(BAD) REPUTATION IN RELATIONAL CONTRACTING

RAHUL DEB[◊], MATTHEW MITCHELL[†], AND MALLESH M. PAI[‡]

MARCH 2, 2021

ABSTRACT: 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.

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

—Malcolm X

Attention is currency for much of the advertising driven Internet. Typically, consumers do not pay for content and instead revenue is generated by their continued attention in the form of clicks. A consequence is that content providers need to sustain interest as consumers can freely withdraw their attention at any time. This creates a dilemma: since 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? Specifically, how do they manage reputation in an environment where

[◊]DEPARTMENT OF ECONOMICS, UNIVERSITY OF TORONTO, RAHUL.DEB@UTORONTO.CA
[†]GRADUATE DEPARTMENT OF MANAGEMENT, UNIVERSITY OF TORONTO, MATTHEW.MITCHELL@UTORONTO.CA
[‡]DEPARTMENT OF ECONOMICS, RICE UNIVERSITY, 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, Marina Halac, 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.

"fake" content can be generated at will? And how does this revenue model (as opposed to traditional subscription payments) affect market functioning?

Although it is a very different context, similar incentives are faced by elected politicians. The ability of the politician determines whether or not he has effective policy ideas to address the changing economic and social needs of his constituents. 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? On balance, are institutions like terms limits (that dampen reputational forces) efficient given that good politicians are forced to leave office?

Finally, similar incentives are also faced by "experts" employed in organizations. To fix ideas, consider a scientist working at a pharmaceutical company. The scientist's ability determines the frequency at which he receives research ideas. These ideas may ultimately lead to new products after successful drug trials. Even a talented scientist, when under pressure to keep his job, may choose to recommend a costly drug trial for a new compound which is very unlikely to succeed. After all, this may help create the impression that he is generating new ideas. How does the company's decision of when to fire the scientist and their bonus structure (for successful trials) balance their dual goals of giving a talented scientist a sufficient amount of time to succeed (and reveal himself to be of high ability), but also prevent costly failed trials?

What is common to these (and numerous other) seemingly disparate examples is that they are environments in which (1) a principal wants to dynamically screen an agent with a privately known type (good or bad) (2) the type determines the *rate* at which she receives information that is payoff relevant for the principal, and, (3) the agent acts strategically based on this information in an effort to manage his *reputation*. 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 description that follows, we link various elements of the model to the first application. In Section 5, we develop in detail the implications of our results for all the three settings mentioned above.

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 (that is, 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, there is a unique Nash equilibrium outcome in which the principal never experiments. This is because 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* (that is, 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 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 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

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

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 above 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 our applications, this corresponds to an unwillingness to visit a website with a consistently low standard of reporting etc. 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.

So what can we say about the set of equilibria in our model? Our first main result (part 1 of Theorem 3) shows that the principal can never completely prevent the agent from choosing inefficient actions in any equilibrium. Specifically, 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, we show that, whenever quality control is necessary for experimentation, the unique equilibrium outcome is one in which the principal never experiments. This result shows that 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. First, observe that 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, it might be natural

[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 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. Returning to our first application, this implies that even a small amount of doubt in the veracity of a content provider (like a news source) can have a large impact for consumers. Specifically, the incentives that drive relationship breakdown in our model mirror the oft-cited criticism of the news media: the need to constantly generate new content drives down the quality of reporting. This in turn engenders public mistrust, which contributes to the environment of dwindling web traffic.

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 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. We examine strategies in which the principal experiments at least for three periods (on path). If the agent generates a success in the first two periods, he is granted "tenure" and first best continuation play follows. If not, the principal experiments for one additional "grace period." Critically, and in violation of the refinement, the principal stops experimenting after this grace period irrespective of whether or not a success arises in this period. In this grace period, the agent acts efficiently—this is a best response because he is indifferent over all strategies (since he is guaranteed to be employed for exactly this grace period regardless of what he does). We identify parameter values for which quality control is necessary, and nevertheless, these strategies are a Nash equilibrium. Further, we show that in this Nash equilibrium, both players get at least a fraction of the first-best payoff (irrespective of their discount factors).

We then develop the applied implications of these results. In Section 5.1, we discuss the online content market and argue that moving from advertising driven revenue to subscription based payment can potentially incentivize higher quality, instead of clickbait, journalism. In Section 5.2, we argue that a slight variant of the model (to which the same insights apply) can be used to demonstrate that term limits can improve policy making not despite, but precisely because, even proven good politicians must leave office at the end of their term. In Section 5.3, we discuss the implications of our results to how, and which, experts should be hired in organizations. Additionally, our model allows us to take a nuanced view of employee job terminations.

Finally, in the discussion Section 6, we argue that our main insight is robust to perturbing various assumptions of the model. Specifically, we allow for one-sided transfers, costly actions and costly information acquisition. Finally, we show that commitment power can benefit the principal.

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 the applications our model speaks to.

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), more recent papers are Atakan and Ekmekci (2012, 2013). As we discuss, our main result (Theorem 3) combines key ideas from both literatures.

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. This 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, since we present the bulk of our results in the absence of transfers, the works of Li, Matouschek, and Powell (2017) and Mitchell (2020) are also related although their problems concern 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.

Finally, our paper is related to a few papers in the literature on dynamic mechanism design which analyze outcomes both with and without commitment. Guo (2016) studies a dynamic setting without transfers where an agent is privately informed about the quality of a risky arm and he prefers greater experimentation than the principal (both prefer the risky arm only when it is good). While our setting differs substantially in terms of the payoff structure (and, hence the application to which we speak), another critical difference is that the agent in our environment strategically acts in response to additional private information she receives over time. 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 to the agent to take "risky" actions that are potentially damaging to the latter's reputation. Deb, Pai, and Said (2018) study a pure adverse selection dynamic environment without transfers where a principal is trying to determine whether a strategic political forecaster has accurate information by comparing the dynamics of his predictions leading up to an election with the eventual winner. Once again, there are some superficial similarities in the payoff structure but the games are otherwise quite different.

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 $t \in \{1, \ldots, \infty\}$; 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.

Agent's Information: In each period, the agent receive 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 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, λ_r 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 $\underline{o}$ 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 (risky) project always (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,

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

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, $h^t = (h_1, \ldots, h_t)$ and $i^t = (i_1, \ldots, i_t)$. 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, referred to simply as 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 $\mathscr{H}$ and $\mathscr{H}^g$ 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 $h^{t-1} \in \mathscr{H}$. 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 $(h^{t-1}, i^t) \in \mathscr{H}^g$. 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 $h^{t-1} \in \mathscr{H}$ 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 $(h^{t-1}, i^t) \in \mathscr{H}^g$ is said to be on path (off path) if it can (cannot) be reached with positive probability when $\theta = \theta_g$.

[3]Note that $\mathscr{H}^g$ 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.

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 $h^{t-1} \in \mathscr{H}$, $(h^{t-1}, i^{t-1}) \in \mathscr{H}^g$ 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 $\tilde{x}'$, $\tilde{a}'_\theta$.

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.

Figure 1 below summarizes the actions and the order of play in the repeated game.

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 extremely 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. As we will discuss below, 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.

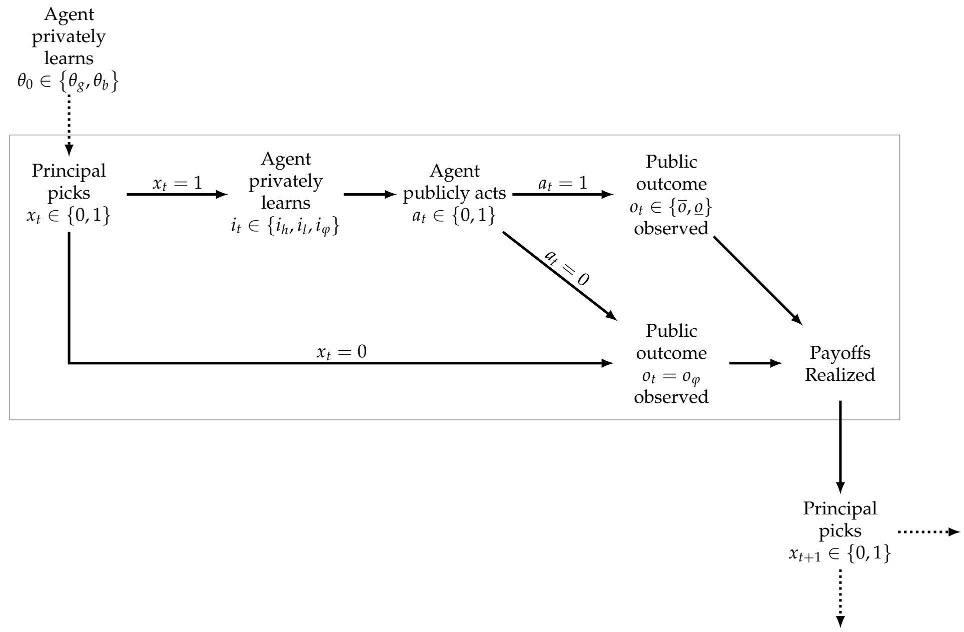

FIGURE 1. Flow chart describing the repeated game; the rectangle contains the stage game.

We show in Section 6 that our main insight is unaffected if we dispense with assumptions (ii), (iii) or (iv). The bulk of the paper deliberately focuses on the above simple version of the model to reduce notation and make the results transparent.

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 sans 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 these mathematical statements also appear in the 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, suppose the agent is known to be the good type. Then, there is a unique Pareto-optimal NE outcome in which 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 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.

[4]Formally, the principal's unique NE strategy when $p_0 = 0$ is $\tilde{x}(h^{t-1}) = 0$ for all $h^{t-1} \in \mathscr{H}$. 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 $h^{t-1} \in \mathscr{H}$ and $\tilde{a}_{\theta_g}(h^{t-1}, i^{t-1}i_t) = 1(0)$ for all $(h^{t-1}, i^{t-1}i_t) \in \mathscr{H}^g$ where $i_t = (\neq) i_g$.

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 one, the continuation play is nontrivial.

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

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]

With this refinement in place, any loss of surplus is not a consequence of the players choosing maximally inefficient strategies under complete information; instead, inefficiency arises exclusively from frictions created by the principal's need to screen.

[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 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 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.

2.3. *The "Static" Benchmark and the Need for Quality Control*

The goal of this section is to present, and provide a foundation for, the main parameter values that we focus on. Essentially, 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 one. 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

$$\tilde{x}(h^0) = 1 \quad \text{and} \quad \tilde{x}(h_1 h^{t-1}) = \begin{cases} 1 & \text{if } h_1 = \overline{h}, \\ 0 & \text{if } h_1 \neq \overline{h}, \end{cases} \qquad \text{for all } h_1 h^{t-1} \in \mathscr{H},\ t \geq 1,$$

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

$$\tilde{a}_{\theta_g}(h_1 h^{t-1}, i^t i_{t+1}) = \begin{cases} 1 & \text{if } h_1 = \overline{h} \text{ and } i_{t+1} = i_g, \\ 0 & \text{otherwise}, \end{cases} \qquad \text{for all } (h_1 h^{t-1}, i^t i_{t+1}) \in \mathscr{H}^g,\ t \geq 1.$$

$$\tilde{a}_{\theta_b}(h^0) > 0 \text{ and } \tilde{a}_{\theta_b}(h^t) = 0, \quad \text{for all } h^t \in \mathscr{H},\ t \geq 1. \tag{1}$$

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 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 (despite the game continuing after a success) 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 non-success outcomes is interior (since $\tilde{p}(\underline{h}), \tilde{p}(h_\varphi) < 1$) and thus the continuation play as prescribed by the above stragies (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:*

(1) 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 $h^t \in \mathscr{H}$ and $(h^{t'}, i^{t'} i_r) \in \mathscr{H}^g$ such that $\tilde{a}_{\theta_b}(h^t) > 0$ and $\tilde{a}_{\theta_g}(h^{t'}, i^{t'} i_r) = 1$.

(2) 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. The discussion also highlights our key theoretical contribution which 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. Since experimentation is costly, there is a belief below which experimentation is not worth it even if uncertainty about the agent's type could somehow 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 non-negative. 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 non-success 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.

4. Correcting (Bad) Reputational Incentives

In this section, we demonstrate how maximally inefficient behavior on path can help relationship functioning. At the most basic level, this can easily 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 $h^{t-1} \in \mathscr{H}$, $1 \leq t \leq T$ (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 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 both players nonetheless get positive average payoffs even as they become patient?

In the remainder of this section, we construct a simple NE and identify parameter values for which the answer to the above question is in the affirmative. Unlike the NE described at the beginning of this section, the following strategies have the feature that both players move to the efficient continuation equilibrium following successes at some histories. We describe strategies informally below in Figure 2 and in words; the formal definitions are in Appendix D.

The principal's strategy is the following:

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

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

When $t \geq T$, strategies are

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

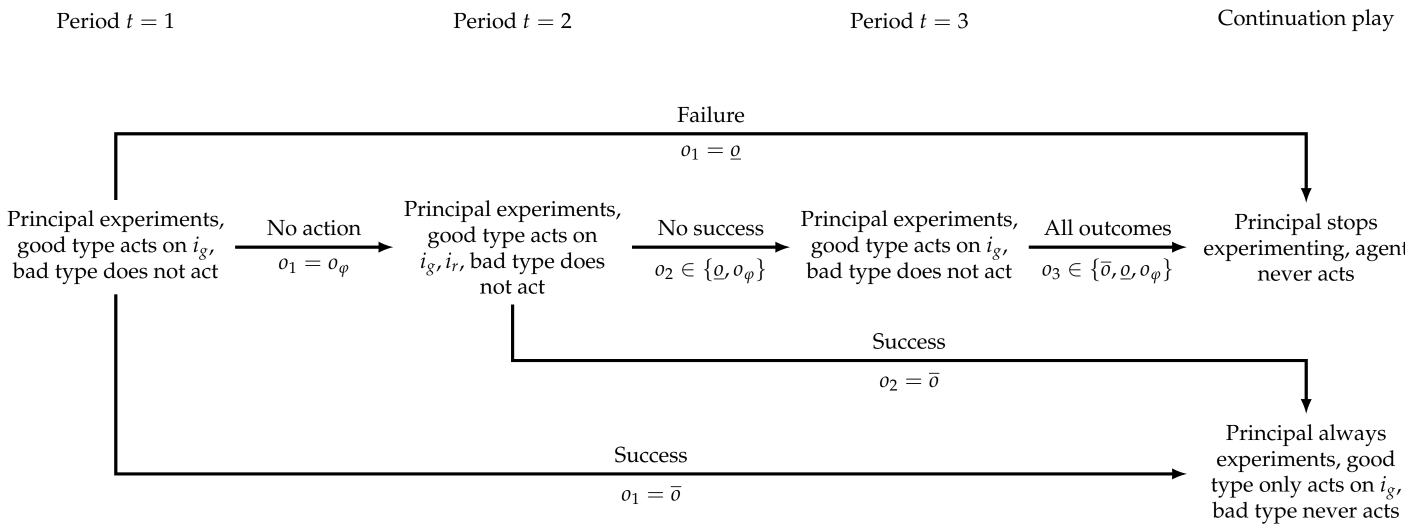

FIGURE 2. A NE with two screening periods when $\underline{\Pi}(p_0) \leq 0$.

