Theoretical Economics 17 (2022), 763–800 1555-7561/20220763

(Bad) reputation in relational contracting

Rahul Deb
Department of Economics, University of Toronto

Matthew Mitchell
Graduate Department of Management, University of Toronto

Mallesh M. Pai
Department of Economics, Rice University

Motivated by markets for "expertise," we study a bandit model where a principal chooses between a safe and risky arm. A strategic agent controls the risky arm and privately knows whether its type is high or low. Irrespective of type, the agent wants to maximize duration of experimentation with the risky arm. However, only the high type arm can generate value for the principal. Our main insight is that reputational incentives can be exceedingly strong unless both players coordinate on maximally inefficient strategies on path. We discuss implications for online content markets, term limits for politicians, and experts in organizations.

Keywords. Expert advice, bad reputation, strategic experimentation, bandits, relational contracting.

JEL classification. D82, D86.

"One forgets that though a clown never imitates a wise man, the wise man can imitate the clown."

—Malcolm X

Several environments in which a principal relies on an "expert" (agent) share the following features: (1) the agent has a privately known type (good or bad) and the principal wishes to dynamically screen, (2) the type determines the *rate* at which the agent receives private information that is payoff relevant for the principal, and (3) the agent

Rahul Deb: rahul.deb@utoronto.ca
Matthew Mitchell: matthew.mitchell@utoronto.ca
Mallesh M. Pai: mallesh.pai@rice.edu
We are very grateful to Alp Atakan and Jeff Ely whose discussions of the paper pushed us to greatly improve it. We would also like to thank Heski Bar-Isaac, Dirk Bergemann, Jernej Copic, Martin Cripps, Joyee Deb, Mehmet Ekmekci, Drew Fudenberg, Bob Gibbons, Johannes Hörner, Matt Jackson, Navin Kartik, Thomas Mariotti, Maher Said, Colin Stewart, Juuso Välimäki, Alex Wolitzky and audiences at numerous seminars for their time and insightful comments. We especially thank Mingzi Niu for excellent research assistance. Deb would like to thank Boston University, the Einaudi Institute and MIT for their hospitality, and the SSHRC for their continued and generous financial support. Pai gratefully acknowledges financial support from NSF grant CCF-1763349.

© 2022 The Authors. Licensed under the Creative Commons Attribution-NonCommercial License 4.0. Available at https://econtheory.org. https://doi.org/10.3982/TE4803

acts strategically based on this information in an effort to manage his *reputation* and prolong his relationship with the principal.

For example, a surfer on the internet searching for content may be faced with a content provider of unknown quality. The provider wishes to sustain the surfer's attention, but this creates a dilemma: genuine content (which varies in quality) can only be generated periodically. How do content providers balance the quality and the frequency of the new content they provide with the aim of retaining both interest and trust? Similar incentives are faced by elected politicians. The ability of the politician determines whether or not he has effective policy ideas, but a politician trying to stay in office could feel pressured into enacting risky policies that are unlikely to succeed in an attempt to appear proactive. How does this reelection incentive affect the politician's policy choices? Finally, similar incentives are also faced by "experts" in organizations. For example, a scientist in a pharmaceutical company who is trying to establish a reputation for being innovative may choose to recommend a costly drug trial, which is very unlikely to succeed.

We develop and study a novel repeated game to analyze such environments. Because both players are long-lived and the lack of commitment implies that incentives are provided only via continuation value, our framework is effectively a relational contracting setting where the agent has persistent (and periodic) private information.

Our main insight is that the agent's reputational incentives are so strong that they can destroy the relationship unless both players can coordinate on "maximally inefficient" strategies. Perhaps paradoxically, this shows that what helps relationship functioning is precisely the players' ability to mutually agree to terminate it at the point where uncertainty is resolved and, as a result, the relationship is at its most valuable.

Summary of model and results

Our model is perhaps easiest to describe as a *bandit model with a strategic arm* and we use this metaphor throughout the paper. In the next paragraph, we link various elements of the model to the first application since this is perhaps the least apparent of the three settings mentioned above. The implications of our results for all three applications are developed in detail in the body of the paper.

In each period, the principal (online content consumer) chooses between experimenting with a costly risky arm of unknown type (visiting the website) and a costless safe arm (not visiting, the outside option). The risky arm's type is privately known by the agent (content provider) who also controls its output. If the arm is the good type, it stochastically receives private project ideas (news stories) that vary in quality (the accuracy of reporting); the bad type never receives any ideas. If the principal decides to experiment, the agent chooses whether or not to costlessly implement the project (publish a story) based on the idea he received (if any). An implemented project generates a public success or failure (the veracity of the story once cross-checked by other news outlets), the probability of which depends on the quality of the idea. In our model, good projects always succeed, bad projects (implemented despite not receiving an idea) never succeed, while implementing a risky project sometimes results in a success (and

sometimes in a failure). The principal wants to simultaneously maximize the number of successes (true stories) and minimize failures, whereas the agent wants to maximize the duration of experimentation by the principal (i.e., maximize the number of website visits). To make the model interesting, we assume that implementing risky projects is myopically inefficient, in that, it generates a negative expected payoff for the principal. This way, there is a tension between the agent's need to establish reputation (generate successes) and the principal's desire to avoid failures.

Since implementing a project is costless to the agent, there are no frictions in our environment if the agent's type is publicly known. In this "first-best" benchmark (Theorem 1), the principal-optimal Nash equilibrium strategy is for her to always experiment with the good type agent, and, in response, the agent (who is indifferent between all strategies) acts efficiently, that is, only implements good projects. These Nash equilibrium strategies also generate the unique Pareto-efficient outcome. Conversely, if the agent is known to be the bad type, the principal never experiments in the unique Nash equilibrium outcome: experimentation is costly and the bad type can never generate positive payoffs for the principal.

Despite the lack of frictions, there are a multiplicity of Nash equilibria even when the agent is known to be the good type. Indeed, there is a Nash equilibrium in which there is complete relationship breakdown: the good type never runs a project and, in response, the principal never experiments. Our main insight is to show that such *maximal inefficiency*, on path, can be necessary for the relationship to function.

To see why, now suppose there is uncertainty about the agent's type. We first examine Nash equilibrium outcomes with the following (mild) refinement, the sole purpose of which is to rule out the maximal inefficiency identified above. The refinement requires that, at all on-path histories where the principal first learns that the agent is the good type (for sure), the continuation equilibrium is *nontrivial* (i.e., the principal experiments at some continuation history with positive probability). To put it differently, whenever the agent first proves himself to be the good type, he receives positive continuation utility because the relationship does not completely break down. Continuation play at these histories can be inefficient as long as it is not maximally so. In what follows, this refinement is implicit whenever we refer to *equilibrium* without using the additional "Nash" qualifier.

When there is type uncertainty, the principal may experiment with the risky arm in order to give the agent a chance to reveal himself to be the good type by generating a success. Of course, whether or not experimentation is worthwhile depends on how many (costly) failures the principal must suffer along the way. The key tradeoff in our model is that, absent any dynamic enforcement, the good type always wants to implement risky projects to generate successes since he does not bear the cost of failures. To make this tradeoff between the agent's action strategy and the principal's decision to experiment explicit, we consider a second "static" benchmark, which captures the highest payoff that the principal can achieve from experimenting for a single period subject to the refinement.[1] Here, the principal experiments for the first period and stops experimenting if no success is generated. Conversely, a success (reveals the agent to be

[1]We term this as static because screening only occurs for one period.

the good type and) is followed by first-best continuation play (which yields the highest possible continuation value for the principal). Since only a success guarantees positive continuation value, the good type (strictly) best responds by implementing both good and risky projects; being indifferent, he does not run bad projects (which avoids unnecessary failures for the principal).

The sign of the principal's payoff in this benchmark determines her tolerance for what we call a lack of *quality control*; that is, when the good type agent always implements risky projects. If, and only if, the principal's payoff in this benchmark is positive, the strategies described in the previous paragraph constitute an equilibrium, and hence, the principal is willing to experiment even when the good type agent always runs risky projects (Theorem 2).[2] Conversely, when the payoff in this benchmark is negative, we say *quality control is necessary* for experimentation. In this case, the principal has to provide dynamic incentives to police the agent's actions in order to receive an overall positive value from experimentation; simple static strategies of the sort described above will not work.

Our first result (part 1 of Theorem 3) shows that the principal can never completely prevent the agent from choosing inefficient actions in any equilibrium—in every nontrivial equilibrium, the agent implements both risky and bad projects on path. Part 2 of Theorem 3 shows that the surplus loss from this inefficiency can be large. Specifically, whenever quality control is necessary for experimentation, the unique equilibrium outcome is one in which the principal never experiments. In short, whenever the principal needs to discipline the agent to make experimentation worthwhile, the agent's need to establish reputation makes this impossible.

What makes our result stark is that breakdown can occur even with arbitrarily small amounts of uncertainty or even if "minimal quality control" can make experimentation profitable. We discuss each of these and their implications in turn. There are parameter values such that quality control might be necessary even when the agent is almost surely known to be the good type. For instance, when failures are very costly to the principal, experimentation will not be profitable even with a high belief if the good type agent always implements risky projects. However, when the principal's belief is high, one might expect that the agent's reputational incentives are weaker and that he can be incentivized to act efficiently sufficiently often to make experimentation worthwhile. An implication of our main result is that this is not the case: the principal's equilibrium payoff discontinuously drops from that of the first-best to zero as there is infinitesimal uncertainty about whether the agent is the good type.

When the principal's payoff in the static benchmark is negative but small, this corresponds to the case where the principal would not experiment if the agent acts inefficiently at every opportunity but experimentation would be profitable if the agent chose

[2]One needs to carefully specify the strategy of the bad type: he must also act with positive probability as otherwise acting alone will cause the principal's belief to jump to one and so, per the refinement, the continuation equilibrium after a failure must also be nontrivial. However, because our refinement only bites if beliefs jump to exactly 1, this probability can be taken to be arbitrarily small so that the principal's loss from the bad type is correspondingly small.

to behave efficiently even a small fraction of the time. Our result implies that the principal cannot even provide the long run incentives to make efficient play occur at a small fraction of histories. In this sense, the loss of surplus due to reputational concerns can be large relative to the first best.

In Section 4, we show how coordination on maximal inefficiency can restore relationship functioning (Theorem 5). In Section 5, we first apply this insight to demonstrate that term limits for politicians can improve policy making not despite, but precisely because, even proven good politicians must leave office at the end of their term. We also discuss the two other applications: the benefits of moving from advertising driven revenue to subscription based payment in online content markets and the hiring of experts in organizations. In our concluding remarks (Section 6), we also argue that our main result is robust to relaxing several assumptions.

Related literature

This paper is most closely related to two strands of the repeated games literature. The first is the literature on "bad reputation," which builds on the work of Ely and Välimäki (2003) (henceforth EV). They consider a two-player repeated game and show that the reputational incentives of a long-lived agent with a privately known type can cause the loss of all surplus when faced with a sequence of short-lived principals. The game we analyze is distinct from theirs in that our model does not not have the payoff structure of a bad reputation game in the sense of Ely, Fudenberg, and Levine (2008). More importantly, our paper is the first to establish a bad reputation result in a model with two long-lived players (one of whom has multiple strategic types) and in which the principal's discount factor is intermediate. Moreover, our result does not depend on the agent's discount factor at all. We postpone a more detailed discussion of the differences to Sections 3 and 4. As we argue, these properties are not just of theoretical interest but are also important to capture relevant applications.

These features also distinguish our result from the broader reputation literature, which demonstrates how a long-lived player with a privately known type, facing a sequence of short-lived players, can attain her Stackelberg payoff for sufficiently high discount factors (by mimicking commitment types). There is a substantially smaller fraction of this literature that identifies classes of games in which this result obtains with two long-lived patient players. Early examples are Schmidt (1993) and Cripps and Thomas (1997), and more recent papers are Atakan and Ekmekci (2012, 2013).

Our paper can also be thought of as an instance of a relational contracting problem. Like our setting, Levin (2003) studies a model with both adverse selection and strategic actions but importantly, his agent draws his private cost type independently in each period. More recently, Li, Matouschek, and Powell (2017) consider a setting without transfers where the agent has independently drawn private information in each period. The absence of persistent private information for the informed player is the main distinction between our paper and the vast majority of this literature. In Halac (2012), the principal has a privately known persistent outside option. This qualitatively differs because this private information does not affect the total surplus and instead determines the amount

that the principal can credibly promise to the agent. Malcomson (2016) is a recent instance of a paper with persistent payoff-relevant private types; his main result is that full separation of the agent's types is not possible via a relational contract. All of the papers in this strand of the literature study substantially different economic settings from this work but an important additional difference is that our model features periodic (in addition to the persistent) private information. This difference is not merely cosmetic and the latter ingredient is critical to generate our main economic insights. Finally, Mitchell (2021) also considers a related problem without transfers but his setting is one of pure moral hazard.

While otherwise very different, the reputational incentives (and the fact that they can distort behavior) in our setting are similar in spirit to those in models where experts with private types choose actions in an attempt to demonstrate competence (Prendergast and Stole (1996), Morris (2001), Ottaviani and Sørensen (2006)). More recently, Backus and Little (2018) consider a single period, extensive-form game of expert advise where they derive conditions under which an expert can admit uncertainty. Our setting shares an essential modeling feature that good types may not be able to provide the principal with positive utility in all periods. Aghion and Jackson (2016) consider a political economy setting where voters (principal) must incentivize a politician (agent). Formally, they consider a setting without transfers where a principal is trying to determine the type of a long-lived agent. While some features of our game are similar (the agent's payoff and the fact that she receives private information in each period), the main driving forces in their model are different. Specifically, signaling is not a source of inefficiency in their setting; instead, the principal wants the agent to take "risky" actions that are potentially damaging to the latter's reputation.

1. The model

We study a discrete time, infinite horizon repeated game of imperfect public monitoring between a principal and an agent. We denote time by $t \in \{1, \dots, \infty\}$; the principal and agent discount the future with discount factors $\delta, \beta \in (0, 1)$, respectively.

Agent's initial type: The agent starts the game with a privately known type θ, which can either be good (θ_g) or bad (θ_b). The agent is the good type θ_g with commonly known prior probability $0 < p_0 < 1$. This initial type determines the rate at which the agent can generate positive payoffs for the principal.

We begin by describing the stage game (summarized in Figure 1) after which we define strategies and Nash equilibrium.

1.1 *The stage game*

At each period t, the principal and agent play the following extensive-form stage game where the order of our description matches the timing of moves.

Principal's action: The principal begins the stage game by choosing whether or not to experiment $x_t \in \{0, 1\}$, where $x_t = 1$ corresponds to experimenting. The cost of action x_t is cx_t with $c > 0$ so experimentation is costly. If the principal chooses not to experiment, the stage game ends. When she experiments, the stage game proceeds as follows.

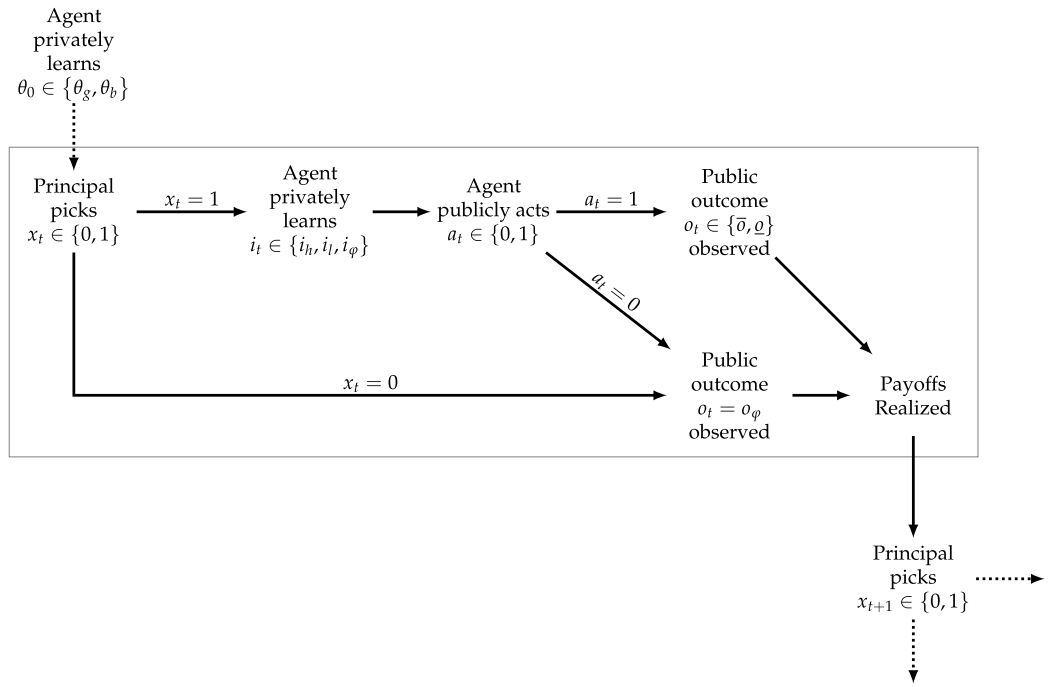

FIGURE 1. Flow chart describing the repeated game; the rectangle contains the stage game.

