$ matrix that contains the values for $\rho_{i_t j}^{y_t}$. Since for all i_t and $y \in Y$, $\rho_{i_t j}^{y_t} > 0$ by Assumption 1 and J columns are linearly independent by Assumption 3, the unique solution to $\Pi_{i_t} x = 0$ is $x = 0$ and there exists a strictly positive constant $k = \inf_{i\in I, x\neq 0}\frac{\|\Pi_i x\|}{\|x\|}$. Thus, $\|\Pi_i x\| \geq k\|x\|$, which implies

$$\gamma_t(1-\gamma_t)\left\|\Pi_{i_t}(s_2 - E^{\cdot n}[\tilde{\tau}_t \mid h_t])\right\| \geq \gamma_t(1-\gamma_t)k\left\|(s_2 - E^{\cdot n}[\tilde{\tau}_t \mid h_t])\right\| \to 0, \qquad Q - \text{a.s.}$$

This implies (3). ■

Note that Lemma 1 holds also $Q^{n\cdot}$ almost surely, since γ_t is also a bounded martingale with respect to $Q^{n\cdot}$, which is the probability measure that describes how the game evolves from the perspective of the strategic type of player 1. Note that any statement that holds Q almost surely, also holds $Q^{n\cdot}$ almost surely.

The immediate implication of Lemma 1 is Corollary 4:

Corollary 4 *At any Nash equilibrium of the incomplete-information game satisfying Assumptions 1, 3 and 5,*

$$\lim_{t\to\infty} \gamma_t \left\| s_2 - E^{\cdot n}[\tilde{\tau}_t \mid h_t] \right\| \to 0, \qquad Q^{\cdot n} - a.s.$$

Corollary 4 says that if player 2 is indeed strategic and the game evolves according to the play of strategic type of player 2, either his reputation for being the commitment type disappears, i.e. $\gamma_t \to 0$, $Q^{\cdot n}$ - almost surely, or he is expected to play the commitment action in the limit.

Proof. We first show that $\frac{\gamma_t}{(1-\gamma_t)}$ is a $Q^{\cdot n}$-martingale (with respect to filtration $\{\mathcal{H}_{1t}\}_t$ and $\{\mathcal{H}_t\}$ due to Assumption 5). For all h_{t+1}, i_t and for all i,

$$\begin{aligned}
E^{\cdot n}[\frac{\gamma_{t+1}}{(1-\gamma_{t+1})}|\mathcal{H}_{t+1}] &= \sum_{y\in Y} \text{prob}[y_t|h_t,n].\frac{\gamma_{t+1}(h_t,i_t,y_t)}{1-\gamma_{t+1}(h_t,i_t,y_t)} \\
&= \sum_{y\in Y} E^{\cdot n}[\sum_{j\in J} \tilde{\tau}_t^j \rho_{ij}^{y_t} \mid h_t] \frac{\gamma_t \sum_{j\in J} s_2^j \rho_{i_t j}^{y_t}}{(1-\gamma_t) E^{\cdot n}[\sum_{j\in J} \tilde{\tau}_t^j \rho_{i_t j}^{y_t} \mid h_t]} \\
&= \sum_{y\in Y} \frac{\gamma_t \sum_{j\in J} s_2^j \rho_{ij}^{y_t}}{(1-\gamma_t)} \\
&= \frac{\gamma_t \sum_{j\in J} s_2^j \sum_{y\in Y} \rho_{ij}^{y_t}}{(1-\gamma_t)} \\
&= \frac{\gamma_t}{(1-\gamma_t)}
\end{aligned}$$

The third equation is due to Assumption 5 and the last step is by Assumption 1. Thus, for all t,

$$E^{\cdot n}[\frac{\gamma_t}{(1-\gamma_t)}] = \frac{\gamma_0}{(1-\gamma_0)} \tag{5}$$

Note that γ_t converges to some random variable Q - a.s. implies γ_t converges $Q^{\cdot n}$ - a.s. (since $Q^{\cdot n}$ is absolutely continuous with respect to Q). Since $\frac{\gamma_0}{(1-\gamma_0)}$ is finite, we must have $\lim_{t\to\infty} \gamma_t < 1 \quad Q^{\cdot n}$ - a.s. Suppose on the contrary, there is a set $D \in \Omega$ with $Q^{\cdot n}(D) > 0$ such that $\gamma_t(\omega) \to 1$ for all $\omega \in D$. Then, $\frac{\gamma_t}{(1-\gamma_t)} \to \infty$ on D, which contradicts to (5). ■

We want to point out that $\lim_{t\to\infty} \gamma_t \| s_2 - E^{\cdot n}[\tilde{\tau}_t \mid h_t] \| \to 0$, also $Q^{nn} -$ a.s.

4.2 Player 2's beliefs about player 1's beliefs

We have shown that if player 1 does not eventually learn that player 2 is strategic (when player 2 is strategic and the histories are induced by the play of strategic player 2), then player 1 must think that strategic type of player 2's strategy should be close to that of commitment type of player 2's. Now, we want to show that player 2 will know that (since reputations are public by Assumption 5) player 1 eventually expects to see the commitment action of player 2 in the continuation game on these histories.

Lemma 2 *Suppose Assumptions 1, 3 and 5 hold and suppose there exists $A \in \Omega$ such that $Q^{\cdot n}(A) > 0$ and $\gamma_\infty(\omega) > 0$ for all $\omega \in A$, i.e. there exists a set of events with strictly positive measure in which reputation of player 2 does not necessarily disappear. Then, there exists $\eta > 0$ and $F \subset A$ with $Q^{\cdot n}(F) > 0$ (and $Q(F) > 0$) such that, for any $\xi > 0$, there exist T for which,*

$$\gamma_t > \eta, \quad \forall t \geq T,$$

$$E\left[\sup_{s \geq t} \|s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s]\| \middle| \mathcal{H}_t\right] < \xi, \quad \forall t \geq T \tag{6}$$

for all $\omega \in F$; and for all $\psi > 0$

$$Q\left(\sup_{s \geq t} \|s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s]\| < \psi \mid \mathcal{H}_t\right) \to 1 \tag{7}$$

where the convergence is uniform on F. [30]

Proof. First observe that on set A, $0 < \lim_{t\to\infty} \gamma_t(\omega) < 1$ by Corollary 4. Since $Q^{\cdot n}(A) > 0$ and $\gamma_\infty(\omega) > 0$ for all $\omega \in A$, there exist sufficiently small $\nu > 0$ and $\eta > 0$ such that $Q^{\cdot n}(D) > 2\nu$, where $D := \{\omega \in A : 2\eta < \lim_{t\to\infty} \gamma_t(\omega) < 1\}$. Note that D has positive measure under Q, i.e. there exists ν such that $Q(D) > 2\nu$ (since $Q^{\cdot n}$ is absolutely continuous with respect to Q.) Then, by Lemma 1, $\|s_2 - E^{\cdot n}[\tilde{\tau}_t \mid \mathcal{H}_t]\|$ converge Q - almost surely to zero on D. [31] So, the random variables $Z_t := \sup_{s\geq t} \|s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s]\|$ also converge Q - almost surely (also $Q^{\cdot n}$ - almost surely) to zero on D. Thus, on D, by an extension of Hart (1985) Lemma 4.24, given in Mailath and Samuelson (2006) [32]

$$E[Z_t|\mathcal{H}_t] \to 0, \quad Q - \text{a. s.} \quad \text{and} \quad Q^{\cdot n} - \text{a. s.}$$

and also $E^{\cdot n}[Z_t|\mathcal{H}_t] \to 0, \quad Q^{\cdot n} - \text{a. s.}$

[30] The same statements of the Lemma hold for $E^{\cdot n}$ and $Q^{\cdot n}$. In subsequent sections, both versions are going to be used.

[31] Note that $\|s_2 - E^{\cdot n}[\tilde{\tau}_t \mid \mathcal{H}_t]\|$ also converge $Q^{\cdot n}$ and $Q^{\cdot c}$ - almost surely to zero on D.

[32] This lemma states that if $\{X_n\}_{n=1}^\infty$ is a bounded sequence of real random variables on some $(\Omega, \mathcal{F}, P)$, converging 0 as $n \to \infty$ and $\{\mathcal{F}_n\}_{n=1}^\infty$ is a nondecreasing sequence of σ - fields, then $E[X_n \mid \mathcal{F}_n] \to 0$ P-a.s.

Egorov's Theorem (Chung (1974)) [33] then implies that there exists an $F \subset D$ such that $Q^{\cdot n}(F) > \nu$ (note that $Q(F) > 0$) on which the convergence of γ_t and $E[Z_t|\mathcal{H}_t]$ (and $E^{\cdot n}[Z_t|\mathcal{H}_t]$) is uniform. The uniform convergence of $E[Z_t|\mathcal{H}_t]$ on F implies that, for any $\xi > 0$, there exist a T such that on F, for all $t > T$, $\gamma_t > \eta$ and

$$E[Z_t|\mathcal{H}_t] = E\left[\sup_{s \geq t} \|s_2 - E[\tilde{\tau}_s \mid \mathcal{H}_s]\| \middle| \mathcal{H}_t\right] < \xi \tag{8}$$

In order to show (7), fix $\psi > 0$. Then, for all $\xi' > 0$ such that $\xi = \xi'\psi$, (8) holds. Hence,

$$E[Z_t|\mathcal{H}_t] = E[Z_t|Z_t < \psi, \mathcal{H}_t].Q(Z_t < \psi \mid \mathcal{H}_t) + E[Z_t|Z_t \geq \psi, \mathcal{H}_t].Q(Z_t \geq \psi \mid \mathcal{H}_t) < \xi'\psi.$$

Since the first part of the sum is greater and equal to 0 and $E[Z_t|Z_t \geq \psi, \mathcal{H}_t] \geq \psi$, we have,

$$Q(Z_t \geq \psi \mid \mathcal{H}_t) < \xi',$$

or $Q(Z_t < \psi \mid \mathcal{H}_t) > 1 - \xi'$ for all $t > T$ on F. This implies (7) and completes the proof. [34] ■

4.3 Player 1's best response to player 2

If player 1 were to be short-lived, as long as she thinks that she is facing a commitment strategy, she gives the myopic best reply to the commitment strategy of the opponent, which is s_1. This may not be true if player 1 is long-lived. She may have an incentive to play something other than the best response to the commitment action of player 2. But, note that in this case, as long as player 1 discounts, any losses from not playing a current best response should be recovered within a finite period of time. However, if player 1 is convinced that the commitment action will be played not only now, but also in the future, there will be no opportunity to accumulate subsequent gains, and hence she might as well play the stage-game best response.