- 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 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. Stops acting thereafter regardless of the period three outcome.

Finally, the bad type never acts.

The key feature of the above strategies is that both players move to efficient continuation play after successes in periods one or two, and (the grace) period three is used to provide incentives for both players. When the conditions

$$q_r \leq \lambda_g + \lambda_r q_r, \tag{C1}$$

$$(\lambda_g + \lambda_r q_r)\left(1 + \frac{\delta}{1-\delta}(\lambda_g - c)\right) - \lambda_r(1-q_r)\kappa = 0 \quad \text{and} \tag{C2}$$

$$\delta(1 - (\lambda_g + q_r\lambda_r))(\lambda_g - c) > c \tag{C3}$$

hold, we can find a high enough initial belief p_0 such that quality control remains necessary and, nonetheless, these strategies are a NE. In Appendix D, we argue that these conditions are not vacuous. Indeed, we show that the set of parameters for which there exists such an equilibrium are not non-generic.

Before discussing the implications of the example, we briefly argue why the above conditions ensure that the strategies are mutual best responses. 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 first inequality guarantees that the good type's strategy is also a best response. Observe that 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). Condition (C1) ensures that the good type always (at least weakly) prefers to not run a risky project in period one; clearly, he strictly prefers to not run a bad project. This is because the condition ensures that the option value of waiting till period two is better than implementing a risky project and potentially ending the relationship if a failure arises. Importantly, observe that, as long as the parameters satisfy condition (C1), the agent's strategy is a best response *irrespective of his discount factor* β.

The second condition (C2) implies that quality control is necessary irrespective of the initial belief p_0. This is because

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

We now argue that third condition (C3) combined with a sufficiently high initial belief p_0 implies that the principal's strategy is a best response. 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 because $\lambda_g > c$, the principal's payoff

$$\tilde{p}(h_\varphi h_\varphi)\lambda_g - c = \frac{p_0(1-\lambda_g)(1-\lambda)}{p_0(1-\lambda_g)(1-\lambda)+(1-p_0)}\lambda_g - c > 0,$$

when the agent does not take an action in the first two periods is positive in period three when p_0 is sufficiently high. Of course, if the agent generates a failure in period two, then the principal's belief jumps to 1 and her payoff is $\lambda_g - c > 0$.

The principal's expected payoff in period two is positive when

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

which, in turn, holds when p_0 is sufficiently high because of condition (C3). To see this, note that the principal's belief that the agent is the good type in period two when no action is taken in period one is

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

which is increasing in p_0 and is 1 when $p_0 = 1$.

Finally, the principal's expected payoff in period one is clearly positive since her period one stage game payoff is positive. To see this, 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.

The example shows that, despite quality control being necessary, *both* players can get a fraction of the first best surplus because continuation play is efficient after successes in period one and two and no failures arise on path in period one. Importantly, conditions (C1)–(C3) can be satisfied for all values of δ, β.

A few further comments about this construction are worth making. First, 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 ($p_0\lambda_g$) of the first-best payoff $\overline{\Pi}$ for all parameter values that satisfy conditions (C1)–(C3); 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 vanishing small fraction of the first best.[9]

Second, note that these strategies obviously do not describe the principal-optimal NE. While it is prohibitively hard to derive the principal-optimal NE in such a general environment, we feel that this example suffices to make our main economic point about the usefulness of maximally inefficient coordination.

Taken together, Theorem 3 and the example in this section 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 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. The next section demonstrates this by developing three very different applications where the forces of our model arise.

[9]For instance take $\delta \to 1$ and simultaneously make κ larger so that (C2) continues to hold. 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. Note that making the principal more patient only slackens the constraint (C3).

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.[10] Conversely, we too get a folk theorem if the principal becomes patient.[11] 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. Implications

The purpose of this section is to show that the main forces in our model are present in a variety of seemingly unconnected economic environments.

[10] Specifically, Theorem 7 in Appendix D shows that if the principal is myopic and the bad type agent is a commitment type that never acts, quality control being necessary implies that the good type's average payoff converges to 0 as he becomes arbitrarily patient ($\beta \to 1$).

[11] Theorem 8 in Appendix D shows that the principal can get an average payoff of $p_0(\lambda_g - c)$ when $\delta \to 1$.

5.1. *Subscriptions for Online Content*

In this first application, we use the model to argue that subscription-based payment (as opposed to advertising-driven revenue) can incentivize higher quality journalism by weakening the need to constantly generate clicks. We begin by reminding the reader how the model maps to this application. Consider a consumer of online content (the principal) deciding whether to visit the website of a content provider (the agent) whose journalistic quality is privately known. Visiting the website is costly (either cognitively or in terms of time) for the consumer. If the content provider is high quality, he periodically receives news stories (projects in the model) that vary in the quality of reporting; low quality providers cannot produce novel content. The provider decides whether or not to publish content knowing that the veracity of the story will be revealed once it is cross checked by other news outlets. The consumer wants to maximize (minimize) the amount of true/breaking (false/reprinted) content she consumes and the provider wants to maximize the number of clicks. Our refinement seems quite natural for this application as there is no obvious reason for, or interpretation of, maximal inefficiency in this market.

When the consumer demands a minimal quality of reporting (that is, 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 $\frac{1-\delta^T}{1-\delta}c$ 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.

So consider the following strategies. The Principal's strategy can be described as:

- Commits to experimenting for the first T periods.
- If a failure is ever observed, or the agent never acts in the first T periods, she never experiments after T (even if a success is additionally observed).
- Otherwise, i.e. if success(es) but no failure(s) are observed in these T periods, she experiments forever after T.

The strategy for the good-type agent is:

- Only runs good projects in the first $T-1$ periods. If a success is generated, he only runs good projects thereafter. If a failure arrives, he stops acting.
- Implements both good and risky projects in period T. If a success is generated, he only runs good projects thereafter.
- Does not act at all other histories.

The bad-type agent acts with small probability at period T, and does not act at all other histories.

In Appendix E, these strategies are formally defined and we provide parameter values for which they constitute an equilibrium despite quality control being necessary. In short, the bad type is clearly best responding since he is indifferent between all strategies.[12] Finally, the good type's strategy is a best response if he does not have an incentive to run risky projects at any period prior to T. A sufficient condition for this is condition (C1) from the previous section since the likelihood of getting a good project in the next period makes the option value of waiting more attractive than the lottery of generating a success today with probability q_r.

For the principal, partial commitment 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. Indeed, when $\underline{\Pi}(p_0)$ is negative but small, there are parameter values such that 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 (that is, 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

[12] Note that the bad type has no incentive to act prior to T (even though the good type acts with positive probability) because failures between periods 1 and $T-1$ are off path and are punished by the principal.

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."

More broadly, the recent financial resurgence of the New York Times due to its paying subscribers appears to be shifting the journalism industry towards a subscription-based model. Invariably such a change in business model is justified by, among other factors, the fact that it will allow a focus on high-quality journalism and away from articles aimed primarily at generating clicks.

5.2. *Term Limits for Politicians*

In this section, we describe a slight variant of our model to argue that term limits can improve policy making by reelection seeking politicians.[13] 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).[14] 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.[15] Consider symmetric equilibria in which every politician follows the same strategy once elected. An implication of our main result is

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

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

[15]Formally, $p_0(\lambda_g + \lambda_r q_r)\left(1 + \frac{\delta}{1-\delta}\lambda_g\right) - p_0\lambda_r(1 - q_r)\kappa \leq 0$.

that, if quality control is necessary, every elected politician never implements any projects in the *voter-optimal* symmetric equilibrium.[16]

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.[17] 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 at the beginning of 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.

5.3. *Relational Contracting in Firms*

Our results 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 5.1). This is because the latter typically have shorter contracts and have stronger incentives to prove expertise to extend employment duration. Additionally, our 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

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

[17]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.

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 6.1. 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. Discussion

To facilitate, presentation, we have deliberately analyzed the simplest instance of the model in the previous sections and we now demonstrate that the main economic insights are not fragile. Specifically, we show that we can incorporate one-sided transfers, costly actions and costly information acquisition. We also show that the principal can strictly benefit from the ability to commit. The reader who is uninterested in the specific details can jump ahead to the concluding remarks.

6.1. *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 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); instead, we impose it because we feel it is realistic for our applications—it would be highly unusual for a firm to be making bonus payments to an expert not in their employ—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 3 below 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 4. *Suppose $p_0 \in (0, 1)$. Then:*

(1) 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.

(2) 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.

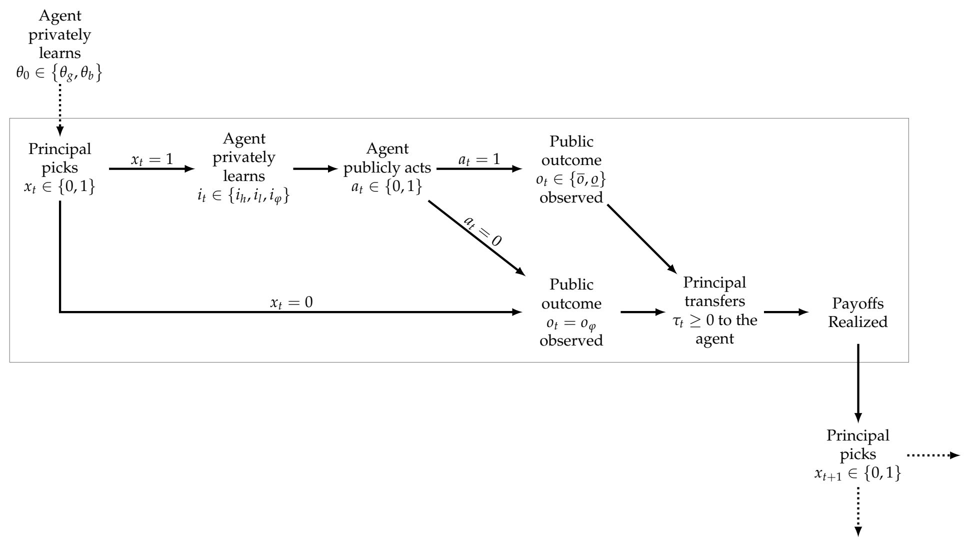

FIGURE 3. Flow chart describing the repeated game with transfers (the rectangle contains the stage game)

Transfers do not help the principal avoid inefficient actions or the resulting relationship breakdown caused by the agent's reputational incentives. 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 two-sided transfers. This essentially allows the agent to 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.

6.2. *Costly Actions*

The careful reader would have noticed that Theorem 3 leveraged the fact that actions were assumed to be costless for the agent. This assumption is not required for our main insight: namely, reputational forces are destructive unless both players coordinate on extremely inefficient strategies despite uncertainty being resolved. The reason for this is also the same: if the cost of acting is small, the agent still has an incentive to generate successes via risky projects before experimentation stops and this causes unraveling. The argument however is more involved and we hence relegate it to the online appendix.

Suppose now that the agent has to pay a cost $\mathcal{C} > 0$ to implement a project irrespective of quality. Costly actions introduce two additional complications to the model. First, the mere fact that the agent was willing to bear the cost of implementing a project is an additional signaling device. Second, even when uncertainty is resolved, incentives are no longer aligned because there is an additional moral hazard problem and the agent needs to be incentivized to act. Nonetheless, we argue that our main theoretical insight (Theorem 3) is not knife-edged and continues to hold if the agent bears a small cost of acting.

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.

We now define the stronger version of the refinement. For any $\underline{u} \in \left(0, \frac{1}{1-\beta}\right)$, 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 $\underline{u}$.

Formally, NE strategies $\tilde{x}$, $\tilde{a}_\theta$ (of the game with costly actions) are a $\underline{u}$-equilibrium if, at all on-path histories $h^{t-1}h_t \in \mathscr{H}$, $(h^{t-1}h_t, i^t) \in \mathscr{H}^g$ such that $\tilde{p}(h^{t-1}) < 1$ and $\tilde{p}(h^{t-1}h_t) = 1$, the good type's continuation payoff satisfies $U_{\theta_g}(h^{t-1}h_t, i^t, \tilde{x}, \tilde{a}_{\theta_g}) \geq \underline{u}$.

Note that this refinement is still significantly weaker than the renegotiation-proofness refinement employed by EV which, as mentioned earlier, requires efficient play after uncertainty is resolved. When $\underline{u}$ is small, it is strictly less than the good type's payoff in any NE (when $p_0 = 1$) which yields payoffs that lie on the Pareto frontier. This is because the principal will always experiment in the first period of a Pareto-efficient nontrivial NE and so the agent's payoff is bounded below by one.[18]

Finally, with costly actions, we say that quality control is necessary when

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

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.

The following theorem shows that our main result is unaffected when costs are small.

THEOREM 5. *For any $\underline{u} \in \left(0, \frac{1}{1-\beta}\right)$, 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 (for all $p_0 \in (0,1)$ such that $\underline{\Pi}^{\mathcal{C}}(p_0) \leq 0$).*

6.3. *Costly Information Acquisition*

Our main insights will continue to hold if, 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). Similar to the previous section, we will need to strengthen the refinement but we can nonetheless conclude 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.

6.4. *Full Commitment*

Since we are working within the generality of Nash equilibria, it is natural to ask whether commitment power helps the principal in this setting. For instance, in dynamic pricing problems, the Coase conjecture disappears in the space of Nash equilibria and the principal can

[18]There is no benefit to either party in delaying experimentation so the principal must experiment with positive probability in the first period. If the principal mixes in the first period, it is easy to construct a Pareto improving equilibrium where the principal experiments with probability one.

achieve static monopoly profits. There are numerous other dynamic mechanism design problems (for instance, Guo (2016) and Deb, Pai, and Said (2018)) where commitment is not required to maximize profits. More generally, Ben-Porath, Dekel, and Lipman (2018) have shown that, for a large class of static mechanism design problems with "evidence," the commitment optimal can be achieved even without commitment (they have ongoing work which shows that similar insights also apply to certain dynamic environments).

We now allow the principal to announce her strategy $\tilde{x}$ in advance and commit to it. In response, the agent's strategy $\tilde{a}_\theta$ is the best response to $\tilde{x}$ that maximizes the principal's payoff (an assumption that is always made in mechanism design).

THEOREM 6. *There are parameter values for which full commitment to $\tilde{x}$ allows the principal to attain a strictly higher payoff than the principal-optimal NE.*

In Appendix C, we state the result more formally by providing sufficient conditions under which commitment makes the principal strictly better off. A NE strategy does not rule out that the principal terminates the relationship even if her belief is 1, however, it is not sequentially rational for her to experiment on path if her belief is sufficiently low. Full commitment additionally allows the principal to commit to continue experimenting even if her expected continuation value is negative. This allows her to provide the agent continuation value even if her belief is low. Theorem 6 shows that this is a more profitable way of screening the agent as opposed to foregoing surplus when the agent is known to be the good type. In a nutshell, the principal can avoid excessive failures from risky projects by committing to continuation value after inaction by the agent.