Agent's information: In each period, the agent receives a project idea i_t. These can either be bad ($i_t = i_b$), risky ($i_t = i_r$), or good ($i_t = i_g$). i_t is drawn independently in each period from a distribution that depends on the agent's type θ.

The bad type only receives bad project ideas, that is, if $\theta = \theta_b$, then $i_t = i_b$ with probability 1.

The good type additionally receives risky and good project ideas stochastically. That is, if $\theta = \theta_g$, the agent gets good ($i_t = i_g$), risky ($i_t = i_r$) and bad ($i_t = i_b$) project ideas with probabilities λ_g, $\lambda_r \in (0, 1)$ and $1 - (\lambda_g + \lambda_r)$, respectively, where $\lambda_g + \lambda_r =: \lambda \in (0, 1)$.

Agent's action: After receiving the project idea, the agent decides whether or not to costlessly implement the project. Formally, he picks a public action $a_t \in \{0, 1\}$ where $a_t = 1(0)$ denotes whether (or not) the project was run.

Public outcomes: If the principal experiments ($x_t = 1$) and the agent acts ($a_t = 1$), a public outcome $o_t \in \{\overline{o}, \underline{o}\}$ is realized from a distribution μ given by

$$\mu(\overline{o} \mid i_t) = \begin{cases} 1 & \text{if } i_t = i_g, \\ q_r & \text{if } i_t = i_r, \\ 0 & \text{if } i_t = i_b, \end{cases}$$

where $q_r \in (0, 1)$ and $\underline{o}$ is realized with the complementary probability.

A success ($\overline{o}$) can only be generated by the good type and the likelihood of a success is determined by the project quality: a good project always generates a success and a risky project sometimes generates a success. The tension in the model arises from the

fact that the good type may want to implement risky projects in an effort to signal his type. Since this can generate failures ($\underline{o}$), the bad type may also act in an attempt to pool even though he only ever receives bad project ideas.

This "good news" assumption (common in bandit models) implies that successes perfectly reveal that the agent is the good type. As we will discuss below, this assumption is deliberately stark (we do not require it for our main insights). It is intended to highlight that relationship breakdown can arise even though the good type can separate perfectly at histories where he receives good project ideas and, hence, one might expect screening to be possible.

If either the principal does not experiment ($x_t = 0$) or the agent does not act ($a_t = 0$), the stage game ends with the agent generating neither a success nor failure. We denote this outcome by $o_t = o_\varphi$. Note that the extensive form implies that the agent does not receive project ideas or get to move if the principal does not experiment; to simplify notation, we define $a_t = 0$ and $i_t = i_b$ when $x_t = 0$.

We use the shorthand notation $h_t = (x_t, a_t, o_t)$, $h_t \in \{\not{h}, h_\varphi, \overline{h}, \underline{h}\}$ to describe the (public) outcome of the stage game. Here,

$$\not{h} := (x = 0, a = 0, o = o_\varphi), \qquad h_\varphi := (x = 1, a = 0, o = o_\varphi),$$
$$\overline{h} := (x = 1, a = 1, o = \overline{o}), \qquad \underline{h} := (x = 1, a = 1, o = \underline{o}).$$

In words, $\not{h}$ denotes the case where principal chooses not to experiment, the remaining three correspond to the separate outcomes that can occur after the principal experiments. h_φ denotes the case where the agent does not act, and $\overline{h}$, $\underline{h}$ denote the cases where the agent acts and a success, failure respectively are observed.

We use time superscripts to denote vectors. Thus, $h^t = (h_1, \dots, h_t)$ and $i^t = (i_1, \dots, i_t)$. Additionally, $h^{t'} h^t$ (and analogously for other vectors) denotes the $t' + t$ length vector where the first t' elements are given by $h^{t'}$ and the $t' + 1^{\text{st}}$ to $t' + t^{\text{th}}$ elements are given by h^t.

Stage game payoffs: The agent wants to maximize the duration of experimentation. Formally, his normalized payoff is $u(x_t) = x_t$, so he receives a unit payoff whenever the principal experiments.

The principal wants to maximize (minimize) the number of successes (failures). Her payoff v is given by

$$v(h_t) = \begin{cases} 1 - c & \text{if } h_t = \overline{h}, \\ -\kappa - c & \text{if } h_t = \underline{h}, \\ -c & \text{if } h_t = h_\varphi, \\ 0 & \text{if } h_t = \not{h}. \end{cases}$$

In words, gross of cost, the principal realizes a normalized payoff of 1 for every success, a loss of $\kappa > 0$ for every failure and 0 otherwise.

Payoff assumptions: We assume that risky projects yield a net loss for the principal: $q_r < (1 - q_r)\kappa$. This assumption creates one of the key tradeoffs in the model: absent signaling value, the principal wants to prevent the agent from running risky projects.

Additionally, we assume that the cost of experimentation is sufficiently low to make the model nontrivial, that is, $c < \lambda_g$. If this assumption is not satisfied, the principal has no incentive to experiment even if the agent is known to be the good type.

1.2 *The repeated game*

Histories: $h^{t-1} \in \{\not{h}, h_\varphi, \overline{h}, \underline{h}\}^{t-1}$ denotes the public history (henceforth, simply a history) at the beginning of period t. This contains all the previous actions of and outcomes observed by both players. The good type agent's private history additionally contains all previous project ideas i^{t-1} and the period-t project idea i_t when he is deciding whether or not to act (the bad type only receives bad project ideas and, therefore, has no additional private history). We use the convention that $h^0 = \varphi$ denotes the start of the game. We use $\mathcal{H}$ and $\mathcal{H}^g$ to respectively denote the set of histories and the set of the good type agent's private histories.[3]

Agent's strategy: We denote the agent's strategy by $\tilde{a}_\theta$. When the agent is the bad type, $\tilde{a}_{\theta_b}(h^{t-1}) \in [0, 1]$ specifies the probability with which the agent acts at each period t as a function of the history $h^{t-1} \in \mathcal{H}$. When the agent is the good type, $\tilde{a}_{\theta_g}(h^{t-1}, i^t) \in [0, 1]$ specifies the probability with which he acts at each period t as a function of his private history $(h^{t-1}, i^t) \in \mathcal{H}^g$. Since the agent can only act when $x_t = 1$, this is implicitly assumed in the notation and we do not add this as an explicit argument of $\tilde{a}_\theta$ for brevity.

Principal's strategy: The principal's strategy $\tilde{x}(h^{t-1}) \in [0, 1]$ specifies the probability of experimenting in each period t as a function of the history.

On- and off-path histories: Given strategies $\tilde{x}$ and $\tilde{a}_\theta$, a history $h^{t-1} \in \mathcal{H}$ is said to be on path (off path) if it can (cannot) be reached with positive probability for either (both) of the agent's possible types $\theta \in \{\theta_g, \theta_b\}$. Similarly, a private history $(h^{t-1}, i^t) \in \mathcal{H}^g$ is said to be on path (off path) if it can (cannot) be reached with positive probability when $\theta = \theta_g$.

Beliefs: A belief $\tilde{p}$ is associated with a pair of strategies $\tilde{x}$, $\tilde{a}_\theta$ (we suppress explicit dependence on the strategies for notational convenience). $\tilde{p}(h^{t-1})$ is principal's belief that the agent is the good type at history h^{t-1}. If this history is on-path, $\tilde{p}(h^{t-1})$ is derived from the agent's strategy $\tilde{a}_\theta$ by Bayes' rule. In the entirety of what follows, we impose no restriction on how off-path beliefs are formed.

Expected payoffs: We use $U_{\theta_g}(h^{t-1}, i^{t-1}, \tilde{x}, \tilde{a}_{\theta_g})$, $U_{\theta_b}(h^{t-1}, \tilde{x}, \tilde{a}_{\theta_b})$, $V(h^{t-1}, \tilde{x}, \tilde{a}_\theta)$ to denote the expected payoff of the good, bad type agent, and principal, respectively, at histories $h^{t-1} \in \mathcal{H}$, $(h^{t-1}, i^{t-1}) \in \mathcal{H}^g$ given strategies $\tilde{x}$, $\tilde{a}_\theta$. Note that the principal's payoff implicitly depends on her belief $\tilde{p}(h^{t-1})$ at this history.

Nash equilibrium: A Nash equilibrium (henceforth referred to as NE) consists of a pair of strategies $\tilde{x}$, $\tilde{a}_\theta$ such that they are mutual best responses. Formally, this implies that $U_{\theta_g}(h^0, i^0, \tilde{x}, \tilde{a}_{\theta_g}) \geq U_{\theta_g}(h^0, i^0, \tilde{x}, \tilde{a}'_{\theta_g})$, $U_{\theta_b}(h^0, \tilde{x}, \tilde{a}_{\theta_b}) \geq U_{\theta_b}(h^0, \tilde{x}, \tilde{a}'_{\theta_b})$ and $V(h^0, \tilde{x}, \tilde{a}_\theta) \geq V(h^0, \tilde{x}', \tilde{a}_\theta)$ for any other strategies $\tilde{x}'$, $\tilde{a}'_\theta$.

[3]Note that $\mathcal{H}^g$ only contains histories where the outcomes for the stage game are compatible with the project ideas received by the good type; recall that successes cannot arise when the good type implements bad projects.

Nontrivial NE: A NE is *nontrivial* if there is an on-path history (h^t) where the principal experiments with positive probability $(\tilde{x}(h^t) > 0)$.

We similarly will also use the phrase "nontrivial" to describe strategies and equilibria (with our refinement in place). It always refers to the fact that the principal experiments on path.

Before beginning the analysis, it is worth discussing the main assumptions of the model: (i) the monitoring structure mapping project type to outcomes, (ii) the lack of transfers, (iii) actions being costless, and (iv) ideas arriving for free. The first is probably the least controversial because it is common in bandit models to assume a "good news" information structure. As mentioned above, successes in our model perfectly reveal that the agent is the good type. This assumption is deliberately stark. It is intended to highlight that relationship breakdown can arise even though the good type can separate perfectly at histories where he receives good project ideas, and hence, one might expect screening to be possible. We argue below that our main insight is unaffected if we dispense with these assumptions.

2. Two benchmarks and the equilibrium refinement

We begin by considering two simple benchmarks, which provide context for our main result. Section 2.1 considers a setting where there is no incomplete information. This motivates the refinement that we define in Section 2.2. Finally, we consider a "static" benchmark in Section 2.3 that provides a foundation for the key parameter values that we focus on.

2.1 *The first-best complete information benchmark*

The purpose of this section is to show that the key friction in the model is the principal's need to screen the good from the bad type. To do so, we consider a benchmark in which the agent's type is publicly known; formally, this corresponds to the case where the prior belief $p_0 \in \{0, 1\}$. The theorem below characterizes the "first-best" NE under complete information. As we often do in the paper, formal statements are presented in words without notation to make them easier to read. Where we feel it is helpful to the reader, we additionally describe the strategies using the notation in footnotes and/or Appendices.

Theorem 1. *If the agent is known to be the bad type, there is a unique NE in which the principal never experiments.*[4] *Conversely, if the agent is known to be the good type, then in the unique Pareto-optimal NE outcome the principal always experiments and the agent only implements good projects.*[5]

This result is obvious so the following discussion serves as the proof. The principal never experiments when the agent is known to be the bad type since experimentation

[4]Formally, the principal's unique NE strategy when $p_0 = 0$ is $\tilde{x}(h^{t-1}) = 0$ for all $h^{t-1} \in \mathcal{H}$. The agent's strategies can be picked arbitrarily.

[5]Formally, the Pareto-optimal NE strategies when $p_0 = 1$ are $\tilde{x}(h^{t-1}) = 1$, $\tilde{a}_{\theta_b}(h^{t-1}) = 0$ for all $h^{t-1} \in \mathcal{H}$ and $\tilde{a}_{\theta_g}(h^{t-1}, i^{t-1}i_t) = 1(0)$ for all $(h^{t-1}, i^{t-1}i_t) \in \mathcal{H}^g$ where $i_t = (\neq) i_g$.

is costly and this type cannot generate any successes, which are the only outcomes that yield a positive payoff to the principal. Conversely, if the agent is known to be the good type, there are no frictions. As long as the agent only runs good projects, it is always profitable for the principal to experiment. Since actions are costless, the agent is indifferent between all strategies and so, in particular, always choosing the principal optimal action is a best response. In what follows, whenever we describe the behavior of either player as *efficient*, we are referring to the strategies described in Theorem 1 above. We use

$$\overline{\Pi} := \frac{\lambda_g - c}{1-\delta} > 0$$

to denote the first-best payoff that the principal obtains from the good type. This is clearly an upper bound for the payoff that the principal can achieve in any NE for $p_0 \in [0, 1]$.

2.2 *The equilibrium refinement*

Despite the lack of frictions when the agent is known to be the good type, there exist other, inefficient Nash equilibria where the principal does not always experiment because the agent acts inefficiently at certain histories. Indeed, complete relationship breakdown is also a NE: the principal never experiments and, in response, the agent never acts—these strategies are mutual best responses. As our main insight is the importance of such on-path *maximal inefficiency* for relationship functioning, we will show that the market can breakdown when we impose the following refinement that rules out these (and only these) strategies on path.

Equilibrium refinement: An *equilibrium* (with no additional qualifier) is a NE in which, at all on-path histories where the principal's belief (first) jumps to 1, the continuation play is nontrivial.

Formally, NE strategies $\tilde{x}, \tilde{a}_\theta$ are an equilibrium if, at all on-path histories $h^{t-1}h_t \in \mathcal{H}$ such that $\tilde{p}(h^{t-1}) < 1$ and $\tilde{p}(h^{t-1}h_t) = 1$, there exists an on-path continuation history $h^{t-1}h_t h^{t'} \in \mathcal{H}$ such that $\tilde{x}(h^{t-1}h_t h^{t'}) > 0$.

Some readers might view this refinement as natural because it can be thought of as an extremely weak form of renegotiation proofness. While we are not aware of other papers that employ this exact refinement, more restrictive versions exist in the literature. Another natural but stronger alternative would require continuation play to be nontrivial at *all* on-path histories $h^{t-1} \in \mathcal{H}$ where the belief is $\tilde{p}(h^{t-1}) = 1$. An even stronger refinement would be to impose efficient continuation play (in the sense of Theorem 1) once uncertainty is resolved. This refinement is employed by EV in whose model similar multiplicity arises; they explicitly refer to it as renegotiation-proofness.[6]

[6]To the best of our knowledge, there is no single, accepted renegotiation-proofness refinement in repeated games with uncertainty. Our refinement (imposed after the agent's type uncertainty is resolved) is substantially weaker than what is implied by many renegotiation-proofness refinements in the literature for complete information games. For instance, the consistency requirements in Farrell and Maskin (1989), Abreu, Pearce, and Stacchetti (1993) and, most recently, the sustainability requirement (without renegotiation frictions) of Safronov and Strulovici (2017) would yield the efficient outcome when the agent's type

Finally, note that the above refinement is, in an informal sense, weaker than standard refinements such as Markov perfection. This is, once again, because the refinement only has bite the first time the principal's belief jumps to 1 on path, whereas Markov perfection restricts on-path behavior to always be measurable with respect to beliefs. That said, observe that Markov strategies are not a subset of our refinement. In particular, Markovian strategies allow for breakdown at all belief 1 histories.[7]

2.3 *The "static" benchmark and the need for quality control*

In this section, we present, and provide a foundation for, the main parameter values that we focus on. We are interested in parameter values where there is a tradeoff between the agent's short term need to establish reputation and the principal's value from experimentation. To do so, we examine a "static" benchmark where the agent only has a single period to prove himself to be the good type. When parameter values are such that the payoff in this benchmark is negative, experimentation is only worthwhile to the principal if she can provide dynamic incentives to prevent the agent from always implementing risky projects.