The next lemma, which follows from Lemma 4 of Cripps, Mailath, and Samuelson (2004), explores this intuition. It shows that if the commitment type and strategic type of player 2 play sufficiently similar from some time on, strategic type of player 1 will be best responding to the commitment type's strategy for arbitrarily many periods.

[33] Egorov's Theorem states that if $\{X_n\}$ converges on the set C, then for any $\epsilon > 0$, there exists $C_0 \subset C$ with measure $\mathcal{P}(C \setminus C_0) < \epsilon$ such that X_n converges uniformly in C_0.

[34] The other way to show this: $Q(Z_t \geq \psi \mid \mathcal{H}_t) \leq \frac{E[Z_t|\mathcal{H}_t]}{\psi} < \frac{\xi}{\psi}$ by Chebyshev-Markov inequality since Z_t has a finite mean and $Z_t \geq 0$. Since $\psi > 0$ and $\xi = \xi'\psi$, we get $Q(Z_t \geq \psi \mid \mathcal{H}_t) < \xi'$ for all $\xi' > 0$.

Lemma 3 *Suppose $\hat{\tau}$ be a simple pure public strategy and $BR_1(\hat{\tau})$ is the set of best replies of strategic type of player 1 to $\hat{\tau}$.*[35] *Let $(\tilde{\sigma}, \tilde{\tau})$ be Nash equilibrium strategies in the incomplete-information game. If $\tilde{\sigma}$ is a pure strategy,* [36] *then for all $T > 0$, there exists $\psi > 0$ such that if the strategic player 1 observes a (public) history h_t so that*

$$Q\left(\sup_{s \geq t} \|s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s]\| < \psi \mid h_t\right) > 1 - \psi \tag{9}$$

then for $\hat{\sigma} \in BR_1(\hat{\tau})$, the continuation strategy of $\tilde{\sigma}$ after the history h_t agrees with $\hat{\sigma}$ for the next T periods.

Proof. Fix $T > 0$ and a (public) history h'_t. Let $\hat{\tau}(h_s) = s_2$ denote the continuation play of committed player 2 after the public history h_s, where h'_t is the initial segment of h_s.

Since player 1 is discounting, there exist $T' \geq T$ and $\epsilon > 0$ such that if for $s = t, ..., t + T'$ and for all h_s with initial segment h'_t

$$\|s_2 - E^{\cdot n}[\tilde{\tau}_s \mid h_s]\| < \epsilon, \tag{10}$$

is satisfied, then the continuation strategy of $\tilde{\sigma}$ after the history h'_t agrees with $\hat{\sigma} \in BR_1(\hat{\tau})$, for the next T periods.

Now, we want to show (10) holds for $s = t, ..., t + T'$ and for all h_s with initial segment h'_t. Suppose not, i.e. there exist h_s, for some $s = t, ..., t + T'$ such that

$$\|s_2 - E^{\cdot n}[\tilde{\tau}_s \mid h_s]\| \geq \epsilon,$$

For a contradiction, define $\bar{\rho} \equiv \min_{y,i,j} \rho^y_{ij}$ and $\psi = \frac{1}{2}\min\{\epsilon, \bar{\rho}^{T'}\}$. Since player 1 is playing a pure strategy, the probability of the continuation history h_s, conditional on the history h'_t, is at least $\bar{\rho}^{T'}$. Thus,

$$Q\left(\|(s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s])\| \geq \epsilon \mid h'_t\right) \geq \bar{\rho}^{T'},$$

Since $\psi < \epsilon$, we get

$$Q\left(\sup_{s \geq t} \|(s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s])\| \geq \psi \mid h'_t\right) \geq \bar{\rho}^{T'},$$

contradicting (9), since $\bar{\rho}^{T'} > \psi$. ■

[35] Remember that $\hat{\tau}$ assigns s_2, which is a pure action, in each period independent of history and the repeated strategy best response $\hat{\sigma} \equiv BR_1(\hat{\tau})$ is a singleton, which assigns r_1 in every period after any history.

[36] We don't need $\tilde{\sigma}$ to be pure. Instead, we could assume that there exists $k > 0$ such that for all h_t, if $\tilde{\sigma}^i(h_t) > 0$, then $\tilde{\sigma}^i(h_t) > k$.

4.4 Proof of disappearance of player 2's reputation

We wave the complications created by the two-sided incomplete information by using the fact that only player 2 is subject to binding moral hazard at the commitment profile. On a subset of states F in Lemma 2, the strategic type of player 1 believes that she should be playing a best response to the commitment strategy of player 2, which also coincides with the strategy of commitment type of player 1. So, the strategic type of player 2, knowing what player 1 thinks about his future behavior, will best respond to both the strategic type and the commitment type of player 1's strategy with high probability. Since, player 2's best response to player 1's strategy is different than his commitment strategy, the strategic and the commitment type of player 2 are expected to play differently, which will provide the desired contradiction to $\gamma_t \nrightarrow 0$ on F.

So, what we need to show is that player 1 eventually assigns high probability to player 2 believing with high probability that player 1 believes player 2's strategy is very close to commitment strategy (when the game evolves according to the play of the strategic type of player 2). And thus player 1 believes that the strategic and the commitment type of player 2 act differently, since he believes that the strategic player 1 is giving a best response, which coincides with the play of the commitment type of player 1. Thus player 1's second order beliefs about the future behavior of player 2 contradicts with her first order beliefs, leading to a contradicting to $\gamma_\infty > 0$ on F.

More specifically, we are going to show that player 1 assigns a probability $1-\zeta$ to player 2 believing with probability at least $1-\eta$ that player 1 thinks player 2's strategy is within ξ of the commitment strategy when the probability measure over the histories are induced by the play of the strategic type of player 2, [37] i.e.

$$Q^{\cdot n}\left(Q^{\cdot n}\left(\sup_{s\geq t}\|s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s]\| < \xi \mid \mathcal{H}_t\right) > 1-\eta \mid \mathcal{H}_t\right) > 1-\zeta$$

We pick ξ, η and ζ such that we arrive to the contradiction we are after. We choose ξ and ζ such that $\zeta < 1$ and $\xi < \min\{\psi, 1-\zeta\}$. First, we make a couple of observations that provide the basis for the contradiction.

Since the commitment strategy of player 2, $\hat{\tau}$ (i.e. playing always s_2), is a simple pure strategy, by Lemma 3, for any $T' > 0$, there exists $\psi > 0$ such that if player 1 observes a public history h_t so that

$$Q\left(\sup_{s\geq t}\|s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s]\| < \psi \mid h_t\right) > 1-\psi, \tag{11}$$

then the continuation strategy of $\tilde{\sigma}$ after the history h_t agrees with $\hat{\sigma} \in BR_1(\hat{\tau})$ for the next T' periods,

[37] In other words, player 1 assigns at most probability ζ to the commitment strategy of player 2. This is because player 2 believes with high probability player 1 thinks he is going to act like a commitment type and best respond to it. Since player 2 is subject to binding moral hazard at the commitment profile, player 1 eventually expects player 2 playing the commitment strategy with no more than ζ probability. But, this contradicts her initial belief of player 2's strategy being ξ close to the commitment strategy if we choose ξ and ζ properly.

where $\hat{\sigma} = \{r_1\}_{t=0}^{\infty}$.

Now, suppose for a contradiction, that there is a set of states A with $Q^{\cdot n}(A) > 0$ and $\gamma_\infty(\omega) > 0$ for all $\omega \in A$ (note that $\gamma_\infty(\omega) < 1$ on A by Corollary 4). Then, by Lemma 2, there is a set $F \subset A$ with $Q^{\cdot n}(F) > 0$ such that for any $\xi > 0$, there exists a T such that for any $t > T$ and $\omega \in F$,

$$Q\left(\sup_{s \geq t} \|s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s]\| < \xi | \mathcal{H}_t\right) \to 1 \tag{12}$$

Hence, there exists a subset $G \in F$ with $Q^{\cdot n}(G) > 0$ such that on G,

$$\|s_2 - E^{\cdot n}[\tilde{\tau}_t \mid \mathcal{H}_t]\| < \xi \quad Q - \text{a.s.}$$

Note that (13) implies, for any $\xi > 0$ and any $t > T$, on G,

$$\|s_2 - E^{\cdot n}[\tilde{\tau}_t \mid \mathcal{H}_t]\| < \xi \quad Q^{\cdot n} - \text{a.s.} \tag{13}$$

Also, by following the same argument in Lemma 2 (through extension of Hart's lemma given in footnote 32), we conclude that for some η and ζ, on G,

$$Q^{\cdot n}\left(\sup_{s \geq t} \|s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s]\| < \xi | \mathcal{H}_t\right) > 1 - \eta\zeta. \tag{14}$$

This shows that with a high probability $(1 - \eta\zeta)$, player 2 believes that player 1 assigns player 2's strategy to be ξ close to the commitment strategy for any $t > T$.