We end this section by noting that commitment to a strategy $\tilde{x}$ is *not* the same as commitment to a direct revelation mechanism. This is because our model is a setting with both adverse selection and strategic actions. A direct mechanism should solicit the agent's type at the beginning of the game and should provide (possibly mixed) action recommendations after the agent reports his period-t project idea. These are both features that are absent from the principal's strategy $\tilde{x}$ in the repeated game. A full characterization of the optimal mechanism is hard and we have only been able to derive it in special cases. This difficulty is to be expected. In order to solve for the optimal mechanism, we have to first derive the most profitable way to implement a given incentive compatible action strategy for the agent and then optimize over all such possible action strategies. Even the first of these two steps is hard as mechanism design problems without transfers can be complex even in static pure adverse selection environments without strategic actions.[19]

7. Concluding Remarks

We analyzed a novel repeated game where a principal interacts with an agent who, in addition to having a private type, has periodic private information and can act strategically. Our main insight is that, despite both players being long-lived and impatient, reputational forces can be destructive. We identify the role of maximally inefficient play on sustaining the relationship. We demonstrate that the main economic forces we model are present in several disparate applications and discuss the unique implications of our findings in each.

In our model, bad types can only generate failures and thus good types can perfectly out themselves by producing a success. As we have argued, this good news assumption makes our results more stark because one might expect that the ability of the good type to perfectly separate facilitates screening. That said, a natural alternative to this assumption is to allow bad types to also generate successes with a (small) positive probability by implementing bad projects. Note that in this case, our refinement would have no bite since there need not be on-path histories where the principal's belief jumps to one. That said, it is possible to analyze this case by using 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.

We end with two natural avenues for future research. There are applications where our assumption of imperfect public monitoring may be inappropriate. Instead, in such environments, the principal could privately observe the outcome which is an imperfect signal (à la Compte (1998)) of the project quality. Here, the principal can communicate with the agent by sending him cheap talk messages. An example of such a setting is a firm hiring a consultant to provide advice whose outcome is privately observed by the firm which then provides feedback.

[19] A related model in which the optimal contract can be solved for is Varas (2018). A key distinction is that the structure of that model allows the author to a priori identify the optimal action strategy for the agent. The mechanism design problem can therefore be reduced to finding the optimal incentive scheme that implements that action strategy.

Another direction to generalize our analysis is to incorporate multiple agents, each of whom is trying to entice the principal to experiment with him (Board (2011) is an instance of such a relational contracting problem). This is particularly relevant in online markets where multiple content providers often compete for a consumer's limited attention (Fainmesser and Galeotti (2019) study a static model of competition by influencers with known quality). These and numerous other variants of our model are not mere extensions as they involve different technical challenges and speak to distinct real-world problems. We hope to address at least some of these questions in future work.

APPENDIX A. PROOFS OF THEOREM 3 AND THEOREM 4

We will 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 6.1.

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 $h_1 \dots h_{t-1} \in \mathscr{H}$ and $\tau_{t'} \geq (=) 0$ when $h_{t'} \neq (=) h\!\!\!/$ 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 $\mathscr{T}$.

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 $\mathscr{T}^g$ to denote the set of the good type agent's private histories.

We use $\tilde{\mathscr{T}}^t$ where $t \geq 1$ to denote the set of on-path histories at the beginning of period $t+1$. Here, 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 $(h^{t-1}, \tau^{t-1}) \in \mathscr{T}$. Similarly, $\tilde{a}_{\theta_g}(h^{t-1}, \tau^{t-1}, i^t) \in [0,1]$ for each private history $(h^{t-1}, \tau^{t-1}, i^t) \in \mathscr{T}^g$.

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 $(h^{t-1}, \tau^{t-1}) \in \mathscr{T}$, $h_t \in \{\overline{h}, \underline{h}, h_\varphi\}$. Note that, by definition, $\tilde{\tau}(h^{t-1}h_t, \tau^{t-1})$ is the Dirac measure at 0 when $h_t = h\!\!\!/$.

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. Once again, 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. Note that, for brevity, we are suppressing dependence on strategies.

We now define "last histories." The necessity of such histories arising on-path will be crucial for our arguments.

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 $(h^t, \tau^t) \in \tilde{\mathscr{T}}^t$ 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,

$$\underline{p} = \inf \left\{ \tilde{p}\left(\hat{h}^{t'}, \hat{\tau}^{t'}\right) \;\middle|\; (\hat{h}^{t'}, \hat{\tau}^{t'}) \in \tilde{\mathscr{T}}^{t'},\ t' \geq 0 \right\} \;<\; 1,$$

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 $(h^{t'}, \tau^{t'}) \in \tilde{\mathscr{T}}^{t'}$ 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 $(\hat{h}^t, \hat{\tau}^t) \in \tilde{\mathscr{T}}^t$ 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 $(\hat{h}^t\hat{h}^s, \hat{\tau}^t\hat{\tau}^s) \in \mathscr{T}$, $\hat{h}^s = (\not{h}, \dots, \not{h})$, $\hat{\tau}^s = (0, \dots, 0)$, $s \geq 1$ 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 $(\hat{h}^s\hat{h}_{s+1}\check{h}^{s'}, \hat{\tau}^s\check{\tau}^{s'+1}) \in \tilde{\mathscr{T}}^{s+s'+1}$ such that $\tilde{x}(\hat{h}^s\hat{h}_{s+1}\check{h}^{s'}, \hat{\tau}^s\check{\tau}^{s'+1}) > 0$ and $\tilde{p}(\hat{h}^s\hat{h}_{s+1}\check{h}^{s'}, \hat{\tau}^s\check{\tau}^{s'+1}) = \tilde{p}(\hat{h}^s\hat{h}_{s+1}, \hat{\tau}^s) = \tilde{p}(\hat{h}^t, \hat{\tau}^t)$ and so $(h^{t'}, \tau^{t'}) = (\hat{h}^s\hat{h}_{s+1}\check{h}^{s'}, \hat{\tau}^s\check{\tau}^{s'+1})$ 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 $U_{\theta_b}(\hat{h}^s\check{h}_{s+1}, \hat{\tau}^s) > 0$ for $\check{h}_{s+1} \in \{\underline{h}, \hat{h}_\varphi\}$, $\check{h}_{s+1} \neq \hat{h}_{s+1}$. 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 $(\hat{h}^s\check{h}_{s+1}, \hat{\tau}^s)$ 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 $s \geq 1$,

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

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

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

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.

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

(1) 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.

(2) 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.

Proof. We prove each part in turn.

Part (1). 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

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

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 (2). 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 $\mathcal{T} \subset \mathscr{T}$ starting at history $(h^t, \tau^t) \in \mathscr{T}$ 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 $\tilde{\mathscr{T}}^t$ into three mutually disjoint sets:

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

(2) The set $\tilde{\mathscr{S}}^t \subset \tilde{\mathscr{T}}^t$ of all histories that follow from a success being generated in a non-last history: this consists of all histories $(h^t, \tau^t) \in \tilde{\mathscr{T}}^t$ 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 $(h^{s-1}, \tau^{s-1}) \notin \tilde{\mathscr{L}}^{s-1}$ is not a last history.

(3) All other remaining histories $\tilde{\mathscr{R}}^t = \tilde{\mathscr{T}}^t \setminus \left(\tilde{\mathscr{S}}^t \cup \tilde{\mathscr{L}}^t\right)$. Note that for all $(h^t, \tau^t) \in \tilde{\mathscr{R}}^t$, 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 $\tilde{\mathscr{R}}^t$ is less than p_0. Formally, for all t, $\tilde{\mathscr{R}}^t$ is nonempty and there is a history $(h^t, \tau^t) \in \tilde{\mathscr{R}}^t$ 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 or

$$p_0 = \mathbb{E}[\tilde{p}(h^t, \tau^t)].$$

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

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

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

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

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

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

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

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

First suppose $\inf_{(\hat{h}^t,\hat{\tau}^t)\in\tilde{\mathscr{R}}^t}\tilde{p}(\hat{h}^t,\hat{\tau}^t)=p_0$. Then, for all histories $(\hat{h}^s,\hat{\tau}^s)\in\mathscr{T}^s$, $0\le s\le t$, we must have $\tilde{p}(\hat{h}^s,\hat{\tau}^s)=p_0$ which implies $(\hat{h}^s,\hat{\tau}^s)\in\tilde{\mathscr{R}}^s$. Suppose this were not true. Take an earliest history $(\hat{h}^s\hat{h}_{s+1},\hat{\tau}^s\hat{\tau}_{s+1})\in\mathscr{T}^{s+1}$ (with the smallest $s<t$) such that $\tilde{p}(\hat{h}^s\hat{h}_{s+1},\hat{\tau}^s\hat{\tau}_{s+1})\neq p_0$. Then, by definition, we must have $\tilde{p}(\hat{h}^s,\hat{\tau}^s)=p_0$ and $(\hat{h}^s,\hat{\tau}^s)\in\tilde{\mathscr{R}}^s$ since $(\hat{h}^s,\hat{\tau}^s)\notin\tilde{\mathscr{L}}^s$ because the beliefs at all last histories are strictly greater than p_0. Then, by Bayes' consistency, there must be a continuation history $(\hat{h}^s\check{h}_{s+1},\hat{\tau}^s\check{\tau}_{s+1})\in\tilde{\mathscr{R}}^{s+1}$ such that $\tilde{p}(\hat{h}^s\check{h}_{s+1},\hat{\tau}^s\check{\tau}_{s+1})<p_0$,[20] which would be a contradiction.

Now note that, if the principal does not experiment at any continuation history following $(\hat{h}^t,\hat{\tau}^t)\in\tilde{\mathscr{R}}^t=\tilde{\mathscr{T}}^t$, 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.

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

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

$$\bar{s} = \max\left\{ s \le t \;\middle|\; \begin{array}{l} (\check{h}^{s-1}\check{h}_s\check{h}^{t-s},\check{\tau}^t) \in \tilde{\mathscr{R}}^t,\ \check{h}_s \in \{\underline{h}, h_\varphi\}, \tilde{p}(\check{h}^{s-1}\check{h}_s\check{h}^{t-s},\check{\tau}^t) \le p_0 \text{ and} \\ \tilde{p}(\check{h}^{s-1}\check{h}_s\check{h}^{t-s},\check{\tau}^t) \le \inf_{(\hat{h}^t,\hat{\tau}^t)\in\tilde{\mathscr{R}}^t}\tilde{p}(\hat{h}^t,\hat{\tau}^t) + \varepsilon \end{array} \right\},$$

to be the last period amongst the low belief histories where the principal observes a non-success 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, $\check{h}^{t-\bar{s}} = (\not{h},\ldots,\not{h})$ for all $(\check{h}^{\bar{s}}\check{h}^{t-\bar{s}},\check{\tau}^t) \in \mathcal{R}^t_{\bar{s}}$. Observe that the maximum is well defined because $\tilde{\mathscr{R}}^t$ 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 $(\check{h}^t,\check{\tau}^t) \in \mathcal{R}^t_{\bar{s}}$, we must have $\tilde{x}(\check{h}^{\bar{s}}\hat{h}^{s-\bar{s}},\check{\tau}^{\bar{s}-1}\hat{\tau}^{s-\bar{s}+1}) = 0$ for all $\bar{s} \le s < t$, $(\check{h}^{\bar{s}}\hat{h}^{s-\bar{s}},\check{\tau}^{\bar{s}-1}\hat{\tau}^{s-\bar{s}+1}) \in \tilde{\mathscr{T}}^s$. In words, this says that the principal does not experiment between periods $\bar{s}+1$ and t at *any* on-path continuation history following $(\check{h}^{\bar{s}},\check{\tau}^{\bar{s}-1})$. 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, $(\check{h}^{\bar{s}}\hat{h}^{t-\bar{s}},\check{\tau}^{\bar{s}-1}\hat{\tau}^{t-\bar{s}+1}) \in \tilde{\mathscr{T}}^t$ implies $(\check{h}^{\bar{s}}\hat{h}^{t-\bar{s}},\check{\tau}^{\bar{s}-1}\hat{\tau}^{t-\bar{s}+1}) \in \mathcal{R}^t_{\bar{s}}$.

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

$$\tilde{p}(\check{h}^{\bar{s}}\hat{h}^{s-\bar{s}+1},\check{\tau}^{\bar{s}-1}\hat{\tau}^{s-\bar{s}+2}) \le \tilde{p}(\check{h}^{\bar{s}}\hat{h}^{s-\bar{s}},\check{\tau}^{\bar{s}-1}\hat{\tau}^{s-\bar{s}+1}) = \tilde{p}(\check{h}^{\bar{s}},\check{\tau}^{\bar{s}-1}\hat{\tau}_s) = \tilde{p}(\check{h}^{\bar{s}},\check{\tau}^{\bar{s}}) \le p_0.$$

The first inequality follows from Bayes' consistency and the equalities follow from the facts that there is no experimentation from periods $\bar{s}+1$ to s and the transfer in period $\bar{s}$ does not change the belief at period $\bar{s}+1$. This combined with the fact that $(\check{h}^{\bar{s}-1},\check{\tau}^{\bar{s}-1}) \in \tilde{\mathscr{R}}^{\bar{s}-1}$, implies that $(\check{h}^{\bar{s}}\hat{h}^{s-\bar{s}+1},\check{\tau}^{\bar{s}-1}\hat{\tau}^{s-\bar{s}+2}) \in \tilde{\mathscr{R}}^{s+1}$ 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 $(\check{h}^{\bar{s}}\hat{h}^{s-\bar{s}+1},\check{\tau}^{\bar{s}-1}\hat{\tau}^{s-\bar{s}+2})$ with associated belief $\tilde{p}(\check{h}^{\bar{s}}\hat{h}^{s-\bar{s}+1},\check{\tau}^{\bar{s}-1}\hat{\tau}^{s-\bar{s}+2})$ (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 $(\check{h}^{\bar{s}}\hat{h}^{t-\bar{s}},\check{\tau}^{\bar{s}-1}\hat{\tau}^{t-\bar{s}+1}) \in \tilde{\mathscr{R}}^t$ satisfying

$$\tilde{p}(\check{h}^{\bar{s}}\hat{h}^{t-\bar{s}},\check{\tau}^{\bar{s}-1}\hat{\tau}^{t-\bar{s}+1}) \le \tilde{p}(\check{h}^{\bar{s}}\hat{h}^{s-\bar{s}+1},\check{\tau}^{\bar{s}-1}\hat{\tau}^{s-\bar{s}+2})$$

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

So finally suppose, to the converse, that the principal stopped experimenting after all histories $(\check{h}^t, \check{\tau}^t) \in \mathcal{R}^t_{\bar{s}}$. An implication is that

$$U_{\theta_b}(\check{h}^{\bar{s}}, \check{\tau}^{\bar{s}-1}) = 0.$$

The equality follows from the above argument which shows that the principal does not experiment between periods $\bar{s}+1$ and t after history $(\check{h}^{\bar{s}}, \check{\tau}^{\bar{s}-1})$ and that the transfer in period $\bar{s}$ 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 $\bar{s}$ non-success outcome $(\check{h}^{\bar{s}-1}\hat{h}_{\bar{s}}, \check{\tau}^{\bar{s}})$, $\hat{h}_{\bar{s}} \in \{\underline{h}, h_\varphi\}$, $\hat{h}_{\bar{s}} \neq \check{h}_{\bar{s}}$ must either be off-path or we must have $U_{\theta_b}(\check{h}^{\bar{s}-1}\hat{h}_{\bar{s}}, \check{\tau}^{\bar{s}-1}) = 0$. 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 $(\check{h}^{\bar{s}-1}\hat{h}_{\bar{s}}, \check{\tau}^{\bar{s}-1})$ which is not possible since this would imply $\tilde{p}(\check{h}^{\bar{s}}, \check{\tau}^{\bar{s}}) = 1$.