Specifically, consider the following strategies:

Principal Experiment in period one. If no success is generated, stop experimenting. If a success is generated, always experiment at all future histories.

Good-type agent Only implement good and risky projects in period 1. If a success is generated, follow the efficient strategy. If no success is generated, stop acting.

Bad-type agent Act with positive probability in period one and then stops acting.

Formally, these "static" strategies are given by

$$\begin{aligned}
\tilde{x}(h^0) &= 1 \quad \text{and} \\
\tilde{x}(h_1 h^{t-1}) &= \begin{cases} 1 & \text{if } h_1 = \overline{h}, \\ 0 & \text{if } h_1 \neq \overline{h}, \end{cases} \quad \text{for all } h_1 h^{t-1} \in \mathcal{H},\ t \geq 1, \\
\tilde{a}_{\theta_g}(h^0, i_1) &= \begin{cases} 1 & \text{if } i_1 \in \{i_g, i_r\}, \\ 0 & \text{if } i_1 = i_b, \end{cases} \qquad (1) \\
\tilde{a}_{\theta_g}(h_1 h^{t-1}, i^t i_{t+1}) &= \begin{cases} 1 & \text{if } h_1 = \overline{h} \text{ and } i_{t+1} = i_g, \\ 0 & \text{otherwise}, \end{cases} \\
&\qquad \text{for all } (h_1 h^{t-1}, i^t i_{t+1}) \in \mathcal{H}^g,\ t \geq 1. \\
\tilde{a}_{\theta_b}(h^0) &> 0 \quad \text{and} \quad \tilde{a}_{\theta_b}(h^t) = 0, \quad \text{for all } h^t \in \mathcal{H},\ t \geq 1.
\end{aligned}$$

To summarize, these strategies correspond to the principal experimenting for a single period in the hope that the agent produces a success. A success is followed by the

is initially known to be good. All three are defined for complete information repeated games with normal form stage games but can be applied to our extensive-form stage game as well.

[7]That said, the only Markov equilibria that are ruled out by our refinement are uninteresting equilibria where there is only one success on path following which the game ends.

first-best continuation payoff. If no success is generated, the principal stops experimenting and the game effectively ends. Faced with this strategy, it is a strict best response for the good type to implement both good and risky projects because only successes generate positive continuation payoffs. The bad type is indifferent between implementing a bad project or not (as is the good type), so acting with positive probability is, in particular, a best response. We term these strategies as static because screening only occurs for one period.

An upper bound for the principal's payoff from these strategies is given by

$$\underline{\Pi}(p_0) := p_0(\lambda_g + \lambda_r q_r)(1 + \delta\overline{\Pi}) - p_0\lambda_r(1 - q_r)\kappa - c.$$

This is the expected payoff that the principal receives from the actions of the good type; it is an upper bound because it does not account for the losses from failures generated by the bad type. The first term corresponds to the payoff after a success (the principal gets a payoff of 1 in period one and the first-best continuation payoff $\overline{\Pi}$) and the second term is the loss from a failure. It is straightforward to observe that a necessary and sufficient condition for these strategies to constitute an equilibrium is that this bound is positive.

THEOREM 2. *There exists an equilibrium in static strategies* (*given by (1)*) *iff* $\underline{\Pi}(p_0) > 0$.

When $\underline{\Pi}(p_0) > 0$, there is always a low enough positive probability $\varepsilon > 0$ of action by the bad type ($\tilde{a}_{\theta_b}(h^0) = \varepsilon$) such that the principal's payoff from experimentation is positive. Note that, as long as the bad type sometimes, but not always, acts ($\tilde{a}_{\theta_b}(h^0) \in (0, 1)$), the principal's belief following both nonsuccess outcomes is interior (since $\tilde{p}(\underline{h})$, $\tilde{p}(h_\varphi) < 1$), and thus the continuation play as prescribed by the above strategies (1) is allowed by the refinement since it has no bite at these histories.

Conversely, when $\underline{\Pi}(p_0) \leq 0$, it is a strict best response for the principal to not experiment in period one. This is because her payoff is strictly lower than this upper bound since the bad type acts with positive probability. Note that there cannot be an equilibrium where the bad type does not act ($\tilde{a}_{\theta_b}(h^0) = 0$). Suppose to the contrary that there was such an equilibrium. In this putative equilibrium, only the good type acts and so, irrespective of outcome, the principal's posterior belief must go to one after an action ($\tilde{p}(h_1) = 1$ for $h_1 \in \{\overline{h}, \underline{h}\}$). Our refinement implies that this will result in positive continuation value for the agent. As a result, acting ($\tilde{a}_{\theta_b}(h^0) = 1$) will be a strict best response for the bad type, which is a contradiction.

When $\underline{\Pi}(p_0) \leq 0$, we say *quality control is necessary* for experimentation to generate a positive expected payoff. In this case, the principal must use the dynamics of her relationship with the agent in order to provide the good type with the necessary incentives to implement risky projects sufficiently infrequently. Note that the monitoring structure we have chosen in our model is intentionally stark. By assuming successes are perfectly revealing, it should be easier for the good type to establish reputation compared to a setting where bad projects could also generate a success with positive probability. In other words, one might think that screening would be relatively easier for the principal compared to the case where bad types, too, could sometimes produce successes. We argue that reputational forces can be destructive *despite* this stark good news monitoring structure.

3. Relationship breakdown

Our main insight is the minimal efficiency requirement imposed by the refinement can cause market breakdown.

Theorem 3. *Suppose $p_0 \in (0, 1)$. Then:*

(i) In every nontrivial equilibrium, the good type agent implements risky projects and the bad type implements bad projects (with positive probability) on path.

Formally, in every nontrivial equilibrium, there exist on-path histories $h^t \in \mathcal{H}$ and $(h^{t'}, i^{t'} i_r) \in \mathcal{H}^g$ such that $\tilde{a}_{\theta_b}(h^t) > 0$ and $\tilde{a}_{\theta_g}(h^{t'}, i^{t'} i_r) = 1$.

(ii) If quality control is necessary ($\underline{\Pi}(p_0) \leq 0$), the unique equilibrium outcome is that the principal never experiments.

Taken together, these results have stark economic implications. The first statement of Theorem 3 implies that a principal who is willing to experiment will necessarily face inefficient actions. The second statement of Theorem 3 argues that the payoff implications of this inefficiency can be large. Specifically, the agent's reputational incentives prevent any quality control whenever it is necessary for experimentation, and hence, the principal can never internalize the long run benefits of experimentation.

A first consequence of this result is that there are two separate discontinuities that arise in the equilibrium payoff set. First, observe that there are parameter values such that $\underline{\Pi}(1) \leq 0$; for instance, this can arise when experimentation is costly (high c), failures yield large losses (high κ) or when risky projects are very unlikely to generate successes (low q_r). In this case, $\underline{\Pi}(p_0) \leq 0$ for all $p_0 \in (0, 1)$ and so the principal will never experiment. Thus, both players' payoffs discontinuously fall from the first-best when $p_0 = 1$ to zero when $p_0 < 1$.

Similarly, if we fix the prior p_0 but instead alter the parameter values so that the payoff upper bound from the static benchmark $\underline{\Pi}(p_0) \downarrow 0$, the agent's payoff set once again discontinuously shrinks to zero. As long as $\underline{\Pi}(p_0) > 0$, there always exists at least one nontrivial equilibrium (described in Section 2.3) where the principal experiments in period one and the good type can attain the first best continuation payoff with probability $\lambda_g + \lambda_r q_r$.

We end this section by providing some intuition for Theorem 3. Our key theoretical contribution establishes a bad reputation result with two long-lived players and intermediate discounting by combining two key ideas from separate strands of the reputation literature.

We first observe that there is a positive belief below, which the principal does not experiment in *any* NE. Experimentation is costly, so there is a belief below which experimentation is not worth it even if uncertainty about the agent's type would surely resolve itself with one period of experimentation (and be followed by the first-best continuation equilibrium).

Now suppose the principal were myopic ($\delta = 0$) so her payoff in every period of the game must be nonnegative. Then, at any history, the principal would only indulge in costly experimentation if the agent generated a success with at least probability c. As a

result, the principal's posterior belief following one of the two nonsuccess outcomes—failure or inaction—must not only fall (because beliefs follow a martingale) but must fall at a *minimal rate*. It is then possible to show that every nontrivial equilibrium must eventually have a "last" on-path history at which the principal's belief p is less than p_0 and at which the agent must generate a success or otherwise the principal stops experimenting (because her belief will fall below the above mentioned cutoff). But note that both types of agent's incentives at such a last history are identical to their incentives in the static benchmark. Here, it will be a best response for the good type to run risky projects and for the bad type to act. To see the latter, note that if the bad type did not act, then acting alone will take the principal's posterior belief to one. Under our refinement, the agent is then guaranteed a positive continuation value, and hence, not acting cannot be a best response. Since $\underline{\Pi}(p_0) \leq 0$ implies $\underline{\Pi}(p) \leq 0$ for all $p \leq p_0$ and δ, it therefore cannot be a best response for the principal to experiment at such a history contradicting the existence of a nontrivial equilibrium.

This is essentially the insightful inductive argument developed by EV adapted to our model. Note that this argument cannot be employed when the principal is forward looking ($\delta > 0$). Simply, this is because the principal is willing to bear periodic losses in order to generate a positive average payoff. A forward looking principal can slow down the rate of learning by experimenting even when the agent is generating successes with low probabilities. In particular, she could continue to experiment along a path where beliefs drop but *asymptote* to a level that is not low enough to generate breakdown. If so, there would be no last on-path history and the above inductive cannot be applied.

We argue in the proof that this cannot happen. We derive an upper bound for the principal's payoff that captures the fact that, were beliefs to asymptote, there must be an on-path history where the principal's payoff from successes drops below the cost of experimentation. This argument, that links the uninformed player's continuation payoffs to her learning, bears some resemblance to a fundamental insight from the reputation literature with two long-lived players. For instance, Schmidt (1993) studies games of "conflicting interests" where the commitment optimal action of the player trying to establish reputation holds his uninformed opponent down to her minimax value. Intuitively, the informed player can guarantee himself the commitment payoff when patient because there cannot be another equilibrium where the uninformed player attempts to learn by playing an action that is not a best response to the commitment action in any period. Because best-responding to the commitment action yields the minimax value to the uninformed player, she will never choose an action in response that makes her worse off or, in other words, learning is too costly in these games. Similarly, in our setting, the principal wants to avoid last histories by prolonging experimentation (to learn the agent's type) but cannot do so because it eventually becomes exceeding costly.

3.1 *Allowing transfers*

For some applications, it is unrealistic to rule out transfers. For instance, in addition to a fixed wage, firms can choose to pay contingent bonuses to experts they employ. We introduce transfers by altering the stage game to allow the principal to make payments

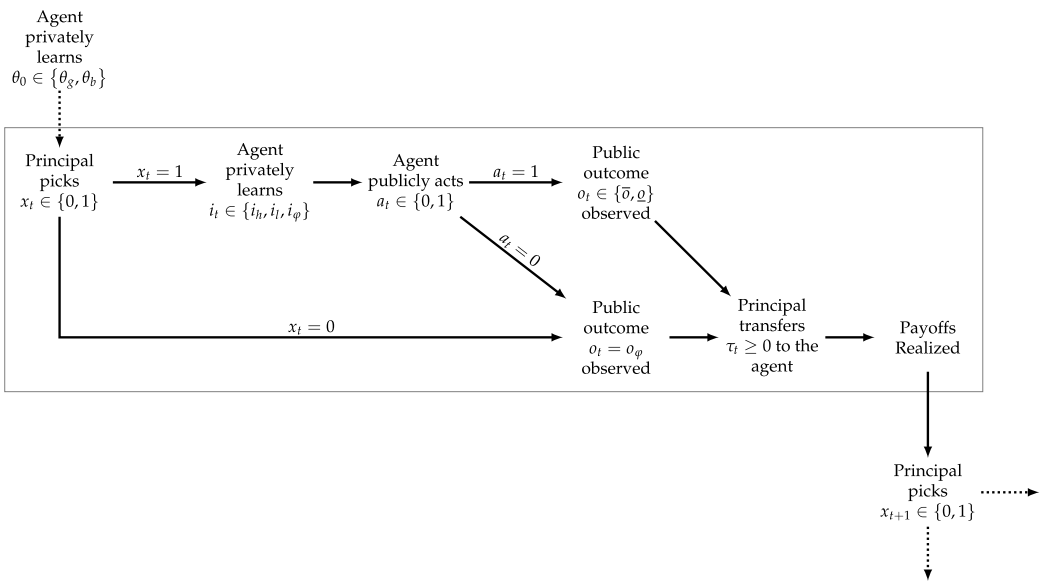

FIGURE 2. Flow chart describing the repeated game with transfers (the rectangle contains the stage game).

to the agent after the public outcome is observed. We assume that the principal can only make transfers in periods where she experiments. This assumption is not required for our result (Theorem 4). We impose it because we feel it is realistic for our applications and it shortens the proof. Formally, we denote the transfer in period t by $\tau_t \geq 0$. These transfers are observed by both players and so become part of the public history. For easy reference, Figure 2 describes how this alters the stage game. The formal description of histories and strategies can be found in the Appendix. The equilibrium refinement applies verbatim. Importantly, note that the refinement applied to this version of the game does not restrict transfers in any way.

Theorem 3 extends verbatim to the game with transfers.

THEOREM 4. *Suppose $p_0 \in (0, 1)$. Then:*

(i) In every nontrivial equilibrium of the game with transfers, the good type agent (surely) implements risky projects and the bad type implements bad projects (with positive probability) on path.

(ii) If quality control is necessary ($\underline{\Pi}(p_0) \leq 0$), the unique equilibrium outcome of the game with transfers is that the principal never experiments.

Transfers do not help the principal avoid inefficient actions or relationship breakdown. In a nutshell, inefficiency is the result of the agent's desire to generate successes and one-sided transfers from the principal can only further reward successes instead of punishing failures. To see this, first note that, once again, every nontrivial equilibrium must have a last on-path history at which the principal experiments (for similar reasons

to the case without transfers). Transfers do not help prevent the agent from running risky projects at this last history. Any payment after a success only exacerbates the problem. Moreover, the principal can never credibly promise payments after a failure since she stops experimenting, and thus her continuation payoff is 0 which, in turn, implies that she has no reason to honor the promise. Thus, $\underline{\Pi}(p_0)$ remains an upper bound for the principal's payoff at such last period histories since one-sided transfers can only lower her continuation utility.

We end this subsection by observing that inefficiency can trivially be eliminated by allowing transfers from the agent to the principal. For example, the principal could ask the agent to reimburse her for the costs of failures; this would align incentives between the principal and the good type by disincentivizing risky actions. Observe that the payments need not necessarily come at the moment of failure and could potentially be moved to later dates, like after the agent proves themselves to be the good type. More generally, the agent could simply buy the experimentation technology from the principal. As is the case in many contracting problems, this complete removes any frictions because only the good type will be willing to pay a sufficiently high amount.

4. Correcting (bad) reputational incentives

In this section, we theoretically demonstrate how maximally inefficient behavior on path can help relationship functioning.

First, observe that, despite quality control being necessary, there are NE in which both players receive positive payoffs. At the most basic level, this can be seen from strategies that *ignore outcomes* and, therefore, in particular, do not condition on (type-revealing) successes. The principal experiments for the first T periods and the agent acts efficiently. At all subsequent periods, the principal never experiments (irrespective of outcomes in the first T periods) and the agent never acts.[8] It is easy to see that there exists a $T \geq 1$ such that these strategies constitute a NE whenever $p_0\lambda_g > c$. This is possible even when quality control is necessary because failures never arise on path. As the outcomes in the first T periods have no effect on continuation play from period $T+1$ onwards, both types of the agent are indifferent between all strategies (irrespective of project ideas) and so acting efficiently is, in particular, a best response. The principal on the other hand has an incentive to experiment for the first T periods as long as her belief $\tilde{p}(h^{t-1})$ never satisfies $\lambda_g \tilde{p}(h^{t-1}) < c$ for any on-path history $h^{t-1} \in \mathcal{H}$, $1 \leq t \leq T$ (by assumption, at least $T = 1$ satisfies this).