Define

$$g_t := Q^{\cdot n}\left(\sup_{s \geq t} \|(s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s]\| < \xi | \mathcal{H}_t\right)$$

$$\kappa_t := Q^{\cdot n}(g_t > 1 - \eta | \mathcal{H}_t)$$

We want to show $\kappa_t > 1 - \zeta$. [38] Since, $E^{\cdot n}[g_t \mid \mathcal{H}_t] > 1 - \eta\zeta$ by condition (14), and

$$\begin{aligned} E^{\cdot n}[g_t \mid \mathcal{H}_t] &= E^{\cdot n}[g_t \mid g_t \leq 1 - \eta, \mathcal{H}_t](1 - \kappa_t) + E^{\cdot n}[g_t \mid g_t > 1 - \eta, \mathcal{H}_t]\kappa_t \\ &\leq (1 - \eta)(1 - \kappa_t) + \kappa_t \end{aligned}$$

we get,

$$1 - \eta\zeta < (1 - \eta)(1 - \kappa_t) + \kappa_t$$

[38] Note that reputations are public and both players use $\mathcal{H}_t$ as their information sets.

which implies $\kappa_t > 1 - \zeta$ on F. So,

$$Q^{\cdot n}\left(Q^{\cdot n}\left(\sup_{s \geq t} \|(s_2 - E^{\cdot n}[\tilde{\tau}_s \mid \mathcal{H}_s]\| < \xi \mid \mathcal{H}_t\right) > 1 - \eta \mid \mathcal{H}_t\right) > 1 - \zeta$$

This says that player 1 assigns a probability of at least $1 - \zeta$ (after observing histories generated by the play of strategic type of player 2) to strategic type player 2 believing with probability at least $1 - \eta$ that player 1 believes player 2's strategy is within ξ of the commitment strategy and thus is going to give a best response to the commitment strategy of player 2 for at least next T' periods if player 1 is strategic type and play the commitment strategy if player 1 is the commitment type, both of which are the same strategy.

At time $t > T$, player 2 believes that with probability μ_t, player 1 is the commitment type who plays s_1 every period, and with probability $1 - \mu_t$, he is the strategic type who believes that player 2's strategy is within ξ of the commitment strategy from then on and hence plays the best reply to it (since we have picked $\xi < \psi$) thereafter. That is why, after $t > T$, both types of player 1 is expected to play r_1 thereafter, there won't be any revision of posterior about player 1's type, and hence $\mu_{t>T} = \mu_T$.

Since s_2 is not a best response to r_1 (the myopic best reply of strategic player 1 to s_2 and also the strategy of the commitment type of player 1) for strategic type of player 2, there exists $\eta_\mu > 0$ such that for any repeated game strategy of the strategic type of player 1 that attaches probability at least $1 - \eta_\mu$ to $\hat{\sigma}$ (to always playing r_1), s_2 is suboptimal for the strategic type of player 2 (by the upper-hemicontinuity of the best response correspondence) in period 0. Note that since player 2 believes that player 1 is commitment type who plays $s_1 = r_1$ with probability μ, η depends on μ. Let $\bar{\eta} \equiv \sup_{\mu \in (0,1)} \eta_\mu$ such that s_2 is suboptimal for the strategic player 2 in period 0 if strategic type of player 1 attaches $1 - \bar{\eta}$ to $\hat{\sigma}$, regardless of player 2's belief about player 1's type. Define $\bar{\rho} := \min_{y,i,j} \rho_{ij}^y$ (> 0 by Assumption 1). So, if player 2 assigns probability at least $1 - \bar{\rho}\bar{\eta} \equiv 1 - \eta$ to $\hat{\sigma}$, then he assigns at least probability $1 - \bar{\eta}$ to $\hat{\sigma}$ after any deviation that leaves the probability of $\hat{\sigma}$ (conditional on any signal) unchanged. Note that s_2 is suboptimal for any belief μ, in particular $\mu_{t>T} = \mu_T$.

Since we have picked ξ such that $\xi < \psi$, and condition (11) holds for all $t > T$, strategic type of player 1 chooses to play r_1, the unique best response to the commitment action thereafter, whenever he believes that player 2's strategy is within ξ of the commitment strategy. Hence, in any period $t > T$, player 1 assigns a probability of at least $1 - \zeta$ to player 2 believing that player 1's subsequent play is r_1 thereafter with at least probability $1 - \eta$. Thus, player 1 assigns probability at least $1 - \zeta$ to player 2's play in period t being a best response to $\hat{\sigma}$. Since s_2 is pure, it specifies an action $\hat{j}$ with probability 1. However, player 1 must believe that that action is played with no more than ζ probability in period t. But since, $1 - \zeta > \xi$, this contradicts (13). Player 1's second order beliefs about strategic player 2's behavior (after observing the relevant game has been evolving and histories have been generated by the play of strategic player 2) contradicts with her first order beliefs. This completes the proof of the first part of Proposition 1, i.e. $\gamma_t \to 0 \quad Q^{\cdot n}$ - almost surely, which implies $\gamma_t \to 0 \quad Q^{nn}$ - almost surely.

4.5 Uniform disappearance of player 2's reputation

Uniform convergence of $\gamma_t \to 0$, $Q^{\cdot n}$-almost surely means that there exists some period T after which reputation converges to zero across all Nash equilibria. Suppose, on the contrary, there is a Nash equilibrium for each T after which reputation of player 2 survives. Then the sequence of these Nash equilibria where the reputation lasts beyond T converges to a limiting Nash equilibrium with a sustainable reputation, which contradicts to disappearance of reputation result for any Nash equilibria.

The uniform disappearance of player 2's reputation, can be proved as the proof of Theorem 3 of Cripps, Mailath, and Samuelson (2007). We present the proof here for the sake of completeness. We need to show that for all $\varepsilon > 0$, there exists T, such that for all Nash equilibria $(\tilde{\sigma}, \tilde{\tau})$ of the incomplete-information game,

$$Q^{\cdot n}_{\sigma,\tilde{\tau}}(\gamma_t(\sigma,\tilde{\tau}) < \varepsilon, \forall t > T) > 1 - \varepsilon,$$

where $Q^{\cdot n}_{\sigma=(\hat{\sigma},\tilde{\sigma}),\tilde{\tau}}$ is the probability measure induced on Ω by $(\sigma, \tilde{\tau})$ and the strategic type of player 2 and $\gamma_t(\sigma, \tilde{\tau})$ is the associated reputation of player 2.

Proof. Suppose for a contradiction, there exists $\varepsilon > 0$ such that for all T, there is a Nash equilibrium profile α_T such that

$$Q^{\cdot n}_{\alpha_T}(\gamma_t(\alpha_T) < \varepsilon, \quad \forall t > T) \leq 1 - \varepsilon,$$

where $Q^{\cdot n}_{\alpha_T}$ is the measure induced by the strategic type of player 2 and Nash equilibrium profile α_T, and $\gamma_t(\alpha_T)$ is the posterior about player 2's type under α_T.

Since the space of strategy profiles is compact in the product topology, there is a convergent subsequence $\{\alpha_{T_k}\}$ with limit α^*. Relabel this sequence so that $\alpha_k \to \alpha^*$ and

$$Q^{\cdot n}_k(\gamma^k_t < \varepsilon, \quad \forall t > k) \leq 1 - \varepsilon, \quad \text{or}$$

$$Q^{\cdot n}_k(\gamma^k_t \geq \varepsilon, \quad \text{for some} \quad t > k) \geq \varepsilon.$$

Since each α_k is a Nash equilibrium, $\gamma^k_t \to 0$, $Q^{\cdot n}_k$ - almost surely. So, there exists $K_k > k$ such that

$$Q^{\cdot n}_k(\gamma^k_t < \varepsilon, \quad \forall t \geq K_k) \geq 1 - \varepsilon/2.$$

So, for all k, we have

$$Q^{\cdot n}_k(\gamma^k_t \geq \varepsilon, \quad \text{for some } t, k < t < K_k) \geq \frac{\varepsilon}{2}.$$

Let χ_k denote the stopping time

$$\chi_k = \min\{t > k : \gamma_t^k \geq \varepsilon\}$$

and Z_t^k be the associated stopped process,

$$p_t^k = \begin{cases} \gamma_t^k & \text{if} \quad t < \chi_k, \\ \varepsilon & \text{if} \quad t \geq \chi_k. \end{cases}$$

Note that p_t^k is a supermartingale under $Q_k^{\cdot n}$ and for $t < k$, $p_t^k = \gamma_t^k$. So, for all k and $t \geq K_k$,

$$\begin{aligned} E_k^{\cdot n} p_t^k &\geq \varepsilon Q_k^{\cdot n}(\chi_k \leq t) \\ &\geq \frac{\varepsilon^2}{2}. \end{aligned}$$

On the other hand, since α^* is a Nash equilibrium, $\gamma_t^* \to 0$, $Q_{\alpha^*}^{\cdot n}$- almost surely. So, there exists s such that

$$Q_{\alpha^*}^{\cdot n}(\gamma_s^* < \varepsilon^2/12) > 1 - \varepsilon^2/12.$$

Then,

$$E_{\alpha^*}^{\cdot n} \gamma_s^* \leq \frac{\varepsilon^2}{12}\left(1 - \frac{\varepsilon^2}{12}\right) + \frac{\varepsilon^2}{12} < \frac{\varepsilon^2}{6}$$