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

Step 4: Final contradiction. We define

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

to be the lowest belief that can arise at a history that does not follow a last history. For $\varepsilon > 0$, we pick an $(h^t, \tau^t) \in \tilde{\mathscr{R}}^t$ such that

$$\tilde{p}(h^t, \tau^t) \leq \underline{p}^{\mathscr{R}} + \varepsilon \text{ and } \tilde{x}(h^t, \tau^t) > 0,$$

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 $s \geq 1$:

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

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 non-last 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

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

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

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

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

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

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

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 a linear function.

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 non-last history in these s periods, (ii) the expected continuation value at last histories, (iii) the cost of experimentation, transfers and failures at histories $(h^t h^{s'},\tau^t\tau^{s'}) \in \tilde{\mathscr{R}}^{t+s'}$, $0 \leq s' \leq s-1$ and (iv) the expected continuation value at all remaining period-$t+s+1$ histories $(h^t h^s,\tau^t\tau^s) \in \tilde{\mathscr{R}}^{t+s}$.

We can now derive a simple upper bound

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

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 non-last histories that arrive between periods $t+1$ and $t+s$ arrive immediately.

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

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

where the expectations are taken with respect to $\tilde{\pi}(\cdot | h^t h^{s'}, \tau^t \tau^{s'})$. Summing the last term 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

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

where the inequality follows from (4) and the fact that $q^{\mathscr{L}} p^{\mathscr{L}} \leq \tilde{p}(h^t, \tau^t) \leq p_0$, $\underline{\Pi}(p_0) \leq 0$. 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 $(\hat{h}^{t'}, \hat{\tau}^{t'}) \in \tilde{\mathscr{R}}^{t'}$ 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.

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

1. *(1) 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 $h^t \in \mathscr{H}$ and $(h^{t'}, i^{t'} i_r) \in \mathscr{H}^g$ such that $\tilde{a}_{\theta_b}(h^t) > 0$ and $\tilde{a}_{\theta_g}(h^{t'}, i^{t'} i_r) = 1$.
2. *(2) If quality control is necessary ($\underline{\Pi}(p_0) \leq 0$), the unique equilibrium outcome is that the principal never experiments.*

PROOF. 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:

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

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

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

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

where $\tilde{\mathscr{H}}^t$ 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 $\tilde{\mathscr{T}}^t$ 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 are 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. ■

Online Appendix

Appendix B. Proof of Theorem 5

While the basic intuition of this result is similar to our main result (Theorem 3), the proof is subtly different. Before we proceed to the formal proof, we orient the reader with a high-level overview of the proof strategy.

To contrast, recall that the proof-idea in the absence of costs was to show that any nontrivial equilibrium must have a "last history" at which the principal experiments but then never experiments again unless a success is generated. Further, there must be a last history at which the principal's belief was than the initial prior. Since acting was without cost, the good type strictly preferred to run both good and risky projects, while the bad type was indifferent between acting and not. The former fact ensured that the value from experimentation is negative (because $\underline{\Pi} \leq p_0$), while the latter ensures that our refinement has no bite in case a failure is generated (as the bad type randomized). However, if acting is costly, this proof strategy does not immediately work because the bad type would no longer be indifferent between acting and not. Therefore we need to take a slightly different approach.

First we prove a couple of "bookkeeping" lemmata. Lemma 2 shows that, when the cost of acting is less than the agent's discount factor, there is always a NE when $p_0 = 1$ in which the good type's payoff is at least 1. Therefore the notion of $\underline{u}$-equilibrium does not become vacuous when costs are small because the highest NE payoff when uncertainty is resolved does not shrink to zero. Then, Lemma 3 derives upper-bounds for the payoffs of the good type and principal as a function of the bad type's payoff. The argument accommodates the fact that the principal and agent discount at different rates.

The proof of the theorem then follows from the following strategy. Assume for the sake of contradiction that there is a nontrivial $\underline{u}$-equilibrium for some small cost $\mathcal{C} > 0$ of acting. First, we show that the $\underline{u}$-equilibrium cannot have a last history. Pick a history at which the principal experiments and her belief is close to the lowest among histories at which she experiments. We argue that among these histories, there must be one in which there is an outcome (inaction or failure) such that the principal's belief drops below this infimum and so she subsequently stops experimenting. Since there cannot be a last history, the principal must experiment again after other non-success outcome which we argue is a failure. Next, roughly, we show that at such a history, the good type must strictly prefer to implement both good and risky projects. Further, the bad type must be indifferent between acting and not, and be randomizing: otherwise, the principal's belief upon observing a failure would either be 1 (impossible in a $\underline{u}$-equilibrium) or 0 (contradicting nontrivial continuation play after failure). This implies that the continuation value of the bad type subsequent to acting must be exactly $\mathcal{C}/\beta$. We then argue that the good type has a strict incentive to implement bad projects. Finally, the small continuation payoff of the bad type implies that the principal's payoff subsequent to failure must also be small (Lemma 3) and, hence, her net payoff at the history must be negative, concluding the contradiction.

We now begin by arguing that, for sufficiently low cost of actions, there is always a NE when $p_0 = 1$ such that the good type's payoff is 1.

LEMMA 2. *Suppose $p_0 = 1$. For any $\mathcal{C} \in (0, \beta)$, there exist NE strategies $\tilde{x}$, $\tilde{a}_\theta$ such that $U_{\theta_g}(h^0, i^0, \tilde{x}, \tilde{a}_{\theta_g}) \geq$ 1.*

PROOF. We prove this lemma constructively by describing NE strategies that yield the good type a payoff greater than 1. Consider the following principal's strategy:

$$\tilde{x}(h^0) = 1,$$
$$\tilde{x}(h^t) = \begin{cases} \frac{\mathcal{C}}{\beta(1+\lambda_g\mathcal{C})} & \text{if } h^t = (\overline{h}, \dots, \overline{h}), \\ 0 & \text{otherwise.} \end{cases}$$

In words, the principal experiments in the first period and continues to experiment with positive probability as long as successes have been generated. As soon as the good type does not generate a success or the principal does not experiment, she stops experimenting in the future.

In response, the good type plays the following strategy:

$$\tilde{a}_{\theta_g}(h^0, i_1) = \begin{cases} 1 & \text{if } i_1 = i_g, \\ 0 & \text{if } i_1 \in \{i_r, i_b\}, \end{cases}$$
$$\tilde{a}_{\theta_g}(h^t, i^{t+1}) = \begin{cases} \frac{c}{\lambda_g} & \text{if } h^t = (\overline{h}, \dots, \overline{h}) \text{ and } i^{t+1} = (i_g, \dots, i_g), \\ 0 & \text{otherwise.} \end{cases}$$

In words, the good type only implements good projects in the first period and, at all subsequent histories where he has generated successes in the past, he continues to only implement good projects. At all other histories, he does not act. Since $p_0 = 1$, we can pick the bad type's strategy arbitrarily because every strategy is a best response.

We now argue that these strategies are mutual best responses. First observe that the good type is indifferent between running good projects and not acting at every history where the principal experiments. This is because, at any history where the principal experiments, the continuation payoff $\widehat{U}_{\theta_g}$ for the good type after generating a success satisfies

$$\widehat{U}_{\theta_g} = \frac{\mathcal{C}}{\beta(1+\lambda_g \mathcal{C})}\left(1 + \beta\lambda_g \widehat{U}_{\theta_g}\right) \quad \implies \widehat{U}_{\theta_g} = \frac{\mathcal{C}}{\beta}.$$

Therefore, he is indifferent between implementing good projects and not acting. Additionally, he strictly prefers not to run risky and bad projects since they are costly and the former generates a success with probability strictly less than 1.

Finally observe that the principal's payoff is $V(h^0, \tilde{x}, \tilde{a}_\theta) = \lambda_h - c$ since the good type implements only good projects in period 1 and does so with probability 1. The stage game payoff to the principal at any subsequent history where she experiments is $\lambda_g \frac{c}{\lambda_g} - c = 0$ and therefore the principal is indifferent between experimenting or not at all histories where successes have been generated in every prior period. ■

The purpose of Lemma 2 is to show that as the cost $\mathcal{C}$ becomes small, the highest NE payoff for the good type once uncertainty is resolved does not shrink to 0. In other words, for low costs, the $\underline{u}$-equilibrium refinement does not become vacuous. We do not explicitly invoke this lemma below but its implication is implicit whenever we take the cost of action $\mathcal{C}$ to be small.

The next lemma establishes a relationship between the equilibrium payoff of the bad type to the good type and the principal. Specifically, it derives upper bounds for the payoffs of the latter two as a function of the payoff of the bad type.

LEMMA 3. *Let $\tilde{x}$, $\tilde{a}_\theta$ be a NE and consider an on-path history $h^t \in \tilde{\mathscr{H}}^t$, $\tilde{x}(h^t) \in (0,1)$ where $U_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta) > 0$. Then,*

$$U_{\theta_g}(h^t, i^t, \tilde{x}, \tilde{a}_\theta) \le \left(U_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta) + \frac{\mathcal{C}}{1-\beta}\right)\frac{1}{1-\beta}$$

$$V(h^t, \tilde{x}, \tilde{a}_\theta) \le \max\left\{U_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta) + \frac{\mathcal{C}}{1-\beta}, \frac{\delta^{\lfloor \hat{t} \rfloor}}{1-\delta}\right\}(1+\delta\overline{\Pi})$$

$$where \quad \hat{t} = \frac{\log\left[(1-\beta)U_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta) + \frac{\mathcal{C}}{1-\beta}\right]}{\log \beta},$$

are upper bounds for the payoffs of the good type and the principal respectively.

PROOF. We introduce the following two pieces of new notation just for the proof of this lemma. $C_{\theta_b}(h^t, \tilde{x}, \ \tilde{a}_\theta)$ denotes the expected discounted sum of the cost of acting that the bad type pays moving forward from history h^t. $X_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta) = U_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta) + C_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta)$ is the expected discounted total duration of experimentation by the principal when the agent is the bad type.

Observe that $C_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta) \leq \frac{\mathcal{C}}{1-\beta}$ because the right side is the highest possible value computed by adding the cost of acting at every period. Therefore

$$X_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta) \leq U_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta) + \frac{\mathcal{C}}{1-\beta}.$$

We derive the upper bound for the principal's and the good type's payoff by assuming that, whenever the principal experiments for the first time, the highest possible payoff arises. Further, we derive an upper bound for

$$\tilde{x}(h^t) + \delta(1-\tilde{x}(h^t))\tilde{x}(h^t\not{h}) + \delta^2(1-\tilde{x}(h^t\not{h}))(1-\tilde{x}(h^t))\tilde{x}(h^t\not{h}\not{h}) + \cdots, \tag{6}$$

the discounted probability that the principal experiments for the first time.

We shorten notation to $y_0 = \tilde{x}(h^t)$, $y_1 = (1-\tilde{x}(h^t))\tilde{x}(h^t\not{h})$, $y_2 = (1-\tilde{x}(h^t\not{h}))(1-\tilde{x}(h^t))\tilde{x}(h^t\not{h}\not{h})$ and so on. The solution to the following problem

$$\overline{Y} = \max_{y_0, y_1, \ldots} \left\{ y_0 + \delta y_1 + \delta^2 y_2 + \cdots \right\},$$

subject to

$$y_0 + \beta y_1 + \beta^2 y_2 + \cdots \leq X_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta),$$

$$y_{t'} \in [0,1] \text{ for all } t' \geq 0.$$

is an upper bound for (6) because the second set of constraints relaxes the relation between the y variables.

Observe first that when $\delta \leq \beta$, it must be that $\overline{Y} \leq X_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta)$ where the right side would be the solution to the maximization problem ignoring the second set of constraints.

When $\delta > \beta$, the solution to the problem is to find the smallest t' such that $\frac{\beta^{t'}}{1-\beta} \leq X_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta)$ and set $y_{t'} = 1(0)$ for all $s > (<) t'$. Solving the equality, we get

$$\frac{\beta^{\hat{t}}}{1-\beta} = X_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta)$$
$$\implies \hat{t} = \frac{\log[(1-\beta)X_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta)]}{\log \beta},$$

which implies $\overline{Y} \leq \frac{\delta^{\lfloor \hat{t} \rfloor}}{1-\delta}$.

The upper bound for the good type immediately follows because he discounts the future at the same rate as the bad type and so $U_{\theta_g}(h^t, i^t) \leq X_{\theta_b}(h^t, \tilde{x}, \tilde{a}_\theta)\frac{1}{1-\beta}$. To see this, note that $\frac{1}{1-\beta}$ is the highest possible total utility for either type.

The upper bound for the principal's payoff follows by multiplying $(1 + \delta\overline{\Pi})$ to the above computed bounds for $\overline{Y}$. This payoff corresponds to observing a success with probability 1 followed by the highest possible continuation payoff whenever the principal first experiments. ■

Before we proceed to the proof of the theorem, we define one additional piece of notation. For a given NE $\tilde{x}$, $\tilde{a}_\theta$, let $\widehat{V}(p)$ be the supremum of payoffs the principal receives in this equilibrium at on-path histories h^t that satisfy $\tilde{p}(h^t) \leq p$. The dependence on the strategies is left implicit. Formally,

$$\widehat{V}(p) = \sup\{V(h^t, \tilde{x}, \tilde{a}_\theta) : h^t \in \tilde{\mathscr{H}}^t,\ \tilde{p}(h^t) \leq p\},$$

where recall that $\tilde{\mathscr{H}}^t$ is the set of on path histories in period $t+1$.

We are now ready to prove the theorem.

THEOREM 5. *For any $\underline{u} \in \left(0, \frac{1}{1-\beta}\right)$, 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 (for all $p_0 \in (0,1)$ such that $\underline{\Pi}^{\mathcal{C}}(p_0) \leq 0$).*

PROOF. Given a nontrivial $\underline{u}$-equilibrium $\tilde{x}$, $\tilde{a}_\theta$, we define

$$\underline{p}^{\mathcal{C}} = \inf\left\{\tilde{p}\left(\hat{h}^{t'}\right) \;\middle|\; \hat{h}^{t'} \in \tilde{\mathscr{H}}^{t'},\ t' \geq 0, \tilde{x}\left(\hat{h}^{t'}\right) > 0\right\}$$

to denote the infimum of the beliefs at on-path histories at which the principal experiments. The notation suppresses dependence on the strategies, this should cause no confusion below.