However, such Nash equilibria have the feature that both players' *average* payoff, while positive, might be very small. This is because, for any prior p_0, there is a maximal T

[8]Formally, $\tilde{x}$, $\tilde{a}$ are defined as follows: for all $0 \leq t \leq T-1$, $h^t \in \mathcal{H}$, $(h^t, i^t i_{t+1}) \in \mathcal{H}^g$, strategies are

$$\tilde{x}(h^t) = 1, \qquad \tilde{a}_{\theta_g}(h^t, i^t i_{t+1}) = \begin{cases} 1 & \text{if } i_{t+1} = i_g, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad \tilde{a}_{\theta_b}(h^t) = 0.$$

When $t \geq T$, strategies are

$$\tilde{x}(h^t) = \tilde{a}_{\theta_g}(h^t, i^{t+1}) = \tilde{a}_{\theta_b}(h^t) = 0 \quad \text{for all } h^t \in \mathcal{H}, (h^t, i^{t+1}) \in \mathcal{H}^g.$$

for which the above strategies can be a NE because the principal's belief falls if the agent does not generate a success. Therefore, the average payoff for the agent approaches 0 as he becomes arbitrarily patient ($\beta \to 1$). EV show that, in their model, the agent's average payoff vanishes in this way in every NE. A natural question in our setting is whether there exist Nash equilibria such that quality control is necessary but even patient players get positive average payoffs? The following result shows that this is indeed the case.

Theorem 5. *There are cutoff values $\underline{\delta}, \underline{p}_0, \underline{\lambda}_g, \overline{\lambda}_g, \overline{c} \in (0,1)$ with $\overline{c} < \underline{\lambda}_g < \overline{\lambda}_g < \frac{1}{2}$ for which the following holds: for all $\delta \in (\underline{\delta}, 1)$, $\beta \in (0,1)$, $p_0 \in (\underline{p}_0, 1)$, $\lambda_g \in (\underline{\lambda}_g, \overline{\lambda}_g)$, $\lambda_r \in (0, \frac{1}{2})$, $q_r \in (0, \underline{\lambda}_g)$, and $c \in (0, \overline{c})$, there exists a nonempty interval $(\underline{\kappa}, \overline{\kappa})$ such that for any $\kappa \in (\underline{\kappa}, \overline{\kappa})$, quality control is necessary but there nonetheless exists a Nash equilibrium in which the payoff to the principal and the good type are at least $p_0\lambda_g\overline{\Pi} - c$ and $\lambda_g/(1-\beta)$, respectively.*

In words, this result states that it is possible to hold the prior p_0, the arrival rates λ_g, λ_r, the success probability from risky projects q_r and the cost c fixed (as long as they lie in a particular range) but nonetheless find NE in which both players' average payoff is bounded from below when they become patient despite quality control being necessary. Before providing intuition for the result, it is worth making a brief comment about the order of the quantifiers. Note that we provide a range for all parameters except κ, the choice of which depends on the remaining parameters. This is because as $\delta \to 1$ (holding the other parameters fixed), we need κ to also grow to ensure that quality control remains necessary ($\underline{\Pi}(p_0) \leq 0$). Importantly, raising κ does not affect the payoff bounds for either player.[9]

The proof of Theorem 5 is constructive and we describe the NE strategies informally below: in Figure 3 and in words. The formal definitions of the strategies are in Appendix B. Unlike the NE described at the beginning of this section in which outcomes are ignored, the strategies we construct have the feature that both players move to the efficient continuation equilibrium following successes at some histories.

The principal's strategy is the following: experiment in period one and experiment forever (stop experimenting) if a success (failure) is observed. If the agent does not act in period one, experiment in period two. If a success is observed, experiment at all subsequent periods. If the agent does not act in period one and no success is observed in period two, experiment in period three. Stop experimenting from period four onward.

In response, the bad type never acts. The good type agent acts as follows: only run good projects in period one. If a success is generated, follow the efficient strategy. If a failure is generated (off path), stop acting. If no project was run in period one, implement both good and risky projects in period two. If a success is generated, follow the efficient strategy. If no project was run in period one and no success is observed in period two, only implement good projects in period three. Agent stops acting thereafter regardless of the period three outcome. The proof in Appendix B shows that these strategies are a NE for parameters that satisfy the conditions in Theorem 5.

[9]Note that we could also raise c to ensure $\underline{\Pi}(p_0) \leq 0$ but doing so causes the first-best payoff $\overline{\Pi}$ to shrink.

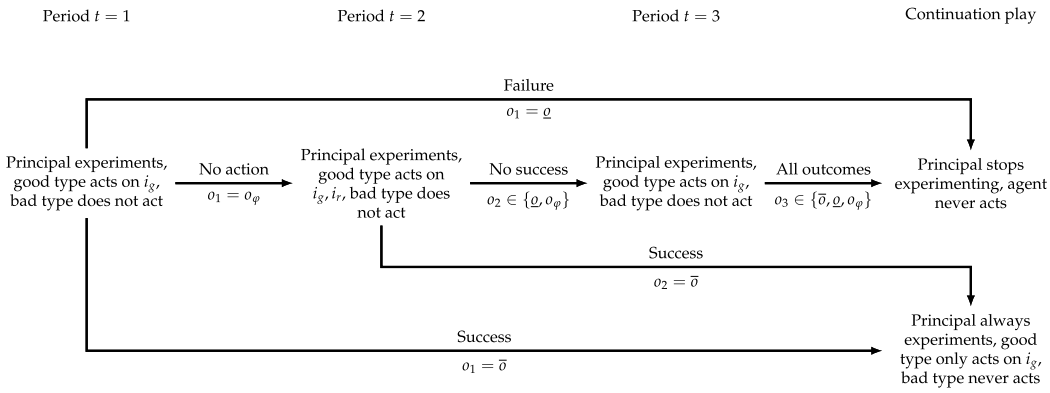

FIGURE 3. A NE with two screening periods when $\underline{\Pi}(p_0) \leq 0$.

Given the above strategies, the payoff bounds can be computed readily by adding the first period payoff with the continuation value from a success being generated (and ignoring the continuation value from other outcomes). Now observe that it is possible to construct an even simpler NE in which the agent gets a positive fraction of the first-best payoff: simply remove the first period and consider the strategies from period two onwards. The reason we add the first period is to guarantee a lower bound for the principal's average payoff as well. The lack of an on-path failure in period one guarantees the principal a minimal fraction of the first-best payoff; if we only considered the NE that started at period two, there would be parameter values that satisfy these conditions but would yield the principal a vanishingly small payoff. Intuitively, if we did not have the first period, then the principal is only getting one period of positive payoff (period three) since quality control is necessary and the agent runs risky projects in period two.

Taken together, Theorems 3 and 5 provide the cleanest comparison of our results with those of EV. They consider a model in which a principal faces an agent whom they want to induce to take an action that matches an underlying stochastic state. The good-type agent observes the state and his payoff function is identical to that of the principal's; conversely, the bad type does not receive any information and, irrespective of the state, strictly prefers to take one of the two actions. In order for experimentation to be profitable for the principal, the good type must always do the "right thing" with a sufficiently high probability: thus, their payoff structure inherently captures the parameter restrictions that we explicitly impose via the quality control condition. Importantly, the principal can only observe the agent's past actions but neither the realizations of the state nor the past payoffs. EV show that, when the agent in their setting gets arbitrarily patient, either (i) the agent's average payoff goes to 0 in every NE when the principal is completely myopic or (ii) first-best payoffs (their folk theorem) can be attained when the principal is also arbitrarily patient.

By contrast, we simultaneously demonstrate the destructive strength of reputational incentives and identify a necessary condition that NE strategies must satisfy if the relationship is to function. We feel that our environment allows us to speak to numerous applications (some described below) that cannot be modeled by standard reputation

theorems that require either myopic or fully patient players or both and a monitoring structure that prevents players from observing their payoffs.

Specifically, our insights require that the principal is *long lived but impatient* and cannot be obtained if the principal was either myopic or fully patient. First, note that quality control being necessary ($\underline{\Pi}(p_0) \leq 0$) implies an upper bound on the principal's level of patience (when all other parameters are fixed). Then observe that, if the principal were myopic ($\delta = 0$), the strategies constructed in this section would not be a NE (because her period two payoff would be negative). Moreover, it is easy to show that, if the principal is myopic, an analogue of EV's main bad reputation result (Theorem 1 in their paper) also arises in our model. Conversely, we too get a folk theorem if the principal becomes patient.[10] Finally, unlike EV, the agent's discount factor β does not play an important role in our setting.

There are a number of key modeling differences that drive these distinct implications. Our agent only cares about extending the relationship (and not about the outcomes from the projects) but this assumption alone would not generate our results in EV's model. It would be impossible to screen such an agent (even with a patient principal) in their setting since their monitoring structure is such that the principal would never learn anything. Conversely, altering their monitoring structure so the principal could observe her past payoffs would make their bad reputation result disappear. This is because rewarding the agent when her action matches the state ensures that "doing the right thing" is a best response for the good type and is a strategy that cannot be mimicked by the bad type (since he receives no information). Instead, our good type has the ability to produce one additional outcome that cannot be mimicked by the bad type but projects that can generate this outcome only arrive stochastically. It is this modeling choice combined with the agent's payoff that generates our novel insights.

5. Applications

The bad reputation force we identify (in Theorems 3 and 4) arises in a variety of different settings. In this section, we discuss how market functioning can be restored in three different applications by altering various assumptions of the model. First, we demonstrate that a consequence of Theorem 5 is that term limits can result in better policy making by reelection seeking politicians. Second, we argue that introducing some commitment for the principal (via subscriptions) can improve the quality of online content. Finally, in the context of organizations, we discuss the role played by an expert's outside option and how the form of transfers matters to provide incentives.

5.1 *Term limits for politicians*

We now argue that our results provide one explanation for why term limits can improve policy making by reelection seeking politicians.[11] To do so, we describe a slight variant

[10] These results are available in the working paper version of this paper.

[11] We are grateful to Navin Kartik for pointing out this application of our result.

of our model that allows for reelection but contains essentially the identical forces to the framework we have analyzed above.

A representative voter (the principal) needs a politician (the agent) to represent her in each period. In each period t, experimentation for the voter is compulsory and costless ($c = 0$), she simply decides whether or not to retain the politician ($x_t = 1$) or to irreversibly replace him ($x_t = 0$) with another ex ante identical candidate (who is the good type with probability p_0).[12] A good type elected politician stochastically receives policy ideas that can either be good or risky; a bad type politician can only enact policies that result in failures. Payoffs to both players are as in our model and have obvious interpretations in this context. The refinement applies almost verbatim: a politician who reveals himself to be the good type cannot be replaced for sure in the next period.

The necessity of quality control in this application implies that the voter will get a negative utility if the politician always enacts risky policies.[13] Consider symmetric equilibria in which every politician follows the same strategy once elected. An implication of our main result is that, if quality control is necessary, every elected politician never implements any projects in the *voter-optimal* symmetric equilibrium.[14]

This result is easy to see. As far as an individual politician is concerned, the reputational incentives are identical to the original model. So suppose, to the converse, that there was a symmetric equilibrium where the voter received a positive payoff. Then, the ex ante expected value of electing each politician must be positive. This introduces an implicit opportunity cost in retaining a current politician as the voter could always get positive utility by replacing him. This opportunity cost would then play an identical role to the cost of experimentation ($c > 0$) in our model and we would then get the same contradiction as that driving Theorem 3.

The above argument demonstrates that reelection incentives can have an extremely perverse effect on policy choices. As we have already argued, reputational incentives can be weakened by strategies that feature maximally inefficient play on path. In this context, such strategies can be interpreted as *term limits*; the politician is replaced even though he has proven himself to be good. Indeed, our analysis is instructive precisely because this is almost always presented as an argument *against* imposing term limits.[15] The voter's strategy can easily incorporate such term limits: every politician is always replaced after a fixed number T of periods but can additionally be replaced before. The argument in Section 4 implies that, despite quality control being necessary, the voter can get a positive payoff from higher quality policy making if she follows such a strategy.

5.2 *Online content markets and the role of commitment*

In the Introduction, we described how our framework can be applied to model the incentives of online content providers. When the consumer demands a minimal quality

[12]As previously mentioned, our results are not affected by assuming the irreversibility of stopping experimentation.

[13]Formally, $p_0(\lambda_g + \lambda_r q_r)(1 + (\delta/(1-\delta))\lambda_g) - p_0\lambda_r(1-q_r)\kappa \leq 0$.

[14]Since there must be an elected representative even if they generate a negative voter payoff, there will be other equilibria in this model.

[15]One (of many) recent such instances is the fourth of "Five reasons to oppose congressional term limits" by Casey Burgat published in Brookings Blog on January 18, 2018.

of reporting (i.e., she wants the provider to not always publish poorly vetted content), Theorem 3 demonstrates that the need to generate clicks can make the market function extremely poorly. An alternate payment model for online content is subscriptions. A natural way to capture a T period subscription in our model is via partial commitment: whenever the principal decides to experiment, she commits to experimenting for $T > 1$ periods and sinks the T period discounted cost of experimentation $((1-\delta^T)/(1-\delta))c$ up front. Thus, the stage game now consists of the principal's experimentation decision followed by the agent receiving project ideas and acting for T periods. Note that our refinement applies verbatim since the only restriction on the principal's strategy is to force her to experiment at histories where she might otherwise not.

It is possible to show that such partial commitment for the principal to T periods of experimentation implies that the good type does not need to always run risky projects and this in turn can generate a positive average profit for the principal.[16] Indeed, when $\underline{\Pi}(p_0)$ is negative but small, we can construct simple equilibria in which the principal receives a positive expected payoff from experimentation in *every period* (not just a positive total payoff) gross of the sunk cost. This latter observation is important because otherwise the principal could choose to "ignore" the agent's action in period T (i.e., not visit the website despite having paid for the subscription).

To summarize, having been paid a subscription fee, a provider can be more judicious about his choice of content as he does not need to incentivize each additional click. Indeed, a recent article ("We Launched a Paywall. It Worked! Mostly." by Nicholas Thompson on May 3, 2019) from the editor-in-chief of technology publication Wired highlights exactly this mechanism in work after they instituted their paywall. One of the lessons they learned was that the content that drove subscriptions ("long-form reporting, Ideas essays, and issue guides") was typically harder and more time consuming to produce (akin to good projects in our model). Importantly, Thompson observes that when

> "your business depends on subscriptions, your economic success depends on publishing stuff your readers love—not just stuff they click. It's good to align one's economic and editorial imperatives! And by so doing, we knew we'd be guaranteeing writers, editors, and designers that no one would be asked to create clickbait crap of the kind all digital reporters dread."

5.3 *Relational contracting in firms*

Our results also have numerous implications for relational contracting in organizations. Consider the hiring of experts. When failure is very costly, for instance for failed drug trials in the biotech sector, firms may benefit from hiring in house researchers as opposed to biotech consultants (for the same reasons that subscriptions were effective in Section 5.2). This is because the latter typically have shorter contracts and have stronger incentives to prove expertise to extend employment duration. Additionally, our

[16]There are numerous dynamic mechanism design problems (for instance, Guo (2016) and Deb, Pai, and Said (2018)) where full commitment for the principal does not yield a higher payoff relative to the principal-optimal NE. In our model, commitment makes the principal strictly better off and details of this argument are in the working paper version.

model provides one rationale for hiring brand name experts who may have higher outside options. The model can easily be generalized to give the agent an outside option $0 \leq \underline{u} \leq 1/(1-\delta)$ and give him the ability to unilaterally resign at the beginning of any period. A higher outside option can ensure that the agent would rather resign than enter a continuation equilibrium where the likelihood of termination is high. Since our refinement does not force the principal to permanently hire the agent even after she knows the agent is the good type for sure, fixed term contracts can arise as equilibrium outcomes when $\underline{u} > 0$ but not when $\underline{u} = 0$ since in the former case the agent is happy to leave once the likelihood of being retained is sufficiently low. Note that in this case, the agent not the principal ends the relationship.

Our model also speaks to how experts should be compensated; the extension which accommodates transfers can be found in Section 3.1. We show (in Theorem 4) that bonuses alone cannot correct bad reputation forces. Instead, it might be more effective to link the compensation of experts directly to firm performance—importantly, they need to be exposed to both the upside and downside—via stock options as this will make it failures costly to the agent. It should be pointed out that there are many papers that describe a variety of different benefits of giving agents "skin in the game;" our dynamic model highlights the trade-offs inherent in determining the optimal vesting period for stock options. If the vesting period is too long, this may create perverse incentives for the agent to prove his worth and extend the duration of his employment at the expense of the firm. Conversely, a short vesting period may end up giving away a share of the firm to an unqualified expert.

