İstanbul Bilgi University

MURAT SERTEL CENTER FOR ADVANCED ECONOMIC STUDIES

BELİS

BİLGİ ECONOMICS LAB OF İSTANBUL

Working Paper Series

Working Paper No: 2021-01

Gradualism in a principal-agent interaction

Shahin Baghirov   Ayça Ebru Giritligil

May 2021

Gradualism in a principal-agent interaction

Shahin Baghirov * Ayça Ebru Giritligil †

Koç University Istanbul Bilgi University

May 2021

Abstract

This paper theoretically and experimentally analyzes a repeated principal-agent game with varying relative stakes. The principal hires an informed agent to observe the state and take action accordingly. There is a high probability that the principal and the agent have misaligned preferences and the principal does not know the agent's type. Repeated play becomes valuable in such a setting by improving the equilibrium payoff of the principal. The agent has reputation incentives that motivate her to take action matching the true state in the initial periods rather than maximize her own period payoff. As Morris (2001) calls, the "discipline effect" of the reputation incentives benefits the principal. We show it is optimal for the principal to start the interaction small and increase the stakes gradually. The agent's reputation incentives are managed so that the reputation evolves slowly. We test these predictions in four treatments via online experiments. Each period receives an equal stake in the first treatment. In the other three treatments, the interaction starts mall, and the stakes increase at different speeds. We show that the smaller the interaction starts, the higher the reputation incentives of the agent. More importantly, we show that the principal earns a higher payoff in starting-small (gradualism) treatments than the equal-stakes treatment. We contribute to the literature on gradualism by showing it is a valuable tool to improve equilibrium payoffs in the principal-agent framework with asymmetric information.

JEL classification: C72, C90, D82, D83

Keywords: Reputation, repeated games, gradualism.

* *Department of Economics, Koç University, Rumelifeneri Yolu, Sariyer 34450, Istanbul, Turkey. sbaghirov14@ku.edu.tr.*

† *Department of Economics, Istanbul Bilgi University, Eski Silahtarağa Elektrik Santrali Kazım Karabekir Cad. No:2, Eyupsultan 34060, Istanbul, Turkey. ayca.ebru@bilgi.edu.tr.*

1 Introduction

We frequently participate in interactions that last longer than a single period. In environments with asymmetric information, repeated play becomes valuable. Reputation incentives arise, and the players can obtain higher equilibrium payoffs via updating beliefs in each other. Furthermore, repeated play enhances cooperation in prisoners' dilemma-like settings where it is hard to reach cooperation in one-time interactions.

This paper theoretically and experimentally analyzes a repeated principal-agent interaction. The principal decides whether to delegate a job to the agent whose preferences are unknown. We focus on the setting where there is a sufficiently small probability of the agent being a good type. If the interaction is one-shot, the principal does not hire the agent with such a low reputation level. However, in a repeated play, the principal can monitor the agent's behavior and act accordingly. The principal hires the agent in the first period and continues hiring as long as the agent displays the desired behavior. The principal's hiring behavior creates reputational incentives for the agent to act according to the principal's expectations.

The extent of the reputational incentives depends on how the stakes are allocated across the interaction periods. In the classical studies of repeated communications, each period receives an equal weight. In this paper, we vary the allocation of stakes to find the principal-optimal one. For instance, the agent does not have reputational incentives if the interaction starts with sufficiently high stakes. In other words, if the future communication is not adequately important compared to the current play, then the agent aims to maximize today's payoff rather than focusing on the reputation. It may be the desired behavior in specific environments. However, the principal does not hire the agent in the absence of reputational incentives if her initial reputation is low. The principal-optimal allocation involves starting small and increasing the stakes gradually.

This paper analyzes a three-period repeated principal-agent game. The agent is either a good or a bad type, and the principal does not know the agent's type. The principal hires the informed agent to observe the state and take action accordingly. The good agent shares the same preferences with the principal and prefers to take the action that matches the state. The bad agent, on the other hand, prefers a particular action irrespective of the state. Consider, for example, a state official hiring a private firm for a public project that will last up to three years. At the beginning of each year, the official decides to hire the agent or not. If the agent is hired, she implements the project. At the end of the year, the official observes both the firm's action and whether the execution was adequate. Next, he makes the hiring decision for the next year.

It is optimal for the principal to start the interaction small in such a setting. Starting small leads to reputation incentives that benefit the principal. In the initial periods, the bad agent does not want to reveal her type and displays the principal's desired behavior. Morris (2001) names this as the "discipline effect" of reputation incentives. Following the example above, the government official may have the freedom to choose the project's relative size each year. He can dynamically decide which proportion of the remaining project to implement in each period. This way, he can control the reputational incentives of the agent. Alternatively, a third party may allocate the stakes across the periods. The optimal allocation of stakes is such that the bad agent is indifferent between investing in reputation or not in each period.

In the experiment performed in this paper, we test the predictions of the theory with four treatments. Each treatment involves a particular allocation of the stakes. In one treatment, each period gets an equal stake. There are three treatments of starting small. Stakes grow at a different speed in each treatment. We show that the smaller the interaction starts, the higher the reputation incentives are. The reputation incentives of the agent benefit the principal. The payoff of the principal is higher in the starting small treatments compared to the equal-stakes treatment. Hence, we can conclude the subjects understand the reputation incentives well in the lab environment. If managed, these incentives can improve the equilibrium payoff of the principal. Reputation incentives become more valuable in a setting as presented in this paper in which there is a high risk of hiring the agent.

2 The model

An uninformed principal and an informed agent interact on a project. The project lasts for n periods, and N is the index set of the periods. In each period $i \in N$, the state $\theta_i \in \{0,1\}$ is realized. Each state is equally likely, i.e. $prob[\theta_i = 0] = prob[\theta_i = 1] = \frac{1}{2}$ and independent across the periods.

At the beginning of each period i, the principal decides whether to terminate the project. Principal's period i termination strategy is defined by λ_i where $\lambda_i \in [0,1]$ is the probability of hiring the agent. Once the agent is not hired, the game ends and the payoffs are realized. If $\lambda_i = 1$, then the agent observes the true state and takes action $a_i \in \{0,1\}$. The principal observes a_i and θ_i, updates his belief in agent's type, and the game proceeds to the next period.

The expected continuation payoff of player $y \in \{g,b,P\}$ (good agent, bad agent and the principal respectively) in period i is given as follows: $V_i^y = \sum_{j=1}^{n} \gamma_j U_j^y$ where U_i^y denotes the stage game payoff in period i and $\gamma_i \in \Gamma$ is the stake of the period i such that $\sum_{i=1}^{n} \gamma_i = 1$. γ_i's are common knowledge and constant for all players.

There are two types of agents: $\{g,b\}$ - *good* and *bad* types. The good agent shares the same preferences with the principal, while the bad agent favors the higher action regardless of the state. The agent is privately informed about his type, and the principal believes that the agent is good with the probability $\rho_1 \in (0,1)$. ρ_1 measures the initial reputation of the agent being good.

For any $i \in N$, let H_i be the set of all histories before decision i is made, i.e. $H_i = (\{\gamma_1, o_1\}, \{\gamma_2, o_2\}, ..., \{\gamma_{i-1}, o_{i-1}\})$ where $o_i = (\lambda_i, a_i, \theta_i)$. A period i strategy for the principal is given by $\Lambda_i : H_i \to [0,1]$ where $\Lambda_i(h)$ is the principal's choice of λ_i. Principal's belief about the type of agent is defined by the probability of agent being good in period i: $\rho_i : H_i \to [0,1]$. The agent moves after histories of this type $I_i = (h, \lambda_i, \theta_i)$ where I_i is the period i information set of the agent. The good agent's (mixed) strategy is given by $\mu_i^{\theta_i} = prob(a_i = 0|\theta_i)$ and the bad agent's strategy is given by $\nu_i^{\theta_i} = prob(a_i = 0|\theta_i)$. $\mu_i \& \nu_i : \Im_i \to \Delta[0,1]$ where $\Im_i$ is the set of all period i information sets.

The timing of the stage game in period i can be summarized as follows:

1) The principal decides whether to hire the agent. If $\lambda_i = 0$, then the project is terminated and the payoffs are realized.
2) If $\lambda_i = 1$, nature chooses the state $\theta_i \in \{0,1\}$.
3) The agent observes θ_i and takes the action $a_i \in \{0,1\}$

The stage game payoffs are given by the table:

Good agent (and the principal)

	$\theta = 0$	$\theta = 1$
$a = 0$	k	$k-1$
$a = 1$	$k-1$	k

Bad agent

	$\theta = 0$	$\theta = 1$
$a = 0$	$k-1$	$k-1$
$a = 1$	k	k

where $k \in (0, 1/2)$ is assumed to be the payoff accruing from the project in each period. This constraint on k is made for the following reason. If the principal thinks that the agent chooses her action independent of the state, he does not hire her for the values of k in the defined range. Not hiring the agent gives the payoff of 0 to all players. In other words, if the payoff received from the project (k) is "not sufficiently high", then the outside option is a better choice for the principal relative to facing an agent whose action does not depend on the state.

Period i stage game strategy of the agents may be summarized with the table below:

	$\theta_i = 0$	$\theta_i = 1$
good agent	$0\, with\, probability\, \mu_i^0$ $1\, with\, probability\, 1-\mu_i^0$	0 with probability μ_i^1 $1\, with\, probability\, 1-\mu_i^1$
bad agent	$0\, with\, probability\, \nu_i^0$ $1\, with\, probability\, 1-\nu_i^0$	$0\, with\, probability\, \nu_i^1$ $1\, with\, probability\, 1-\nu_i^1$

We focus on the perfect Bayesian equilibria with Markovian property. In that regard, history matters only in terms of its effect on the agent's reputation. Thus, if the agent has the same reputation level after two different histories, then the principal's action is the same in both cases. The principal's decision depends on the past observations only through their effect on the reputation of the agent. More formally, for any $i \in N$ and $h, h' \in H_i$: $\rho_i(h) = \rho_i(h')$ implies $\lambda_i(a_i|h) = \lambda_i(a_i|h')$ and $\mu_i(h) = \mu_i(h')$ and $\nu_i^{\theta_i}(h) = \nu_i^{\theta_i}(h')$.

Proposition 1. *In any perfect Bayesian equilibrium:*

1) *There is a reputation incentive for the agent to play action* 0 *when the state is* 0*. More precisely,*

$$\rho_{i+1}(a_i = 0, \theta_i = 0) > \rho_{i+1}(a_i = 1, \theta_i = 0) \;\; \forall i \text{ and } \forall h \in H_i$$

$\rho_{i+1}(0,0) > \rho_{i+1}(1,0)$ *implies* $\mu_i^0 = 1$.

2) *If* μ_i^1 *and/or* ν_i^1 *is greater than zero, then* $\rho_{i+1}(0,1) > \rho_{i+1}(1,1)$ *and* $\nu_i^1 < \mu_i^1$ *by Bayes' rule.*

Proof. All proofs are relegated to Appendix B. □

Note that the good agent has no incentive to act differently from the state observed when $\theta_i = 0$. Thus, playing action 1 with a positive probability when the state is 0 can not be an equilibrium strategy for the good agent. It is easy to see as playing action 1 hurts both the stage game payoff and the reputation, hence $\mu_i^0 = 1$ in any equilibrium. For ease of exposition, we will write μ_i to refer to μ_i^1.