Since $\alpha_k \to \alpha^*$ in the product topology, there exists a $k' > s$ such that for all $k \geq k'$,

$$E_k^{\cdot n} \gamma_s^k < \frac{\varepsilon^2}{3}.$$

But since $k' > s$, $p_s^k = \gamma_s^k$ for $k \geq k'$ and so for any $t \geq K_k$,

$$\frac{\varepsilon^2}{3} > E_k^{\cdot n} \gamma_s^k = E_k^{\cdot n} p_s^k \geq E_k^{\cdot n} p_t^k \geq \frac{\varepsilon^2}{2},$$

which is a contradiction. ■

5 Player 1's reputation disappears uniformly

In this section, we provide the proof for Proposition 2. Suppose that player 2's reputation disappears uniformly in any Nash equilibrium of the two-sided incomplete information game, i.e. for all $\varepsilon > 0$, there exists T_2 such that for all Nash equilibria $(\tilde{\sigma}, \tilde{\tau})$,

$$Q^{\cdot n}_{\sigma,\tilde{\tau}}(\gamma_t(\sigma, \tilde{\tau}) < \varepsilon, \forall t > T_2) > 1 - \varepsilon,$$

where $Q^{\cdot n}_{\sigma,\tilde{\tau}}$ is the probability measure induced on Ω by $(\sigma, \tilde{\tau})$ and the strategic type of player 2 and $\gamma_t(\sigma, \tilde{\tau})$ is the associated reputation of player 2. So, after T_2 on, player 1 attaches a very high probability to be facing the strategic type of player 2, i.e. facing the commitment type no more than ε probability,

$$Q^{\cdot n}_{\sigma,\tilde{\tau}}(\gamma_t(\sigma, \tilde{\tau}) \geq \varepsilon, \quad \text{for some} \quad t > T_2) \leq \varepsilon,$$

Player 1 thinks she will be seeing a strategy by the strategic type of player 2 after T_2. We proceed by a similar argument as the one provided for the proof of Proposition 1 in Section 4. The counterparts of Lemma 1, 2 and 3 hold for player 1. With these results at hand and the above assumption, we derive that player 1's reputation disappears (uniformly) as well.

The following Lemma argues that either player 2's expectation of the strategy played by the strategic type of player 1 is in the limit the same as the strategy played by the commitment type of player 1, or player 2's posterior probability that player 1 is the commitment type converges to zero (given that player 1 is indeed strategic). The key idea is the same: Strictly positive beliefs about player 1's types can exist in the long-run only if both types of player 1 play identically in the limit provided that the public signals are statistically informative about player 1's actions.

Lemma 4 *Suppose Assumptions 1, 2 and 5 are satisfied. In any Nash equilibrium of the incomplete-information game,*

$$\lim_{t\to\infty} \mu_t(1 - \mu_t)\|\hat{\sigma}_t - E^{n\cdot}[\tilde{\sigma}_t|\mathcal{H}_t]\| = 0, \quad Q - a.s. \tag{15}$$

Note that since $\hat{\sigma}_t$ is a simple commitment strategy, it can be replaced by s_1. The proof is the same as the one given for Lemma 1.

Corollary 5 *At any Nash equilibrium of the incomplete-information game satisfying Assumptions 1, 2 and 5,*

$$\lim_{t\to\infty} \mu_t\|s_1 - E^{n\cdot}[\tilde{\sigma}_t|\mathcal{H}_t]\| = 0, \quad Q^{n\cdot} - a.s.$$

Note that $\lim_{t\to\infty} \mu_t\|s_1 - E^{n\cdot}[\tilde{\sigma}_t|\mathcal{H}_t]\| = 0$, also Q^{nn} - a.s.

Corollary 5 says that if player 2 does not eventually learn that player 1 is strategic, then player 2 must think that strategic type of player 1's strategy should be close to that of commitment type since the distributions of public signals induced by the two types are not distinguishable. We now show that strategic type of player 1 will know that player 2 believes this, since reputations are public by Assumption 5.

Lemma 5 *Suppose Assumptions 1, 2 and 5 hold and suppose there exists $A \in \Omega$ such that $Q^{n\cdot}(A) > 0$ and $\mu_\infty(\omega) > 0$ for all $\omega \in A$, i.e. there exists a set of events with strictly positive measure in which reputation of player 1 does not necessarily disappear. Then, there exists $\eta > 0$ and $F \subset A$, with $Q^{n\cdot}(F) > 0$, such that, for any $\xi > 0$, there exist T_1 for which,*

$$\mu_t > \eta, \quad \forall t \geq T_1,$$

$$E\left[\sup_{s\geq t} \|s_1 - E^{n\cdot}[\tilde{\sigma}_s \mid \mathcal{H}_s]\| \middle| \mathcal{H}_t\right] < \xi, \quad \forall t \geq T_1 \tag{16}$$

for all $\omega \in F$; and for all $\psi > 0$

$$Q\left(\sup_{s\geq t} \|s_1 - E^{n\cdot}[\tilde{\tau}_s \mid \mathcal{H}_s]\| < \psi \mid \mathcal{H}_t\right) \to 1 \tag{17}$$

where the convergence is uniform on F.

The next lemma shows that if the commitment type and strategic type of player 1 play sufficiently similar, strategic type of player 2 will be best responding to the commitment type's strategy for arbitrarily many periods.

Lemma 6 *Suppose $\hat{\sigma}$ be a simple pure public strategy and $BR_2(\hat{\sigma})$ is the set of best replies of strategic type of player 1 to $\hat{\sigma}$.[39] Let $(\tilde{\sigma}, \tilde{\tau})$ be Nash equilibrium strategies in the incomplete-information game. If $\tilde{\tau}$ is a pure strategy, then for all $T > 0$, there exists $\psi > 0$ such that if player 2 observes a (public) history h_t so that*

$$Q\left(\sup_{s\geq t} \|s_1 - E^{n\cdot}[\tilde{\sigma}_s \mid \mathcal{H}_s]\| < \psi \mid h_t\right) > 1 - \psi \tag{18}$$

then for $\tau' \in BR_2(\hat{\sigma})$, the continuation strategy of $\tilde{\tau}$ after the history h_t agrees with τ' for the next T periods.

We argue for a contradiction that if there is a set of states with positive measure that is induced by the play of the strategic type of player 1 on which $\mu_t \nrightarrow 0$, then, on a subset of states F in Lemma 5, the strategic type of player 2 believes that player 1's strategy is very close to her commitment strategy and thus he should be playing a best response to the commitment strategy of player 1 (after T_1), which is different than the strategy of the commitment type of player 2. Since the both players can compute what the other player believes about themselves and their future play, the strategic type of player 1 will know how player 2 thinks her future behavior is going to be and act accordingly. By Proposition 1, after some

[39] Note that $\hat{\sigma}$ assigns s_1 in every period independent of history. The repeated game best response of player 2 in the complete-information game is $BR_2(\hat{\sigma})$ is a singleton that assigns r_2 in each period.

T_2, the reputation of player 2 will disappear in all Nash equilibria, and thus after $T = \max\{T_1, T_2\}$, player 1 expects to see a best reply to her commitment strategy from the strategic type of player 2 with a high probability. Since strategic type of player 2's strategy is different than the commitment type's strategy, his reputation will disappear even more after T. More precisely, the histories for which player 2's reputation can be rebuilt would have measure 0. Thus, the strategic player 1 best responds to the strategic type of player 2 (who gives a best reply to the commitment strategy of player 1 which is different than commitment type of player 2's strategy) with high probability. However, since the strategic type player 1's best response to player 2's strategy is different than her commitment strategy, the strategic and the commitment type of player 1 are expected to play differently, which will provide the contradiction to $\mu_t \nrightarrow 0$ on F.

More specifically, we are going to show that player 2 assigns a probability $1-\zeta$ to player 1 believing with probability at least $1-\eta$ that player 2 thinks player 1's strategy is within ξ of the commitment strategy when the probability measure over the histories are induced by the play of the strategic type of player 1,

$$Q^{n\cdot}\left(Q^{n\cdot}\left(\sup_{s\geq t}\|s_1 - E^{n\cdot}[\tilde{\sigma}_s \mid \mathcal{H}_s]\| < \xi \mid \mathcal{H}_t\right) > 1-\eta \mid \mathcal{H}_t\right) > 1-\zeta.$$

We choose ξ and ζ such that $\xi < \min\{\psi, 1-\zeta\}$ and $\zeta < 1$ to arrive the desired contradiction. Suppose that there is a set of states A with $Q^{n\cdot}(A) > 0$ and $\mu_\infty(\omega) > 0$ for all $\omega \in A$ (note that $\mu_\infty(\omega) < 1$ on A by Corollary 5.) Then, by Lemma 5, there is a set $F \in A$ with $Q^{n\cdot}(F) > 0$ (also $Q(F) > 0$) such that for any $\xi > 0$, there exists a T_1 such that for any $t > T_1$ and $\omega \in F$,

$$Q\left(\sup_{s\geq t}\|s_1 - E^{n\cdot}[\tilde{\sigma}_s \mid \mathcal{H}_s]\| < \xi \mid \mathcal{H}_t\right) \to 1 \tag{19}$$

Then, there exists a subset $G \in F$ with $Q^{n\cdot}(G) > 0$ such that for any $t > T_1$, on G,

$$\|s_1 - E^{n\cdot}[\tilde{\sigma}_t \mid \mathcal{H}_t]\| < \xi, \quad Q-\text{a.s.} \tag{20}$$

Note that (20) implies $\|s_1 - E^{n\cdot}[\tilde{\sigma}_t \mid \mathcal{H}_t]\| < \xi \quad Q^{n\cdot}$- a.s., Q^{nn}-a.s. and $Q^{\cdot n}$-a.s. Also, by the extension of Hart's lemma given in footnote 32), we conclude that for some η and ζ,

$$Q^{n\cdot}\left(\sup_{s\geq t}\|s_1 - E^{\cdot n}[\tilde{\sigma}_s \mid \mathcal{H}_s]\| < \xi|\mathcal{H}_t\right) > 1-\eta\zeta. \tag{21}$$