The maximal inefficiency we highlight also arises in many employment relationships. Note that our model allows for different types of firm-worker separations. Firing after a failure disincentivizes the agent from implementing risky or bad projects. Such separations can be interpreted as firing "with cause." Even in the absence of failures, information is acquired about the quality of the firm-worker match as the relationship matures and the firm's belief that the worker is the good type may drop over time. This may lead to separation of the sort common in the labor literature that incorporates learning about match value (an early work is Jovanovic (1979)). Such a "without cause" separation often legally requires a notice period; this appears in the form of the third period employment in the example of Section 4. Indeed, such grace periods are a feature of academic contracts in which professors who are denied tenure are typically given an additional year of employment. Importantly, in the academic context, separation is typically irreversible and the professor's performance in the grace year does not lead to reversals of tenure decisions.

6. Concluding remarks

In this concluding section, we address the robustness of our main insight. The model we described in Section 1 makes some assumptions that may give the impression that our main result (Theorem 3) is knife-edged. This is not the case and we briefly describe some natural extensions here; formal statements and proofs are in the working paper version.

Perhaps the most stark is that we assumed costless actions; so instead suppose that the agent has to pay a cost $\mathcal{C} > 0$ to implement a project irrespective of quality. As should be unsurprising, we need to impose a stronger version of our refinement. Recall that Theorem 3 leveraged the observation that, if the principal stops experimenting following inaction, the agent would always act if doing so resulted in a nontrivial continuation play (irrespective of magnitude of the continuation payoff). Of course, when actions are costly, this will only happen if, at the very least, the continuation payoff from the latter compensates the agent for the cost of the action.

Consider the following stronger version of the refinement. For any $\underline{u} \in (0, 1/(1-\beta))$, a *$\underline{u}$-equilibrium* is a NE in which, at all on-path histories where the principal's belief (first) jumps to one, the continuation payoff for the good type is at least $\underline{u}$.

With costly actions, we say that quality control is necessary when

$$\underline{\Pi}^{\mathcal{C}}(p_0) := p_0(\lambda_g + \lambda_r q_r)\big(1 + \delta\overline{\Pi}^{\mathcal{C}}\big) - p_0\lambda_r(1-q_r)\kappa - c \leq 0,$$

where $\overline{\Pi}^{\mathcal{C}} \leq \overline{\Pi}$ is the highest NE payoff that the principal can obtain when the agent is known to be the good type ($p_0 = 1$). Note that this is a weaker condition because the principal cannot achieve payoff $\overline{\Pi}$ even when uncertainty is resolved because of the moral hazard problem introduced by costly actions.

Our main insight extends to this environment. Specifically, we can show that, for any $\underline{u} \in (0, 1/(1-\beta))$, there exists a cost $\overline{\mathcal{C}} > 0$ such that, when the cost of actions $\mathcal{C} \in (0, \overline{\mathcal{C}})$ is lower, the unique $\underline{u}$-equilibrium outcome is that the principal never experiments whenever quality control is necessary.

Now suppose that, instead of costly actions, the good type needs to pay a small cost to draw good, risky ideas (with the same probabilities λ_g, λ_r, respectively). Here, too, we can show that, for any $\underline{u} \in (0, 1)$, relationship breakdown is the unique $\underline{u}$-equilibrium outcome for sufficiently small costs of information acquisition. As in the case with costly actions, the good type will need to be incentivized to acquire information even when uncertainty is resolved.

Finally, we can also allow for successes to not be perfectly revealing: implementing bad projects can also generate successes with a (small) positive probability and so successes can be generated by both agent types. Consequently, there need not be on-path histories where the principal's belief jumps to one. Hence, we will once again require a stronger variant of our refinement: when the principal belief first jumps above a threshold, the continuation play must be nontrivial. Our main insight extends to this monitoring structure.

Appendix A: Proofs of Theorem 3 and Theorem 4

We prove Theorem 4 for the game with transfers and then argue that it implies Theorem 3. Before we proceed with the proofs, we want to formally define histories and strategies for the game with transfers which we chose to omit from Section 3.1.

Transfers: After the public outcome is realized in period t, the principal makes a one-sided transfer $\tau_t \in \mathbb{R}_+$ to the agent.

Histories: A (public) history $(h_1 \dots h_{t-1}, \tau_1 \dots \tau_{t-1})$ at the beginning of period $t \geq 2$ satisfies $h_1 \dots h_{t-1} \in \mathcal{H}$ and $\tau_{t'} \geq (=) 0$ when $h_{t'} \neq (=) \not{h}$ for all $1 \leq t' \leq t$. In addition to the previous actions and outcomes, it also contains the previous transfers (that are zero whenever the principal does not experiment). (h^0, τ^0) denotes the beginning of the game. We denote the set of histories using $\mathcal{T}$.

As before, the good type agent's period-t private history $(h^{t-1}, \tau^{t-1}, i^t)$ additionally contains the current and previous project ideas i^t. We use $\mathcal{T}^g$ to denote the set of the good type agent's private histories.

We use $\tilde{\mathcal{T}}^t$ where $t \geq 1$ to denote the set of on-path histories at the beginning of period $t+1$. The notation reflects the fact that this set depends on the (equilibrium) strategies $(\tilde{x}, \tilde{\tau}, \tilde{a}_\theta)$.

Agent's strategy: The bad type's strategy $\tilde{a}_{\theta_b}(h^{t-1}, \tau^{t-1}) \in [0, 1]$ specifies the probability of acting at each period t public history $(h^{t-1}, \tau^{t-1}) \in \mathcal{T}$. Similarly, $\tilde{a}_{\theta_g}(h^{t-1}, \tau^{t-1}, i^t) \in [0, 1]$ for each private history $(h^{t-1}, \tau^{t-1}, i^t) \in \mathcal{T}^g$.

Principal's strategy: The principal's strategy consists of two functions: an experimentation decision $\tilde{x}(h^{t-1}, \tau^{t-1}) \in [0, 1]$ and a transfer strategy $\tilde{\tau}(h^{t-1}h_t, \tau^{t-1}) \in \Delta(\mathbb{R}_+)$, which specifies the distribution over transfers for each $(h^{t-1}, \tau^{t-1}) \in \mathcal{T}$, $h_t \in \{\overline{h}, \underline{h}, h_\varphi\}$. Note that, by definition, $\tilde{\tau}(h^{t-1}h_t, \tau^{t-1})$ is the Dirac measure at 0 when $h_t = \not{h}$.

Beliefs: Finally, the principal's beliefs $\tilde{p}(h^{t-1}, \tau^{t-1})$, $\tilde{p}(h^{t-1}h_t, \tau^{t-1})$ at the moment she makes her experimentation and transfer decision, respectively, also depend on the past history of transfers. These beliefs are derived by Bayes' rule on path and are not restricted off-path.

Expected payoffs: Expected payoffs now additionally also have as an argument the history of transfers and are denoted by $U_{\theta_g}(h^{t-1}, \tau^{t-1}, i^t)$, $U_{\theta_b}(h^{t-1}, \tau^{t-1})$, $V(h^{t-1}, \tau^{t-1})$ to denote the expected payoff of the good, bad type agent, and principal, respectively. We sometimes also refer to the agent's expected payoff $U_{\theta_g}(h^t, \tau^{t-1}, i^t)$, $U_{\theta_b}(h^t, \tau^{t-1})$ after the outcome but before the principal's transfer in period-t is realized. For brevity, we suppress dependence on strategies.

Last history: A last history is an on-path history at which the principal experiments such that, if no success is generated, the principal stops experimenting (almost surely) at all future histories. Formally, a last history is an on-path $(h^t, \tau^t) \in \tilde{\mathcal{T}}^t$ such that $\tilde{x}(h^t, \tau^t) > 0$, $\tilde{p}(h^t, \tau^t) < 1$ and $U_{\theta_b}(h^t h_{t+1}, \tau^t) = 0$ for all on-path $h_{t+1} \in \{\underline{h}, h_\varphi\}$. Note that this also implies $U_{\theta_g}(h^t h_{t+1}, \tau^t, i^{t+1}) = 0$ for $h_{t+1} \in \{\underline{h}, h_\varphi\}$.

Note that we define a last history in terms of the agent's payoff (as opposed to the principal's experimentation decision) so that we can avoid qualifiers about the principal's actions at on-path continuation histories that are reached with probability 0 (after the principal's transfer is realized via her mixed strategy). Also observe that $U_{\theta_b}(h^t h_{t+1}, \tau^t \tau_{t+1}) = 0$ implies that $\tau_{t+1} = 0$ since it cannot be a best response for the principal to make a positive transfer to the agent in period $t+1$ when her continuation value is 0.

We now argue that every nontrivial NE must have such a last history.

Lemma 1. *Suppose $p_0 \in (0, 1)$. Every nontrivial NE $(\tilde{x}, \tilde{\tau}, \tilde{a}_\theta)$ of the game with transfers must have a last history.*

Proof. We first define

$$\underline{p} = \inf\{\tilde{p}(\hat{h}^{t'}, \hat{\tau}^{t'}) \mid (\hat{h}^{t'}, \hat{\tau}^{t'}) \in \tilde{\mathcal{T}}^{t'}, t' \geq 0\} < 1,$$

to denote the infimum of the beliefs at on-path histories. We will now assume to the converse that there is no last history. With this assumption in place, the following steps yield the requisite contradiction.

Step 1: Experimentation at low beliefs. For all $\varepsilon > 0$, there exists an on-path history $(h^{t'}, \tau^{t'}) \in \tilde{\mathcal{T}}^{t'}$ such that the principal experiments $\tilde{x}(h^{t'}, \tau^{t'}) > 0$ and the belief is close to the infimum $\tilde{p}(h^{t'}, \tau^{t'}) < 1$, $\tilde{p}(h^{t'}, \tau^{t'}) - \underline{p} < \varepsilon$.

By definition, there exists a history $(\hat{h}^t, \hat{\tau}^t) \in \tilde{\mathcal{T}}^t$ such that the belief satisfies $\tilde{p}(\hat{h}^t, \hat{\tau}^t) < 1$ and $\tilde{p}(\hat{h}^t, \hat{\tau}^t) - \underline{p} < \varepsilon$. It remains to be shown that there is such a history at which the principal experiments. We consider the case where $\tilde{x}(\hat{h}^t, \hat{\tau}^t) = 0$ as otherwise $(h^{t'}, \tau^{t'}) = (\hat{h}^t, \hat{\tau}^t)$ would be the required history.

If $U_{\theta_b}(\hat{h}^t, \hat{\tau}^t) > 0$, then this implies that there is an on-path continuation history $(\hat{h}^t\hat{h}^s, \hat{\tau}^t\hat{\tau}^s) \in \mathcal{T}$, $\hat{h}^s = (\not{h}, \dots, \not{h})$, $\hat{\tau}^s = (0, \dots, 0)$, $s \geq 1$ where the principal eventually experiments $\tilde{x}(\hat{h}^t\hat{h}^s, \hat{\tau}^t\hat{\tau}^s) > 0$; this is because the principal is not allowed to make transfers unless she experiments first. But since the belief does not change without experimentation, this implies that $\tilde{p}(\hat{h}^t\hat{h}^s, \hat{\tau}^t\hat{\tau}^s) = \tilde{p}(\hat{h}^t, \hat{\tau}^t)$, and thus $(h^{t'}, \tau^{t'}) = (\hat{h}^t\hat{h}^s, \hat{\tau}^t\hat{\tau}^s)$ is the requisite history.

So suppose instead that $U_{\theta_b}(\hat{h}^t, \hat{\tau}^t) = 0$. Let $0 < s+1 \leq t$ be the last period in the history $(\hat{h}^t, \hat{\tau}^t)$ at which the principal experiments. Formally, $\tilde{x}(\hat{h}^s, \hat{\tau}^s) > 0$ and $(\hat{h}^t, \hat{\tau}^t) = (\hat{h}^s h_{s+1}\hat{h}^{t-s-1}, \hat{\tau}^s\hat{\tau}^{t-s})$ where $\hat{h}^{t-s-1} = (\not{h}, \dots, \not{h})$ when $s < t-1$ and $\hat{\tau}^{t-s} = (0, \dots, 0)$. Such a history $(\hat{h}^s, \hat{\tau}^s)$ must exist because the NE is nontrivial. If principal's realized choice was not to experiment at period $s+1$, then we have $h_{s+1} = \not{h}$. Since the principal's belief could not have changed without experimentation being realized, this implies $\tilde{p}(\hat{h}^s, \hat{\tau}^s) = \tilde{p}(\hat{h}^t, \hat{\tau}^t)$ and $(h^{t'}, \tau^{t'}) = (\hat{h}^s, \hat{\tau}^s)$ is the requisite history.

So finally suppose that the realized action of the principal was to experiment at $s+1$. Then we must have $\hat{h}_{s+1} \in \{\underline{h}, h_\varphi\}$ (note that if $\hat{h}_{s+1} = \overline{h}$, this would imply $\tilde{p}(\hat{h}^t, \hat{\tau}^t) = 1$ which is a contradiction). If $U_{\theta_b}(\hat{h}^s\hat{h}_{s+1}, \hat{\tau}^s) > 0$, this would imply the existence of a first continuation history $(\hat{h}^s\hat{h}_{s+1}\breve{h}^{s'}, \hat{\tau}^s\breve{\tau}^{s'+1}) \in \tilde{\mathcal{T}}^{s+s'+1}$ such that $\tilde{x}(\hat{h}^s\hat{h}_{s+1}\breve{h}^{s'}, \hat{\tau}^s\breve{\tau}^{s'+1}) > 0$ and $\tilde{p}(\hat{h}^s\hat{h}_{s+1}\breve{h}^{s'}, \hat{\tau}^s\breve{\tau}^{s'+1}) = \tilde{p}(\hat{h}^s\hat{h}_{s+1}, \hat{\tau}^s) = \tilde{p}(\hat{h}^t, \hat{\tau}^t)$ and so $(h^{t'}, \tau^{t'}) = (\hat{h}^s\hat{h}_{s+1}\breve{h}^{s'}, \hat{\tau}^s\breve{\tau}^{s'+1})$ would be the requisite history.

Therefore, the only remaining case to analyze is when $U_{\theta_b}(\hat{h}^s\hat{h}_{s+1}, \hat{\tau}^s) = 0$. Since we have assumed that there is no last history, it must be the case that the principal experiments after the other non-success outcome or that $U_{\theta_b}(\hat{h}^s\breve{h}_{s+1}, \hat{\tau}^s) > 0$ for $\breve{h}_{s+1} \in \{\underline{h}, \hat{h}_\varphi\}$, $\breve{h}_{s+1} \neq \hat{h}_{s+1}$. But then it cannot be a best response for type θ_b to reach history $(\hat{h}^s\hat{h}_{s+1}, \hat{\tau}^s)$ since he can always reach history $(\hat{h}^s\breve{h}_{s+1}, \hat{\tau}^s)$ costlessly with probability 1. This implies that $\tilde{p}(\hat{h}^s\hat{h}_{s+1}, \hat{\tau}^s) = 1$ (since this history is on path) which once again yields the contradiction $\tilde{p}(\hat{h}^t, \hat{\tau}^t) = 1$. Thus, $\hat{h}_{s+1} \in \{\underline{h}, h_\varphi\}$ is not possible and the proof of this step is complete.

Step 2: Upper bound for the principal's payoff. Fix an $\varepsilon > 0$ and a history $(h^t, \tau^t) \in \tilde{\mathcal{T}}^t$ such that the principal experiments $\tilde{x}(h^t, \tau^t) > 0$ and the belief satisfies $\tilde{p}(h^t, \tau^t) < 1$,

$\tilde{p}(h^t, \tau^t) - \underline{p} < \varepsilon$. For any integer $s \geq 1$,

$$\frac{\varepsilon}{1-\underline{p}} s - c + \delta^s \overline{\Pi} \geq V\big(h^t, \tau^t\big) \tag{2}$$

is an upper bound for the principal's payoff at history (h^t, τ^t).

For the remainder of this step, we assume that the principal's realized action choice is to experiment at (h^t, τ^t). Since $\tilde{x}(h^t, \tau^t) > 0$, experimenting at (h^t, τ^t) must be a (weak) best response and so it suffices to compute an upper bound for the principal's payoff assuming she experiments for sure at (h^t, τ^t).