Given the above-defined strategies, the principal's belief about the type of the agent in period $i+1$ - $\rho_{i+1}(a_i, \theta_i)$ - is updated by Bayes' rule as follows:

$$\rho_{i+1}(0,0) = \frac{\rho_i}{\rho_i + (1-\rho_i)\nu_i^0}$$

$$\rho_{i+1}(1,0) = 0$$

$$\rho_{i+1}(0,1) = \frac{\rho_i \mu_i}{\rho_i \mu_i + (1-\rho_i)\nu_i^1}$$

$$\rho_{i+1}(1,1) = \frac{\rho_i (1-\mu_i)}{\rho_i (1-\mu_i) + (1-\rho_i)\left(1-\nu_i^1\right)}$$

Prediction A.

1) The principal's belief that the agent is good type will be 0 after observing $\{a_i, \theta_i\} = \{1,0\}$ in any treatment.

2) The principal's belief that the agent is good type will be the highest after $\{0,0\}$. More precisely, $\rho_{i+1}(0,0) \geq \rho_{i+1}(1,1) > \rho_{i+1}(1,0)$.

3) The principal will hire the agent more frequently after $\{0,0\}$.

Note that 2 implies 3.

3 Experimental Design.

The interaction between the principal and the agent in a three-period interaction ($n = 3$) is analyzed in the experiment. The experiments are conducted online via oTree (Chen et al., 2016).We set $\rho_1 = 0.1$ and $k = \frac{1}{3}$. The experiment comprises four treatments. The independent variable varying across treatments is the allocation of stakes.

- Treatment 1 - ×1 treatment: $\Gamma_1 = \left\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right\}$. It is the baseline treatment where each period receives an equal stake. In the allocation Γ_1, stakes remain the same for all the periods. It is baseline allocation because it corresponds to the classical repeated game analysis where each stage game is the same.

- Treatment 2 - ×2 treatment: $\Gamma_2 = \left\{\frac{1}{7}, \frac{2}{7}, \frac{4}{7}\right\}$. It is one of the three "starting small" treatments. In this treatment, the interaction starts small but increases at a slow speed. The relative stakes increase by multiplying 2.

- Treatment 3 - ×3 treatment: $\Gamma_3 = \left\{\frac{1}{13}, \frac{3}{13}, \frac{9}{13}\right\}$. This is the starting small treatment where the stakes evolve at a higher speed. The relative stakes increase by multiplying 3.

- Treatment 4 - ×4 treatment: $\Gamma_4 = \left\{\frac{1}{21}, \frac{4}{21}, \frac{16}{21}\right\}$. In this treatment, the stakes start small and increase fast. The relative stakes increase by multiplying 4.

Each treatment is tested in a different session, so we have a between-subject design. There are 20 subjects in each session. Each subject plays 15 super-games (3-period games), each with a new opponent (stranger matching). Half of the subjects are randomly selected to be principal at the beginning of the experiment. One of the ten agents ($\rho_1 = 0.1$) is randomly selected to be the good type and the remaining is allocated to the bad type. The subject's role as the principal or the agent, and type remains the same throughout the session. Neutral labels are used for types of agents. Type 1 and 2 correspond to good and bad types, respectively. After the roles are distributed, all subjects were informed about the distribution of types and the difference in different types' payoff scheme. 0 and 1 in the state and the action sets are labeled as green and blue, respectively.

The timing of each super-game is as follows:

1) The subjects are matched.

2) The principal is asked about his belief in the probability of the agent being good (non-incentivized).

3) Principal chooses λ_1 - decides whether to hire the agent.

4) (If $\lambda_1 = 1$) Computer randomly chooses green or blue. (θ_1 is set)

5) The agent observes the computer's choice and takes action - green or blue.

6) The period ends, principal observes both the agent's and the computer's color choice. The next period starts, and the same steps 2-5 are repeated.

The above steps repeat as long as the principal continues to hire the agent. If the principal hires the agent for all the periods, the game ends after period 3. In each period, the principal observes past choices of the agent, states, and his payoff before making the decision. The principal does not observe the action of the agent from previous super-games. Thus, we avoid building a reputation between the super-games. In each session, 15 super-games are played to enable learning. 2 of 15 super-games are randomly selected, and the subject's earnings from these two periods are paid. 20TL participation fee is paid to each subject. The average session length was 100 minutes, and the average earning was 52TL.

The timing of the experiment is as follows:

1) Instructions explained.

2) Incentivized quiz (2 TL per correct answer).

3) 15 Super-games.

4) Incentivized Holt-Laury (Holt and Laury, 2002) task.

5) Survey, demographics (10 TL is paid for completing the survey).

Exchange rates are arranged across treatments so that maximum and minimum possible earnings in each super-game corresponds to 10 TL and -20 TL in all treatments. Detailed comparison of parameters across treatments is relegated to Appendix A.

4 Theory and Predictions.

We are trying to find the principal-optimal allocation of γ_i when there is a sufficiently small probability that the agent is the good type. Theory predicts $\gamma_i = \frac{1-k}{1-k^n}k^{n-i}$. We test $k = \frac{1}{3}$ experimentally, hence, the predicted optimal stakes allocation is as follows: $\gamma_i = \frac{3^{i-1}}{27}$. This corresponds to Γ_3 where the stakes evolve by multiplying by 3. For this allocation of stakes, the bad agent is indifferent between both actions in each period, and the good agent prefers action 0. The principal hires the agent after observing action 0 and does not hire after action 1. In this formation, the reputation of the agent increases gradually through the interaction. We assume that the principal sets the relative stakes for convenience to define the principal-optimal allocation. Note that this assumption does not alter the equilibrium analysis as any third party trying to maximize the principal's expected continuation payoff can set the stakes.

There are two possible equilibrium behavior: (1) Babbling when $\theta_i = 1$ and (2) reputation incentive when $\theta_i = 1$. In the first case, both types of agent play action 1 when $\theta_i = 1$, i.e. there is no reputation update. The principal's belief $\rho_{i+1}(0,1)$ is sufficiently low so that the agent has no incentive to play action 0. This behavior is mostly observed in the experiment; hence we will focus on the babbling equilibrium through the main text. In the alternative equilibrium, there may be reputation incentives given $\theta_i = 1$. If so, then $\mu_i^1 > \nu_i^1$ by Bayes rule so that $\rho_{i+1}(0,1) > \rho_{i+1}(1,1)$. Alternative equilibrium strategies and payoff are detailed in the proof of each proposition.

The bad agent plays action 0 with a positive probability in the first and the second periods when the state is 0. The mixing probability of the agent must in a certain range to satisfy the equilibrium beliefs. The reputation incentives of the bad agent has two-dimensional effect on the expected payoff of the principal. The principal's period i payoff is increasing in ν_i^0. It corresponds to the discipline effect of reputation incentives. On the other hand, for large enough k, the expected continuation payoff of the principal is decreasing in ν_i^0. The reason is that, higher ν_i^0 implies the bad agent survives till the next period with a higher probability. We focus on large k values in this analysis. We use principal-optimality as the equilibrium selection criteria. Principal-optimality is used for determining ν_i^0. As the expected continuation payoff of the principal is decreasing in ν_i^0, the agent will play action 0 with the lowest probability satisfying the equilibrium conditions. The agent is indifferent between different probabilities, hence this is also the agent-optimal equilibrium.

The following proposition formalizes the principal-optimal equilibrium.

Proposition 2. *If we start with a sufficiently low initial reputation ($\rho_1 << (1-2k)^2$) and $k \geq \frac{1}{4}$, $\gamma_i = \frac{1-k}{1-k^3}k^{3-i}$ is the principal-optimal allocation of stakes and the following strategies constitute the principal-and-agent-optimal equilibrium:*

$$v_1^0 = \frac{4k\rho_1(1-k)}{(1-\rho_1)(1-2k)^2}, \; v_2^0 = \frac{2k\rho_1}{(1-\rho_1)(1-2k)}$$

$$v_1^1 \& v_2^1 = 0$$

$$\mu_1^1 \& \mu_2^1 = 0$$

$$\lambda_{i+1}(0,0) = 1, \; \lambda_{i+1}(1,\theta_i) \& \lambda_{i+1}(0,1) = 0 \; for \, \theta_i \in \{0,1\} \, and \, i \in \{2,3\}$$

Proof. Appendix. □

The principal's ex-ante expected payoff is the highest if the stakes are allocated according to Γ_3. The principal hires the agent if $\{a_i, \theta_i\} = \{0,0\}$ is observed.

Prediction B. The principal gets the highest payoff in the treatment Γ_3**.**

In the following subsections, we will define equilibrium behaviors in each treatment. Refer to Appendix for the comparison of treatment parameters.

4.1 Equal stakes - Γ_1.

Proposition 3. *Set $\rho_1 = 0.1$ and $k = \frac{1}{3}$ where the allocation of stakes is given by $\Gamma_1 = \left\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right\}$. Assuming $\lambda_1 = 1$, the following assessment constitutes the unique Perfect Bayesian equilibrium:*

$v_i^{\theta_i} = 0 \, for \; \theta_i \in \{0,1\}$ *and* $\mu_i = 0 \; \forall i \in \{1,2,3\}$.

$\lambda_2(0,0) = 1, \; \lambda_2(1,0) = \lambda_2(1,1) = \lambda_2(0,0) = 0$

Proof. Appendix. □

The principal earns an expected continuation payoff of -0.027. Hence, the principal does not hire the agent in the first period. We expected the principal to hire the agent with a positive frequency in the experiments because of either home-made priors or risk preferences. Thus, the equilibrium behavior is defined by assuming $\lambda_1 = 1$.

This allocation of stakes implies there is no reputational incentive in any period. The evolution of reputation is following in the above equilibrium:

- $\rho_2(0,0) = 1$ & $\rho_2(1,1) = \rho_1$
- $\rho_2(1,0) = \rho_2(0,1) = 0$

If $(0,0)$ is observed, then the principal knows the agent is the good type. The principal hires the agent only after this observation.

Prediction C

1) In Treatment 1, There is no reputational incentive.
2) The smaller the interaction starts, the higher the reputational incentives.

4.2 Starting small - stakes increase slowly: Treatment 2 - Γ_2.

Proposition 4. *Set $\rho_1 = 0.1$ and $k = \frac{1}{3}$ where the allocation of stakes is given by $\Gamma_2 = \left\{\frac{1}{7}, \frac{2}{7}, \frac{4}{7}\right\}$. The following assessment constitutes the principal-and-agent-optimal equilibrium:*

$$v_1^0 = \tfrac{2\rho_1}{1-\rho_1} = \tfrac{2}{9}\,,\ v_2^0 = 0$$

$$v_i^1 = \mu_i = 0\ for\ \forall i \in \{1,2,3\}.$$

$$\lambda_2(0,0) = 1,\ \lambda_2(1,0) = \lambda_2(1,1) = \lambda_2(0,1) = 0$$

$$\lambda_3(0,0) = 1,\ \lambda_3(1,1) = \tfrac{1}{2},\ \lambda_3(1,0) = \lambda_3(0,1) = 0$$

Proof. Appendix. □

The principal has an expected continuation payoff of 0.0024. The agent has a reputation incentive only in the first period.