Observe that $\underline{p}^{\mathcal{C}} \in (0, p_0]$. Clearly $\underline{p}^{\mathcal{C}}$ must be bounded above by p_0 since the $\underline{u}$-equilibrium is nontrivial and beliefs do not change until the principal experiments. Additionally, $\underline{p}^{\mathcal{C}}$ cannot be 0 since there is a minimal positive belief below which the principal would never experiment even if she could choose the agent's strategy.

The proof now proceeds in four steps. The first three steps establish general properties of nontrivial $\underline{u}$-equilibria and the final step provides the requisite contradiction.

Step 1: Suppose $\mathcal{C} < \beta q_r \underline{u}$ and let $\tilde{x}$, $\tilde{a}_\theta$ be a nontrivial $\underline{u}$-equilibrium. Then, there is no last history.[21]

Suppose, for contradiction, that there is a last history h^t. Then the good type would strictly prefer to run risky projects $(\tilde{a}_{\theta_g}(h^t, i^t i_r) = 1)$ because the continuation payoff from inaction and failure is 0 and the expected payoff from the risky project is bounded below by $\beta q_r \underline{u} - \mathcal{C} > 0$ because the continuation payoff after a success must be at least $\underline{u}$.

Therefore failure is on path ($h^t \underline{h} \in \tilde{\mathscr{H}}^{t+1}$) which in turn implies that we must have $\tilde{p}(h^t \underline{h}) < 1$. The latter follows from the fact that, if $\tilde{p}(h^t \underline{h}) = 1$, the refinement implies that continuation play must be nontrivial which contradicts h^t being a last history. But $\tilde{p}(h^t \underline{h}) < 1$ implies that the bad type acts with positive probability $(\tilde{a}_{\theta_b}(h^t) > 0)$ which is impossible since it is not a best response to pay the cost $\mathcal{C}$ of implementing the bad project for 0 continuation payoff. This completes the proof of this step.

Step 2: Let $\tilde{x}$, $\tilde{a}_\theta$ be a nontrivial $\underline{u}$-equilibrium. For any $\varepsilon \in \left(0, 1 - \underline{p}^{\mathcal{C}}\right)$, there is an on-path history $h^t \in \tilde{\mathscr{H}}^t$ such that $\tilde{p}\left(h^t\right) - \underline{p}^{\mathcal{C}} < \varepsilon$, $\tilde{x}\left(h^t\right) > 0$ and there is an on path outcome $h_{t+1} \in \{h_\varphi, \underline{h}\}$ at which the belief $\tilde{p}\left(h^t h_{t+1}\right) < \underline{p}^{\mathcal{C}}$.

In words, there is an on-path history where the belief is arbitrarily close to $\underline{p}^{\mathcal{C}}$ such that there is an on-path continuation history where the belief falls below $\underline{p}^{\mathcal{C}}$ and so the principal stops experimenting.

Suppose as a contradiction that for some ε, there is no such history. Formally, there is $\varepsilon \in \left(0, 1 - \underline{p}^{\mathcal{C}}\right)$ such that for any $h^t \in \tilde{\mathscr{H}}^t$ for which, $\tilde{p}(h^t) - \underline{p}^{\mathcal{C}} < \varepsilon$ and $\tilde{x}(h^t) > 0$, the continuation histories satisfy $\tilde{p}(h^t h_{t+1}) \geq \underline{p}^{\mathcal{C}}$ for all $h^t h_{t+1} \in \tilde{\mathscr{H}}^{t+1}$.

[21]Recall that a last history $h^t \in \tilde{\mathscr{H}}^t$ is such that $\tilde{x}(h^t) > 0$, $\tilde{p}(h^t) < 1$ and $U_{\theta_b}(h^t h_{t+1}, \tilde{x}, \tilde{a}_\theta) = 0$ for $h_{t+1} \in \{\underline{h}, h_\varphi\}$.

Now consider a history $h^t \in \tilde{\mathscr{H}}^t$ such that such that $\tilde{p}\left(h^t\right) - \underline{p}^{\mathcal{C}} < \frac{\varepsilon}{n^s}$, $\tilde{x}\left(h^t\right) > 0$ where $n, s > 1$ are both positive integers. Let $\hat{q}$ be the probability that, upon experimenting, the principal expects to reach an on-path continuation history $h^t h_{t+1} \in \tilde{\mathscr{H}}^{t+1}$ at which the belief $\tilde{p}\left(h^t h_{t+1}\right) \geq \underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^{s-1}}$. Since the beliefs follow a martingale, they must average to $\tilde{p}(h^t)$. This immediately yields an upper bound for $\hat{q}$ since

$$\hat{q}\left(\underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^{s-1}}\right) + (1-\hat{q})\underline{p}^{\mathcal{C}} \;\leq\; \tilde{p}(h^t) \;\implies\; \hat{q} \;\leq\; \frac{\tilde{p}(h^t) - \underline{p}^{\mathcal{C}}}{\frac{\varepsilon}{n^{s-1}}} \;\leq\; \frac{1}{n}.$$

The first inequality follows from the fact that the highest value of $\hat{q}$ can be derived by taking the lowest possible values for the beliefs that multiply $\hat{q}$ and $1-\hat{q}$ on the left side of the first inequality.

By definition, an upper bound for the principal's payoff $V(h^t, \tilde{x}, \tilde{a}_\theta)$ is $\widehat{V}\left(\underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^s}\right)$. Now take an arbitrary $\rho > 0$. Observe that if $\widehat{V}\left(\underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^s}\right) > V(h^t, \tilde{x}, \tilde{a}_\theta)$, then there must be must be an on-path history $h^{t'} \in \tilde{\mathscr{H}}^{t'}$ such that $V(h^{t'}, \tilde{x}, \tilde{a}_\theta) \geq \widehat{V}\left(\underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^s}\right) - \rho$, $V(h^{t'}, \tilde{x}, \tilde{a}_\theta) \geq V(h^t, \tilde{x}, \tilde{a}_\theta)$, $\tilde{x}(h^{t'}) > 0$ and $\tilde{p}(h^{t'}) \leq \underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^s}$. Conversely, if $\widehat{V}\left(\underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^s}\right) = V(h^t, \tilde{x}, \tilde{a}_\theta)$, then $h^{t'} = h^t$ is such a history.

This in turn implies

$$\begin{aligned} V(h^t, \tilde{x}, \tilde{a}_\theta) \leq \widehat{V}\left(\underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^s}\right) &\leq V(h^{t'}, \tilde{x}, \tilde{a}_\theta) + \rho \\ &\leq \frac{1}{n}(1 + \delta\widehat{V}(1)) - c + \delta\widehat{V}\left(\underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^{s-1}}\right) + \rho. \end{aligned}$$

The third inequality can be understood as follows. Since $\tilde{x}(h^{t'}) > 0$, the principal's payoff can be calculated by assuming she experiments at $h^{t'}$ for sure. This accounts for the cost c of experimentation. $\hat{q}$ is the highest likelihood that the belief in an on-path continuation history is above $\underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^{s-1}}$. The first term is then an upper bound for the principal's payoff in this case because it assumes that $\hat{q}$ takes its highest value, the agent generates a success immediately and the principal gets the highest possible continuation value. Finally, the last term is an upper bound for the principal's continuation payoff from on-path continuation histories at which the belief is below $\underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^{s-1}}$. This is because this term assumes that the principal reaches these histories with probability 1 and $\widehat{V}$ is increasing

We can now recursively repeat this argument to get

$$\widehat{V}\left(\underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^s}\right) \leq \frac{1}{n}(1 + \delta\widehat{V}(1)) - c + \delta\widehat{V}\left(\underline{p}^{\mathcal{C}} + \frac{\varepsilon}{n^{s-1}}\right) + \rho$$

$$\leq \left(\frac{1}{n}+\frac{\delta}{n}\right)(1+\delta\widehat{V}(1))-(1+\delta)\,c+\delta^2\widehat{V}\left(\underline{p}^{\mathcal{C}}+\frac{\varepsilon}{n^{s-2}}\right)+(1+\delta)\,\rho$$
$$\leq \frac{2}{n}(1+\delta\widehat{V}(1))-c+\delta^2\widehat{V}\left(\underline{p}^{\mathcal{C}}+\frac{\varepsilon}{n^{s-2}}\right)+2\rho$$
$$\leq \frac{s}{n}(1+\delta\widehat{V}(1))-c+\delta^s\widehat{V}\left(\underline{p}^{\mathcal{C}}+\varepsilon\right)+s\rho.$$

The proof of step then follows by contradiction as taking s, n large, ρ small such that $\frac{s}{n}$ and $s\rho$ are small implies $V(h^t,\tilde{x},\tilde{a}_\theta)\leq\widehat{V}\left(\underline{p}^{\mathcal{C}}+\frac{\varepsilon}{n^s}\right)<0$.

Step 3: There exists a $\hat{\mathcal{C}}>0$ such that, for every $\mathcal{C}\leq\hat{\mathcal{C}}$ and every nontrivial $\underline{u}$-equilibrium $\tilde{x}$, $\tilde{a}_\theta$ the following is true: at any on-path history $h^t\in\tilde{\mathscr{H}}^t$ such that $\tilde{x}(h^t)>0$ and the bad type's payoff is 0 prior to acting (so $U_{\theta_b}(h^th_\varphi,\tilde{x},\tilde{a}_\theta)=0$ and $U_{\theta_b}(h^t\underline{h},\tilde{x},\tilde{a}_\theta)\leq\mathcal{C}/\beta$), implementing good and risky projects ($\tilde{a}_{\theta_g}(h^t,i^ti_g)=\tilde{a}_{\theta_g}(h^t,i^ti_r)=1$) is a strict best response for the good type.

Fix such a history h^t. Because the principal experiments, there must be nontrivial continuation play after inaction or failure because, by Step 1, there are no last histories.

Since $U_{\theta_b}(h^th_\varphi,\tilde{x},\tilde{a}_\theta)=0$, it must be the case that the principal stops experimenting after inaction. Therefore, there must be nontrivial continuation play after failure.

Next observe that for $\mathcal{C}$ small enough, it must be the case that $h^t\overline{h}\in\tilde{\mathscr{H}}^{t+1}$ or that successes are on path. Suppose not. Then, an upper bound for the principal's at h^t is

$$V(h^t,\tilde{x},\tilde{a}_\theta)\leq -c+\delta V(h^t\underline{h},\tilde{x},\tilde{a}_\theta).$$

The first term is just the cost of experimentation at h^t and the second term is an upper bound for the principal's continuation utility since $V(h^th_\varphi,\tilde{x},\tilde{a}_\theta)=0$. Since the bad type's continuation utility satisfies $U_{\theta_b}(h^t\underline{h},\tilde{x},\tilde{a}_\theta)\leq\frac{\mathcal{C}}{\beta}$, by Lemma 3, when we take $\mathcal{C}$ small enough, the principal's payoff becomes $V(h^t\underline{h},\tilde{x},\tilde{a}_\theta)<\frac{c}{\delta}$ which provides the requisite contradiction.

Finally, when $\mathcal{C}$ is sufficiently small, the good type will have a strict incentive to implement both good and risky projects. This is because $q_r\underline{u}-\mathcal{C}>0$ where the right side is continuation payoff from not acting.

Step 4: Suppose, for contradiction, that for every $\overline{\mathcal{C}}>0$, there is a $\mathcal{C}\in\left(0,\overline{\mathcal{C}}\right)$ and $p_0\in(0,1)$ such that quality control is necessary ($\underline{\Pi}^{\mathcal{C}}(p_0)\leq 0$) but there is nonetheless a nontrivial $\underline{u}$-equilibrium.

We will argue that this cannot happen for sufficiently small $\overline{\mathcal{C}}$.

Let $\tilde{x}$, $\tilde{a}_\theta$ be such a nontrivial $\underline{u}$-equilibrium. From Step 2, there is an on-path history $h^t \in \tilde{\mathscr{H}}^t$ such that $\tilde{p}\left(h^t\right) - \underline{p}^{\mathcal{C}} < \varepsilon$, $\tilde{x}\left(h^t\right) > 0$ and there is an on path outcome $h_{t+1} \in \{h_\varphi, \underline{h}\}$ at which the belief $\tilde{p}\left(h^t h_{t+1}\right) < \underline{p}^{\mathcal{C}}$. This implies that $h_{t+1} = h_\varphi$ since the bad type would not pay the cost of action to get continuation utility $U_{\theta_b}(h^t h_{t+1}, \tilde{x}, \tilde{a}_\theta) = 0$. Additionally, this implies that $U_{\theta_b}(h^t \underline{h}, \tilde{x}, \tilde{a}_\theta) \leq \frac{\mathcal{C}}{\beta}$. Hence, for small $\mathcal{C}$, Step 3 implies that the good type will implement good and risky projects ($\tilde{a}_{\theta_g}(h^t, i^t i_g) = \tilde{a}_{\theta_g}(h^t, i^t i_r) = 1$) and so, in particular, failures will be on path.

We now argue that the good type will also have a strict incentive to run bad projects ($\tilde{a}_{\theta_g}(h^t, i^t i_b) = 1$) for small $\mathcal{C}$. First observe that it must be the case that the belief $\tilde{p}(h^t \underline{h}) < 1$ which in turn implies $U_{\theta_b}(h^t \underline{h}, \tilde{x}, \tilde{a}_\theta) = \frac{\mathcal{C}}{\beta}$. Otherwise, the refinement would imply $U_{\theta_g}(h^t \underline{h}, i^t i_b, \tilde{x}, \tilde{a}_\theta) \geq \underline{u}$ which, from Lemma 3, cannot happen for sufficiently small $\mathcal{C}$ (since $U_{\theta_b}(h^t \underline{h}, \tilde{x}, \tilde{a}_\theta) \leq \frac{\mathcal{C}}{\beta}$). By repeating Steps 2 and 3, for the continuation game beginning at history $h^t \underline{h}$, we can find a continuation history where the good type once again has a *strict* incentive to run good and risky projects and so $U_{\theta_g}(h^t \underline{h}, i^t i_b, \tilde{x}, \tilde{a}_\theta) > U_{\theta_b}(h^t \underline{h}, \tilde{x}, \tilde{a}_\theta) = \frac{\mathcal{C}}{\beta}$.

This means that an upper bound of the principal's payoff from experimenting at history h^t is

$$V(h^t, \tilde{x}, \tilde{a}_\theta) \leq \underline{\Pi}^{\mathcal{C}}(\tilde{p}\left(h^t\right)) - \tilde{p}\left(h^t\right)(1-\lambda)\kappa + \delta V(h^t \underline{h}, \tilde{x}, \tilde{a}_\theta).$$

Since $\underline{\Pi}^{\mathcal{C}}(p_0) \leq 0$, $\underline{\Pi}^{\mathcal{C}}$ is continuous and $\tilde{p}\left(h^t\right) < p_0 + \varepsilon$, we can pick $\varepsilon, \mathcal{C}$ small enough so that the right side is strictly negative. To see this, note that $U_{\theta_b}(h^t \underline{h}, \tilde{x}, \tilde{a}_\theta) = \frac{\mathcal{C}}{\beta}$ and so Lemma 3 implies that $V(h^t \underline{h}, \tilde{x}, \tilde{a}_\theta)$ becomes arbitrarily small as $\mathcal{C}$ shrinks.