We use q to denote the probability that at least one success arrives in the periods between $t+1$ and $t+s$. Now consider the principal's beliefs at on-path histories $(h^t h^s, \tau^t \tau^s) \in \tilde{\mathcal{T}}^{t+s}$. Since the beliefs follow a martingale, they must average to the belief $\tilde{p}(h^t, \tau^t)$. This immediately yields an upper bound for q since

$$q + (1-q)\underline{p} \leq \tilde{p}\big(h^t, \tau^t\big) \quad \Longrightarrow \quad q \leq \frac{\tilde{p}\big(h^t, \tau^t\big) - \underline{p}}{1-\underline{p}} \leq \frac{\varepsilon}{1-\underline{p}}.$$

Since the belief jumps to 1 after a success, the maximal probability of observing a success (subject to Bayes' consistency) can be obtained by assuming that the belief is the lowest possible ($\underline{p}$) when a success does not arrive (which happens with probability $1-q$).

We can write the principal's payoff $V(h^t, \tau^t)$ by summing three separate terms: (i) the expected payoff from the outcomes in the s periods following t, (ii) the expected cost of experimentation and transfers in the s periods, and (iii) the expected continuation value at $t+s+1$.

To derive the upper bound for the principal's payoff from (2), we label each of the terms of

$$\underbrace{\frac{\varepsilon}{1-\underline{p}} s}_{\text{(i)}} - \underbrace{c}_{\text{(ii)}} + \underbrace{\delta^s \overline{\Pi}}_{\text{(iii)}},$$

so that they individually correspond to a bound for each component (i)–(iii) of the payoffs.

(i) qs is an upper bound for the expected payoff that the principal can receive from outcomes in the s periods following history (h^t, τ^t). This corresponds to getting a success in every one of the s periods (with no loss from discounting) whenever at least one success arrives (which occurs with probability $q \leq \varepsilon/(1-\underline{p})$) and no losses from failures.

(ii) Since the principal experiments for sure at (h^t, τ^t), her expected cost of experimentation must be greater than c. Additionally, her cost from transfers must be at weakly greater than 0.

(iii) The principal's expected continuation value must be less than $\overline{\Pi}$ since this is the first-best payoff corresponding to the case where the agent is known to be the good type for sure.

Step 3: Final contradiction. By simultaneously taking s large and εs small, we can find a history $(h^t, \tau^t) \in \tilde{\mathcal{T}}^t$ where the principal experiments $\tilde{x}(h^t, \tau^t) > 0$ but the maximal payoff she can get (given by the bound (2)) is negative. Since this cannot be true in any NE, this contradicts the assumption that there is no last history and completes the proof of the lemma. □

We are now in a position to prove Theorem 4.

Proof of Theorem 4. We prove each part in turn.

Part (i). Suppose $(\tilde{x}, \tilde{\tau}, \tilde{a}_\theta)$ is a nontrivial equilibrium. Then Lemma 1 shows that there must be a last history. At this history, implementing risky projects $\tilde{a}_{\theta_g}(h^t, \tau^t, i^t i_r) = 1$ is a strict best response for type θ_g since our refinement implies that

$$q_r U_{\theta_g}\big(h^t \overline{h}, \tau^t\big) + (1 - q_r) U_{\theta_g}\big(h^t \underline{h}, \tau^t\big) \geq q_r U_{\theta_g}\big(h^t \overline{h}, \tau^t\big) > 0 = U_{\theta_g}\big(h^t h_\varphi, \tau^t\big).$$

This additionally implies that failure at this history $(h^t \underline{h}, \tau^t)$ must be on path. A final consequence is that we must have $\tilde{a}_{\theta_b}(h^t, \tau^t) > 0$ as otherwise $\tilde{p}(h^t \underline{h}, \tau^t) = 1$ which, due to the refinement, would contradict the fact that (h^t, τ^t) is a last history. This proves the first part of the theorem.

Part (ii). Suppose $(\tilde{x}, \tilde{\tau}, \tilde{a}_\theta)$ is a nontrivial equilibrium. Then Lemma 1 implies that there must be a last history. Since the principal's expected payoff at any on-path history must be nonnegative, this implies that her beliefs at *all* last histories must be strictly greater than p_0 (since $\underline{\Pi}(p_0) \leq 0$). We will show this is not possible via the following sequence of steps.

We first define $\tilde{\pi}(\mathcal{T}|h^t, \tau^t)$ which denotes the probability of reaching the set of histories $\mathcal{T} \subset \mathcal{T}$ starting at history $(h^t, \tau^t) \in \mathcal{T}$ via equilibrium strategies $(\tilde{x}, \tilde{\tau}, \tilde{a}_\theta)$.

Step 1: Partitioning the histories. We can partition the set of on-path period-$t+1$ histories $\tilde{\mathcal{T}}^t$ into three mutually disjoint sets:

(i) The set $\tilde{\mathcal{L}}^t \subset \tilde{\mathcal{T}}^t$ of all histories that follow from a last history: this consists of all histories $(h^s h^{t-s}, \tau^s \tau^{t-s}) \in \tilde{\mathcal{T}}^t$ such that (h^s, τ^s) is a last history for some $1 \leq s \leq t$.

(ii) The set $\tilde{\mathcal{S}}^t \subset \tilde{\mathcal{T}}^t$ of all histories that follow from a success being generated in a non-last history: this consists of all histories $(h^t, \tau^t) \in \tilde{\mathcal{T}}^t$ such that $h_s = \overline{h}$ for some $1 \leq s \leq t$, $h_{s'} \neq \overline{h}$ for all $1 \leq s' < s$ and $(h^{s-1}, \tau^{s-1}) \notin \tilde{\mathcal{L}}^{s-1}$ is not a last history.

(iii) All other remaining histories $\tilde{\mathcal{R}}^t = \tilde{\mathcal{T}}^t \backslash (\tilde{\mathcal{S}}^t \cup \tilde{\mathcal{L}}^t)$. Note that for all $(h^t, \tau^t) \in \tilde{\mathcal{R}}^t$, we must have $h_s \neq \overline{h}$ for all $1 \leq s \leq t$.

Step 2: The lowest belief amongst histories in the set $\tilde{\mathcal{R}}^t$ is less than p_0. Formally, for all t, $\tilde{\mathcal{R}}^t$ is nonempty and there is a history $(h^t, \tau^t) \in \tilde{\mathcal{R}}^t$ such that $\tilde{p}(h^t, \tau^t) \leq p_0$.

Since beliefs follow a martingale, the expected value of the beliefs at the beginning of period $t+1$ must equal the prior at the beginning of the game, that is, $p_0 = \mathbb{E}[\tilde{p}(h^t, \tau^t)]$. Note that the expectation above and those that follow in the proof of this step are taken with respect to the distribution $\tilde{\pi}(\cdot|h_0, \tau_0)$ over period $t+1$ histories induced by the

equilibrium strategies. We can rewrite this expression as

$$p_0 = \tilde{\pi}(\tilde{\mathcal{S}}^t|h^0,\tau^0)\mathbb{E}\big[\tilde{p}(h^t,\tau^t) \mid (h^t,\tau^t) \in \tilde{\mathcal{S}}^t\big] + \tilde{\pi}(\tilde{\mathcal{L}}^t|h^0,\tau^0)\mathbb{E}\big[\tilde{p}(h^t,\tau^t) \mid (h^t,\tau^t) \in \tilde{\mathcal{L}}^t\big]$$
$$+ \tilde{\pi}(\tilde{\mathcal{R}}^t|h^0,\tau^0)\mathbb{E}\big[\tilde{p}(h^t,\tau^t) \mid (h^t,\tau^t) \in \tilde{\mathcal{R}}^t\big].$$

Since $\tilde{p}(h^t,\tau^t) = 1$ for all $(h^t,\tau^t) \in \tilde{\mathcal{S}}^t$, the above equation becomes

$$p_0 = \tilde{\pi}(\tilde{\mathcal{S}}^t|h^0,\tau^0) + \tilde{\pi}(\tilde{\mathcal{L}}^t|h^0,\tau^0)\mathbb{E}\big[\tilde{p}(h^t,\tau^t) \mid (h^t,\tau^t) \in \tilde{\mathcal{L}}^t\big]$$
$$+ \tilde{\pi}(\tilde{\mathcal{R}}^t|h^0,\tau^0)\mathbb{E}\big[\tilde{p}(h^t,\tau^t) \mid (h^t,\tau^t) \in \tilde{\mathcal{R}}^t\big]. \tag{3}$$

Now recall that the beliefs at all last histories must be strictly greater than p_0. This implies that $\mathbb{E}[\tilde{p}(h^t,\tau^t) \mid (h^t,\tau^t) \in \tilde{\mathcal{L}}^t] > p_0$ whenever $\tilde{\pi}(\tilde{\mathcal{L}}^t|h^0,\tau^0) > 0$ since these are the histories that follow last histories and beliefs follow a martingale. Of course, this in turn implies that $\tilde{\pi}(\tilde{\mathcal{R}}^t|h^0,\tau^0) > 0$ and that $\mathbb{E}[\tilde{p}(h^t,\tau^t) \mid (h^t,\tau^t) \in \tilde{\mathcal{R}}^t] \le p_0$, which shows that $\tilde{\mathcal{R}}^t$ is nonempty and that there must be a history $(h^t,\tau^t) \in \tilde{\mathcal{R}}^t$ such that $\tilde{p}(h^t,\tau^t) \le p_0$.

Step 3: The principal experiments following a history in $\tilde{\mathcal{R}}^t$ where the belief is low. Formally, for all t and $\varepsilon > 0$, there is a history $(h^t,\tau^t) \in \tilde{\mathcal{R}}^t$ such that

$$\tilde{p}(h^t,\tau^t) \le \inf_{(\hat{h}^t,\hat{\tau}^t)\in\tilde{\mathcal{R}}^t} \tilde{p}(\hat{h}^t,\hat{\tau}^t) + \varepsilon,\ \tilde{p}(h^t,\tau^t) \le p_0 \quad \text{and} \quad \tilde{x}(h^t h^{t'}, \tau^t\tau^{t'}) > 0,$$

for some $(h^t h^{t'}, \tau^t\tau^{t'}) \in \tilde{\mathcal{T}}^{t+t'}$, $t' \ge 0$.

First, suppose $\inf_{(\hat{h}^t,\hat{\tau}^t)\in\tilde{\mathcal{R}}^t} \tilde{p}(\hat{h}^t,\hat{\tau}^t) = p_0$. Then, for all histories $(\hat{h}^s,\hat{\tau}^s) \in \mathcal{T}^s$, $0 \le s \le t$, we must have $\tilde{p}(\hat{h}^s,\hat{\tau}^s) = p_0$, which implies $(\hat{h}^s,\hat{\tau}^s) \in \tilde{\mathcal{R}}^s$. Suppose this were not true. Take an earliest history $(\hat{h}^s\hat{h}_{s+1}, \hat{\tau}^s\hat{\tau}_{s+1}) \in \mathcal{T}^{s+1}$ (with the smallest $s < t$) such that $\tilde{p}(\hat{h}^s\hat{h}_{s+1}, \hat{\tau}^s\hat{\tau}_{s+1}) \ne p_0$. Then, by definition, we must have $\tilde{p}(\hat{h}^s,\hat{\tau}^s) = p_0$ and $(\hat{h}^s,\hat{\tau}^s) \in \tilde{\mathcal{R}}^s$ since $(\hat{h}^s,\hat{\tau}^s) \notin \tilde{\mathcal{L}}^s$ because the beliefs at all last histories are strictly greater than p_0. Then, by Bayes' consistency, there must be a continuation history $(\hat{h}^s\breve{h}_{s+1}, \hat{\tau}^s\breve{\tau}_{s+1}) \in \tilde{\mathcal{R}}^{s+1}$ such that $\tilde{p}(\hat{h}^s\breve{h}_{s+1}, \hat{\tau}^s\breve{\tau}_{s+1}) < p_0$,[17] which would be a contradiction.

Now note that, if the principal does not experiment at any continuation history following $(\hat{h}^t,\hat{\tau}^t) \in \tilde{\mathcal{R}}^t = \tilde{\mathcal{T}}^t$, it cannot be a best response for her to experiment at any history before period $t+1$ either because no successes are ever generated on path. This contradicts the fact that the equilibrium is nontrivial.

Next, suppose $\inf_{(\hat{h}^t,\hat{\tau}^t)\in\tilde{\mathcal{R}}^t} \tilde{p}(\hat{h}^t,\hat{\tau}^t) < p_0$. We define

$$\bar{s} = \max\Big\{s \le t \mid \big(\breve{h}^{s-1}\breve{h}_s\breve{h}^{t-s}, \breve{\tau}^t\big) \in \tilde{\mathcal{R}}^t,\ \breve{h}_s \in \{\underline{h}, h_\varphi\},\ \tilde{p}\big(\breve{h}^{s-1}\breve{h}_s\breve{h}^{t-s}, \breve{\tau}^t\big) \le p_0 \text{ and}$$
$$\tilde{p}\big(\breve{h}^{s-1}\breve{h}_s\breve{h}^{t-s}, \breve{\tau}^t\big) \le \inf_{(\hat{h}^t,\hat{\tau}^t)\in\tilde{\mathcal{R}}^t} \tilde{p}(\hat{h}^t,\hat{\tau}^t) + \varepsilon\Big\},$$

to be the last period among the low belief histories where the principal observes a nonsuccess outcome after experimenting. Let the set of period-$t+1$ histories that yield the above maximum be $\mathcal{R}^t_{\bar{s}}$. Note that, by definition, $\breve{h}^{t-\bar{s}} = (\not{h}, \dots, \not{h})$ for all $(\breve{h}^{\bar{s}}\breve{h}^{t-\bar{s}}, \breve{\tau}^t) \in$

[17] Clearly, if $\tilde{p}(\hat{h}^s\hat{h}_{s+1}, \hat{\tau}^s\hat{\tau}_{s+1}) < p_0$, then $(\hat{h}^s\breve{h}_{s+1}, \hat{\tau}^s\breve{\tau}_{s+1}) = (\hat{h}^s\hat{h}_{s+1}, \hat{\tau}^s\hat{\tau}_{s+1})$ is such a history.

$\mathcal{R}^t_{\overline{s}}$. Observe that the maximum is well-defined because $\tilde{\mathcal{R}}^t$ is nonempty (from Step 2) and the principal experiments before t (otherwise the infimum of the beliefs cannot be strictly lower than p_0).

We now show that, for any history $(\breve{h}^t, \breve{\tau}^t) \in \mathcal{R}^t_{\overline{s}}$, we must have $\tilde{x}(\breve{h}^{\overline{s}}\hat{h}^{s-\overline{s}}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{s-\overline{s}+1}) = 0$ for all $\overline{s} \leq s < t$, $(\breve{h}^{\overline{s}}\hat{h}^{s-\overline{s}}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{s-\overline{s}+1}) \in \tilde{\mathcal{T}}^s$. In words, this says that the principal does not experiment between periods $\overline{s}+1$ and t at *any* on-path continuation history following $(\breve{h}^{\overline{s}}, \breve{\tau}^{\overline{s}-1})$. A consequence of this statement is that the principal's belief does not change until period $t+1$ and, since the beliefs at all last histories are strictly greater than p_0, $(\breve{h}^{\overline{s}}\hat{h}^{t-\overline{s}}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{t-\overline{s}+1}) \in \tilde{\mathcal{T}}^t$ implies $(\breve{h}^{\overline{s}}\hat{h}^{t-\overline{s}}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{t-\overline{s}+1}) \in \mathcal{R}^t_{\overline{s}}$.