The evolution of reputation is as follows in the above equilibrium:

- $\rho_2(0,0) = \frac{1}{3}$ & $\rho_2(1,1) = \rho_1$
- $\rho_2(1,0) = \rho_2(0,1) = 0$
- $\rho_3(0,0) = 1$ & $\rho_3(1,1) = \rho_2$
- $\rho_3(1,0) = \rho_3(0,1) = 0$

Note that there are higher reputational incentives in Treatment 2 compared to Treatment 1. Hence, observing $(0,0)$ does not increase the reputation as much. In other words, the reputation of the agent evolves slower in Treatment 2 compared to Treatment 1.

4.3 Starting small - stakes increase in medium speed: Treatment 3 - Γ_3.

Proposition 5. *Set* $\rho_1 = 0.1$ *and* $k = \dfrac{1}{3}$ *where the allocation of stakes is given by* $\Gamma_3 = \left\{\dfrac{1}{13}, \dfrac{3}{13}, \dfrac{9}{13}\right\}$*. The following assessment constitutes the principal-optimal equilibrium:*

$v_1^0 = \frac{8\rho_1}{1-\rho_1} = \frac{8}{9}$, $v_2^0 = \frac{2\rho_2}{1-\rho_2}$

v_i^1 & $\mu_i = 0$ *for* $i \in \{1,2,3\}$

$\lambda_i(0,0) = 1,\ \lambda_i(1,0) = \lambda_i(1,1) = \lambda_i(0,1) = 0$ *for* $i \in \{2,3\}$

Proof. Appendix. □

The principal has an expected continuation payoff of 0.0218. Note that the agent has a reputation incentive in all periods except the last one. The reputation will evolve slower under this allocation compared to Treatment 1 and Treatment 2. From Proposition 2, Γ_3 is the optimal allocation of the stakes for the principal.

As there are more reputation incentives, the reputation of the agent will evolve slower in this setup. The evolution of reputation is following in the above equilibrium:

- $\rho_2(0,0) = \frac{1}{9}$ & $\rho_2(0,1) < \frac{1}{9}$
- $\rho_2(1,0) = 0$ & $\rho_2(1,1) = \rho_1 = \dfrac{1}{10}$

- $\rho_3(0,0) = \frac{1}{3} \& \rho_3(0,1) < \frac{1}{3}$
- $\rho_3(1,0) = 0 \& \rho_3(1,1) = \rho_2$

4.4 Starting small - stakes increase fast: Treatment 4 - Γ_4.

Proposition 6. *Set $\rho_1 = 0.1$ and $k = \frac{1}{3}$ where the allocation of stakes is given by $\Gamma_4 = \left\{\frac{1}{21}, \frac{4}{21}, \frac{16}{21}\right\}$. The following assessment constitutes the principal-optimal equilibrium:*

$v_1^0 = \frac{8\rho_1}{1-\rho_1} = \frac{8}{9}$, $v_2^0 = \frac{2\rho_2}{1-\rho_2}$

$v_i^1 \& \mu_i = 0$ *for* $i \in \{1,2,3\}$

$\lambda_2(0,0) = \lambda_3(0,0) = \frac{3}{4}$,

$\lambda_i(0,1) = \lambda_i(1,0) = \lambda_i(1,1) = 0$ *for* $i \in \{2,3\}$.

Proof. Appendix. □

The first thing to realize is that this allocation of stakes implies high reputation incentives. Hence, if the principal hires the agent after observing $(0,\theta_i)$,then the agent's best response will be to play action 0 with a probability of 1 Then, the reputation is not updated. As the initial reputation is sufficiently low, this behavior can not constitute an equilibrium. Thus, the principal plays a mixed strategy. The principal gets the expected payoff of 0.013 in this setup.

The evolution of reputation is as follows in the above equilibrium:

- $\rho_2(0,0) = \frac{1}{9}$ & $\rho_2(0,1) < \frac{1}{9}$
- $\rho_2(1,0) = 0 \& \rho_2(1,1) = \rho_1$
- $\rho_3(0,0) = \frac{1}{3} \& \rho_3(0,1) < \frac{1}{3}$
- $\rho_3(1,0) = 0 \& \rho_3(1,1) = \rho_2$

Prediction D.

1) In Treatment 3&4, the principal will not hire the agent project after action 1 in each period.

2) The principal will hire the agent more often after observing action 0 Treatment 3 compared to Treatment 4.

4.5 Payoff comparison.

Theory predicts that when the agent's initial reputation is sufficiently low, then the principal-optimal allocation of relative stakes is such that the bad agent is indifferent between two actions in each period. The stakes increase gradually in the equilibrium; hence our findings comply with the literature on gradualism. If the interaction starts with a higher stake and the stakes evolve slower (Treatment 2), then there will not be sufficient reputation incentives for the agent. On the other hand, if the interaction starts very small and the stakes increase rapidly (Treatment 4, $\times 4$ speed), the reputation incentive will be so high that the principal will be unable to update his belief in the type of agent. Thus, gradualism with the right speed ($\times 3$) is expected to lead to the highest payoff.

Prediction E. The ranking of the allocations in terms of the expected payoff of the principal is as follows: $T3 > T4 > T2 > T1$ [1].

It states that starting the interaction small is better for the principal if the initial reputation is sufficiently low. However, if the stakes do not increase at the right speed (x3), there will be an excess of or a deficit of reputation incentive. Hence, the result will be sub-optimal for the principal. If the interaction starts smaller than the equilibrium stake and evolves faster, there will be high reputational incentives. However, three types of gradualism are preferred to starting small.

5 Results

The behavior of the subject is in accordance with the predictions of theory to some extent. The reputational incentives and the principal's average payoff are observed to vary across treatments. As expected, the smaller the interaction starts, the more reputational incentives are there.

The first thing to note is that the good agent never plays the action different from the state. They follow the babbling strategy.

The main divergence from the predictions occurs when the bad agent plays action 1 when the state is equal to 0. The good agent has no motivation to act accordingly; hence the principal is expected to not hire the agent after observing such an outcome. However, the principal hires the agent in 64.6% of all such occurrences. This behavior is be explained with risk incentives and subjective beliefs about the next period's state in the following subsections.

[1] $U^P(T3) = 0.0218$, $U^P(T4) = 0.013$, $U^P(T2) = 0.0024$, $U^P(T1) = -0.027$

5.1 Agent Behavior and Reputation Incentives.

The reputation incentive is measured by the frequency of the bad agent playing action 0, more precisely, ν_i^0. Note that there is no reputational incentive in the last period, hence we expect $\nu_3^{\theta_3} = 0$ in any treatment. It is observed in the data that no bad agent plays action 0 in the last period. Hence, we can conclude that $\nu_i^0 > 0$ stems from reputation incentives rather than altruism of the bad agent. The smaller the interaction starts, the more reputational incentives are. Hence, we expect $\nu_i(T1) < \nu_i(T2) < \nu_i(T3) < \nu_i(T4)$ for $i = \{1,2\}$. We focus on first-period reputation incentives because second-period reputation incentives depend on first period outcomes in addition to the allocation of stakes. A similar trend is also observed in the second period, although insignificant; hence we do not report them. The observed ν_1^0 frequencies and the results of logit regression are reported in Table 1. Prediction C2 is satisfied.

Note that, theory predicts two possible equilibrium behavior:

1) Babbling when $\theta_i = 1$. In this equilibrium, both types of agent play action 1 if $\theta_i = 1$ is observed. The principal's out-of-equilibrium path belief in this equilibrium must be $\rho_{i+1}(0,1) < \bar{\rho}_{i+1}$ so that the agent has no reputational incentive to play action 0. This equilibrium behavior is observed in all treatments expect for some occurances in treatment 3.

2) Reputation when $\theta_i = 1$. We observe this equilibrium behavior only with the bad agent in Treatment 3 (in the 16 of the total 72 occurrences of $\theta_1 = 1$). $\nu_2^1 = 0.22$ is observed and $\rho_2(0,1)$ is reported to be 49.1 on average.

As the second behavior is rarely observed and only in one treatment, we refine our focus on $\theta_1 = 0$ for the analysis of reputational incentives.

5.2 Beliefs.

When a period starts, the principal is asked about his belief in the agent's probability of being the good type. At the beginning of the first period, they are informed that there is a 10% chance of matching with a good agent. 85% of the principal reported belief of 10% in the first period, and the mean belief is 13%, which is reasonably close to the objective probability of the agent being good type.

If $a_i = 1$ and $\theta_i = 0$ is observed in any period, then it implies that the agent is bad type, i.e. $\rho_{i+1}(1,0) = 0$. If $a_1 = 1$ and $\theta_1 = 0$ is observed, 87% of the principals reported belief equal or below 10% (54,5% being equal to zero). Even though it is significantly away from the expected belief of 0, we observe that $\{a_1, \theta_1\} = \{1,0\}$ decreases the

Table 1

Reputation incentives across treatments.

	(1) v_1^0	(2) *rep*
Treatment 1	0.098	
	(0.04)	
Treatment 2	0.324	0.83^{***}
	(0.06)	(0.27)
Treatment 3	0.492	1.27^{***}
	(0.06)	(0.27)
Treatment 4	0.81	2.18^{***}
	(0.05)	(0.28)
Constant		-1.29^{***}
		(0.22)
N		263

(1) Observed average probabilities of bad agent playing action 0 given the state of 0. Standard errors in parentheses.
(2) Logit regression of probability of bad agent playing action 0. Standard errors in parentheses.
Dependent variable is reputation. It takes value 1 if the bad agent plays action 0, 0 otherwise given state is equal to 0 in the first period.
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

agent's reputation. If we define $\Delta\rho_{21}$ as $\rho_2 - \rho_1$, the mean $\Delta\rho_{21}$ is equal to $-3,2$, and $91,84\%$ of observed $\Delta\rho_{21}$ values are non-positive. In percentage terms, the average change in beliefs is -112%. $\Delta\rho_{32}(1,0)$ has also a negative mean of -4.5.

As expected, $\{a_1, \theta_1\} = \{0,0\}$ increases the agent's reputation for all treatments. $\Delta\rho_{21}(0,0)$ has an average of 28.65 across all treatments where the average $\rho_{21}(0,0) = 42.76$. We expect it to be higher in the treatments with lower reputational incentives; however, no such trend is observed. $\Delta\rho_{32}(0,0)$ has a mean of 18, which again holds with our expectation of reputation concerns at force.

The principals seem to be more optimistic than the theory predicts when $\{a_1, \theta_1\} = \{1,1\}$ is observed. At best, we expect there to be no reputation loss, i.e., no reputation update. However, the average $\Delta\rho_{21}(1,1)$ is equal to 5.3 across all treatments, significantly different from 0. On the other hand, $\Delta\rho_{32}(1,1)$ is equal to 4.9, which is not different from 0 at 99% significance level.

The belief updating is far away from Bayesian updating; however, the sign of change in beliefs follows what theory predicts. Table 2 shows that Prediction A2 holds.