This provides the final contradiction and the theorem follows. ∎

APPENDIX C. PROOF OF THEOREM 6

We begin by defining an *optimal mechanism* $\tilde{x}^*$ as

$$\tilde{x}^* \in \underset{\tilde{x}}{\operatorname{argmax}}\, V(h^0, \tilde{x}, \tilde{a}_\theta),$$
$$\text{such that } \tilde{a}_\theta \text{ is a best response to } \tilde{x}.$$

Since the principle can commit, note that the above implies that $\tilde{x}^*$ does not have to be sequentially rational at all on-path histories.

We now define $\underline{\kappa}$ as the solution to the equation

$$\lambda_g - c + \max\left\{\lambda_r(q_r - (1-q_r)\underline{\kappa}),\ -(1-\lambda_g-\lambda_r)\underline{\kappa}\right\} + \delta\overline{\Pi} = 0, \tag{7}$$

This value of $\underline{\kappa}$ is high enough that a single inefficient action by an agent (who otherwise acts efficiently) makes the principal's payoff negative even when the belief assigns probability one to the good type. Note that clearly this also implies that the principal's payoff is not positive at any history (where the belief is less than one) where the good type implements either risky or bad projects with probability one.

Recall that we had defined a last on-path history $h^t \in \mathscr{H}$ to be a history where the principal experiments ($\tilde{x}(h^t) > 0$, $\tilde{p}(h^t) < 1$) such that she stops experimenting if no success is generated ($\tilde{x}(h^t h_{t+1} h^{t'}) = 0$ for all $h_{t+1} \in \{\underline{h}, h_\varphi\}$ and $h^t h_{t+1} h^{t'} \in \mathscr{H}$). Also note that, while Lemma 1 was for the game with transfers, it also applies to the baseline game defined in Section 1. This is because the outcome of every NE of the game without transfers can also be achieved in NE when transfers are allowed (since when transfers are off-path, they can be punished).

We now restate Theorem 6 more precisely.

THEOREM (Formal Restatement of Theorem 6)**.** *Suppose the prior belief p_0 satisfies $p_0\lambda_g - c > 0$ and that both players have the same discount factor $\delta = \beta$. Then, for all $\kappa > \underline{\kappa}$ and every optimal mechanism $\tilde{x}^*$ with associated agent best response $\tilde{a}_\theta$, there exists an on-path history $h^t \in \mathscr{H}$ such that $V(h^t, \tilde{x}^*, \tilde{a}_\theta) < 0$.*

This theorem implies that whenever $p_0\lambda_g - c > 0$ and $\kappa > \underline{\kappa}$, the principal optimal NE generates a strictly lower payoff for the principal than an optimal mechanism since $V(h^t, \tilde{x}^*, \tilde{a}_\theta) < 0$ is not possible at any on-path history in any NE. The condition on the belief p_0 is sufficient to ensure that all optimal mechanisms are nontrivial.

PROOF. We begin by arguing that it is without loss to restrict attention to pure strategy best responses. Formally, consider the pure strategy $\tilde{a}^p_\theta$ constructed from $\tilde{a}_\theta$ as follows: at any on-path history where the agent mixes, $\tilde{a}^p_\theta$ chooses the action that benefits the principal. The agent never acts in $\tilde{a}^p_\theta$ at any off-path history.

Note that, under $\tilde{a}^p_\theta$, the set of on-path histories is a subset of those under $\tilde{a}_\theta$. Moreover, for any on-path history h^t under $\tilde{a}^p_\theta$, we must have $V(h^t, \tilde{x}^*, \tilde{a}^p_\theta) = V(h^t, \tilde{x}^*, \tilde{a}_\theta)$; we cannot have

$V(h^t, \tilde{x}^*, \tilde{a}^p_\theta) > V(h^t, \tilde{x}^*, \tilde{a}_\theta)$ as this would imply $V(h^0, \tilde{x}^*, \tilde{a}^p_\theta) > V(h^0, \tilde{x}^*, \tilde{a}_\theta)$ contradicting the optimality of the latter.

We now assume to the converse that there is an optimal mechanism $\tilde{x}^*$ with associated pure strategy best response $\tilde{a}^p_\theta$ such that $V(h^t, \tilde{x}^*, \tilde{a}^p_\theta) \geq 0$ for all on-path histories $h^t \in \tilde{\mathscr{H}}^t$. We will now show that this is not possible which, in turn, proves the theorem because of the above argument.

Step 1: The good type never implements risky or bad projects. Formally, $\tilde{a}^p_{\theta_g}(h^{t'}, i^{t'} i_{t'+1}) = 0$ for all on-path $(h^{t'}, i^{t'} i_{t'+1}) \in \mathscr{H}^g$ with $i_{t+1} \in \{i_r, i_b\}$.

Suppose he did. Then, we would immediately have the contradiction $V(h^{t'}, \tilde{x}^*, \tilde{a}_\theta) < 0$ since we have assumed that $\kappa > \underline{\kappa}$ which implies that the principal's continuation value is negative whenever the good type runs a risky or bad project.

Step 2: The principal does not experiment at any on-path history where a failure has been generated in the past. Formally, $\tilde{x}^*(h^t) = 0$ at every on-path $h^t = (h_1, \ldots, h_t) \in \tilde{\mathscr{H}}^t$, $t \geq 1$ such that $h_{t'} = \underline{h}$ for some $1 \leq t' \leq t$. Consequently, it is without loss to consider $\tilde{a}_{\theta_b}(h^t) = 0$ for all $h^t \in \tilde{\mathscr{H}}^t$.

From Step 1, at any such on-path history h^t, the principal's belief must assign probability one to the bad type $\tilde{p}(h^t) = 0$. If the principal experimented at such a history, it would imply that $V(h^t, \tilde{x}^*, \tilde{a}^p_\theta) < 0$ which is a contradiction. This immediately implies that $\tilde{a}^p_{\theta_b}(h^t) = 0$ for all $h^t \in \tilde{\mathscr{H}}^t$ is a best response to the optimal mechanism $\tilde{x}^*$ since experimentation stops after failure.

Following Steps 1 and 2, it is without loss to consider the following best responses for the agent

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

for all (and not just on-path) histories $h^t \in \mathscr{H}$.

Step 3: There must be a last history.

Suppose not. Then, from the proof of Lemma 1 (which allows for the case where the principal never makes transfers), there must exist an on-path history $h^t \in \tilde{\mathscr{H}}^t$ such that $\tilde{x}^*(h^t) > 0$ and $V(h^t, \tilde{x}^*, \tilde{a}^p_\theta) < 0$ which is a contradiction.

Step 4: At every last history, the principal stops experimenting even after a success. Formally, following every last history $h^t \in \tilde{\mathscr{H}}^t$, it must be the case that $\tilde{x}^*(h^t\overline{h}h^{t'}) = 0$ for all $h^t\overline{h}h^{t'} \in \tilde{\mathscr{H}}^{t+t'+1}$, $t' \geq 0$.

First observe that $\tilde{p}(h^t) > 0$ as otherwise $\tilde{x}^*(h^t) > 0$ implies the contradiction that $V(h^t, \tilde{x}^*, \tilde{a}^p_\theta) < 0$. Now suppose to the converse that the principal experimented following a success at some last history h^t. Then, $\tilde{a}^p_{\theta_g}(h_t, i_t i_r) = 1$ is a strict best response which contradicts Step 1.

Step 5: Final contradiction. The principal's payoff can be strictly raised by altering the mechanism at any last history h^t.

We now define an alternate mechanism $\tilde{x}$ and show that it can strictly raise the principal's payoff. We begin by taking $\tilde{x} = \tilde{x}^*$ and then alter $\tilde{x}$ in what follows; at any history where $\tilde{x}$ has not been explicitly altered, it is the same as $\tilde{x}^*$.

We first change the optimal mechanism $\tilde{x}^*$ by reducing the amount of experimentation at h^t but increasing the experimentation in the continuation mechanism (so h^t is no longer a last history).

We define

$$\tilde{x}(h^t) = \tilde{x}^*(h^t) - \varepsilon, \quad \tilde{x}(h^t\overline{h}) = \gamma \quad \text{and} \quad \tilde{x}(h^t h_\varphi) = q_r\gamma,$$

where $\varepsilon, \gamma > 0$ are chosen to satisfy

$$(\tilde{x}^*(h^t) - \varepsilon)(1 + (\lambda_g + (1-\lambda_g)q_r)\delta\gamma) = \tilde{x}^*(h^t). \tag{8}$$

As should be clear from the equation above, the good type's payoff when the principal decides to experiment at h^t is the same under $\tilde{x}$ (the left side) as it is under $\tilde{x}^*$ (the right side).

We will now alter the continuation mechanism when the principal's realized choice at h^t is not to experiment. The aim is to ensure that the payoff of the good type at h^t is the same in $\tilde{x}$ and $\tilde{x}^*$. Note that this is necessary because under mechanism $\tilde{x}$ there is an ε greater likelihood that the continuation history $h^t\not{h}$ is reached.[22]

We first use $\not{h}^{t'} = (\not{h}, \dots, \not{h})$ to denote length $t' \geq 1$ vector where each element corresponds to the principal not experimenting; $\not{h}^{t'} = \varphi$ by convention when $t' = 0$. We now recursively

[22] If we had assumed that there was a public randomization device, this could have been done in a very straightforward way: with probability ε at history h^t, the principal could stop experimenting forever and with probability $1 - \tilde{x}^*(h^t)$ the principal's continuation strategy at histories $h^t\not{h}h^{t'} \in \mathscr{H}$ is given by $\tilde{x}(h^t\not{h}h^{t'}) = \tilde{x}^*(h^t\not{h}h^{t'})$. We instead replicate the above described mechanism without resorting to a public randomization device as follows.

define $\tilde{x}$ for all $t' \geq 1$, $h^t \not{h}^{t'} \in \mathscr{H}$ as

$$\tilde{x}(h^t \not{h}^{t'}) \prod_{s=0}^{t'-1} (1 - \tilde{x}(h^t \not{h}^{s})) = \tilde{x}^*(h^t \not{h}^{t'}) \prod_{s=0}^{t'-1} (1 - \tilde{x}^*(h^t \not{h}^{s})). \tag{9}$$

These terms capture the probability that the principal experiments at period $t + t' + 1$ but does not experiment between periods $t + 1$ and $t + t'$ in mechanisms $\tilde{x}$ and $\tilde{x}^*$ respectively. Note that, conditional on experimenting at any period $t + 2$ onwards, mechanisms $\tilde{x}$ and $\tilde{x}^*$ are identical.

We now argue that $\tilde{a}^p_\theta$ is also a best response to $\tilde{x}$. First observe, that $\tilde{x}$ has not been altered at any history where a failure has occurred in the past. Thus, like $\tilde{x}^*$, the principal stops experimenting in $\tilde{x}$ whenever a failure is observed (Step 2) and thus $\tilde{a}^p_{\theta_b}$ remains a best response.

We finally argue that $\tilde{a}^p_{\theta_g}$ is a best response to $\tilde{x}$. First observe that the agent is indifferent between acting or not on low-quality information at h^t. Then observe that the good type's payoff at h^t (before the principal's experimentation decision is realized) is given by

$$\begin{aligned}
& U_{\theta_g}(h^t, i^t, \tilde{x}, \tilde{a}^p_{\theta_g}) \\
= \;& \tilde{x}(h^t) \left[1 + (\lambda_g + (1 - \lambda_g) q_r) \delta \gamma\right] \\
& + \sum_{t'=1}^{\infty} \delta^{t'} \left(\prod_{s=0}^{t'-1} (1 - \tilde{x}(h^t \not{h}^{s})) \right) \tilde{x}(h^t \not{h}^{t'}) \Big[1 + \lambda_g \delta U_{\theta_g}(h^t \not{h}^{t'} \overline{h}, i^{t+t'} i_g, \tilde{x}, \tilde{a}^p_{\theta_g}) + \lambda_r \delta U_{\theta_g}(h^t \not{h}^{t'} h_\varphi, i^{t+t'} i_r, \tilde{x}, \tilde{a}^p_{\theta_g}) \\
& + (1 - \lambda_g - \lambda_r) \delta U_{\theta_g}(h^t \not{h}^{t'} h_\varphi, i^{t+t'} i_b, \tilde{x}, \tilde{a}^p_{\theta_g})\Big] \\
= \;& \tilde{x}^*(h^t) \\
& + \sum_{t'=1}^{\infty} \delta^{t'} \left(\prod_{s=0}^{t'-1} (1 - \tilde{x}^*(h^t \not{h}^{s})) \right) \tilde{x}^*(h^t \not{h}^{t'}) \Big[1 + \lambda_g \delta U_{\theta_g}(h^t \not{h}^{t'} \overline{h}, i^{t+t'} i_g, \tilde{x}^*, \tilde{a}^p_{\theta_g}) + \lambda_r \delta U_{\theta_g}(h^t \not{h}^{t'} h_\varphi, i^{t+t'} i_r, \tilde{x}^*, \tilde{a}^p_{\theta_g}) \\
& + (1 - \lambda_g - \lambda_r) \delta U_{\theta_g}(h^t \not{h}^{t'} h_\varphi, i^{t+t'} i_b, \tilde{x}^*, \tilde{a}^p_{\theta_g})\Big] \\
= \;& U_{\theta_g}(h^t, i^t, \tilde{x}^*, \tilde{a}^p_{\theta_g}),
\end{aligned}$$

where the equalities follow from (8), (9) and the fact that $\tilde{x}$ and $\tilde{x}^*$ are identical at histories $h^t \not{h}^{t'} h_{t+t'+1}$ for $h_{t+t'+1} \in \{\overline{h}, h_\varphi\}$. Hence, the incentive at any history prior to period t remains unaffected. Also, note that at any history after period $t + 1$ where the good type can act, his incentives are identical under $\tilde{x}$ and $\tilde{x}^*$ and so $\tilde{a}^p_{\theta_g}$ is also a best response to $\tilde{x}$.