Suppose this were not the case, and consider the smallest $\overline{s} \leq s < t$ such that there is an on-path history $(\breve{h}^{\overline{s}}\hat{h}^{s-\overline{s}}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{s-\overline{s}+1}) \in \tilde{\mathcal{T}}^s$ where the principal experiments $\tilde{x}(\breve{h}^{\overline{s}}\hat{h}^{s-\overline{s}}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{s-\overline{s}+1}) > 0$. Then there must be an on-path continuation history $(\breve{h}^{\overline{s}}\hat{h}^{s-\overline{s}+1}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{s-\overline{s}+2}) \in \tilde{\mathcal{T}}^{s+1}$, $\hat{h}_{s-\overline{s}+1} \in \{\underline{h}, h_\varphi\}$ such that

$$\tilde{p}\big(\breve{h}^{\overline{s}}\hat{h}^{s-\overline{s}+1}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{s-\overline{s}+2}\big) \leq \tilde{p}\big(\breve{h}^{\overline{s}}\hat{h}^{s-\overline{s}}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{s-\overline{s}+1}\big) = \tilde{p}\big(\breve{h}^{\overline{s}}, \breve{\tau}^{\overline{s}-1}\hat{\tau}_s\big) = \tilde{p}\big(\breve{h}^{\overline{s}}, \breve{\tau}^{\overline{s}}\big) \leq p_0.$$

The first inequality follows from Bayes' consistency and the equalities follow from the facts that there is no experimentation from periods $\overline{s}+1$ to s and the transfer in period $\overline{s}$ does not change the belief at period $\overline{s}+1$. This combined with the fact that $(\breve{h}^{\overline{s}-1}, \breve{\tau}^{\overline{s}-1}) \in \tilde{\mathcal{R}}^{\overline{s}-1}$, implies that $(\breve{h}^{\overline{s}}\hat{h}^{s-\overline{s}+1}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{s-\overline{s}+2}) \in \tilde{\mathcal{R}}^{s+1}$ since the beliefs at all last histories are strictly greater than p_0. But then, we can once again use the argument in Step 2 starting at the history $(\breve{h}^{\overline{s}}\hat{h}^{s-\overline{s}+1}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{s-\overline{s}+2})$ with associated belief $\tilde{p}(\breve{h}^{\overline{s}}\hat{h}^{s-\overline{s}+1}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{s-\overline{s}+2})$ (instead of the beginning of the game (h^0, τ^0) and belief p_0) to conclude that there must be a period-$t+1$ history $(\breve{h}^{\overline{s}}\hat{h}^{t-\overline{s}}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{t-\overline{s}+1}) \in \tilde{\mathcal{R}}^t$ satisfying

$$\tilde{p}\big(\breve{h}^{\overline{s}}\hat{h}^{t-\overline{s}}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{t-\overline{s}+1}\big) \leq \tilde{p}\big(\breve{h}^{\overline{s}}\hat{h}^{s-\overline{s}+1}, \breve{\tau}^{\overline{s}-1}\hat{\tau}^{s-\overline{s}+2}\big)$$

which would contradict the maximality of $\overline{s}$.

So finally suppose, to the converse, that the principal stopped experimenting after all histories $(\breve{h}^t, \breve{\tau}^t) \in \mathcal{R}^t_{\overline{s}}$. An implication is that $U_{\theta_b}(\breve{h}^{\overline{s}}, \breve{\tau}^{\overline{s}-1}) = 0$. The equality follows from the above argument, which shows that the principal does not experiment between periods $\overline{s}+1$ and t after history $(\breve{h}^{\overline{s}}, \breve{\tau}^{\overline{s}-1})$ and that the transfer in period $\overline{s}$ must be 0 since the principal's continuation value after this history is 0. Then it must be the case that the history corresponding to the other period $\overline{s}$ nonsuccess outcome $(\breve{h}^{\overline{s}-1}\hat{h}_{\overline{s}}, \breve{\tau}^{\overline{s}})$, $\hat{h}_{\overline{s}} \in \{\underline{h}, h_\varphi\}$, $\hat{h}_{\overline{s}} \neq \breve{h}_{\overline{s}}$ must either be off-path or we must have $U_{\theta_b}(\breve{h}^{\overline{s}-1}\hat{h}_{\overline{s}}, \breve{\tau}^{\overline{s}-1}) = 0$. To see this, note that if this was not the case, then it is a strict best response for the bad type to costlessly reach history $(\breve{h}^{\overline{s}-1}\hat{h}_{\overline{s}}, \breve{\tau}^{\overline{s}-1})$, which is not possible since this would imply $\tilde{p}(\breve{h}^{\overline{s}}, \breve{\tau}^{\overline{s}}) = 1$.

But then, $\tilde{x}(\breve{h}^{\overline{s}-1}, \breve{\tau}^{\overline{s}-1}) > 0$ is not possible as otherwise we would get the contradiction that $(\breve{h}^{\overline{s}-1}, \breve{\tau}^{\overline{s}-1}) \in \tilde{\mathcal{L}}^{\overline{s}-1}$ is a last history. Thus, the principal must experiment after at least one history $(\breve{h}^t, \breve{\tau}^t) \in \mathcal{R}^t_{\overline{s}}$ and the proof of this step is complete.

Step 4: Final contradiction. We define

$$\underline{p}^{\mathcal{R}} = \inf\big\{\tilde{p}\big(h^t, \tau^t\big) \mid \big(h^t, \tau^t\big) \in \tilde{\mathcal{R}}^t, t \geq 0\big\} \leq p_0,$$

to be the lowest belief that can arise at a history that does not follow a last history. For $\varepsilon > 0$, we pick an $(h^t, \tau^t) \in \tilde{\mathcal{R}}^t$ such that $\tilde{p}(h^t, \tau^t) \leq \underline{p}^{\mathcal{R}} + \varepsilon$ and $\tilde{x}(h^t, \tau^t) > 0$, and we will argue that the principal's payoff must be negative at such a history if ε is small enough.

First, observe that a consequence of Step 3 is that such a history must exist. We now define a few additional terms for any arbitrary $s \geq 1$:

$$q^{\mathcal{S}} := \tilde{\pi}\big(\tilde{\mathcal{S}}^{t+s} | h^t, \tau^t\big), \qquad q^{\mathcal{L}} := \tilde{\pi}\big(\tilde{\mathcal{L}}^{t+s} | h^t, \tau^t\big) \quad \text{and}$$
$$p^{\mathcal{L}} := \mathbb{E}\big[\tilde{p}\big(h^t\hat{h}^s, \tau^t\hat{\tau}^s\big) \mid \big(h^t\hat{h}^s, \tau^t\hat{\tau}^s\big) \in \tilde{\mathcal{L}}^{t+s}\big].$$

The first two terms are the probabilities that, after s periods, the players reach a history that follows from a success being generated in a nonlast history or a history that follows a last history respectively. The third term is the expected belief at the latter set of histories where the expectation is taken with respect to the distribution $\tilde{\pi}(\cdot|h^t, \tau^t)$.

The martingale property of beliefs implies

$$\begin{aligned}\tilde{p}\big(h^t, \tau^t\big) &= q^{\mathcal{S}} + q^{\mathcal{L}} p^{\mathcal{L}} + \big(1 - q^{\mathcal{S}} - q^{\mathcal{L}}\big)\mathbb{E}\big[\tilde{p}\big(h^t\hat{h}^s, \tau^t\hat{\tau}^s\big) \mid \big(h^t\hat{h}^s, \tau^t\hat{\tau}^s\big) \in \tilde{\mathcal{R}}^{t+s}\big] \\ &\geq q^{\mathcal{S}} + q^{\mathcal{L}} p^{\mathcal{L}} + \big(1 - q^{\mathcal{S}} - q^{\mathcal{L}}\big)\underline{p}^{\mathcal{R}},\end{aligned}$$

where, once again, the expectation is taken with respect to the distribution $\tilde{\pi}(\cdot|h^t, \tau^t)$. This implies that

$$q^{\mathcal{S}} \leq \frac{\varepsilon}{1 - \underline{p}^{\mathcal{R}}} \quad \text{and} \quad q^{\mathcal{L}} p^{\mathcal{L}} \leq \tilde{p}\big(h^t, \tau^t\big), \tag{4}$$

where the first inequality follows from the fact that either $q^{\mathcal{L}} p^{\mathcal{L}} = 0$ or $p^{\mathcal{L}} > p_0 \geq \underline{p}^{\mathcal{R}}$ (because beliefs follow a martingale and the beliefs at all last histories are strictly greater than p_0) and $\tilde{p}(h^t, \tau^t) - \underline{p}^{\mathcal{R}} < \varepsilon$.

Now note that at any last history $(h^{t'}, \tau^{t'})$, we must have $\underline{\Pi}(\tilde{p}(h^{t'}, \tau^{t'})) \geq 0$ (because otherwise $\tilde{x}(h^{t'}, \tau^{t'}) > 0$ cannot be an equilibrium action by the principal) and

$$\begin{aligned}\underline{\Pi}\big(\tilde{p}\big(h^{t'}, \tau^{t'}\big)\big) &= \underline{\Pi}\big(\mathbb{E}\big[\tilde{p}\big(h^{t'}\hat{h}^{s'-t'}, \tau^{t'}\hat{\tau}^{s'-t'}\big) \mid \big(h^{t'}\hat{h}^{s'-t'}, \tau^{t'}\hat{\tau}^{s'-t'}\big) \in \tilde{\mathcal{L}}^{s'}\big]\big) \\ &= \mathbb{E}\big[\underline{\Pi}\big(\tilde{p}\big(h^{t'}\hat{h}^{s'-t'}, \tau^{t'}\hat{\tau}^{s'-t'}\big)\big) \mid \big(h^{t'}\hat{h}^{s'-t'}, \tau^{t'}\hat{\tau}^{s'-t'}\big) \in \tilde{\mathcal{L}}^{s'}\big],\end{aligned} \tag{5}$$

where the expectation is taken with respect to $\tilde{\pi}(\cdot|h^{t'}, \tau^{t'})$. The first equality is a consequence of the martingale property of beliefs and the second follows from the fact that $\underline{\Pi}(\cdot)$ is linear.

We can write the principal's payoff at (h^t, τ^t), for a $s \geq 1$ when her realized action choice is to experiment by summing four separate terms: (i) the expected continuation value after a success arrives from a nonlast history in these s periods, (ii) the expected continuation value at last histories, (iii) the cost of experimentation, transfers and failures at histories $(h^t h^{s'}, \tau^t \tau^{s'}) \in \tilde{\mathcal{R}}^{t+s'}$, $0 \leq s' \leq s - 1$, and (iv) the expected continuation value at all remaining period-$t + s + 1$ histories $(h^t h^s, \tau^t \tau^s) \in \tilde{\mathcal{R}}^{t+s}$.

We can now derive a simple upper bound

$$\underbrace{q^{\mathcal{S}}(1+\delta\overline{\Pi})}_{\text{(i)}}+\underbrace{\delta q^{\mathcal{L}}\mathbb{E}\big[\underline{\Pi}\big(\tilde{p}\big(h^t\hat{h}^s,\tau^t\hat{\tau}^s\big)\big)\mid\big(h^t\hat{h}^s,\tau^t\hat{\tau}^s\big)\in\tilde{\mathcal{L}}^{t+s}\big]}_{\text{(ii)}}-\underbrace{c}_{\text{(iii)}}+\underbrace{\delta^s\overline{\Pi}}_{\text{(iv)}},$$

for the principal's payoff when she experiments at (h^t,τ^t) and each term is labeled to individually correspond to a bound for each above mentioned component (i)–(iv) of the principal's payoff. The expectation is taken with respect to the distribution $\tilde{\pi}(\cdot|h^t,\tau^t)$.

(i) This term is an upper bound because it assumes that the successes from nonlast histories that arrive between periods $t+1$ and $t+s$ arrive immediately.

(ii) This term upper bounds the sum of payoffs at last histories by assuming that these payoffs are not discounted beyond period $t+1$. Consider a last history $(h^th^{s'},\tau^t\tau^{s'})$, $1\le s'\le s$. From the perspective of period $t+1$, the payoff from this history is bounded above by

$$\begin{aligned}0&\le\delta^{s'}\tilde{\pi}\big(h^th^{s'},\tau^t\tau^{s'}|h^t,\tau^t\big)\underline{\Pi}\big(\tilde{p}\big(h^th^{s'},\tau^t\tau^{s'}\big)\big)\\&=\delta^{s'}\tilde{\pi}\big(h^th^{s'},\tau^t\tau^{s'}|h^t,\tau^t\big)\underline{\Pi}\big(\mathbb{E}\big[\tilde{p}\big(h^th^{s'}h^{s-s'},\tau^t\tau^{s'}\tau^{s-s'}\big)\mid\big(h^th^{s'}h^{s-s'},\tau^t\tau^{s'}\tau^{s-s'}\big)\in\tilde{\mathcal{L}}^{t+s}\big]\big)\\&\le\delta\tilde{\pi}\big(h^th^{s'},\tau^t\tau^{s'}|h^t,\tau^t\big)\mathbb{E}\big[\underline{\Pi}\big(\tilde{p}\big(h^th^{s'}h^{s-s'},\tau^t\tau^{s'}\tau^{s-s'}\big)\big)\mid\big(h^th^{s'}h^{s-s'},\tau^t\tau^{s'}\tau^{s-s'}\big)\in\tilde{\mathcal{L}}^{t+s}\big],\end{aligned}$$

where the expectations are taken with respect to $\tilde{\pi}(\cdot|h^th^{s'},\tau^t\tau^{s'})$. Summing over all last histories and using the law of iterated expectations gives us the required bound.

(iii) This term only accounts for the cost of experimentation at (h^t,τ^t) but no subsequent costs of experimentation, transfers, or losses due to failures.

(iv) This term is trivially an upper bound for the continuation payoff as it assumes that the principal gets the first best payoff at period $t+s+1$ for sure.

Now observe that

$$\begin{aligned}&q^{\mathcal{S}}(1+\delta\overline{\Pi})+\delta q^{\mathcal{L}}\mathbb{E}\big[\underline{\Pi}\big(\tilde{p}\big(h^t\hat{h}^s,\tau^t\hat{\tau}^s\big)\big)\mid\big(h^t\hat{h}^s,\tau^t\hat{\tau}^s\big)\in\tilde{\mathcal{L}}^{t+s}\big]-c+\delta^s\overline{\Pi}\\&\quad=q^{\mathcal{S}}(1+\delta\overline{\Pi})+\delta q^{\mathcal{L}}\underline{\Pi}\big(p^{\mathcal{L}}\big)-c+\delta^s\overline{\Pi}\\&\quad=q^{\mathcal{S}}(1+\delta\overline{\Pi})+\delta\underline{\Pi}\big(q^{\mathcal{L}}p^{\mathcal{L}}\big)+\delta\big(1-q^{\mathcal{L}}\big)c-c+\delta^s\overline{\Pi}\\&\quad\le\frac{\varepsilon}{1-\underline{p}^{\mathcal{R}}}(1+\delta\overline{\Pi})-\big(1-\delta\big(1-q^{\mathcal{L}}\big)\big)c+\delta^s\overline{\Pi},\end{aligned}$$

where the inequality follows from (4) and the fact that $q^{\mathcal{L}}p^{\mathcal{L}}\le\tilde{p}(h^t,\tau^t)\le p_0$, $\underline{\Pi}(p_0)\le 0$. This last term is negative is we take ε sufficiently small and s sufficiently large. This shows that there must exist an on-path history $(\hat{h}^{t'},\hat{\tau}^{t'})\in\tilde{\mathcal{R}}^{t'}$ where $\tilde{x}(\hat{h}^{t'},\hat{\tau}^{t'})>0$ and the principal's payoff is less than 0, which contradicts the existence of a nontrivial equilibrium and completes the proof. □

We now show that Theorem 4 implies Theorem 3.

Proof of Theorem 3. Take a NE $(\tilde{x}, \tilde{a}_\theta)$ of the original game and consider the following strategies $(\tilde{x}', \tilde{\tau}', \tilde{a}'_\theta)$ in the game with transfers:

$$\tilde{x}'(h^t, \tau^t) = \begin{cases} \tilde{x}(h^t, \tau^t) & \text{if } h^t \in \tilde{\mathcal{H}}^t \text{ and } \tau^t = (0, \dots, 0), \\ 0 & \text{otherwise,} \end{cases}$$

$$\tilde{\tau}'(h^{t+1}, \tau^t) = 0 \quad \text{(the Dirac measure at 0)},$$

$$\tilde{a}'_{\theta_b}(h^t, \tau^t) = \begin{cases} \tilde{a}_{\theta_b}(h^t, \tau^t) & \text{if } h^t \in \tilde{\mathcal{H}}^t \text{ and } \tau^t = (0, \dots, 0), \\ 0 & \text{otherwise,} \end{cases}$$

$$\tilde{a}'_{\theta_g}(h^t, \tau^t, i^t) = \begin{cases} \tilde{a}_{\theta_g}(h^t, \tau^t, i^t) & \text{if } h^t \in \tilde{\mathcal{H}}^t \text{ and } \tau^t = (0, \dots, 0), \\ 0 & \text{otherwise,} \end{cases}$$

where $\tilde{\mathcal{H}}^t$ is the set of period-$t+1$ histories that are on path when players use strategies $(\tilde{x}, \tilde{a}_\theta)$ in the game without transfers. In words, these strategies are identical to $(\tilde{x}, \tilde{a}_\theta)$ at histories where the principal has not made a transfer in the past. Hence, the only on-path histories $\tilde{\mathcal{T}}^t$ when players use strategies $(\tilde{x}', \tilde{\tau}', \tilde{a}'_\theta)$ in the game with transfers will be those where the principal never makes a transfer. Note that $(\tilde{x}', \tilde{\tau}', \tilde{a}'_\theta)$ will be a NE because, if either player deviates off-path, they receive a payoff of 0 since the principal stops experimenting and the agent stops acting. Finally, note that every equilibrium outcome of the original game is also an equilibrium of the game with transfers; indeed we can use the identical construction above. This is because the refinement does not restrict transfers after a success is generated. Thus, Theorem 3 is an immediate consequence of Theorem 4. □

Appendix B: Proof of Theorem 5

Before proving the result, we first formally define the strategies that were informally described in Figure 3.