Table 2

Belief update

$\{a_{i-1}, \theta_{i-1}\}$	ρ_2	ρ_3
$\{0,1\}$	46.41^{***}	16.66^{**}
	(6.31)	(7.92)
$\{0,0\}$	42.77^{***}	47.78^{***}
	(2.33)	(4.28)
$\{1,1\}$	18.61^{***}	31.76^{***}
	(1.27)	(1.94)
$\{1,0\}$	10.61^{***}	19.74^{***}
	(1.84)	(2.8)
Constant	No	No
Subject fixed effects	Yes	Yes
R-squared	0.534	0.451

OLS regression results of reputation on previous period's observation. Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

In addition to the principal's belief about the agent's type, there seems to be another subjective belief in force. In the survey at the end of the experiment, the subjects were asked about their beliefs in the next period's state being 0, given that the current period's state is 0. Most of them reported beliefs ($\pi(\theta_{i+1} = 0|\, \theta_i = 0)$) less than 50%. This belief, alongside risk behavior, can explain why some of the principals continued hiring the agent after observing $\{a_1, \theta_1\} = \{1, 0\}$. Even if they knew that there is a small or no probability that the agent is a good type, they subjectively believed that the probability of two consecutive period's states being 0 is sufficiently low; hence they hired the agent. Furthermore, they reported lower belief on the third period's state being 0 given state 0 is observed in the first two periods ($\pi(\theta_3 = 0|\, \theta_1 = 0\,\&\,\theta_2 = 0)$).

Note that $\pi(\theta_{i+1} = 0|\, \theta_i = 0)$ first-order stochastic dominates $\pi(\theta_3 = 0|\, \theta_1 = 0\,\&\,\theta_2 = 0)$, which is no surprise (Table 3). The subjects form a subjective belief in the state such that there is a less than $1/2$ probability of observing state 0 in two consecutive periods. Observing 0 in all three periods is expected to be even less likely.

Table 3

Frequency of subjective beliefs in the state.

	0.5	0.4-0.5	0.3-0.4	0.2-0.3	<0.2	N
$\pi(\theta_3 = 0 \vert\, \theta_1 = 0 \,\&\, \theta_2 = 0)$	45%	10%	18.75%	10%	16.25%	80
Cumulative	45%	55%	73.75%	83.75%	100%	
$\pi(\theta_{i+1} = 0 \vert\, \theta_i = 0)$	55%	10%	15%	10%	10%	80
Cumulative	55%	65%	80%	90%	100%	

$\pi(\theta_{i+1}|x)$ denotes reported subjective belief in θ_{i+1} given x is realized.

5.3 Principal behavior.

The principal hires the agent if he expects a non-negative payoff. A risk-neutral principal with no subjective belief in the state does not hire the agent if he knows that the agent is a bad type with a high probability. In the particular case where the agent plays action 1 while the state is 0, the principal should not hire the agent. However, we observe a positive hiring rate after such an occurrence. In general, the hiring rates are higher than the prediction of the theory. It was expected. Home-made priors about either the type of the agent or the state in the next period combined with risk preferences may lead the principal to hire the agent even in circumstances where a risk-neutral principal is expected not to hire the agent.

The behavior pattern, on the other hand, is in line with the predictions of the theory. The higher the principal's reported belief in the agent's type, the higher the hiring rate is observed. Moreover, the hiring rate is the highest after $\{a_1, \theta_1\} = \{0,0\}$. The agent is the least hired after $\{a_1, \theta_1\} = \{1,0\}$. As discussed earlier, we expected this rate to be 0. Although this prediction is not met, the hiring rate is still sufficiently low compared to other observations in both the first and second periods. Table 4 summarizes the logit regression results. Prediction A.3 is mostly met, as can be seen from the table.

5.4 Payoff Comparison.

The principal prefers to start the interaction small if the agent is bad type with a sufficiently high probability. The intuition is that the principal can benefit from the reputation incentive of the agent. Morris (2001) defines three effects of the reputation incentives. First is the "discipline effect," which is responsible for the biased agent to play the action matching the state in the initial periods. It is the most prominent effect in our setup. The smaller the interaction starts, the higher the reputation incentives, hence the higher the discipline effect. The second is the "sorting effect," which

Table 4

Hiring behavior

	T1		T2		T3		T4	
$\{a_{i-1}, \theta_{i-1}\}$	$i=2$	$i=3$	$i=2$	$i=3$	$i=2$	$i=3$	$i=2$	$i=3$
$\{0,0\}$	1.79***	1.79***	2.6***	1.7***	2.14***	1.87***	3.46***	1.25***
	(0.54)	(0.54)	(0.52)	(0.54)	(0.37)	(0.54)	(0.51)	(0.33)
$\{1,1\}$	1.13***	0.69***	1.18***	0.31	2.08***	0.44**	3.46***	1.02***
	(0.19)	(0.19)	(0.2)	(0.2)	(0.28)	(0.19)	(0.51)	(0.19)
$\{1,0\}$	0.89***	0.82***	0.25	0.41*	0.33	−0.17	1.7***	0.27
	(0.21)	(0.26)	(0.21)	(0.23)	(0.26)	(0.21)	(0.54)	(0.22)
N	294	220	290	212	296	238	290	278

Logit regression. Standard errors in parentheses.

Dependent variable is hiring. It takes value 1 if the agent is hired, 0 otherwise.

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

enables the principal to update his belief about the agent's type. In the babbling equilibrium - which is observed in most of the games - this effect is reversed. The principal can not update his belief after $\theta_i = 1$ if both agents play action 1 with a probability of 1. On the other hand, if $\theta_i = 0$,the principal would know the agent's type in the absence of reputation incentives. In such an environment, the good agent will play $a_i = 0$, and the bad agent will play $a_i = 1$. The strategies imply $\rho_{i+1}(0,0) = 1$ and $\rho_{i+1}(1,0) = 1$ by Bayes' rule. The third effect is the "political correctness effect," which arises when the good agent plays action 0 for the sake of increasing reputation when the state is 1. However, we observe that the good agent always plays the action matching the true state. Hence, this effect also does not apply to the setting of this paper.

To compare the principals' payoff across treatments, We collect the strategies across all periods and run a simulation based on the strategies. Approximately 1000 observations per treatment are obtained. Two simulations are conducted. In the first one, all the strategies are gathered. We observe that the payoffs in Treatment 2 and Treatment 4 are significantly higher than those in Treatment 1. There is no other significant difference. With this data, we can conclude that starting small is preferred to the equal-stakes treatment. (Table 5)

In the second simulation, we focus on the strategies of slightly risk-neutral subjects. We have a self-assessment risk question and an incentivized version of the Holt-Laury task (Holt and Laury, 2002) to measure the risk preferences. In this simulation, we exclude the subjects in the extremes in either of these two risk assessment methods. More precisely, the subjects who reported self-riskiness less than 2 and higher than 8 (on a scale between 0 and 10) and the subjects whose switching point is greater than 9 or less than 2 at the Holt-Laury task are excluded. The reason for making a weak restriction such as relative risk neutrality is not to lose many observations. Here we observe all

Table 5

Payoff comparison

General	$T4 > T2 > T3 > T1, \quad T2 >^{**} T1, \quad T4 >^{**} T1$
Slightly risk neutral	$T4 > T3 > T2 >^{***} T1, \quad T4 >^{***} T1, \quad T3 >^{*} T1,$

Wilcoxon rank-sum (Mann-Whitney) test.

${}^{*}p < 0.10$, ${}^{**}p < 0.05$, ${}^{***}p < 0.01$.

three starting small treatments to lead to a higher payoff to the principal than the equal stakes treatment. No significant difference among the starting small treatments is observed. As the expected payoffs in these treatments are close, it is different to identify a significant difference among them through experiments.

We run an additional simulation where the "consistency" restriction is added to the above-defined slightly risk-neutral agents. Consistency is defined as not hiring the agent if action 1 is observed given the state is 0. The reason for using observations rather than beliefs to define consistency is that a considerable amount of principals report a positive belief after such outcomes. It may stem from home-made priors or a belief that a good agent may have made a mistake. Moreover, another belief that may lead the principal to hire the agent is the subjective belief in the next period's state. As discussed earlier, the subjects tend to believe the next period's state will be 1 with a probability higher than 0.5 given $\theta = 0$ is observed in the current period. Hence, by focusing on the hiring outcome for the restriction, we can control for both these beliefs. We do not exclude the whole super-game in which $\lambda_{i+1}(1,0) = 1$ for some i. Rather, We include the super-game strategies in the simulation and make correction for $\lambda_{i+1}(1,0)$ as $\lambda_{i+1}(1,0) = 0$. In this simulation, starting small treatments lead to higher payoffs again. Moreover, Treatment 3 leads to a significantly higher payoff compared to Treatment 2. As the expected payoff in Treatment 3 and 4 is close, we cannot conclude on the comparison.[2]

6 Concluding remarks

Repeated play becomes valuable in settings of uncertainty and asymmetric information. In a repeated interaction between a principal and an agent, the agent holds reputation incentives. Attaining a high reputation level is vital for the agent because it determines the principal's future hiring decisions. We show that reputation incentives depend on the relative importance of periods of interaction. Investing in reputation may be costly. In such circumstances, the agent faces a trade-off between taking the costly action that will increase her reputation and the action that maximizes the period payoff. For the agent to invest in reputation, the future interaction's stake should be sufficiently high. The main

[2] $T4 > T3 >^{**} T2 >^{***} T1, \quad T4 >^{***} T1$, where ${}^{*}p < 0.10$, ${}^{**}p < 0.05$, ${}^{***}p < 0.01$.

focus of this paper is whether the principal benefits from the reputation incentives of the agent. We theoretically and experimentally show that it is optimal for the principal to start the interaction small in an environment where there is a low probability that the agent is a good type. Reputation incentives motivate the bad agent to play the action matching the real state. This paper shows that we can measure reputation incentives via online experiments. We observe that both principals and agents comprehend the reputation incentives and act accordingly. The higher the future stakes, the higher the reputation incentives are. In the baseline treatment, each period receives an equal stake. It corresponds to the classic repeated interaction without discounting. In this setup, the importance of decisions does not vary across the periods. If the interaction lasts for three periods, we expected there to be no reputation incentives. We observed small reputation incentives. In the starting small treatments, higher reputation incentives are observed. More precisely, the faster the stakes increase, the higher the reputation concerns are there.

Principals' belief updates and hiring frequencies are also observed to align with the theory's predictions. The hiring frequency is the highest after the agent and the computer choose $\{$green, green$\}$. As expected, the principal hires the agent the least after the agent chooses blue when the computer's choice is green. This outcome signals that the agent is a bad type for sure. However, we observed a positive hiring rate. We explain this with two phenomena. First, our theoretical results rely on the risk neutrality assumption. The subjects might want to take a risk and hire the agent even if they knew that she is the bad type. Second, the subjects might put a probability less than 1/2 to the possibility of the computer choosing green in two consecutive periods. If a principal knows that the agent is a bad type and believes that the computer will choose blue with sufficiently high probability in the next period, it is rational to hire the agent. Answers to the related survey questions justify this statement. Many subjects reported that they believed the computer would choose green with a probability less than 1/2 in the next period given that it has chosen green in the current period.

We conclude that starting the interaction small improves the principal's payoff in an environment where there is high uncertainty. It helps the principal in three ways. First, as the principal is cautious about the agent, starting with small stakes means lower risk. Second, starting small and gradually increasing the stakes "disciplines" the bad agent to choose the same color as the computer. Lastly, the principal can observe the agent's behavior and decrease the uncertainty in the further periods where there are higher stakes.

Appendix A.

Comparison of Treatment parameters.

Table 6

Treatment parameters.