This shows that with a high probability ($1-\eta\zeta$), player 1 believes that player 2 assigns player 1's strategy to be ξ close to the commitment strategy. Define,

$$g_t := Q^{n\cdot}\left(\sup_{s\geq t}\|(s_1 - E^{n\cdot}[\tilde{\sigma}_s \mid \mathcal{H}_s]\| < \xi|\mathcal{H}_t\right)$$

$$\kappa_t := Q^{n\cdot}(g_t > 1-\eta | \mathcal{H}_t)$$

We want to show $\kappa_t > 1-\zeta$. Since, $E^{n\cdot}[g_t \mid \mathcal{H}_t] > 1-\eta\zeta$ by condition (21), and

$$\begin{aligned} E^{n\cdot}[g_t \mid \mathcal{H}_t] &= E^{n\cdot}[g_t \mid g_t \leq 1-\eta, \mathcal{H}_t](1-\kappa_t) + E^{n\cdot}[g_t \mid g_t > 1-\eta, \mathcal{H}_t]\kappa_t \\ &\leq (1-\eta)(1-\kappa_t) + \kappa_t \end{aligned}$$

we have,

$$1-\eta\zeta < (1-\eta)(1-\kappa_t) + \kappa_t$$

which implies $\kappa_t > 1-\zeta$ on F. So,

$$Q^{n\cdot}\left(Q^{n\cdot}\left(\sup_{s \geq t} \|(s_1 - E^{n\cdot}[\tilde{\sigma}_s \mid \mathcal{H}_s]\| < \xi \mid \mathcal{H}_t\right) > 1-\eta \mid \mathcal{H}_t\right) > 1-\zeta$$

This says that player 2 assigns a probability of at least $1-\zeta$ (after observing histories generated by the play of the strategic type of player 1) to player 1 believing with probability at least $1-\eta$ that player 2 believes player 1's strategy is within ξ of the commitment strategy. Note that, by Lemma 6, for all $T' > 0$, there exists $\psi > 0$ such that if player 2 observes a (public) history h_t so that

$$Q\left(\sup_{s \geq t} \|s_1 - E^{n\cdot}[\tilde{\sigma}_s \mid \mathcal{H}_s]\| < \psi \mid h_t\right) > 1-\psi$$

then for $\tau' \in BR_2(\hat{\sigma})$, the continuation strategy of $\tilde{\tau}$ after the history h_t agrees with τ' for the next T' periods. We have picked $\xi < \psi$, hence strategic player 2 best responds to the commitment strategy of player 1 for the next T' periods.

Define $T := \max\{T_1, T_2\}$. Note that after time $t > T$, by uniform disappearance of player 2's reputation, player 1 believes that player 2 is the commitment type with $\gamma_t < \varepsilon$ who plays s_2 every period, and with probability $1-\gamma_t > \varepsilon$ he is the strategic type who believes that player 2's strategy is within ξ of the commitment strategy from then on and hence plays the best reply to it (since we have picked $\xi < \psi$) thereafter. That is why, after $t > T$, the strategic player 2 will give a best response to s_1, which is different than the commitment strategy of player 2, thus it is unlikely for player 2 to rebuilt after T_2. We can stop the process γ_t at T.

Since s_1 is not a best response to $r_2 \neq s_2$ (i.e. the myopic best reply of strategic player 2 to s_1), there exists $\eta > 0$ such that for any repeated game strategy of the strategic type of player 1 that attaches probability at least $1-\eta$ to $\hat{\tau}$ (to always playing r_2), s_1 is suboptimal for the strategic type of player 1 (by the upper-hemicontinuity of the best response correspondence) in period 0. Let $\bar{\eta}$ such that s_1 is suboptimal for the strategic player 1 in period 0 if strategic type of player 2 attaches $1-\bar{\eta}$ to $\hat{\sigma}$, regardless of player 2's belief about player 1 being commitment type. Define $\bar{\rho} := \min_{y,i,j} \rho_{ij}^{y}$ (> 0 by Assumption 1). So, if player 2 assigns probability at least $1 - \bar{\rho}\bar{\eta} \equiv 1 - \eta$ to $\hat{\sigma}$, then he assigns at least probability $1 - \bar{\eta}$ to $\hat{\sigma}$ after any deviation that leaves the probability of $\hat{\sigma}$ (conditional on any signal) unchanged.

Since we have picked ξ such that $\xi < \psi$, for all $t > T$, strategic type of player 2 chooses to play r_2, the unique best response to the commitment action thereafter, whenever he believes that player 1's strategy is within ξ of the commitment strategy. Hence, in any period $t > T$, player 2 assigns a probability of at least $1 - \zeta$ to player 1 believing that player 2's subsequent play is r_2 thereafter with at least probability $1 - \eta$. Thus, player 2 assigns probability at least $1 - \zeta$ to player 1's play in period t being a best response to $\hat{\tau}$. Since s_1 is pure, it specifies an action $\hat{i}$ with probability 1. However, player 2 must believe that that action is played with no more than ζ probability in period t. But since, $1 - \zeta > \xi$, this contradicts (20). Player 2's second order beliefs about strategic player 1's behavior (after observing the relevant game has been evolving and histories have been generated by the play of strategic player 1) contradicts with his first order beliefs.

The uniform disappearance of player 2's reputation follows the same argument as in Section 4.5. Hence, for all $\varepsilon > 0$, there exists T, such that for all Nash equilibria $(\tilde{\sigma}, \tilde{\tau})$ of the incomplete-information game,

$$Q^{nn}_{\tilde{\sigma},\tilde{\tau}}(\mu_t(\tilde{\sigma}, \tilde{\tau}) < \varepsilon, \forall t > T) > 1 - \varepsilon,$$

where $Q^{nn}_{\tilde{\sigma},\tilde{\tau}}$ is the probability measure induced on Ω by $(\tilde{\sigma}, \tilde{\tau})$ and the strategic types of players and $\mu_t(\tilde{\sigma}, \tilde{\tau})$ is the associated reputation of player 1.

6 Concluding Remarks

The main result of this paper is that the reputations of players for playing a strategy that is not part of an equilibrium of the stage game can not be sustainable in the long-run for games with one-sided binding moral hazard (at the commitment profile) under imperfect public monitoring. The way we prove our result is by first showing that the reputation of the player who is subject to binding moral hazard at the commitment profile disappears (uniformly) and then after the type of that player is almost known, the reputation of the other player should disappear as well. The implication of this result on the regulatee-regulator game is: First the reputation of being tough for the regulator disappears. After his true type is almost known, the regulatee starts to take advantage of regulator's uncertainty over her type and the regulatee's reputation of being virtuous disappears eventually, too. Moreover, the continuation equilibrium of the incomplete-information game converges to an equilibrium of the complete-information game in the limit.

There are some interesting related questions we would like to study such as how the rate of disappearance is affected by different priors. For instance, we believe that the existence of a tough regulator postpones the revelation of the true type of the regulatee; whereas the existence of a virtuous regulatee speeds up the revelation of the type of the regulator. So, a regulator whose goal is to understand the type of the regulatee should not pretend to be the tough type.

The other important observation we can make about the regulatee-regulator game is that the reputations are sustainable for more complicated commitment strategies that are equilibria of the repeated complete-information game. For instance, if there is a grim trigger type for the regulator; we believe that the reputations would be sustainable and the equilibrium would be almost efficient, in the sense that players achieve the highest total payoff (very close to the efficient frontier of the feasible and individually rational payoff set). Hence, if a regulator could choose to establish a reputation for a type in the presence of a grim trigger and a tough type, he should choose to mimic the grim trigger type.

The possible applications of such games may help to analyze some important problems such as asset market manipulation and tax evasion. For a discussion about the asset market manipulation, we refer the reader to Ozdogan (2009).

7 Appendix

7.1 Posteriors are bounded martingales.

Lemma 7 *Suppose Assumption 1 holds. Then γ_t is a bounded martingale with respect to the measure Q and filtration $\{\mathcal{H}_{1t}\}_t$. Moreover, Assumption 5 implies that γ_t is also a bounded martingale with respect to the measure Q and filtration $\{\mathcal{H}_t\}_t$. Similarly, random variable μ_t is a bounded martingale with respect to the measure Q and filtration $\{\mathcal{H}_{2t}\}_t$ and $\{\mathcal{H}_t\}_t$.*

Proof. We first show γ_t is a martingale with respect to measure Q and filtration $\{\mathcal{H}_{1t}\}_t$. Let $\gamma_{t+1}(h_t; i_t, y_t)$ denote player 1's belief in period $t+1$ after playing i_t and observing public signal y_t in period t, given public history h_t. Note that $\gamma_{t+1} \equiv \text{prob}[c \mid h_t, i_t, y_t]$. So, for all $h_{1,t+1}$,

$$
\begin{aligned}
E[\gamma_{t+1}|\mathcal{H}_{1,t+1}] &= \sum_{y \in Y} \text{prob}[y_t|h_t, i_t].\gamma_{t+1}(h_t; i_t, y_t) \\
&= \sum_{y \in Y} \text{prob}[y_t|h_t, i_t].\frac{\gamma_t \text{prob}[y_t|h_t, i_t, c]}{\text{prob}[y_t|h_t, i_t]} \\
&= \sum_{y \in Y} \gamma_t \sum_{j \in J} \varsigma_2^j \rho_{i_t j}^{y_t} \\
&= \gamma_t.
\end{aligned}
$$