We end the proof by showing that the principal's payoff from $\tilde{x}$ is strictly higher than $\tilde{x}^*$. In the algebra that follows we use the shorthand notation $\eta = \lambda_g - c$ and $p = \tilde{p}(h^t)$. Now the principal's payoff from $\tilde{x}^*$ at history h^t is:

$$V(h^t, \tilde{x}^*, \tilde{a}^p_\theta) = \tilde{x}^*(h^t)\left(p\lambda_g - c\right) + (1 - \tilde{x}^*(h^t))V(h^t \not{h}, \tilde{x}^*, \tilde{a}^p_\theta). \tag{10}$$

The principal's payoff from the altered mechanism $\tilde{x}$ is:

$$\begin{aligned} V(h^t, \tilde{x}, \tilde{a}^p_\theta) =& (\tilde{x}^*(h^t) - \varepsilon)\left[p\lambda_g(1 + \delta\gamma(\lambda_g - c)) + p(1-\lambda_g)(\delta q_r \gamma(\lambda_g - c)) - (1-p)\delta q_r \gamma c - c\right] \\ &+ (1 - \tilde{x}(h^t))V(h^t \not{h}, \tilde{x}, \tilde{a}^p_\theta) \\ =& (\tilde{x}^*(h^t) - \varepsilon)\left[p\lambda_g(1 + \delta\gamma(\lambda_g - c)) + p(1-\lambda_g)(\delta q_r \gamma(\lambda_g - c)) - (1-p)\delta q_r \gamma c - c\right] \\ &+ (1 - \tilde{x}^*(h^t))V(h^t \not{h}, \tilde{x}^*, \tilde{a}^p_\theta) \end{aligned} \tag{11}$$

We finally show that the perturbed mechanism gives the principal a strictly higher continuation utility. To see this, take the difference between (11) and (10) to get

$$\begin{aligned} V(h^t, \tilde{x}, \tilde{a}^p_\theta) - V(h^t, \tilde{x}^*, \tilde{a}^p_\theta) =& (\tilde{x}^*(h^t) - \varepsilon)(p\eta\delta\gamma(\lambda_g + (1-\lambda_g)q_r) - c(1-p)\delta q_r \gamma) - \varepsilon(p\lambda_g - c) \\ =& (\tilde{x}^*(h^t) - \varepsilon)(p\eta\delta\gamma(\lambda_g + (1-\lambda_g)q_r) - c(1-p)\delta q_r \gamma) - \varepsilon p\eta + \varepsilon(1-p)c. \end{aligned}$$

Substituing in (8), we get

$$\begin{aligned} V(h^t, \tilde{x}, \tilde{a}^p_\theta) - V(h^t, \tilde{x}^*, \tilde{a}^p_\theta) =& -(\tilde{x}^*(h^t) - \varepsilon)c(1-p)\delta q_r \gamma + \varepsilon(1-p)c \\ =& (1-p)c(\varepsilon - (\tilde{x}^*(h^t) - \varepsilon)\delta q_r \gamma) \\ >& 0. \end{aligned}$$

To see the last inequality observe that we can rewrite (8) as

$$\begin{aligned} & (\tilde{x}^*(h^t) - \varepsilon)(\delta(\lambda_g + (1-\lambda_g)q_r)\gamma) = \varepsilon \\ \Longrightarrow\ & (\tilde{x}^*(h^t) - \varepsilon)(\delta(\lambda_g(1-q_r) + q_r)\gamma) = \varepsilon \\ \Longrightarrow\ & (\tilde{x}^*(h^t) - \varepsilon)\delta q_r \gamma < \varepsilon. \end{aligned}$$

■

APPENDIX D. OMITTED MATERIAL FROM SECTION 4

D.1. *Details for the Example*

In this subsection, we formally define the strategies corresponding to the NE example in Section 4 and argue that conditions (C1)–(C3) are not vacuous.

The principal follows the following strategy.

- Experiment in period one and experiment forever (stop experimenting) if a success (failure) is observed. Formally,

$$\tilde{x}(h^0) = 1,\ \tilde{x}(\overline{h}h^t) = 1 \text{ and } \tilde{x}(\underline{h}h^t) = 0 \quad \text{for all } h^t \in \mathscr{H}.$$

- If the agent does not act in period one, experiment in period two. If a success is observed, experiment at all subsequent periods. Formally,

$$\tilde{x}(h_\varphi) = 1, \text{ and } \tilde{x}(h_\varphi \overline{h} h^t) = 1 \quad \text{for all } h^t \in \mathscr{H}.$$

- 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,

$$\tilde{x}(h_\varphi h_2) = 1,\ \tilde{x}(h_\varphi h_2 h^t) = 0 \quad \text{for } h_2 \in \{h_\varphi, \underline{h}\} \text{ and for all } h^t \in \mathscr{H},\ t \geq 1.$$

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,

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

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

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

- 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,

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

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

- 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,

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

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

Finally, the bad type never acts or, formally, that

$$\tilde{a}_{\theta_b}(h^t) = 0 \ \text{ for all } h^t \in \mathscr{H}.$$

We now argue that there are parameter values that satisfy conditions (C1)–(C3); consequently, for these parameter values the above strategies constitute a NE. First note that the initial belief p_0 does not appear in these conditions so, obviously, choosing a sufficiently high p_0 will not violate them. Then, observe that there are enough free parameters to satisfy these conditions. Conditions (C1), (C3) are satisfied by sufficiently small q_r, c respectively for any values of the other parameters. Finally, one can always pick a κ to satisfy equation (C2) since it does not appear in the other two equations.[23]

D.2. *Omitted Results Referred to in Section 4*

In this section, we formally state and prove results that we referred to in Section 4. The first theorem shows that EV's main result (Theorem 1 in their paper) also obtains in our model

[23] Indeed it is easy to show that the set of parameters for which this form of NE exists is not a non-generic set as equation (C2) does not need to equal 0 but being close to 0 suffices.

when the principal is myopic. Let $\mathcal{U}_{\theta_g}(p,\beta)$ denote the highest *average* NE payoff that the good-type agent can obtain when the principal's belief at the beginning of the game is p.[24]

THEOREM 7. *Suppose the bad-type is a commitment type that never acts, the principal is myopic ($\delta = 0$) and (strict) quality control is necessary ($\underline{\Pi}(p_0) < 0$).*[25] *Then,*

$$\lim_{\beta \to 1} \mathcal{U}_{\theta_g}(p,\beta) = 0 \quad \textit{for all } p \leq p_0.$$

PROOF. First observe that because the principal is myopic, the agent must generate successes with sufficiently high probability in every period. That is, if the principal's belief is p, the good type must generate a success with at least probability c/p in order for the myopic principal to pay the cost of experimentation.

Note that this also implies that for p sufficiently low, the principal will stop experimenting because the maximal probability with which the good type can generate a success is bounded above by $\lambda_g + q_r\lambda_r$. We use p_b to denote the belief at which

$$\frac{c}{p_b} = \lambda_g + q_r\lambda_r.$$

Clearly, the principal will never experiment for any belief $p < p_b$.

Now observe that, for any belief $p \geq p_b$ at the start of the game, the belief following inaction must drop sufficiently low. Let $p_{o_\varphi}(p) := \tilde{p}(h_\varphi)$ be shorthand notation for the continuation beliefs generated after updating belief p from period-1 outcome o_φ. Since the bad type never acts, the posterior belief following either a success or failure must jump to 1.

We now use the martingale property of beliefs to derive an upper bound for $p_{o_\varphi}(p)$. Observe that

$$\begin{aligned}
&\text{Prob of success} \times 1 + \text{Prob of failure} \times 1 + \text{Prob of inaction} \times p_{o_\varphi}(p) = p \\
\implies &(\text{Prob of failure} + \text{Prob of inaction})p_{o_\varphi}(p) \leq p - \text{Prob of success} \\
\implies &p_{o_\varphi}(p) \leq \frac{p - \text{Prob of success}}{1 - \text{Prob of success}} \leq \frac{p-c}{1-c},
\end{aligned}$$

[24]Given strategies $\tilde{x}$ and $\tilde{a}_{\theta_g}$, the average payoff of the good type at history (h^{t-1}, i^{t-1}) is defined to be $(1-\beta)U_{\theta_g}(h^{t-1}, i^{t-1}, \tilde{x}, \tilde{a}_{\theta_g})$.

[25]We need this inequality to be strict because the bad type is assumed to never act.

where the last inequality follows from the fact that

$$\text{Prob of success} \geq \frac{c}{p} \geq c.$$

Inverting this mapping this implies that we can divide the space of beliefs as follows

$$p^1 = p_b, \quad p^{n+1} = p^n(1-c) + c, \qquad \text{for all } n \geq 1$$

and observe that

$$p^{n+1} > p^n \text{ and } \lim_{n\to\infty} p^n = 1.$$

Note that the above argument implies that, for all beliefs $p \in (p^n, p^{n+1}]$, we have $p_{o_\varphi}(p) \leq p^n$.

We now prove our result by induction. The base case holds trivially as there is a unique NE with no experimentation when $p \leq p_b$.

Induction hypothesis: $\lim_{\beta\to 1} \mathcal{U}_{\theta_g}(p,\beta) = 0$ for $p \leq p^n$ where n is such that $p^n < p_0$.

Induction Step: We will argue that

$$\lim_{\beta\to 1} \mathcal{U}_{\theta_g}(p,\beta) = 0 \qquad \text{for } p^n < p \leq \min\{p_0, p^{n+1}\}.$$

So consider the NE that assigns the good type the highest payoff $\mathcal{U}_{\theta_g}(p,\beta)$ and let the payoffs from the continuation NE that supports this be $\widehat{\mathcal{U}}_{\theta_g}(h_1,\beta)$ where $h_1 \in \{\not{h}, h_\varphi, \overline{h}, \underline{h}\}$. First suppose $\mathcal{U}_{\theta_g}(p,\beta) = 0$. Then the only on path period-2 history is $h_1 = \not{h}$. It is then without loss to set the off-path continuation payoffs to their lowest values $\widehat{\mathcal{U}}_{\theta_g}(h_\varphi,\beta) = \widehat{\mathcal{U}}_{\theta_g}(\overline{h},\beta) = \widehat{\mathcal{U}}_{\theta_g}(\underline{h},\beta) = 0$.

Now consider the less trivial case where $\mathcal{U}_{\theta_g}(p,\beta) > 0$. First observe that the principal must immediately experiment (or $\tilde{x}(h^0) = 1$) in the agent optimal NE since delaying reveals no information and only lowers the good type's payoffs. Also note that the period-2 history h_φ is on path because the bad type never acts.

We first argue that

$$\widehat{\mathcal{U}}_{\theta_g}(\underline{h},\beta) \leq \widehat{\mathcal{U}}_{\theta_g}(h_\varphi,\beta) \leq \widehat{\mathcal{U}}_{\theta_g}(\overline{h},\beta).$$

To see this, first note that $\widehat{\mathcal{U}}_{\theta_g}(\overline{h},\beta) \geq \widehat{\mathcal{U}}_{\theta_g}(h_\varphi,\beta)$ because otherwise the good type would never implement good projects which would contradict the optimality of the principal's experimentation decision ($\tilde{x}(h^0) = 1$).

Now note that, since $\underline{\Pi}(p) < 0$, the good type cannot have a strict incentive to run risky projects because then it would not be optimal for the principal to experiment. In other words,

$$\widehat{\mathcal{U}}_{\theta_g}(h_\varphi,\beta) \geq q_r \widehat{\mathcal{U}}_{\theta_g}(\overline{h},\beta) + (1-q_r)\widehat{\mathcal{U}}_{\theta_g}(\underline{h},\beta).$$

Since $\widehat{\mathcal{U}}_{\theta_g}(\overline{h},\beta) \geq \widehat{\mathcal{U}}_{\theta_g}(h_\varphi,\beta)$, the above condition immediately implies that $\widehat{\mathcal{U}}_{\theta_g}(h_\varphi,\beta) \geq \widehat{\mathcal{U}}_{\theta_g}(\underline{h},\beta)$ and $\widehat{\mathcal{U}}_{\theta_g}(\underline{h},\beta) \geq 0$ implies

$$\widehat{\mathcal{U}}_{\theta_g}(\overline{h},\beta) \leq \frac{\widehat{\mathcal{U}}_{\theta_g}(h_\varphi,\beta)}{q_r}$$

We complete the proof by observing that

$$\mathcal{U}_{\theta_g}(p,\beta) \leq (1-\beta) + \beta\widehat{\mathcal{U}}_{\theta_g}(\overline{h},\beta) \leq (1-\beta) + \beta\,\frac{\widehat{\mathcal{U}}_{\theta_g}(h_\varphi,\beta)}{q_r}$$

which in turn implies

$$\lim_{\beta\to 1}\mathcal{U}_{\theta_g}(p,\beta) \leq \lim_{\beta\to 1}\left[(1-\beta) + \beta\frac{\widehat{\mathcal{U}}_{\theta_g}(h_\varphi,\beta)}{q_r}\right] = 0$$

because $p_{o_\varphi}(p) \leq p^n$ (irrespective of β) and so we can invoke the induction hypothesis to conclude $\lim_{\beta\to 1}\widehat{\mathcal{U}}_{\theta_g}(h_\varphi,\beta) = 0$. ■

The second result is the analogue of EV's folk theorem (Theorem 4 in their paper) which also obtains in our model when the principal becomes patient.

THEOREM 8. *For any value of the agent's discount factor β, there is a sequence of equilibria such that the principal's average NE payoff converges to the full information payoff $p_0(\lambda_g - c)$ as she becomes arbitrarily patient ($\delta \to 1$).*

PROOF. We will explicitly construct a sequence of such equilibria. Consider the following strategies that depend on the principal's discount factor δ.

Principal:

- Experiment for $T(\delta)$ periods (irrespective of outcomes).
- If a success arrives, experiment forever.
- If no success arrives for $T(\delta)$ periods, stop experimenting.

Formally, for all $h_1 \dots h_t \in \mathscr{H}$,

$$\tilde{x}(h_1 \dots h_t) = \begin{cases} 1 & \text{if } t \leq T(\delta) \text{ or } h_{t'} = \overline{h} \text{ for some } t' \leq T(\delta), \\ 0 & \text{otherwise.} \end{cases}$$

Good type:

- Implement both good and risky projects for the first $T(\delta)$ periods unless a success arrives.
- If a success arrives in the first $T(\delta)$ periods, continue to only implement good projects.
- If no success arrives for $T(\delta)$ periods, stop acting.

Formally, for all $(h_1 \dots h_t, i^t i_{t+1}) \in \mathscr{H}^g$,

$$\tilde{a}_{\theta_g}(h_1 \dots h_t, i^t i_{t+1}) = \begin{cases} 1 & \text{if } t \leq T(\delta), h_{t'} \neq \overline{h} \text{ for any } t' \leq t, \text{ and } i_{t+1} \in \{i_g, i_r\}, \\ 1 & \text{if } i_{t+1} = i_g \text{ and } h_{t'} = \overline{h} \text{ for some } t' \leq T(\delta), \\ 0 & \text{otherwise.} \end{cases}$$

Bad type:

- Never acts.

Formally, for all $h^t \in \mathscr{H}$,

$$\tilde{a}_{\theta_b}(h^t) = 0.$$

Consider the belief $p_f(\delta)$ that satisfies

$$\begin{aligned} &\underline{\Pi}(p_f(\delta)) = 0, \\ \text{i.e. } &p_f(\delta)\left((\lambda_g + q_r\lambda_r)[(1-\delta) + \delta(\lambda_g - c)] - (1-q_r)\lambda_r\kappa(1-\delta)\right) - (1-\delta)c = 0, \\ \text{i.e., } &p_f(\delta) = \frac{(1-\delta)c}{(\lambda_g + q_r\lambda_r)[(1-\delta) + \delta(\lambda_g - c)] - (1-q_r)\lambda_r\kappa(1-\delta)}. \end{aligned}$$

This the lowest belief at which the principal finds it worthwhile to experiment for one period when she gets the first-best continuation value after a success (and stops experimenting if no success arrives), the good type implements both good and risky projects and the bad type does not act. Clearly,

$$p_f(\delta) \to 0 \text{ as } \delta \to 1.$$

Given the agent's strategy, the principal's belief when no project has been run for T periods is

$$\tilde{p}(h_\varphi^T) = \frac{p_0(1-\lambda)^T}{p_0(1-\lambda)^T + 1 - p_0}$$

where $h_\varphi^T = (h_\varphi, \ldots, h_\varphi)$ is the shorthand notation for this history. Set $\overline{T}(\delta)$ be solve $\tilde{p}(h_\varphi^T) = p_f(\delta)$ and observe that $\overline{T}(\delta) \to \infty$ as $\delta \to 1$. In words, as the principal becomes patient, it is sequentially rational for her to experiment for arbitrarily long in search of a success when the agent follows the above strategy.