	Treatment 1	Treatment 2	Treatment 3	Treatment 4
Stakes	$\left\{\frac{1}{3},\frac{1}{3},\frac{1}{3}\right\}$	$\left\{\frac{1}{7},\frac{2}{7},\frac{4}{7}\right\}$	$\left\{\frac{1}{13},\frac{3}{13},\frac{9}{13}\right\}$	$\left\{\frac{1}{21},\frac{4}{21},\frac{16}{21}\right\}$
Show-up fee	20TL	20TL	20TL	20TL
Exchange rate - e	$1/3$	$1/7$	$1/13$	$1/21$
Max. ECU/super-game	30	70	130	210
Max. TL/super-game	10	10	10	10
Min. ECU/super-game	-60	-140	-260	-420
Min. TL/super-game	-20	-20	-20	-20
Quiz ECU/correct answer	6	14	26	42
Survey payment	10TL	10TL	10TL	10TL
Halt-Laury task (ECU)	A: 36 vs 27 B: 60 vs 9	A: 90 vs 60 B: 140 vs 20	A: 160 vs 120 B: 260 vs 40	A: 250 vs 190 B: 420 vs 60

Treatment 1.

The subjects' payoff matrix in each period:

Type 1 agent (and the principal)

	$\theta = 0$	$\theta = 1$
$a = 0$	10	-20
$a = 1$	-20	10

Type 2 agent

	$\theta = 0$	$\theta = 1$
$a = 0$	-20	-20
$a = 1$	10	10

The payoff table above implies that the rational agent has no incentive to differentiate between her actions depending on the state. However, if the agent has home-made other-regarding preferences, then we expect to observe the bad agent playing action 0 more frequently when $\theta = 0$ compared to $\theta = 1$.

The ECU is converted to TRY with exchange rate $e = 1/3$.

Treatment 2.

The subjects' payoff matrix in each period:

Period 1

Type 1 agent (and the principal)

	$\theta = 0$	$\theta = 1$
$a = 0$	10	-20
$a = 1$	-20	10

Type 2 agent

	$\theta = 0$	$\theta = 1$
$a = 0$	-20	-20
$a = 1$	10	10

Period 2

Type 1 agent (and the principal)

	$\theta = 0$	$\theta = 1$
$a = 0$	20	-40
$a = 1$	-40	20

Type 2 agent

	$\theta = 0$	$\theta = 1$
$a = 0$	-40	-40
$a = 1$	20	20

Period 3

Type 1 agent (and the principal)

	$\theta = 0$	$\theta = 1$
$a = 0$	40	-80
$a = 1$	-80	40

Type 2 agent

	$\theta = 0$	$\theta = 1$
$a = 0$	-80	-80
$a = 1$	40	40

The ECU is converted to TRY with exchange rate $e = 1/7$.

Treatment 3.

The subjects' payoff matrix in each period:

Period 1

Type 1 agent (and the principal)

	$\theta = 0$	$\theta = 1$
$a = 0$	10	-20
$a = 1$	-20	10

Type 2 agent

	$\theta = 0$	$\theta = 1$
$a = 0$	-20	-20
$a = 1$	10	10

Period 2

Type 1 agent (and the principal)

	$\theta = 0$	$\theta = 1$
$a = 0$	30	-60
$a = 1$	-60	30

Type 2 agent

	$\theta = 0$	$\theta = 1$
$a = 0$	-60	-60
$a = 1$	30	30

Period 3

Type 1 agent (and the principal)

	$\theta = 0$	$\theta = 1$
$a = 0$	90	-180
$a = 1$	-180	90

Type 2 agent

	$\theta = 0$	$\theta = 1$
$a = 0$	-180	-180
$a = 1$	90	90

The ECU is converted to TRY with exchange rate $e = 1/13$.

Treatment 4.

The subjects' payoff matrix in each period:

Period 1

Type 1 agent (and the principal)

	$\theta = 0$	$\theta = 1$
$a = 0$	10	-20
$a = 1$	-20	10

Type 2 agent

	$\theta = 0$	$\theta = 1$
$a = 0$	-20	-20
$a = 1$	10	10

Period 2

Type 1 agent (and the principal)

	$\theta = 0$	$\theta = 1$
$a = 0$	40	-80
$a = 1$	-80	40

Type 2 agent

	$\theta = 0$	$\theta = 1$
$a = 0$	-80	-80
$a = 1$	40	40

Period 3

Type 1 agent (and the principal)

	$\theta = 0$	$\theta = 1$
$a = 0$	160	-320
$a = 1$	-320	160

Type 2 agent

	$\theta = 0$	$\theta = 1$
$a = 0$	-320	-320
$a = 1$	160	160

The ECU is converted to TRY with exchange rate $e = 1/21$.

Appendix B.

We first define parameter $\delta_i \in (0,1)$ as the relative weight of the future periods while $1-\delta_i$ is the weight of period i. In other words, in each period i the proportion of remaining stakes allocated to period i is $(1-\delta_i)$ and the proportion to leave to subsequent periods is δ_i. Thus, $\gamma_i = (1-\delta_i)\prod_{j}^{i-1}\delta_j$. Defining the relative stakes this way allows us to set stakes dynamically.

Proof. [Proposition 1] We will provide the proof in 2 steps. In first step, we will show the if V_{i+1}^P is non-decreasing in ρ_{i+1}, then $\rho_{i+1}(a_i=0,\theta_i=0) \geq \rho_{i+1}(a_i=1,\theta_i=0)$. In the second step, we will show that V_{i+1}^P is non-decreasing in ρ_{i+1} in any equilibrium for $i=1,2$.

Step 1.

If V_{i+1}^P is non-decreasing in ρ_{i+1}, then $\rho_{i+1}(a_i=0,\theta_i=0) \geq \rho_{i+1}(a_i=1,\theta_i=0)$.

Suppose $\rho_{i+1}(a_i=0,\theta_i=0) < \rho_{i+1}(a_i=1,\theta_i=0)$ for some $i \in \{1,2\}$ in an equilibrium. As V_{i+1}^P is non-decreasing in ρ_{i+1}, then $\lambda_{i+1}(0,0)=1$ implies $\lambda_{i+1}(1,0)=1$. Hence, in period i, the bad agent has no incentive to play action 0 (when $\theta_i=0$) as action 1 leads to both a higher stage payoff and a higher continuation payoff, i.e. $\nu_i^0=0$. It implies $\rho_{i+1}(a_i=0,\theta_i=0) \geq \rho_{i+1}(a_i=1,\theta_i=0)$ for any μ_i^0 by Bayes rule, hence we reach a contradiction.

Step 2.

V_{i+1}^P is non-decreasing in ρ_{i+1} in any equilibrium for $i=1,2$.

In period 3, $V_3^P = k-\frac{1}{2}+\frac{1}{2}\rho_3$ which is clearly increasing in ρ_3. Combined with Step 1, this implies $\rho_3(0,0) \geq \rho_3(1,0)$ in any equilibrium. Playing action 0 when the state is 0 leads to both a higher stage game payoff and a higher continuation payoff to the good agent, hence $\mu_2^0=0$. Moreover $\nu_2^0<1$ and $\nu_2^1<\mu_2^1$ by Bayes' rule.

In period 2, there are two possibilities depending on the range of the reputation.

Case 1. $\rho_2 \geq 1-2k$.

It implies, if the reputation is not update in the second period, then $\lambda_3=1$. Hence, $\mu_2^1=\nu_2^1=0$ for any δ_2 and $\rho_3(1,1)=\rho_2$. Suppose for the contradiction that $\nu_2^1>0$. The best response of the good agent is to play action 1 with probability 1, i.e. $\mu_2^1=0$ which implies $\rho_3(1,1)>\rho_2$. Hence, there is no equilibrium where $\nu_2^1>0$ when $\rho_2 \geq 1-2k$. $\mu_2^1=0$ is the best response to $\nu_2^1=0$.

As $\mu_2^0=0$, $\rho_3(1,0)=0$ and $\lambda_3(1,0)=0$. Hence, the bad agent has reputational incentives for high enough δ_2.

Sequential rationality constraint of the bad agent implies that to play action 0 with a positive probability, the following must hold:

$$(1-\delta_2)(k-1)+\delta_2 k \geq (1-\delta_2)k$$

$$\delta_2 \geq \frac{1}{k+1}$$

When the inequality is strict, then $v_2^0 = 1$ if $\lambda_3(0,0) = 1$.

1) $\delta_2 > \dfrac{1}{k+1}$:

$v_2^0 = 1$ and $V_2^P = (1-\delta_2)k + \delta_2\left(k - \frac{1}{2} + \frac{1}{2}\rho_2\right)$ as there is no update in the reputation after $(a_2, \theta_2) = (0,0)$ and $(1,1)$. V_2^P is clearly increasing in ρ_2

2) $\delta_2 < \dfrac{1}{k+1}$:

$v_2^0 = 0$ and

$$V_2^P = (1-\delta_2)\left(k - \frac{1}{2} + \frac{1}{2}\rho_2\right) + \underbrace{\frac{1}{2}}_{\theta_2=0}\delta_2\rho_2 k + \underbrace{\frac{1}{2}}_{\theta_2=1}\delta_2\left(k - \frac{1}{2} + \frac{1}{2}\rho_2\right)$$

which is also increasing in ρ_2.

3) $\delta_2 = \dfrac{1}{k+1}$:

$v_2^0 \in [0,1]$ and

$$\begin{aligned} V_2^P = {} & (1-\delta_2)\left(\rho_2 k + \frac{1}{2}(1-\rho_2)k + \frac{1}{2}(1-\rho_2)\left(k-1+v_2^0\right)\right) \\ & + \underbrace{\frac{1}{2}}_{\theta_2=0}\delta_2\underbrace{\left(\rho_2 + (1-\rho_2)v_2^0\right)}_{prob[a_2=0|\theta_2=0]}\left(k - \frac{1}{2} + \frac{1}{2}\rho_3(0,0)\right) \\ & + \underbrace{\frac{1}{2}}_{\theta_2=1}\delta_2\left(k - \frac{1}{2} + \frac{1}{2}\rho_2\right) \end{aligned}$$

where $\rho_3(0,0) = \dfrac{\rho_2}{\rho_2 + (1-\rho_2)v_2^0}$ by Bayes' rule. Simplification shows that V_2^P is increasing in ρ_2.

Case 2. $\rho_2 < 1-2k$.