The random variable γ_t is also a martingale with respect to measure Q and filtration $\{\mathcal{H}_t\}_t$ under Assumption 5. In this case, for all h_{t+1} and for all i_t and i,

$$
\begin{aligned}
E[\gamma_{t+1}|\mathcal{H}_{t+1}] &= \sum_{y\in Y} \text{prob}[y_t|h_t].\gamma_{t+1}(h_t; i_t, y_t) \\
&= \sum_{y\in Y} \text{prob}[y_t|h_t].\frac{\gamma_t \text{prob}[y_t|h_t, i_t, c]}{\text{prob}[y_t|h_t, i_t]} \\
&= \sum_{y\in Y} \sum_{j\in J} \rho_{ij}^{y_t}\left(\gamma_t \varsigma_2^j + (1-\gamma_t)E^{\cdot n}[\tilde{\tau}_t^j \mid h_t]\right) \frac{\gamma_t \sum_{j\in J} \varsigma_2^j \rho_{i_t j}^{y_t}}{\sum_{j\in J} \rho_{i_t j}^{y_t}\left(\gamma_t \varsigma_2^j + (1-\gamma_t)E^{\cdot n}[\tilde{\tau}_t^j \mid h_t]\right)} \\
&= \sum_{y\in Y} \sum_{j\in J} \frac{\text{prob}[y_t|i, j]\gamma_t \varsigma_2^j \rho_{i_t j}^{y_t}}{\text{prob}[y_t|i_t, j]} \\
&= \sum_{y\in Y} \sum_{j\in J} \frac{\rho_{ij}^{y_t}\gamma_t \varsigma_2^j \rho_{i_t j}^{y_t}}{\rho_{i_t j}^{y_t}} \\
&= \sum_{y\in Y} \gamma_t \sum_{j\in J} \varsigma_2^j \rho_{ij}^{y_t} \\
&= \gamma_t
\end{aligned}
$$

Note that the forth equality follows from Assumption 5 and the last line is due to Assumption 1. [40]

By the same argument, we can show that μ_t is a bounded martingale with respect to the measure Q and filtration $\{\mathcal{H}_{2t}\}_t$ (and also with respect to filtration $\{\mathcal{H}_t\}_t$ by Assumption 5), thus converges Q-almost surely (and hence $Q^{n\cdot}$ and Q^{nn} - almost surely) to a random variable μ_∞ on Ω. ■

7.2 Proof of Theorem 2

We modify the proof of Cripps, Mailath, and Samuelson (2004) for two long-lived player with uncertainty over the types of both players. Since $(\tilde{\vartheta}_1^{h_t}, \tilde{\vartheta}_2^{h_t})$ are continuation equilibrium profile, for each public history h_t and pure

[40]For the case of product structure,

$$
\begin{aligned}
E[\gamma_{t+1}|\mathcal{H}_{t+1}] &= \sum_{y\in Y} \text{prob}[y_t|h_t].\gamma_{t+1}(h_t; i_t, y_t) \\
&= \sum_{y\in Y} \text{prob}[y_t|h_t].\frac{\gamma_t \text{prob}[y_t|h_t, i_t, c]}{\text{prob}[y_t|h_t, i_t]} \\
&= \sum_{y\in Y} \rho_{ij}^{y} \frac{\gamma_t \text{prob}[y_t|h_t, i_t, c]}{\rho_{i_t j}^{y}} \\
&= \sum_{(y_1,y_2)\in Y} \rho_i^{y_1}\rho_j^{y_2} \frac{\gamma_t \sum_{j\in J} \varsigma_2^j \rho_{i_t}^{y_1}\rho_j^{y_2}}{\rho_{i_t}^{y_1}\rho_j^{y_2}} \\
&= \gamma_t
\end{aligned}
$$

.

strategies $\varsigma_1' \in S_1$ and $\varsigma_2' \in S_2$, the continuation expected payoffs should satisfy:

$$E_{(\tilde{\vartheta}_1^{h_t}, \gamma_t \hat{\vartheta}_2^{h_t} + (1-\gamma_t)\tilde{\vartheta}_2^{h_t})}[U_1(\varsigma_1, \varsigma_2)] \geq E_{\gamma_t \hat{\vartheta}_2^{h_t} + (1-\gamma_t)\tilde{\vartheta}_2^{h_t}}[U_1(\varsigma_1', \varsigma_2)] \tag{22}$$

$$E_{(\mu_t \hat{\vartheta}_1^{h_t} + (1-\mu_t)\tilde{\vartheta}_1^{h_t}, \tilde{\vartheta}_2^{h_t})}[U_2(\varsigma_1, \varsigma_2)] \geq E_{\mu_t \hat{\vartheta}_1^{h_t} + (1-\mu_t)\tilde{\vartheta}_1^{h_t}}[U_1(\varsigma_1, \varsigma_2')] \tag{23}$$

where $\hat{\vartheta}_1^{h_t}$ and $\hat{\vartheta}_2^{h_t}$ are the commitment mixed strategies corresponding to commitment behavior strategies $\hat{\sigma}_{h_t}$ and $\hat{\tau}_{h_t}$ in the continuation game. By Theorem 1, $\mu_t \to 0 \quad Q^{n\cdot}$-almost surely and $\gamma_t \to 0 \quad Q^{\cdot n}$-almost surely which imply $\gamma_t \to 0$ and $\mu_t \to 0 \quad Q^{nn}$- almost surely, by absolute continuity of Q^{nn} with respect to $Q^{n\cdot}$ and $Q^{\cdot n}$. Suppose $\{h_t\}_t$ is a sequence of public histories on which $\gamma_t, \mu_t \to 0$ and $\{(\tilde{\vartheta}_1^{h_t}, \tilde{\vartheta}_2^{h_t})\}_{t=1}^{\infty} \to (\tilde{\vartheta}_1^*, \tilde{\vartheta}_2^*)$ on this sequence. We need to show $(\tilde{\vartheta}_1^*, \tilde{\vartheta}_2^*)$ satisfies (22) and (23), which suffices to show expectation $E_{(\vartheta_1, \vartheta_2)}$ is continuous in $(\vartheta_1, \vartheta_2)$. The continuity of this expectation is given by Theorem 4.4 of Fudenberg and Tirole (1991) and it is due to discounting (since $\delta_1, \delta_2 < 1$).

7.3 Equilibrium of the one-shot regulatee-regulator game

Let $I = \{T, U\}$ be the action space of the regulatee and $J = \{D, L\}$ be that of the regulator. Suppose $\alpha_1(T)$ denote the probability the regulatee puts on $Truthful$ and $\alpha_2(D)$ denotes the probability the regulator puts on $Diligence$. Let μ be the prior belief of the regulator about the regulatee being virtuous; and γ be the prior that the regulatee thinks the regulator is tough, where the prior beliefs are common knowledge. The equilibrium of the one-shot game is given by Proposition 3.

Proposition 3 *The equilibrium $(\alpha_1, \alpha_2) \in \Delta(I) \times \Delta(J)$ is given as follows:*

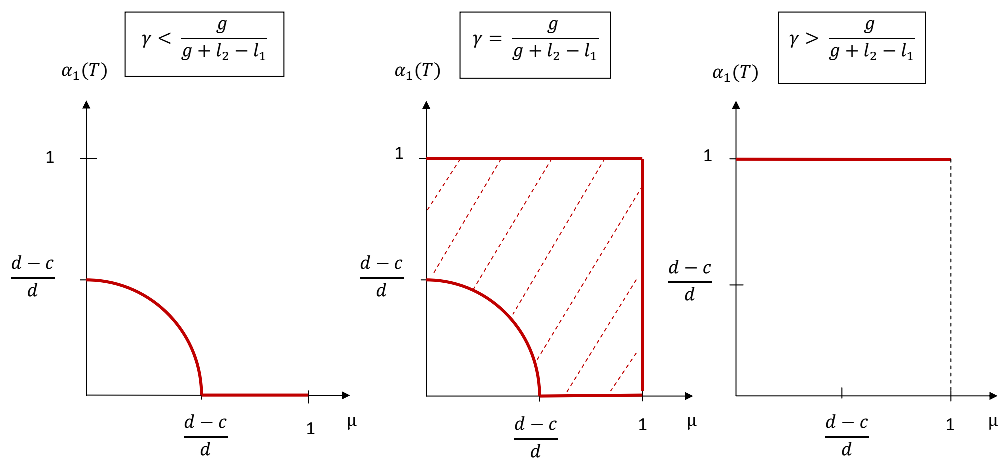
1. *If $\gamma > \frac{g}{g+l_2-l_1}$ and $\mu \in [0,1]$,* $\quad \alpha_1(T) = 1$ *and* $\alpha_2(D) = 0$ *;*
2. *If $\gamma = \frac{g}{g+l_2-l_1}$ and*
   (a) *If $\mu > \frac{d-c}{d}$,* $\quad \alpha_1(T) \in [0,1]$ *and* $\alpha_2(D) = 0$*;*
   (b) *If $\mu = \frac{d-c}{d}$,* $\quad \alpha_1(T) \in [0,1]$ *and* $\alpha_2(D) = 0$*;*
   (c) *If $\mu < \frac{d-c}{d}$,* $\quad \alpha_1(T) \in [1 - \frac{c}{(1-\mu)d}, 1]$ *and* $\alpha_2(D) = 0$*.*
3. *If $\gamma < \frac{g}{g+l_2-l_1}$ and*
   (a) *If $\mu > \frac{d-c}{d}$,* $\quad \alpha_1(T) = 0$ *and* $\alpha_2(D) = 0$*;*
   (b) *If $\mu = \frac{d-c}{d}$,* $\quad \alpha_1(T) = 0$ *and* $\alpha_2(D) \in [0, 1 - \frac{l_2-l_1}{(1-\gamma)(g+l_2-l_1)}]$*;*
   (c) *If $\mu < \frac{d-c}{d}$,* $\quad \alpha_1(T) = 1 - \frac{c}{(1-\mu)d}$ *and* $\alpha_2(D) = 1 - \frac{l_2-l_1}{(1-\gamma)(g+l_2-l_1)}$*;*