The principal follows the following strategy:

- Experiment in period one and experiment forever (stop experimenting) if a success (failure) is observed. Formally, $\tilde{x}(h^0) = 1$, $\tilde{x}(\overline{h}h^t) = 1$ and $\tilde{x}(\underline{h}h^t) = 0$ for all $h^t \in \mathcal{H}$.
- If the agent does not act in period one, experiment in period two. If a success is observed, experiment at all subsequent periods. Formally, $\tilde{x}(h_\varphi) = 1$, and $\tilde{x}(h_\varphi \overline{h} h^t) = 1$ for all $h^t \in \mathcal{H}$.
- If the agent does not act in period one and no success is observed in period two, experiment in period three. Stop experimenting from period four onward. Formally, $\tilde{x}(h_\varphi h_2) = 1$, $\tilde{x}(h_\varphi h_2 h^t) = 0$ for $h_2 \in \{h_\varphi, \underline{h}\}$ and for all $h^t \in \mathcal{H}$, $t \geq 1$.

In response, the good type's strategy is the following:

- Only run good projects in period one. If a success is generated, follow the efficient strategy. If a failure is generated (off-path), stop acting. Formally,

$$\tilde{a}_{\theta_g}(h^0, i_1) = \begin{cases} 1 & \text{if } i_1 = i_g, \\ 0 & \text{if } i_1 \in \{i_r, i_b\}, \end{cases}$$

$$\tilde{a}_{\theta_g}(\overline{h}h^{t-1}, i^t i_{t+1}) = \begin{cases} 1 & \text{if } i_{t+1} = i_g, \\ 0 & \text{otherwise,} \end{cases} \quad \text{for all } (\overline{h}h^{t-1}, i^t i_{t+1}) \in \mathcal{H}^g,$$

$$\tilde{a}_{\theta_g}(\underline{h}h^{t-1}, i^{t+1}) = 0 \quad \text{for all } (\underline{h}h^{t-1}, i^{t+1}) \in \mathcal{H}^g.$$

- If no project was run in period one, implement both good and risky projects in period two. If a success is generated, follow the efficient strategy. Formally,

$$\tilde{a}_{\theta_g}(h_\varphi h_2, i_1 i_2) = \begin{cases} 1 & \text{if } i_2 \in \{i_g, i_r\}, \\ 0 & \text{if } i_2 = i_b, \end{cases}$$

$$\tilde{a}_{\theta_g}(h_\varphi \overline{h}h^{t-2}, i_1 i_2 i^{t-2} i_{t+1}) = \begin{cases} 1 & \text{if } i_{t+1} = i_g, \\ 0 & \text{otherwise,} \end{cases} \quad \text{for all } (h_\varphi \overline{h}h^{t-2}, i_1 i_2 i^{t-2} i_{t+1}) \in \mathcal{H}^g.$$

- If no project was run in period one and no success is observed in period two, only implement good projects in period three. Stops acting thereafter regardless of the period three outcome. Formally,

$$\tilde{a}_{\theta_g}(h_\varphi h_2, i^2 i_3) = \begin{cases} 1 & \text{if } i_3 = i_g, \\ 0 & \text{otherwise,} \end{cases} \quad \text{for all } h_2 \in \{\underline{h}, h_\varphi\} \text{ and } (h_\varphi h_2, i^2 i_3) \in \mathcal{H}^g,$$

$$\tilde{a}_{\theta_g}(h_\varphi h_2 h^t, i^{t+3}) = 0 \quad \text{for all } h_2 \in \{\underline{h}, h_\varphi\} \text{ and } (h_\varphi h_2 h^t, i^{t+3}) \in \mathcal{H}^g,\ t \geq 1.$$

Finally, the bad type never acts or, formally, that $\tilde{a}_{\theta_b}(h^t) = 0$ for all $h^t \in \mathcal{H}$.

We now proceed to proof of the result.

Proof of Theorem 5. Fix any $\underline{\delta} \in (0, 1)$ and $\underline{\lambda}_g < \overline{\lambda}_g < \frac{1}{2}$. First, observe that we can always find a low enough $\overline{c} \in (0, 1)$ to satisfy

$$\big(1 - (\overline{\lambda}_g + 1/2)\big)(\underline{\lambda}_g - \overline{c}) > \frac{1 + \underline{\delta}}{\underline{\delta}}\overline{c}. \tag{6}$$

Note that the above expression implies that $\underline{\lambda}_g > c$. In the remainder of the proof, we implicitly assume that $\delta \in (\underline{\delta}, 1)$, $\lambda_g \in (\underline{\lambda}_g, \overline{\lambda}_g)$, $\lambda_r \in (0, \frac{1}{2})$, and $c \in (0, \overline{c})$.

Now note that there is always a set of values of κ, $\underline{\kappa}$, $\overline{\kappa}$ that satisfy the condition

$$0 < (\lambda_g + \lambda_r q_r)(1 + \delta\overline{\Pi}) - \lambda_r(1 - q_r)\kappa < c. \tag{7}$$

Observe that, when the above condition holds, quality control is necessary irrespective of the initial belief p_0. This is because

$$\begin{aligned} \underline{\Pi}(p_0) &= p_0(\lambda_g + \lambda_r q_r)(1 + \delta\overline{\Pi}) - p_0\lambda_r(1 - q_r)\kappa - c \\ &< (\lambda_g + \lambda_r q_r)(1 + \delta\overline{\Pi}) - \lambda_r(1 - q_r)\kappa - c < 0. \end{aligned}$$

Note, (7) also implies $\underline{\Pi}(p_0) > -c$. Henceforth, we implicitly assume that κ satisfies (7).

We now derive the belief bound $\underline{p}_0$ such that the principal's strategy defined in Figure 3 is a best response for all prior beliefs $p_0 \in (\underline{p}_0, 1)$. Clearly, her continuation strategy is a best response to a success being observed in either periods one or two and to a failure being observed in period one. It remains to be shown that experimenting is a best response in each of the remaining histories in periods one to three.

First, observe that the principal's payoff in period three when the agent does not take an action in the first two periods satisfies

$$\begin{aligned}\tilde{p}(h_\varphi h_\varphi)\lambda_g - c &= \frac{p_0(1-\lambda_g)(1-\lambda)}{p_0(1-\lambda_g)(1-\lambda)+(1-p_0)}\lambda_g - c \\ &> \frac{p_0(1-\overline{\lambda}_g)\big(1-(\overline{\lambda}_g+1/2)\big)}{p_0(1-\overline{\lambda}_g)\big(1-(\overline{\lambda}_g+1/2)\big)+(1-p_0)}\underline{\lambda}_g - \overline{c}.\end{aligned}$$

Clearly, there exists a $\underline{p}' \in (0, 1)$ such that

$$\frac{p_0(1-\overline{\lambda}_g)\big(1-(\overline{\lambda}_g+1/2)\big)}{p_0(1-\overline{\lambda}_g)\big(1-(\overline{\lambda}_g+1/2)\big)+(1-p_0)}\underline{\lambda}_g - \overline{c} > 0$$

for all $p_0 \geq \underline{p}'$ or, in words, that the principal's payoff at this history is positive for high enough initial belief.

The principal's expected payoff in period two when the agent does not act in period one is positive if

$$\begin{aligned}&\underline{\Pi}\big(\tilde{p}(h_\varphi)\big) + \delta\tilde{p}(h_\varphi)\big(1-(\lambda_g+q_r\lambda_r)\big)(\lambda_g - c) - \delta\big(1-\tilde{p}(h_\varphi)\big)c > 0, \\ \iff\quad &\tilde{p}(h_\varphi)\big(1-(\lambda_g+q_r\lambda_r)\big)(\lambda_g - c) > \big(1-\tilde{p}(h_\varphi)\big)c - \frac{\underline{\Pi}\big(\tilde{p}(h_\varphi)\big)}{\delta}.\end{aligned}$$

We now argue that this condition will be true for sufficiently high p_0. First, observe that

$$\tilde{p}(h_\varphi) = \frac{p_0(1-\lambda_g)}{p_0(1-\lambda_g)+(1-p_0)},$$

is increasing in p_0 and is 1 when $p_0 = 1$. Then note that

$$\begin{aligned}&\tilde{p}(h_\varphi)\big(1-(\lambda_g+q_r\lambda_r)\big)(\lambda_g - c) \\ &\quad > \tilde{p}(h_\varphi)\big(1-(\overline{\lambda_g}+1/2)\big)(\underline{\lambda}_g - \overline{c}) \\ &\quad > \frac{1+\underline{\delta}}{\underline{\delta}}\overline{c} > \big(1-\tilde{p}(h_\varphi)\big)c + \frac{c}{\delta} > \frac{1+\delta\big(1-\tilde{p}(h_\varphi)\big)}{\delta}c - \frac{\underline{\Pi}\big(\tilde{p}(h_\varphi)\big)}{\delta}\end{aligned}$$

for sufficiently high p_0. To see this, note that the second inequality holds when p_0 is high because of (6) and the last inequality is a consequence of the fact that $\underline{\Pi}(p_0) > -c$. Let $\underline{p}''$ be a value of p_0 such that the second inequality is satisfied and note that the principal's expected payoff in period two when the agent does not act in period one is positive for all $p_0 \geq \underline{p}''$.

Finally, the principal's expected payoff in period one is positive since her period one stage game payoff is positive: observe that the prior belief $p_0 > \tilde{p}(h_\varphi h_\varphi)$ and the latter was assumed above to be sufficiently high to make the period three stage game payoff positive.

So set $\underline{p}_0 = \max\{\underline{p}', \underline{p}''\}$ and note that for all values $p_0 \in (\underline{p}_0, 1)$, $q_r \in (0, 1)$, $\delta \in (\underline{\delta}, 1)$, $\lambda_g \in (\underline{\lambda}_g, \overline{\lambda}_g)$, $\lambda_r \in (0, \frac{1}{2})$, $c \in (0, \overline{c})$, there is a range of values of κ (satisfying (7)) such that the principal's strategy is a best response.

It remains to be argued that the agent's strategy is a best response. First, note that the bad type never has an incentive to act since he can never generate a success and this is the only way to get the principal to experiment after period three. The good type is indifferent between acting or not in period three and strictly prefers to run both good and risky projects in period two (while being indifferent about whether or not to implement a bad project).

Finally, we argue that when $q_r \in (0, \underline{\lambda}_g)$ the good type strictly prefers to not run a risky project in period one; clearly, he strictly prefers to not run a bad project. To see this, observe that his expected continuation payoff from running a risky project is $q_r/(1-\beta)$. If he chooses to not run a project in period one, his expected continuation payoff is

$$\frac{\lambda_g + q_r\lambda_r}{1-\beta} + \big(1 - (\lambda_g + q_r\lambda_r)\big)(1+\beta).$$

In words, the good type has the option value of waiting till period two and, if he generates a success he receives the efficient continuation payoff (if not, the principal experiments for one more period). Clearly, this option value is strictly greater when $q_r < \lambda_g$ *irrespective of his discount factor* β.

To summarize, the above argument derives bounds such that for any $p_0 \in (\underline{p}_0, 1)$, $\delta \in (\underline{\delta}, 1)$, $\beta \in (0, 1)$, $\lambda_g \in (\underline{\lambda}_g, \overline{\lambda}_g)$, $\lambda_r \in (0, \frac{1}{2})$, $q_r \in (0, \underline{\lambda}_g)$ and $c \in (0, \overline{c})$, there exist values of κ such that the proposed strategies constitute a NE.

We complete the proof by noting that the principal's payoff satisfies

$$V\big(h^0, \tilde{x}, \tilde{a}_\theta\big) = p_0\lambda_g(1+\delta\overline{\Pi}) - c + \delta(1-p_0\lambda_g)V(h_\varphi, \tilde{x}, \tilde{a}_\theta) \geq p_0\lambda_g(1+\delta\overline{\Pi}) - c$$

and the good type's payoff satisfies

$$U_{\theta_g}\big(h^0, i^0, \tilde{x}, \tilde{a}_{\theta_g}\big) = \frac{\lambda_g}{1-\beta} + (1-\lambda_g)\big(1 + \beta U_{\theta_g}(h_\varphi, i_1, \tilde{x}, \tilde{a}_{\theta_g})\big) \geq \frac{\lambda_g}{1-\beta}$$

where $i_1 = i_b$ (we can equivalently plug in $i_1 = i_r$ since the good type's strategy is identical in both cases). □

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1993), "Renegotiation and symmetry in repeated games." *Journal of Economic Theory*, 60, 217–240. [773]

Aghion, Philippe and Matthew O. Jackson (2016), "Inducing leaders to take risky decisions: Dismissal, tenure, and term limits." *American Economic Journal: Microeconomics*, 8, 1–38. [768]

Atakan, Alp E. and Mehmet Ekmekci (2012), "Reputation in long-run relationships." *Review of Economic Studies*, 79, 451–480. [767]

Atakan, Alp E. and Mehmet Ekmekci (2013), "A two-sided reputation result with long-run players." *Journal of Economic Theory*, 148, 376–392. [767]

Backus, Matthew and Andrew T. Little (2018), "I don't know." Unpublished manuscript, Columbia University. [768]

Cripps, Martin W. and Jonathan P. Thomas (1997), "Reputation and perfection in repeated common interest games." *Games and Economic Behavior*, 18, 141–158. [767]

Deb, Rahul, Mallesh M. Pai, and Maher Said (2018), "Evaluating strategic forecasters." *American Economic Review*, 108, 3057–3103. [784]

Ely, Jeffrey, Drew Fudenberg, and David K. Levine (2008), "When is reputation bad?" *Games and Economic Behavior*, 63, 498–526. [767]

Ely, Jeffrey C. and Juuso Välimäki (2003), "Bad reputation." *Quarterly Journal of Economics*, 118, 785–814. [767]

Farrell, Joseph and Eric Maskin (1989), "Renegotiation in repeated games." *Games and Economic Behavior*, 1, 327–360. [773]

Guo, Yingni (2016), "Dynamic delegation of experimentation." *American Economic Review*, 106, 1969–2008. [784]

Halac, Marina (2012), "Relational contracts and the value of relationships." *American Economic Review*, 102, 750–779. [767]

Jovanovic, Boyan (1979), "Job matching and the theory of turnover." *Journal of Political Economy*, 87, 972–990. [785]

Levin, Jonathan (2003), "Relational incentive contracts." *American Economic Review*, 93, 835–857. [767]

Li, Jin, Niko Matouschek, and Michael Powell (2017), "Power dynamics in organizations." *American Economic Journal: Microeconomics*, 9, 217–241. [767]

Malcomson, James M. (2016), "Relational incentive contracts with persistent private information." *Econometrica*, 84, 317–346. [768]

Mitchell, Matthew (2021), "Free (ad)vice." *RAND Journal of Economics*, 52, 3–21. [768]

Morris, Stephen (2001), "Political correctness." *Journal of Political Economy*, 109, 231–265. [768]

Ottaviani, Marco and Peter Norman Sørensen (2006), "Reputational cheap talk." *RAND Journal of Economics*, 37, 155–175. [768]

Prendergast, Canice and Lars Stole (1996), "Impetuous youngsters and jaded old-timers: Acquiring a reputation for learning." *Journal of Political Economy*, 1105–1134. [768]

Safronov, Mikhail and Bruno Strulovici (2017), "From self-enforcing agreements to self-sustaining norms." Unpublished manuscript, Northwestern University. [773]

Schmidt, Klaus M. (1993), "Reputation and equilibrium characterization in repeated games with conflicting interests." *Econometrica*, 61, 325–351. [767, 777]

Co-editor Marina Halac handled this manuscript.

Manuscript received 19 March, 2021; final version accepted 19 July, 2021; available online 22 July, 2021.