1) $\delta_2 > \dfrac{1}{k+1}$:

First, we will show that $\lambda_3(0,\theta_2) = 1$ can not hold. Suppose $\lambda_3(0,\theta_2) = 1$, which implies $v_2^0 = 1$. Then, $\rho_3(0,0) = \rho_2 < 1-2k$, contradiction. Thus, we need to introduce mixing strategy to the principal. Principal mixing between hiring the agent and not after both θ_2 implies $V_3^P = 0$. Thus,

$$
\begin{aligned}
V_2^P &= (1-\delta_2)\left(\frac{1}{2}\rho_2 k + \frac{1}{2}\rho_2\left(k-\mu_2^1\right) + \frac{1}{2}(1-\rho_2)\left(k-v_2^1\right) + \frac{1}{2}(1-\rho_2)\left(k-1+v_2^0\right)\right) \\
&= (1-\delta_2)\left(k - \frac{1}{2} - \frac{1}{2}\rho_2\mu_2^1 + \frac{1}{2}\rho_2 - \frac{1}{2}(1-\rho_2)v_2^1 + \frac{1}{2}(1-\rho_2)v_2^0\right)
\end{aligned}
$$

For the principal to mix between hiring and not after θ_2, $\rho_3(0,\theta_2) = 1-2k$. By Bayes' rule:

$$
\begin{aligned}
\rho_3(0,0) &= \frac{\rho_2}{\rho_2 + (1-\rho_2)v_2^0} = 1-2k \\
\Longrightarrow v_2^0 &= \frac{2\rho_2 k}{(1-\rho_2)(1-2k)}
\end{aligned}
$$

$$
\begin{aligned}
\rho_3(0,1) &= \frac{\rho_2\mu_2^1}{\rho_2\mu_2^1 + (1-\rho_2)v_2^1} \\
\Longrightarrow v_2^1 &= \frac{2\rho_2 k\mu_2^1}{(1-\rho_2)(1-2k)}
\end{aligned}
$$

V_2^P becomes:

$$
V_2^P = (1-\delta_2)\left(\left(k-\frac{1}{2}\right) + \frac{1}{2}\rho_2\left(1-\mu_2^1\right) + \frac{\rho_2 k}{(1-2k)}\left(1-\mu_2^1\right)\right)
$$

V_2^P is clearly non-decreasing in ρ_2 as $\mu_2^1 \leq 1$.

2) $\delta_2 < \dfrac{1}{k+1}$:

$\nu_2^0 = 0$ and

$$V_2^P = (1-\delta_2)\left(k - \frac{1}{2} + \frac{1}{2}\rho_2\right) + \underbrace{\frac{1}{2}}_{\theta_2=0}\delta_2\rho_2 k$$

which is clearly increasing in ρ_2.

3) $\delta_2 = \dfrac{1}{k+1}$:

$\nu_2^0, \mu_2^1, \nu_2^1 \in [0,1]$ and

$$\begin{aligned}
V_2^P = &(1-\delta_2)\left(k - \frac{1}{2} - \frac{1}{2}\rho_2\mu_2^1 + \frac{1}{2}\rho_2 - \frac{1}{2}(1-\rho_2)\nu_2^1 + \frac{1}{2}(1-\rho_2)\nu_2^0\right) \\
&+ \underbrace{\frac{1}{2}}_{\theta_2=0}\delta_2\underbrace{\left(\rho_2 + (1-\rho_2)\nu_2^0\right)}_{prob[a_2=0|\theta_2=0]}\left(k - \frac{1}{2} + \frac{1}{2}\rho_3(0,0)\right) \\
&+ \underbrace{\frac{1}{2}}_{\theta_2=1}\delta_2\underbrace{\left(\rho_2\mu_2^1 + (1-\rho_2)\nu_2^1\right)}_{prob[a_2=0|\theta_2=1]}\left(k - \frac{1}{2} + \frac{1}{2}\rho_3(0,1)\right)
\end{aligned}$$

given $\delta_2 = \dfrac{1}{k+1}$:

$$\frac{\partial V_2^P}{\partial \rho_2} = \frac{1}{2}\frac{1}{1+k}\left(2k\left(1-\nu_2^0\right) + \frac{1}{2}\left(\nu_2^0 + \nu_2^1\right) + \frac{1}{2}\left(1-\mu_2^1\right)\right)$$

which is always positive. Hence, V_2^P is increasing in ρ_2.

We showed that both V_2^P and V_3^P are non-decreasing in ρ_2 and ρ_3, respectively in any equilibrium. Hence, $\rho_{i+1}(0,0) \geq \rho_{i+1}(1,0)$ in any equilibrium. Moreover, if μ_i^1 and/or ν_i^1 is greater than zero, then $\rho_{i+1}(0,1) \geq \rho_{i+1}(1,1)$ must hold. Suppose otherwise. As the higher reputation is preferred, both type of agent will have an incentive to deviate to play action 1. Hence, contradiction.

$\rho_{i+1}(0,0) \geq \rho_{i+1}(1,0)$ implies $\nu_i^0 < 0$ and if μ_i^1 and/or ν_i^1 is greater than zero, then $\rho_{i+1}(0,1) \geq \rho_{i+1}(1,1)$ implies $\nu_i^1 < \mu_i^1$ by Bayes' rule. □

Proof. [Proposition 2] We will provide the proof by backward induction. First, we will provide two periods equilibrium. □

Lemma 1. *In two-periods interaction,the principal-optimal equilibrium:*

1) *If* $\rho_1 \geq 1-2k$*:*

   $\delta_1 = \frac{1}{1+k}$

   $v_1^1 = \mu_1^1 = 0 \,\&\, v_1^0 = 1$

   $\lambda_2(0,0) = \lambda_2(1,1) = 1 \text{ \& } \lambda_2(1,0) = \lambda_2(0,1) = 0$

2) *If* $1-4k+4k^2 \leq \rho_1 < 1-2k$*:*

   $\delta_1 = \frac{1}{1+k}$

   $v_1^1 = \mu_1^1 = 0 \,\&\, v_1^0 = \dfrac{2k\rho_1}{(1-\rho_1)(1-2k)}$

   $\lambda_2(0,0) = 1, \text{ \& } \lambda_2(1,1) = \lambda_2(1,0) = \lambda_2(0,1) = 0$

Proof. [Lemma 1] The expected continuation payoff of the principal in the last period ($i = 2$) is given by:

$$V_2^P = k - \frac{1}{2} + \frac{1}{2}\rho_2$$

hence, $\lambda_2(a_1, \theta_1) = 1$ iff $\rho_2(a_1, \theta_1) \geq 1-2k$

1) $\rho_1 \geq 1-2k$:

   First note that $v_1^1 = \mu_1^1 = 0$ for any $\delta_1$following Proposition 1.

   Principal chooses $\delta_1 = \frac{1}{1+k}$. For contradiction, suppose that $\delta_1 > \frac{1}{1+k}$. Then, $v_1^0 = 1$ and $V_1^P = (1-\delta_1)k + \delta_1\left(k - \frac{1}{2} + \frac{1}{2}\rho_1\right)$ which is decreasing in δ_1. Hence, the principal will deviate to $\delta_1 = \frac{1}{1+k}$. Suppose $\delta_1 < \frac{1}{1+k}$, then $v_1^0 = 0$ and

$$V_1^P = (1-\delta_1)\left(\rho_1 k + (1-\rho_1)\left(k - \frac{1}{2}\right)\right) + \underbrace{\frac{1}{2}}_{\theta_1=0}\delta_1\rho_1 k + \underbrace{\frac{1}{2}}_{\theta_1=1}\delta_1\left(k - \frac{1}{2} + \frac{1}{2}\rho_1\right)$$

   which is increasing in δ_1. Hence, the principal will deviate to $\delta_1 = \frac{1}{1+k}$.

$$V_1^P = \frac{k}{1+k}\left(\rho_1 k + \underbrace{\frac{1}{2}(1-\rho_1)}_{\theta_1=1\&b}(k-1) + \underbrace{\frac{1}{2}(1-\rho_1)}_{\theta_1=0\&b}\left(k-1+v_1^0\right)\right)$$
$$+ \underbrace{\frac{1}{2}}_{\theta_2=0}\frac{1}{1+k}\underbrace{\left(\rho_1+(1-\rho_1)v_1^0\right)}_{prob[a_1=0|\theta_1=0]}\left(k-\frac{1}{2}+\frac{1}{2}\rho_2(0,0)\right)$$
$$+ \underbrace{\frac{1}{2}}_{\theta_2=1}\frac{1}{1+k}\left(k-\frac{1}{2}+\frac{1}{2}\rho_1\right)$$

For $k > \frac{1}{4}, \frac{\partial V_1^P}{\partial v_1^0} = \frac{1}{2}\frac{1}{1+k}(1-\rho_1)\left(2k-\frac{1}{2}\right) > 0$. Hence, V_1^P is increasing in v_1^0 and the principal-optimal $v_1^0 = 1$. $\rho_2(0,0) = \rho_2(1,1) = \rho_1 \geq 1-2k \,\&\, \rho_2(1,0) = \rho_2(0,1) = 0$, thus $\lambda_2(0,0) = \lambda_2(1,1) = 1$ & $\lambda_2(1,0) = \lambda_2(0,1) = 0$. V_1^P becomes:

$$V_1^P = \frac{1}{1+k}\left(k^2+k-\frac{1}{2}+\frac{1}{2}\rho_1\right)$$

2) $1-4k+4k^2 \leq \rho_1 < 1-2k$:

Note that $\lambda_2(1,\theta_1) = 1$ can not be an equilibrium strategy because it implies $v_1^{\theta_1} = 0$ and $\rho_2(1,\theta_1) \leq \rho_1 < 1-2k$ contradiction.

Principal again chooses $\delta_1 = \frac{1}{1+k}$. For contradiction, suppose that $\delta_1 > \frac{1}{1+k}$. Then,

a) $\lambda_2(0,\theta_1) = 1$ for some $\theta_1 \in \{0,1\}$. Sequential rationality of bad agent implies $v_1^{\theta_1} = 1$. But then, $\rho_2(0,\theta_1) = \rho_1 < 1-2k$, contradicting $\lambda_2(0,\theta_1) = 1$.

b) Hence, $\lambda_2(0,\theta_1) < 1$ for both $\theta_1 \in \{0,1\}$. $\lambda_2(0,\theta_1) < 1$ implies $V_2^P = 0$. As$V_2^P = 0$, V_1^P is decreasing in δ_1, hence the principal will deviate to $\delta_1 = \frac{1}{1+k}$.

Now suppose suppose that $\delta_1 < \frac{1}{1+k}$:

$$V_1^P = (1-\delta_1)\left(\rho_1 k + (1-\rho_1)\left(k-\frac{1}{2}\right)\right) + \underbrace{\frac{1}{2}}_{\theta_1=0}\delta_1\rho_1 k$$

which is increasing in δ_1. Hence, the principal will deviate to $\delta_1 = \frac{1}{1+k}$.

For $\delta_1 = \frac{1}{1+k}$, the expected continuation payoff of the principal:

$$
\begin{aligned}
V_1^P = & \frac{k}{1+k}\left(\frac{1}{2}\rho_1 k + \frac{1}{2}\rho_1\left(k-\mu_1\right) + \frac{1}{2}\left(1-\rho_1\right)\left(k-\nu_1^1\right) + \frac{1}{2}\left(1-\rho_1\right)\left(k-1+\nu_1^0\right)\right) \\
& + \underbrace{\frac{1}{2}}_{\theta_2=0}\frac{1}{1+k}\underbrace{\left(\rho_1+\left(1-\rho_1\right)\nu_1^0\right)}_{prob[a_1=0|\theta_1=0]}\left(k-\frac{1}{2}+\frac{1}{2}\rho_2(0,0)\right) \\
& + \underbrace{\frac{1}{2}}_{\theta_2=1}\frac{1}{1+k}\underbrace{\left(\rho_1\mu_1+\left(1-\rho_1\right)\nu_1^1\right)}_{prob[a_1=0|\theta_1=1]}\left(k-\frac{1}{2}+\frac{1}{2}\rho_2(0,1)\right)
\end{aligned}
$$

V_1^P is independent of μ_1 thus the principal-optimal $\mu_1 \in [0,1]$.

$\frac{\partial V_1^P}{\partial \nu_1^1} = -\frac{1}{2}\frac{1}{1+k}\left(1-\rho_1\right) < 0$, hence V_1^P is decreasing in ν_1^1 and the principal-optimal $\nu_1^1 = 0$.