Corresponding equilibrium expected payoff vectors are: [41]

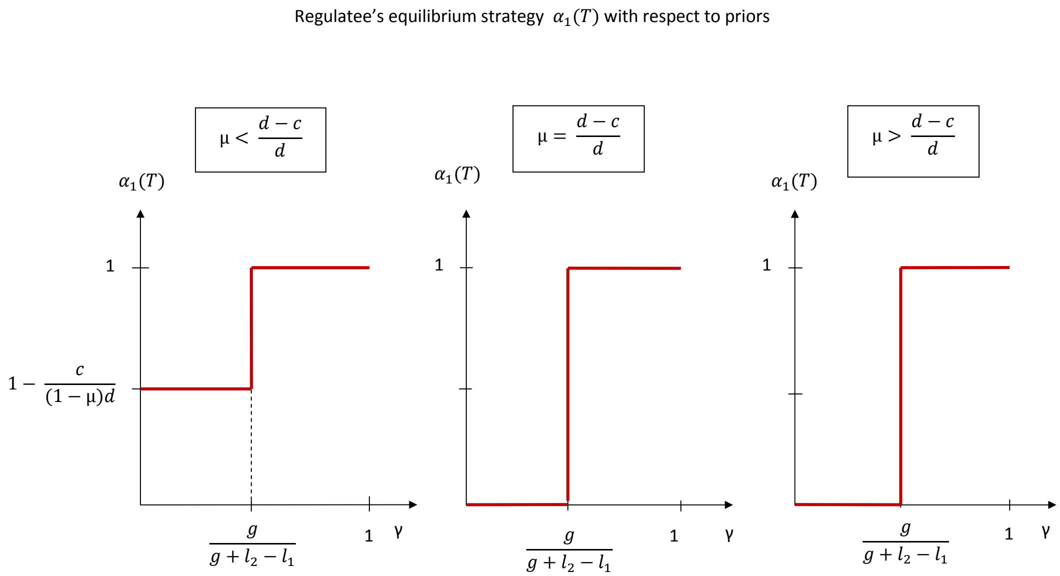
$$(\pi_1, \pi_2) = \begin{cases} (x - \gamma l_1, y) & \textit{Case 1:} \quad \gamma > \frac{g}{g+l_2-l_1} \quad \textit{and} \quad \mu \in [0,1] \\ (x - \frac{g.l_1}{g+l_2-l_1}, [\mu y + (1-\mu)z, y]) & \textit{Case 2.a :} \quad \gamma = \frac{g}{g+l_2-l_1} \quad \textit{and} \quad \mu > \frac{d-c}{d} \\ (x - \frac{g.l_1}{g+l_2-l_1}, [y - \frac{c(y-z)}{d}, y]) & \textit{Case 2.b :} \quad \gamma = \frac{g}{g+l_2-l_1} \quad \textit{and} \quad \mu = \frac{d-c}{d} \\ (x - \frac{g.l_1}{g+l_2-l_1}, [y - \frac{c(y-z)}{d}, y]) & \textit{Case 2.c :} \quad \gamma = \frac{g}{g+l_2-l_1} \quad \textit{and} \quad \mu < \frac{d-c}{d} \\ (x + g - \gamma(g+l_2), \mu y + (1-\mu)z) & \textit{Case 3.a :} \quad \gamma < \frac{g}{g+l_2-l_1} \quad \textit{and} \quad \mu > \frac{d-c}{d} \\ ([x - \frac{g.l_1}{g+l_2-l_1}, x + g - \gamma(g+l_2)], y - \frac{c(y-z)}{d}) & \textit{Case 3.b :} \quad \gamma < \frac{g}{g+l_2-l_1} \quad \textit{and} \quad \mu = \frac{d-c}{d} \\ (x - \frac{g.l_1}{g+l_2-l_1}, y - \frac{c(y-z)}{d}) & \textit{Case 3.c :} \quad \gamma < \frac{g}{g+l_2-l_1} \quad \textit{and} \quad \mu < \frac{d-c}{d} \end{cases}$$

We can summarize the equilibrium strategies depending on the priors by the following graphs.

Regulatee's equilibrium strategy $\alpha_1(T)$ with respect to priors



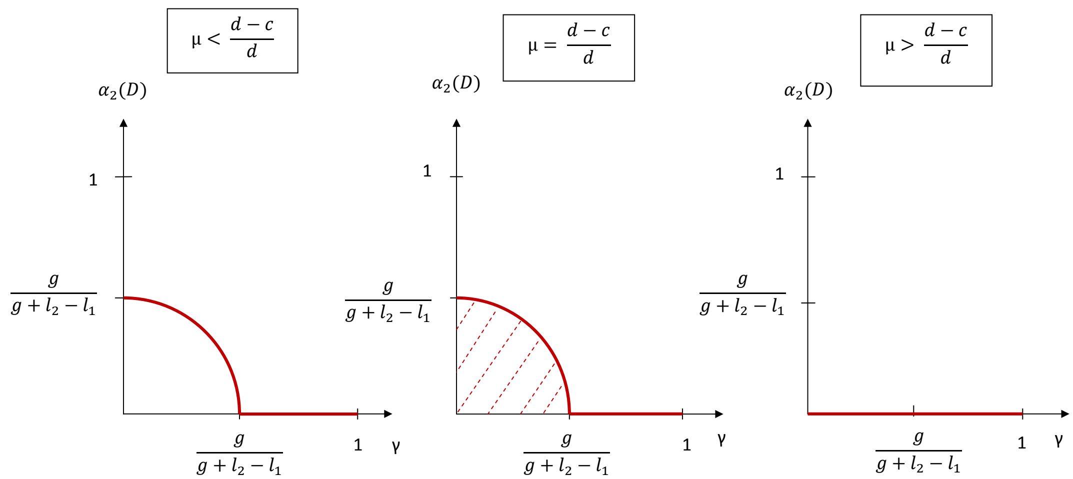
[41]We denote the payoff correspondence for a player with a bracket [ ].



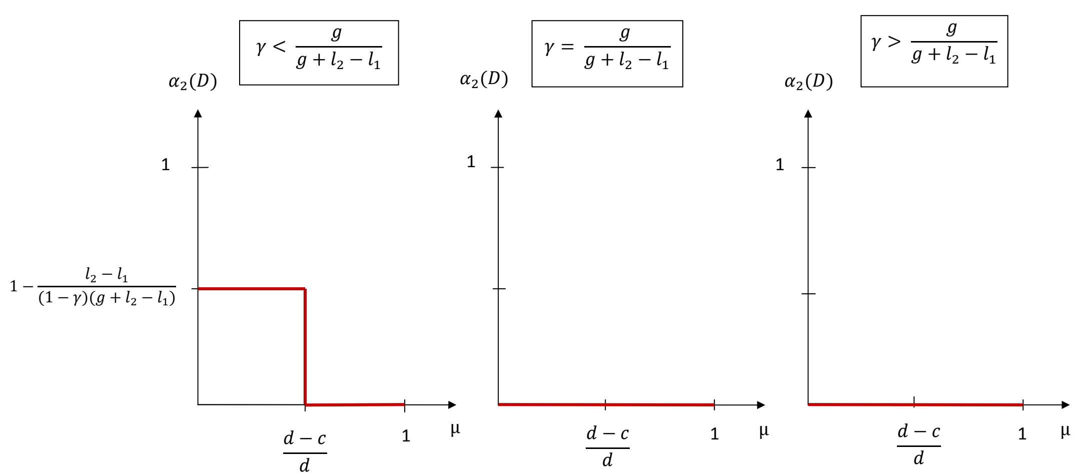
So, the equilibrium strategy for the regulatee (for being truthful) is nonincreasing in the prior on her type of being virtuous μ, for any given prior on the regulator's type; and it is nondecreasing with in the prior on the toughness of the regulator γ for any fixed prior μ. Moreover, if the prior on the toughness of the regulator is sufficiently high (greater than $\frac{g}{g+l_2-l_1}$), the regulatee always chooses to be truthful as she thinks she faces a tough regulator with a high enough probability. Note that that threshold decreases as as the expected gain from being untruthful (g) decreases or the expected loss (l_2) increases.

The regulatee can get her highest possible payoff $x + g$ only if she thinks that it is more likely she faces a *strategic* regulator and the prior on her being virtuous is high enough ($\mu > \frac{d-c}{d}$) so that a strategic regulator chooses to be lazy.

Regulator's equilibrium strategy $\alpha_2(D)$ with respect to priors



Regulator's equilibrium strategy $\alpha_2(D)$ with respect to priors



The regulator's strategy of being diligent is nonincreasing both in the prior γ of being tough type (for any fixed prior on the regulatee being virtuous) and in the prior μ on the regulatee's type (for a given prior on the regulator's type). It is intuitive in the sense that the regulator chooses to be diligent less and less often as he thinks that with a high probability he faces a virtuous regulatee who is truthful; or as he thinks that the regulatee thinks he is tough, so that she chooses to be truthful. Note that the regulator never (for no prior beliefs) chooses to be diligent for sure and when the prior on his toughness is above $\frac{g}{g+l_2-l_1}$, he always chooses to be lazy, knowing that the regulatee is going to choose to be truthful. He randomizes between being lazy and diligent only when $\gamma < \frac{g}{g+l_2-l_1}$ and $\mu \leq \frac{d-c}{d}$, i.e. he thinks that it is more likely he faces a strategic regulatee and the regulatee also thinks that it is more likely she faces a strategic regulator.

Proof. We find the best responses of players for each case.