We now argue that there are constants $\rho_1 < 0$, $\rho_2 \in \mathbb{R}$ and δ^+ such that for all $\delta > \delta^+$, we have

$$\overline{T}(\delta) \geq \rho_1 \log(1-\delta) + \rho_2.$$

This follows from

$$\begin{aligned}
\frac{p_0(1-\lambda)^{\overline{T}(\delta)}}{p_0(1-\lambda_g)^{\overline{T}(\delta)} + 1 - p_0} &= \frac{(1-\delta)c}{(\lambda_g + q_r\lambda_r)[(1-\delta) + \delta(\lambda_g - c)] - (1-q_r)\lambda_r\kappa(1-\delta)}. \\
\implies 1 + \frac{1-p_0}{p_0}(1-\lambda)^{-\overline{T}(\delta)} &= \frac{\lambda_g + \lambda_r(q_r - (1-q_r)\kappa)}{c} + \frac{\delta(\lambda_g + q_r\lambda_r)}{1-\delta}\frac{\lambda_g - c}{c} \\
\implies \frac{1-p_0}{p_0}(1-\lambda)^{-\overline{T}(\delta)} &= \frac{\lambda_g + \lambda_r(q_r - (1-q_r)\kappa) - c}{c} + \frac{\delta(\lambda_g + q_r\lambda_r)}{1-\delta}\frac{\lambda_g - c}{c} \\
&= \frac{\lambda_g - c}{c} + \frac{\delta(\lambda_g + q_r\lambda_r)}{1-\delta}\frac{\lambda_g - c}{c} + \frac{\lambda_r(q_r - (1-q_r)\kappa) - c}{c} \\
&= \frac{\lambda_g - c}{c}\left(1 + \frac{\delta(\lambda_g + q_r\lambda_r)}{1-\delta}\right) + \frac{\lambda_r(q_r - (1-q_r)\kappa) - c}{c} \\
&\geq \frac{\lambda_g - c}{c}\left(\frac{\lambda_g + q_r\lambda_r}{1-\delta}\right) + \frac{\lambda_r(q_r - (1-q_r)\kappa) - c}{c} \\
\implies \ln\frac{1-p_0}{p_0} - \overline{T}(\delta)\log(1-\lambda) &\geq -\log(1-\delta) + \log\left(\frac{(\lambda_g + q_r\lambda_r)(\lambda_g - c)}{c} + \frac{\lambda_r(q_r - (1-q_r)\kappa) - c}{c}(1-\delta)\right) \\
&\geq -\log(1-\delta) + \log\left(\frac{(\lambda_g + q_r\lambda_r)(\lambda_g - c)}{c} + \frac{\lambda_r(q_r - (1-q_r)\kappa) - c}{c}(1-\delta^+)\right)
\end{aligned}$$

The last inequality follows since $\lambda_r(q_r - (1-q_r)\kappa) - c < 0$ and δ^+ can be chosen large enough to make the overall term in the logarithm positive. The constants ρ_1 and ρ_2 can be found by matching terms, note that ρ_1 is negative as desired.

So consider strategies in which the duration of experimentation (without a success) is

$$T(\delta) = \frac{\rho_3}{(\log\delta)^{1/3}} + \rho_4 + \rho_5(\delta),$$

where $\rho_3 < 0$, $\rho_4 \in \mathbb{R}$ and the final term $\rho_5(\delta) \in (-1, 0)$ ensures that $T(\delta)$ is rounded down to the closest positive integer.

Now note that there are parameter values ρ_3–ρ_5 and a discount factor $\underline{\delta} < 1$, $\underline{\delta} > \delta^+$.,such that

$$0 \leq T(\delta) \leq \rho_1 \log(1-\delta) + \rho_2 \leq \overline{T}(\delta) \quad \text{for all} \quad \delta \geq \underline{\delta}.$$

To see this, first observe that, for high enough δ, $\rho_1 \log(1-\delta) + \rho_2 > 0$ and indeed can be made arbitrarily large. Then, we can find a high enough $\underline{\delta}$ and pick ρ_3–ρ_5 so that: [26]

$$0 \leq T(\underline{\delta}) \leq \rho_1 \log(1-\underline{\delta}) + \rho_2, \text{ and,}$$
$$\frac{d}{d\delta}\left(\frac{\rho_3}{(\log\delta)^{1/3}}\right) = -\frac{1}{3}\frac{\rho_3}{\delta(\log\delta)^{4/3}} \leq -\frac{\rho_1}{1-\delta} = \frac{d}{d\delta}(\rho_1 \log(1-\delta)) \text{ for } \delta \geq \underline{\delta}.$$

To see that we can find a sufficiently high $\underline{\delta}$ to satisfy the second inequality, observe that this inequality is equivalent to $\frac{1}{3}\frac{\rho_3}{\rho_1}\frac{1-\delta}{\delta} \geq (\log\delta)^{4/3}$ and that the left side decreases faster (the derivative satisfies $-\frac{1}{3}\frac{\rho_3}{\rho_1}\frac{1}{\delta^2} \to -\frac{1}{3}\frac{\rho_3}{\rho_1}$ as $\delta \to 1$) than the right side (the derivative satisfies $\frac{4}{3}\frac{(\log\delta)^{1/3}}{\delta} \to 0$ as $\delta \to 1$) for sufficiently high δ.

Taken together, these imply that $T(\delta) \leq \rho_1 \log(1-\delta) + \rho_2$ for all $\underline{\delta} \leq \delta < 1$ or, in other words, that the strategies corresponding to this choice of $T(\delta)$ constitute a NE.

This particular $T(\delta)$ was chosen to ensure that $\lim_{\delta\to 1} \delta^{T(\delta)} = 1$. To see this, observe that

$$\lim_{\delta\to 1} \delta^{T(\delta)} = \lim_{\delta\to 1} \delta^{\frac{\rho_3}{(\log\delta)^{1/3}}} \delta^{\rho_4} \delta^{\rho_5(\delta)} = \lim_{\delta\to 1} \delta^{\frac{\rho_3}{(\log\delta)^{1/3}}} = 1,$$

where the second equality follows from the fact that $\rho_5(\delta)$ is bounded. We end the proof by writing the expression for principal's average payoff $\mathcal{V}(\delta)$ which is given by

$$\mathcal{V}(\delta) = p_0\left((\lambda_g + q_r\lambda_r)[(1-\delta) + \delta(\lambda_g - c)] - (1-q_r)\lambda_r\kappa(1-\delta) - (1-\delta)c\right)\frac{1-(\delta(1-(\lambda_g+q_r\lambda_r)))^{T(\delta)}}{1-\delta(1-(\lambda_g+q_r\lambda_r))}$$
$$-(1-p_0)\left(1-\delta^{T(\delta)}\right)c$$

and therefore

$$\lim_{\delta\to 1} \mathcal{V}(\delta) = p_0(\lambda_g - c). \qquad \blacksquare$$

[26]Pick a ρ_3 to satisfy the second inequality and then a ρ_4 to satisfy the first.

Appendix E. Omitted Material from Section 5.1

In this section, we present the strategies that we described informally in Section 5.1. Formally, we are restricting the principal's strategy to be of the form

$$\hat{x}(h_1 \dots h_t) \in \begin{cases} [0,1] & \text{if } h_{t-T+2} = \cdots = h_t = \not{h}, \\ \{1\} & \text{otherwise.} \end{cases}$$

In words, the probability of experimentation must be one if she has experimented in any of the past $T-1$ periods. Here we use the convention that $h_{t'} = \not{h}$ if $t' \leq 0$. A different way of interpreting this particular form of partial commitment is that when a principal decides to experiment, she sinks the T period discounted cost of experimentation $\frac{1-\delta^T}{1-\delta}c$ up front.

Now consider the following strategies. The principal commits to experimenting for the first T periods

$$\hat{x}(h^{t-1}) = 1 \quad \text{for all} \quad t \leq T,\ h^{t-1} \in \mathscr{H}.$$

If a failure or no success arrives over these T periods, the principal stops experimenting. If the agent generates at least one success (and no failures) over the first T periods, the principal experiments forever. Formally, for $t \geq T+1$ and $h_1 \dots h_{t-1} \in \mathscr{H}$,

$$\hat{x}(h_1 \dots h_{t-1}) = \begin{cases} 1 & \text{if } h_{t'} \neq \underline{h} \text{ for all } 1 \leq t' \leq T \text{ and } h_{t''} = \overline{h} \text{ for some } 1 \leq t'' \leq T, \\ 0 & \text{otherwise.} \end{cases}$$

In response, the good type agent only acts on high-quality information for $1 \leq t \leq T-1$ periods

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

and, additionally, on low-quality information at period T if no success has arrived,

$$\tilde{a}_{\theta_g}(h^{T-1}, i_1 \dots i_T) = \begin{cases} 1 & \text{if } i_T = i_g, \\ 1 & \text{if } i_T = i_r \text{ and } i_{t'} \neq i_g \text{ for all } 1 \leq t' \leq T-1, \\ 0 & \text{otherwise,} \end{cases}$$

for all $(h^{T-1}, i^{T-1}i_T) \in \mathscr{H}^g$. If a success arrives in the first T periods, he acts efficiently thereafter. Formally, for $t \geq T+1$ and $(h_1 \dots h_{t-1}, i^{t-1}i_t) \in \mathscr{H}^g$,

$$\tilde{a}_{\theta_g}(h_1 \dots h_{t-1}, i^{t-1}i_t) = \begin{cases} 1 & \text{if } i_t = i_g,\ h_{t'} \neq \underline{h} \text{ for all } 1 \leq t' \leq T \text{ and } h_{t''} = \overline{h} \text{ for some } 1 \leq t'' \leq T, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, the bad type only acts with positive probability $\varepsilon \in (0,1)$ in period T so

$$\tilde{a}_{\theta_b}(h^{T-1}) = \varepsilon \text{ for all } h^{T-1} \in \mathscr{H} \text{ and } \tilde{a}_{\theta_b}(h^{t-1}) = 0 \text{ for all } t \neq T,\ h^{t-1} \in \mathscr{H}.$$

We now argue that there are parameter values such that the above strategies constitute an equilibrium even when quality control is necessary. First note that experimentation can be profitable. For instance, if $p_0\lambda_g > c$, then the principal's payoff taking $T \to \infty$ converges to $\frac{p_0\lambda_g - c}{1-\delta} > 0$ so experimentation is also profitable for some finite T irrespective of the value of $\underline{\Pi}(p_0)$ (the optimal such T is finite).

Clearly, the bad type is best responding since he is indifferent between all strategies.[27] Finally, the good type's strategy is a best response if he does not have an incentive to act on low-quality information at any period prior to T. A sufficient condition for this is $\lambda_g \geq q_r$ since the likelihood of getting high quality information (and thereby a sure success) in the next period makes the option value of waiting more attractive than the lottery of generating a success today with probability q_r.

[27] Note that the bad type has no incentive to act prior to T (even though the good type acts with positive probability) because failures between periods 1 and $T-1$ are off path and are punished by the principal.

REFERENCES

ABREU, DILIP, P. D., AND E. STACCHETTI (1993): "Renegotiation and Symmetry in Repeated Games," *Journal of Economic Theory*, 60(2), 217–240.

AGHION, P., AND M. O. JACKSON (2016): "Inducing Leaders to Take Risky Decisions: Dismissal, Tenure, and Term Limits," *American Economic Journal: Microeconomics*, 8(3), 1–38.

ATAKAN, A. E., AND M. EKMEKCI (2012): "Reputation in Long-Run Relationships," *Review of Economic Studies*, 79(2), 451–480.

——— (2013): "A Two-Sided Reputation Result with Long-Run Players," *Journal of Economic Theory*, 148(1), 376–392.

BACKUS, M., AND A. T. LITTLE (2018): "I Don't Know," Unpublished manuscript, Columbia University.

BEN-PORATH, E., E. DEKEL, AND B. L. LIPMAN (2018): "Mechanisms with Evidence: Commitment and Robustness," *Econometrica*, forthcoming.

BOARD, S. (2011): "Relational Contracts and the Value of Loyalty," *American Economic Review*, 101(7), 3349–67.

COMPTE, O. (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66(3), 597–626.

CRIPPS, M. W., AND J. P. THOMAS (1997): "Reputation and Perfection in Repeated Common Interest Games," *Games and Economic Behavior*, 18(2), 141–158.

DEB, R., M. M. PAI, AND M. SAID (2018): "Evaluating Strategic Forecasters," *American Economic Review*, 108(10), 3057–3103.

ELY, J., D. FUDENBERG, AND D. K. LEVINE (2008): "When is Reputation Bad?," *Games and Economic Behavior*, 63(2), 498–526.

ELY, J. C., AND J. VÄLIMÄKI (2003): "Bad Reputation," *Quarterly Journal of Economics*, 118(3), 785–814.

FAINMESSER, I. P., AND A. GALEOTTI (2019): "The Market for Online Influence," Working Paper.

FARRELL, J., AND E. MASKIN (1989): "Renegotiation in Repeated Games," *Games and Economic Behavior*, 1(4), 327–360.

GUO, Y. (2016): "Dynamic Delegation of Experimentation," *American Economic Review*, 106(8), 1969–2008.

HALAC, M. (2012): "Relational Contracts and the Value of Relationships," *American Economic Review*, 102(2), 750–79.

JOVANOVIC, B. (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87(5), 972–990.

LEVIN, J. (2003): "Relational Incentive Contracts," *American Economic Review*, 93(3), 835–857.

LI, J., N. MATOUSCHEK, AND M. POWELL (2017): "Power Dynamics in Organizations," *American Economic Journal: Microeconomics*, 9(1), 217–41.

MALCOMSON, J. M. (2016): "Relational Incentive Contracts with Persistent Private Information," *Econometrica*, 84(1), 317–346.

MITCHELL, M. (2020): "Free (Ad)vice," *RAND Journal of Economics*, forthcoming.

MORRIS, S. (2001): "Political Correctness," *Journal of Political Economy*, 109(2), 231–265.

OTTAVIANI, M., AND P. N. SØRENSEN (2006): "Reputational Cheap Talk," *RAND Journal of Economics*, 37(1), 155–175.

PRENDERGAST, C., AND L. STOLE (1996): "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning," *Journal of Political Economy*, pp. 1105–1134.

SAFRONOV, M., AND B. STRULOVICI (2017): "From Self-Enforcing Agreements to Self-Sustaining Norms," Unpublished manuscript, Northwestern University.

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

VARAS, F. (2018): "Managerial Short-Termism, Turnover Policy, and the Dynamics of Incentives," *Review of Financial Studies*, 31(9), 3409–3451.