$\frac{\partial V_1^P}{\partial \nu_1^0} = \frac{1}{2}\frac{1}{1+k}\left(1-\rho_1\right)\left(2k-\frac{1}{2}\right) > 0$ for $k > \frac{1}{4}$, hence V_1^P is increasing in ν_1^0. The principal-optimal ν_1^0 is the maximum probability in which the equilibrium believes are satisfied. More precisely, $\rho_2(0,0) \geq 1-2k$ so that $\lambda_2\left(0,0\right) = 1$. $\rho_2(0,0)$ is decreasing in ν_1^0 and the maximum ν_1^0 satisfying equilibrium conditions is $\dfrac{2k\rho_1}{\left(1-\rho_1\right)\left(1-2k\right)}$ and $\rho_2(0,0) = 1-2k$.

V_1^P becomes:

$$
V_1^P = \frac{k}{\left(1+k\right)\left(1-2k\right)}\left(2k-2k^2-\frac{1}{2}+\frac{1}{2}\rho_1\right)
$$

which is non-negative for $\rho_1 \geq 1-4k+4k^2$. Thus, for lower initial reputation levels, $\lambda_1 = 0$

□

Proof. [Proposition 2]

From Lemma 6, we know last two periods equilibria for two different ranges of ρ_2. Note that, by Proposition 1, $\rho_2\left(1,\theta_1\right) \leq \rho_1$. As $\rho_1 < 1-4k+4k^2$, $\lambda_2\left(1,\theta_1\right) = 0$ in any equilibrium. In the first period of the three periods interaction, there are two possibilities: (1) $\rho_2\left(0,\theta_1\right) \geq 1-2k$ or (2)$\rho_2\left(0,\theta_1\right) < 1-2k$. The reputation of the agent must be in the same range after both states. Otherwise, suppose $\rho_2\left(0,\theta_1\right) \geq 1-2k$ and $\rho_2\left(0,\theta_1^{'}\right) < 1-2k$ for $\theta_1 \neq \theta_1^{'}$. For δ_1, if the bad agent is indifferent between both actions for $\theta_1^{'}$, then $\nu_1^{\theta_1} = 1$ as $\left(0,\theta_1\right)$ leads to a higher continuation payoff than $\left(0,\theta_1^{'}\right)$. It implies $\rho_2\left(0,\theta_1\right) = \rho_1 < 1-2k$, contradiction. For δ_1, if the bad agent is indifferent between both actions for θ_1, then $\nu_1^{\theta_1^{'}} = 0$. It implies $\rho_2\left(0,\theta_1^{'}\right) = 1 > 1-2k$, contradiction.

Following the similar calculations as in Lemma 1, it is easy to show that V_1^P:

- is increasing in δ_1 if δ_1 is such that $v_1^{\theta_1} = 0$
- is decreasing in δ_1 if δ_1 is such that $v_1^{\theta_1} \geq 0$

hence, the principal will set δ_1 so that the bad agent is indifferent between playing both actions.

Case 1. $\rho_2(0, \theta_1) \geq 1 - 2k$

The sequential rationality constraint of the bad agent implies:

$$(1-\delta_1)k = (1-\delta_1)(k-1) + \delta_1 \left(\frac{1}{2}k + \frac{k^2}{1+k} \right)$$

$$\Leftrightarrow \delta_1 = \frac{2(1+k)}{2+3k+3k^2}$$

for given δ_1, $\mu_1 = 1$ as $V_2^g > V_2^b$. It is easy to show that V_1^P is decreasing in v_1^1, hence the principal optimal $v_1^1 = 0$. The expected continuation payoff of the principal:

$$\begin{aligned} V_1^P = {} & \frac{k+3k^2}{2+3k+3k^2} \left(\rho_1 \left(k - \frac{1}{2} \right) + (1-\rho_1) \left(\frac{1}{2}k + \frac{1}{2} \left(k - 1 + v_1^0 \right) \right) \right) \\ & + \underbrace{\frac{1}{2}}_{\theta_2=0} \frac{2(1+k)}{2+3k+3k^2} \underbrace{\left(\rho_1 + (1-\rho_1) v_1^0 \right)}_{prob[a_1=0|\theta_1=0]} \frac{1}{1+k} \left(k^2 + k - \frac{1}{2} + \frac{1}{2}\rho_2(0,0) \right) \\ & + \underbrace{\frac{1}{2}}_{\theta_2=1} \frac{2(1+k)}{2+3k+3k^2} \rho_1 k \end{aligned}$$

which is increasing in v_1^0 for $k > \frac{1}{4}$. Hence the maximum v_1^0 satisfying equilibrium conditions is the principal optimal, which corresponds to $v_1^0 = \dfrac{2k\rho_1}{(1-\rho_1)(1-2k)}$ and $\rho_2(0,0) = 1-2k$. V_1^P becomes:

$$V_1^P = \frac{k \left(1 + 12k^3 - 2\rho_1 - 2k^2(4+\rho_1) - k(1-2\rho_1) \right)}{2(-2+k+3k^2+6k^3)} \tag{1}$$

Case 2. $\rho_2(0,\theta_1) < 1-2k$

The sequential rationality constraint of the bad agent implies:

$$(1-\delta_1)k = (1-\delta_1)(k-1) + \delta_1\left(\frac{k^2}{1+k}\right)$$

$$\Leftrightarrow \delta_1 = \frac{1+k}{1+k+k^2}$$

for given δ_1, $\mu_1 = 1$ as $V_2^g > V_2^b$. The expected continuation payoff of the principal:

$$V_1^P = \frac{k^2}{1+k+k^2}\left(\rho_1\left(k-\frac{1}{2}\right) + (1-\rho_1)\left(\frac{1}{2}\left(k-v_1^1\right) + \frac{1}{2}\left(k-1+v_1^0\right)\right)\right)$$
$$+ \underbrace{\frac{1}{2}}_{\theta_2=0}\frac{1+k}{1+k+k^2}\underbrace{\left(\rho_1+(1-\rho_1)v_1^0\right)}_{prob[a_1=0|\theta_1=0]}\frac{k}{(1+k)(1-2k)}\left(2k-2k^2-\frac{1}{2}+\frac{1}{2}\rho_2(0,0)\right)$$
$$+ \underbrace{\frac{1}{2}}_{\theta_2=1}\frac{1+k}{1+k+k^2}\underbrace{\left(\rho_1+(1-\rho_1)v_1^1\right)}_{prob[a_1=0|\theta_1=1]}\frac{k}{(1+k)(1-2k)}\left(2k-2k^2-\frac{1}{2}+\frac{1}{2}\rho_2(0,1)\right)$$

V_1^P is decreasing in v_1^1, hence the minimum v_1^1 satisfying equilibrium conditions is the principal optimal, which corresponds to $v_1^1 = \frac{2k\rho_1}{(1-\rho_1)(1-2k)}$ and $\rho_2(0,1) = 1-2k$. V_1^P is increasing in v_1^0 for $k > \frac{1}{4}$. Hence the maximum v_1^0 satisfying equilibrium conditions is the principal optimal, which corresponds to $v_1^0 = \frac{4k\rho_1(1-k)}{(1-\rho_1)(1-2k)}$ and $\rho_2(0,0) = 1-4k+4k^2$. V_1^P becomes:

$$V_1^P = \frac{k^2\left(-1+6k^3-12k^2+8k^3+\rho_1\right)}{2(1-2k)^2(1+k+k^2)} \tag{2}$$

Comparison shows that for sufficiently small ρ_1, $(2) > (1)$. Hence the principal sets $\delta_1 = \frac{1+k}{1+k+k^2}$. □

Proof. [Proposition 3]

Note that there is reputational incentive in any equilibria. Suppose that $v_i^{\theta_i} > 0$ for some θ_i. The maximum continuation payoff the bad agent can get is $\frac{1}{3}\left(-\frac{2}{3}\right)+\frac{1}{3}\frac{1}{3}+\frac{1}{3}\frac{1}{3}=0$. The agent deviates to play action 1 and get $\frac{1}{3}\frac{1}{3}$. The same reasoning also applies to the good agent. Hence, $\rho_2(0,0)=1$ and $\rho_2(1,1)=\rho_1$. Thus, $\lambda_2(0,0)=1$ and $\lambda_2(1,1)$ depends on V_2^P. For the lowest ρ_2, $V_2^P=\frac{1}{9}\left(\rho_2+\frac{1}{2}\rho_2-\frac{1}{2}(1-\rho_2)\right)=\frac{1}{9}\left(2\rho_2-\frac{1}{2}\right)$. As $\rho_2(1,1)=\rho_1=0.1$, $\lambda_2(1,1)=0$. The agent is hired in the second period only after $(0,0)$. Thus the principal's expected continuation payoff:

$$V_1^P=\frac{1}{9}\left(\rho_1-\frac{1}{2}(1-\rho_1)+\frac{1}{2}\rho_1 2\right)=\frac{1}{9}\left(\frac{5}{2}\rho_1-\frac{1}{2}\right)$$

which is equal to 0 for $\rho_1=0.1$. □

Proof. [Proposition 4]

There is no reputation incentive in the second period. The reasoning is that playing action 0 can lead a maximum continuation payoff $\frac{1}{21}(-4+4)=0$, the agent will deviate to lay action 1 and get $\frac{1}{21}2$. Thus, $v_2^0\&v_2^1\&\mu_2=0$.

First we will show that $\lambda_3(1,1)\in(0,1)$ in any equilibrium. Suppose that $\lambda_3(1,1)=0$. Then, $V_2^b=\frac{1}{21}2$. Sequential rationality constraint of the bad agent implies $v_1^0\&v_1^1=0$. But then, $\rho_2(0,\theta_1)=1$, hence $\lambda_3(1,1)=1$, contradiction.

Suppose that $\lambda_3(1,1)=1$. Then, $V_2^b=\frac{1}{21}(2+4)$. Sequential rationality constraint of the bad agent implies $v_1^0\&v_1^1=1$. But then, $\rho_2(0,\theta_1)=\rho_1$ as $v_1^1=1$ implies $\mu_1=1$ by Proposition 1. As $\rho_1=0.1$, $\rho_3(1,1)=\rho_2=\rho_1=0.1$, then $\lambda_3(1,1)=0$, contradiction.

For $\lambda_3(1,1)\in(0,1)$, $\rho_3(1,1)=\rho_2=\frac{1}{3}$. For $\rho_2(0,0)=\frac{1}{3}$, $\frac{\rho_1}{\rho_1+(1-\rho_1)v_1^0}=\frac{1}{3}$. Hence, $v_1^0=\frac{2\rho_1}{1-\rho_1}=\frac{2}{9}$.

There are two possibilities when $\theta_1=1$. The principal may hire or not hire the agent after $(0,1)$. Note that the agent will never be hired after $(1,1)$ because $\rho_2(1,1)=\rho_1$. If the principal is hiring the agent after $(0,1)$, then $v_1^1>0$ and $\mu_1=1$ by Proposition 1. Moreover, the principal may never hire after observing $\theta_1=1$ irrespective of the action. The agent's best response is $\mu_1=0$ and $v_1^1=0$.