The expected utility of the regulatee from being truthful, i.e. choosing $\alpha_1 = 1$ is

$$\pi_1(T, \alpha_2) = \gamma(x - l_1) + (1-\gamma)[\alpha_2(x - l_1) + (1 - \alpha_2)x]$$

Her expected utility from being untruthful, i.e. $\alpha_1 = 0$, is

$$\pi_1(U, \alpha_2) = \gamma(x - l_2) + (1-\gamma)[\alpha_2(x - l_2) + (1 - \alpha_2)(x + g)]$$

Hence, the regulatee's best response is given by the following:

$$BR_1(\alpha_2) = \begin{cases} 1 & \text{if} \quad \alpha_2 > 1 - \frac{l_2-l_1}{(1-\gamma)(g+l_2-l_1)} \\ [0,1] & \text{if} \quad \alpha_2 = 1 - \frac{l_2-l_1}{(1-\gamma)(g+l_2-l_1)} \\ 0 & \text{if} \quad \alpha_2 < 1 - \frac{l_2-l_1}{(1-\gamma)(g+l_2-l_1)} \end{cases} \tag{24}$$

Note that the strategy of the regulator that makes the regulatee indifferent between being truthful and untruthful, $\alpha_2 = 1 - \frac{l_2-l_1}{(1-\gamma)(g+l_2-l_1)}$, is greater than 0 if $\gamma < \frac{g}{g+l_2-l_1}$ and equals to 0 if $\gamma = \frac{g}{g+l_2-l_1}$. If $\gamma > \frac{g}{g+l_2-l_1}$, then $BR_1(\alpha_2) = 1$ for any α_2.

The expected utility of the regulator by choosing to be diligent, i.e. choosing $\alpha_2 = 1$, is

$$\pi_2(\alpha_1, D) = \mu(y - c) + (1-\mu)[\alpha_1(y - c) + (1 - \alpha_1)(z - c + d)]$$

The expected utility of the regulator by choosing to be lazy, i.e. choosing $\alpha_2 = 0$, is

$$\pi_2(\alpha_1, L) = \mu y + (1-\mu)[\alpha_1 y + (1 - \alpha_1)z]$$

Regulator's best response is given by the following:

$$BR_2(\alpha_1) = \begin{cases} 1 & \text{if} \quad \alpha_1 < 1 - \frac{c}{(1-\mu)d} \\ [0,1] & \text{if} \quad \alpha_1 = 1 - \frac{c}{(1-\mu)d} \\ 0 & \text{if} \quad \alpha_1 > 1 - \frac{c}{(1-\mu)d} \end{cases} \tag{25}$$

Note that the strategy of the Sender that makes the Receiver indifferent between choosing to be diligent and lazy, $\alpha_1 = 1 - \frac{c}{(1-\mu)d}$, is greater than 0 if $\mu < \frac{d-c}{d}$ and equals to 0 if $\mu = \frac{d-c}{d}$. If $\mu > \frac{d-c}{d}$, then $BR_2(\alpha_1) = 0$ for any α_1.

Case 1 $\gamma > \frac{g}{g+l_2-l_1}$ and $\mu \in [0,1]$.

In this case, $BR_1(\alpha_2) = 1$ for any τ. The unique fixed point of the best response correspondence $BR_1 \times BR_2$ is $\alpha_1 = 1$ and $\alpha_2 = 0$ (Regulator is truthful and Receiver is lazy). The equilibrium payoff vector (π_1, π_2) corresponding to this strategy profile is $(x - \gamma l_1, y)$.

Case 2 $\gamma = \frac{g}{g+l_2-l_1}$. The strategy that makes the Sender indifferent between telling the truth and lying is $\tau = 0$. For $\alpha_2 > 0$, $BR_1(\alpha_2) = 1$.

2.a $\mu > \frac{d-c}{d}$: $BR_2(\alpha_1) = 0$ for any α_1. So, the equilibria in this case are $\alpha_1 \in [0,1]$ and $\alpha_2 = 0$. Corresponding equilibrium payoffs for the regulatee is $x - \frac{g.l_1}{g+l_2-l_1}$ and for the regulator lies in $[\mu y + (1-\mu)z, y]$.

2.b $\mu = \frac{d-c}{d}$: Regulator is indifferent between choosing to be lazy and diligent when $\alpha_1 = 0$. Otherwise, $BR_2(\alpha_1) = 0$. The equilibria in this case are $\alpha_1 \in [0,1]$ and $\alpha_2 = 0$. Corresponding equilibrium payoffs for the regulatee is $x - \frac{g.l_1}{g+l_2-l_1}$ and for the regulator lies in $[y - \frac{c(y-z)}{d}, y]$.

2.c $\mu < \frac{d-c}{d}$: Best response correspondence of the regulator is given by the expression 25. Sender is indifferent if $\alpha_2 = 0$ and $BR_1(\alpha_2) = 1$ if $\alpha_2 > 0$. The fixed point of $BR_1 \times BR_2$ are $\alpha_1 \in [1 - \frac{c}{(1-\mu)d}, 1]$ and $\alpha_2 = 0$, providing a payoff of $x - \frac{g.l_1}{g+l_2-l_1}$ to regulatee and $[y - \frac{c(y-z)}{d}, y]$ to regulator.

Case 3 $\gamma < \frac{g}{g+l_2-l_1}$. The best response correspondence of the regulatee is given by expression 24.

3.a $\mu > \frac{d-c}{d}$: $BR_2(\alpha_1) = 0$ for any α_1. The unique equilibrium for these values of beliefs is $\alpha_1 = 0$ and $\alpha_2 = 0$, providing a payoff of $\gamma(x - l_2) + (1-\gamma)(x+g) = x + g - \gamma(g + l_2)$ to regulatee and $\mu y + (1-\mu)z$ to regulator.

3.b $\mu = \frac{d-c}{d}$: The regulator is indifferent between choosing to be diligent and lazy when $\alpha_1 = 0$. Otherwise, $BR_2(\alpha_1) = 0$. The equilibria are $\alpha_1 = 0$ and $\alpha_2 \in [0, 1 - \frac{l_2-l_1}{(1-\gamma)[g+l_2-l_1]}]$ with the associated equilibrium payoff in $[x - \frac{g.l_1}{g+l_2-l_1}, x + g - \gamma(g+l_2)]$ to the regulatee and $y - \frac{c(y-z)}{d}$ to the regulator.

3.c $\mu < \frac{d-c}{d}$: Best response correspondence of the regulator is given by the expression 25. The unique equilibrium strategy profile for these values of beliefs is $\alpha_1 = 1 - \frac{c}{(1-\mu)d}$ and $\alpha_2 = 1 - \frac{l_2-l_1}{(1-\gamma)[g+l_2-l_1]}$, providing equilibrium payoffs $x - \frac{g.l_1}{g+l_2-l_1}$ to the regulatee and $y - \frac{c(y-z)}{d}$ to the regulator.

■

References

BENABOU, R., AND G. LAROQUE (1992): “Using Privileged Information to Manipulate Markets: Insiders, Gurus and Credibility,” *Quarterly Journal of Economics*, 107(3), 921–958.

CELENTANI, M., D. FUDENBERG, D. K. LEVINE, AND W. PESENDORFER (1996): “Maintaining A Reputation Against A Long-Lived Opponent,” *Econometrica*, 64(3), 691–704.

CHUNG, K. L. (1974): *A Course in Probability Theory*. New York: Academic Press.

CRIPPS, M. W., G. J. MAILATH, AND L. SAMUELSON (2004): “Imperfect monitoring and impermanent reputations,” *Econometrica*, 72(2), 407–432.

——— (2007): “Disappering private reputations in long-run relationships,” *Journal of Economic Theory*, 134, 287–316.

FUDENBERG, D., AND D. K. LEVINE (1989): “Reputation and Equilibrium Selection in Games with a Patient Player,” *Econometrica*, 57(4), 759–778.

——— (1992): “Maintaining a Reputation when Strategies are Imperfectly Observed,” *Review of Economic Studies*, 59, 561–579.

——— (1994): “Efficiency and Observability with Long-Run and Short-Run Players,” *Journal of Economic Theory*, 62(1), 103–135.

FUDENBERG, D., D. K. LEVINE, AND E. MASKIN (1994): “The Folk Theorem with Imperfect Public Information,” *Econometrica*, 62(5), 997–1039.

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. Cambridge, MA: MIT Press.

HARSANYI, J. C. (1967): “Games of incomplete information played by Bayesian players, Parts I,” *Manag. Sci.*, 14, 159–182.

HART, S. (1985): “Nonzero-sum two person repeated games with incomplete information,” *Mathematics of Operations Research*, 10, 117–153.

KOREN, G. (1992): “Two-Person Repeated Games where Players Know Their Own Payoffs,” *Courant Institute, preprint*.

MAILATH, G. J., AND L. SAMUELSON (2001): “Who Wants a Good Reputation?,” *Review of Economic Studies*, 68, 415–441.

MAILATH, G. J., AND L. SAMUELSON (2006): *Repeated Games and Reputations*. Oxford University Press.

OZDOGAN, A. (2009): "Strategic Communication vs. Strategic Auditing in Asset Markets with Reputation Concerns," *work in progress*.

PHELAN, C. (2006): "Public trust and government betrayal," *Journal of Economic Theory*, 130, 27–43.

RUDOLF, A., V. S. BERNHARD, AND S. ZAMIR (2002): "Chapter 51: Inspection games," in *Handbook of Game Theory with Economic Applications*, ed. by R. Aumann, and S. Hart, vol. 3. Elsevier.

SCHMIDT, K. M. (1993): "Reputatition and Equilibrium Characterization in Repeated Games of Conflicting Interests," *Econometrica*, 61(2), 325–351.

SOBEL, J. (1985): "A Theory of Credibility," *Review of Economic Studies*, pp. 557–573.

SORIN, S. (1999): "Merging, Reputation, and Repeated Games with Incomplete Information," *Games and Economic Behavior*, 29, 274 – 308.