Case 1. $\lambda_2(0,1)=1$

As $\mu_1 = 1$, $\nu_1^1 = \nu_1^0 = \frac{2\rho_1}{1-\rho_1} = \frac{2}{9}$.

$$V_2^P = \frac{1}{21}\left(2\rho_2 + (1-\rho_2)(-1) + \frac{1}{2}\rho_2 4\right) = \frac{1}{21}(5\rho_2 - 1) = \frac{1}{21}\frac{2}{3}$$

$$V_1^P = \frac{1}{21}\left(-\frac{1}{2} + \left(\rho_1 + (1-\rho_1)\frac{2\rho_1}{(1-\rho_1)}\right)\frac{2}{3}\right) = \frac{1}{21}\left(2\rho_1 - \frac{1}{2}\right) = -\frac{1}{70}$$

Case 2. $\lambda_2(0,1) = 0$

Then, $\mu_1 \& \nu_1^1 = 0$, $\nu_1^0 = \frac{2\rho_1}{1-\rho_1} = \frac{2}{9}$

$$V_1^P = \frac{1}{21}\left\{\rho_1 + \frac{1}{2}(1-\rho_1)1 + \frac{1}{2}(1-\rho_1)\left(\nu_1^0 1 - \left(1-\nu_1^0\right)2\right)\right.$$
$$\left. + \left(\rho_1 + (1-\rho_1)\frac{2\rho_1}{(1-\rho_1)}\right)\frac{2}{3}\right\} = \frac{1}{21}\left(\frac{11}{2}\rho_1 - \frac{1}{2}\right) = \frac{1}{21}\frac{5}{100} \simeq 0.0024$$

As the second case leads to a higher expected continuation payoff, the principal will not hire after observing $\theta_1 = 1$. □

Proof. [Proposition 5]

First we will show that $\lambda_2(1,1) \& \lambda_3(1,1) = 0$ in any equilibrium.

1) Suppose that $\lambda_2(1,1) = 0$ and $\lambda_3(1,1) > 0$. Then, $V_2^b > \frac{1}{39}3$, hence $\nu_1^{\theta_1} = 1$ for both θ_1. But then $\rho_2(0,\theta_1) = \rho_1 = \rho_3(1,1)$, contradicting $\lambda_3(1,1) > 0$.

2) Suppose that $\lambda_2(1,1) > 0$ and $\lambda_3(1,1) > 0$. Note that $\rho_2(1,1) \leq \rho_1$, hence $\lambda_3(1,1) = 0$ after $(a_1,\theta_1) = (1,1)$. Suppose $\lambda_3(1,1) > 0$ after after $(a_1,\theta_1) = (0,\theta_1)$ for any θ_1. Then, $\nu_1^{\theta_1} = 1$ and $\rho_2(0,\theta_1) = \rho_1$, contradicting $\lambda_3(1,1) > 0$.

3) Suppose that $\lambda_2(1,1) > 0$ and $\lambda_3(1,1) = 0$. $\lambda_3(1,1) = 0$ implies $V_2^b = \frac{1}{39}3$. If $\lambda_2(1,1) > 0$, then $\nu_1^1 = 0$. But then, $\mu_1 = 1$ and $\rho_2(0,1) = 1$. The bad agent will deviate to action 0. Hence there is no such equilibrium.

Note that the principal's expected continuation payoff is decreasing in ν_i^0. Hence, $\nu_2^1 = 0$ in the principal-optimal equilibrium.

$$
\begin{aligned}
39V_2^P = & \left(\frac{1}{2}\rho_2 3 + \frac{1}{2}\rho_2\left(-6\mu_2 + 3\left(1-\mu_2\right)\right) + \frac{1}{2}(1-\rho_2)3 + \frac{1}{2}(1-\rho_2)\left(3v_2^0 - 6\left(1-v_2^0\right)\right)\right) \\
& + \frac{1}{2}\left(\rho_2 + (1-\rho_2)v_2^0\right)\left(\frac{27}{2}\rho_3(0,0) - \frac{9}{2}\right) + \frac{1}{2}\rho_2\mu_2 9
\end{aligned}
$$

which is independent off μ_2, hence the principal-optimal $\mu_2 \in [0,1]$.

$$
\frac{\partial 39V_2^P}{\partial v_2^0} = \frac{9}{2}(1-\rho_2) - \frac{9}{4}(1-\rho_2) > 0 \, for \, \rho_2 < 1
$$

hence, V_2^P is increasing in v_2^0. The principal-optimal v_2^0 is the highest probability satisfying the equilibrium condition so that $\lambda_3(0,0) = 1$. In particular, $\rho_3(0,0) = \dfrac{\rho_2}{\rho_2 + (1-\rho_2)v_2^0} = \frac{1}{3}$ and $v_2^0 = \frac{2\rho_2}{1-\rho_2}$. V_2^P becomes $\frac{1}{26}(9\rho_2 - 1)$.

$\lambda_3(1,1) = 0$ and $\lambda_2(0,\theta_1) = 1$ implies that $\frac{1}{9} \leq \rho_2(0,\theta_1) \leq \frac{1}{3}$.

Note that $v_1^1 = 0$ can not hold because the good agent's best response would be $\mu_1 = 1$ and $\rho_2(0,1) = 1$. The bad agent will deviate to action 0. From Proposition 1, $v_1^1 > 0$ implies $\mu_1 = 1$. Again, V_1^P is decreasing in v_1^1, hence the principal-optimal v_1^1 is the lowest probability satisfying the equilibrium condition. In particular, $\rho_2(0,1) = \dfrac{\rho_1}{\rho_1 + (1-\rho_1)v_1^1} = \frac{1}{3}$ and $v_1^1 = \frac{2\rho_1}{1-\rho_1} = \frac{2}{9}$.

$$
\begin{aligned}
39V_1^P = & \left(-\frac{1}{2}\rho_1 + \frac{1}{2}(1-\rho_1)\left(-2v_1^1 + \left(1-v_1^1\right)\right) + \frac{1}{2}(1-\rho_1)\left(v_1^0 - 2\left(1-v_1^0\right)\right)\right) \\
& + \frac{1}{2}\left(\rho_1 + (1-\rho_1)v_1^0\right)\left(\frac{3}{2}\left(9\rho_2(0,0) - 1\right)\right) + \frac{1}{2}\left(\rho_1 + (1-\rho_1)v_1^1\right)\left(\frac{3}{2}\left(9\rho_2(0,1) - 1\right)\right)
\end{aligned}
$$

$$
\frac{\partial 39V_1^P}{\partial v_1^0} = \frac{3}{2}(1-\rho_1) - \frac{3}{4}(1-\rho_1) > 0
$$

hence, V_1^P is increasing in v_1^0. The principal-optimal v_1^0 is the highest probability satisfying the equilibrium condition. In particular, $\rho_2(0,0) = \dfrac{\rho_1}{\rho_1 + (1-\rho_1)v_1^0} = \frac{1}{9}$ and $v_1^0 = \frac{8\rho_1}{1-\rho_1} = \frac{8}{9}$. V_1^P becomes:

$$V_1^P = \frac{1}{39}\left(\frac{27}{2}\rho_1 - \frac{1}{2}\right) = \frac{1.7}{78} \simeq 0.0218$$

□

Proof. [Proposition 6]

We will first show that $\lambda_2(0,\theta_1) < 1$ for both θ_1. Suppose for contradiction that $\lambda_2(0,\theta_1) = 1$ for some θ_1. Then $v_1^{\theta_1} = 1$ which implies $\rho_2(0,\theta_1) = \rho_1$. After $(0,\theta_1)$, if $\lambda_3(0,\theta_2) = 1$ for some θ_2, then $v_2^{\theta_2} = 1$ and $\rho_3(0,\theta_2) = \rho_2(0,\theta_1) = \rho_1 < \frac{1}{3}$, contradicting $\lambda_3(0,\theta_2) = 1$. Hence, if $\lambda_2(0,\theta_1) = 1$ for some θ_1, then $\lambda_3(0,\theta_2) \in (0,1)$ for both θ_2 after observing $(0,\theta_1)$. Then:

$$63V_2^P = \left(\frac{1}{2}\rho_2 4 + \frac{1}{2}\rho_2\left(-8\mu_2 + 4\left(1-\mu_2\right)\right) + \frac{1}{2}\left(1-\rho_2\right)\left(-8v_2^1 + 4\left(1-v_2^1\right)\right) + \frac{1}{2}\left(1-\rho_2\right)\left(4v_2^0 - 8\left(1-v_2^0\right)\right)\right)$$

$\lambda_3(0,\theta_2) \in (0,1)$ implies $\rho_3(0,\theta_2) = \frac{1}{3}$. Hence, $v_2^0 = \frac{2\rho_2}{1-\rho_2}$ and $v_2^1 = \frac{2\mu_2\rho_2}{1-\rho_2}$.

$$63V_2^P = 18\rho_2 - 18\rho_2\mu_2 - 2$$

which is negative for $\rho_2 < \frac{1}{9}$. Thus, $\rho_1 = 0.1$ implies $\lambda_2(0,\theta_1) < 1$ for both θ_1. The lowest ρ_2 for which the agent is hired in the second period is $\rho_2 = \frac{1}{9}$ where $\mu_2 = 0$ and $\lambda_3(a_2,1) = 0$ for both a_2.

$\lambda_2(0,\theta_1) \in (0,1)$ implies $\rho_2(0,\theta_1) = \frac{1}{9}$ and $v_1^{\theta_1} = \frac{8\rho_1}{1-\rho_1} = \frac{8}{9}$ for both θ_1. It in turn implies $\mu_1 = 1$ by Proposition 1. $\lambda_3(0,0) \in (0,1)$ implies $v_2^0 = \frac{2\rho_2}{1-\rho_2} = \frac{1}{4}$. $\lambda_{i+1}(0,\theta_i)$ probabilities follow from the sequential rationality constraints of the bad agent.

$\lambda_2(0,\theta_1) \in (0,1)$ implies $V_2^P = 0$. Hence, $V_1^P = -\frac{1}{126} \simeq -0.00794$ □

Proof. [Proposition 7]

Note that there is reputational incentive in any equilibria. Hence, $\rho_2(0,0) = 1$ and $\rho_2(1,1) = \rho_1$. Thus, $\lambda_2(0,0) = 1$ and $\lambda_2(1,1)$ depends on V_2^P. For the lowest ρ_2, $V_2^P = \frac{1}{18}\left(2\rho_2 + \frac{1}{2}\rho_2 - (1-\rho_2)\right) = \frac{1}{18}\left(\frac{7}{2}\rho_2 - 1\right)$. As $\rho_2(1,1) = \rho_1 = 0.1$, $\lambda_2(1,1) = 0$. The agent is hired in the second period only after $(0,0)$. Thus the principal's expected continuation payoff:

$$V_1^P = \frac{1}{18}\left(3\rho_1 - \frac{3}{2}(1-\rho_1) + \frac{3}{2}\rho_1\right) = \frac{1}{18}\left(6\rho_1 - \frac{3}{2}\right) = -\frac{1}{20} = -0.05$$

□

References

[1] Hossain, T., & Okui, R. (2013). The binarized scoring rule. Review of Economic Studies, 80(3), 984-1001.

[2] Chen, D.L., Schonger, M., Wickens, C., (2016). oTree - An open-source platform for laboratory, online and field experiments. Journal of Behavioral and Experimental Finance, vol 9: 88-97

[3] Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. American economic review, 92(5), 1644-1655.