**REWARDING MEDIOCRITY?
OPTIMAL REGULATION
OF R&D MARKETS
WITH REPUTATION CONCERNS**

Chia-Hui Chen
Junichiro Ishida

March 2017

Rewarding Mediocrity? Optimal Regulation of R&D Markets with Reputation Concerns

Chia-Hui Chen* and Junichiro Ishida†

March 23, 2017

Abstract

In this paper, we consider a dynamic signaling model of an R&D market in which a researcher can choose either a safe project (exploitation) or a risky project (exploration) at each instance. We argue that there are substantial efficiency gains from rewarding minor innovations above their social value and further that it is indeed superior to rewarding major innovations directly, even when those minor innovations are intrinsically valueless in themselves. When only major innovations are rewarded, the R&D market eventually shuts down due to a version of the lemons problem. Rewarding minor innovations is actually conducive to major innovations as it induces self-sorting among researchers, which is essential in providing time and resources necessary for more productive ones to take riskier but more ambitious approaches. This result draws clear contrast to the static counterpart where such a scheme can never be optimal. Our model also exhibits reputation dynamics which capture a pervasive view in academia that "no publications are better than a few mediocre publications" at an early stage of one's career.

---
*Institute of Economic Reserach, Kyoto University. Email: chchen@kier.kyoto-u.ac.jp
†Institute of Social and Economic Research, Osaka University. Email: jishida@iser.osaka-u.ac.jp

1 Introduction

Since the seminal work of Price (1965), it has been well recognized that the distribution of citations is heavily skewed to the right, with a substantial fraction of scientific papers receiving no citations at all. Even many decades later, the situation has remained roughly the same. More recently, based on the citation data from Web of Science, Wallace et al. (2009) document that although "uncitedness" has declined somewhat in recent years, roughly 60% of papers published in 1996 are cited less than 5 times in the 10-year window.[1] On the right tail of the distribution, Redner (1998) documents that in the samples of papers published in journals catalogued by the Institute of Scientific Information (ISI), only 64 and 2103 out of 783,339 papers (0.008% and 0.268%!) are cited more than 1000 and 200 times, respectively, and shows that the distribution of citations $N(x)$ has a power law form $x^{-\alpha}$ for a large x with α close to 3.[2] Granted that the number of citations is only one way to measure the social value of a scientific work, there is little doubt that this is one of those "superstar industries" where output is heavily concentrated among a few top individuals, and most of the effort is bound to fail in the end.

The outlook is quite different, however, once we turn to how researchers are compensated for their achievements: the distribution of reward, broadly defined, is typically much more gradual and compressed than that of scientific impact.[3] Although there are presumably many reasons for why this is so, this compensation practice raises a concern from the efficiency point of view if we take into account the inherent nature of knowledge creation. As widely accepted in the innovation literature, different types of innovation typically require fundamentally different approaches and methods from the outset, where more radical and path-breaking innovations (hereafter, major innovations) often necessitate the exploration of new ideas while more incremental and obscure ones (minor innovations) build on the exploitation of old ideas. Rewarding minor innovations disproportionately may then distort researchers' attention away from exploration to exploitation, thereby systematically reducing the frequency of major innovations that can advance our scientific knowledge. At a glance, it seems to make more sense to reward major innovations more heavily in order to encourage researchers to take

---

[1] A similar pattern is observed in economics as well. For instance, Hamermesh (2015) finds that the distribution of citations is highly right-skewed even among those paper published in "top 5" journals.

[2] Redner (1998) also shows that the distribution has a stretched exponential form $exp[-(x/x_0)^\beta]$ with $\beta \approx 0.44$ for a small x.

[3] In a seminal analysis of the economics of superstars, Rosen (1981) argues that "in certain kinds of economic activity there is concentration of output among a few individuals, marked skewness in the associated distributions of income and very large reward at the top." The contrast is perhaps clear with other superstar industries such as sports and entertainment where top performers earn enormous reward while the vast majority can barely survive.

riskier but more ambitious approaches.

In light of this concern, this paper aims to explore welfare consequences of a compensation practice under which reward is not proportionally aligned with scientific impact. We argue that although the above reasoning sounds intuitive and even convincing, this "reward compression" can be welfare-enhancing under certain conditions, for it is actually conducive to major innovations. Two features of our setup characterize the R&D market of our focus and play a crucial role in deriving this conclusion.

- Reputation matters, especially at an early stage of a researcher's career: good reputation helps attract external resources such as research grants and productive work environments, constituting a crucial factor of production for R&D activities.

- There are different approaches to innovation with different degrees of risk involved: to generate a major innovation, one must try untested ideas and methods which necessarily entail more uncertainty and are most likely to fail.

The former gives rise to reputation concerns on the part of researchers, while the latter defines the information structure of our setting. We show that when these two features are combined together, there arise substantial efficiency gains from rewarding minor innovations above their social value, even when they are intrinsically valueless and any attempt to generate them is socially wasteful.

More specifically, we analyze a dynamic signaling model in which an agent (researcher) undertakes R&D activities over time. The agent's productivity, which is either high or low, is his private information. To capture the route through which reputation affects productivity, we assume that the agent needs an upfront payment, or an investment flow, from the market to undertake R&D activities, as often assumed in the experimentation literature (Bergemann and Hege, 1998, 2005; Hörner and Samuelson, 2013). The market forms a belief by observing a sequence of innovation outcomes and provides an investment flow based on its belief. At each point in time, the agent chooses one of the three possible actions: safe project (exploitation), risky project (exploration), and no effort (shirking). The return to the safe project is small but predictable in that the agent can come up with a minor innovation for sure by working on it. In contrast, the return to the risky project is unpredictable and sporadic: if the agent chooses the risky project, he generates a major innovation only with some (possibly very small) probability while he fails and ends up with no visible output with the remaining probability. Moreover, we assume that the risky project is ability-intensive in that its success probability depends on the agent's ability type, and in particular that only the high type can succeed in it.

A snapshot of our argument goes as follows. The resource constraint faced by the agent implies that he must establish reputation to secure the requisite investment flow from the market. This process of reputation building becomes tricky when only output, but not input, is publicly observable because the risky project is most likely to fail and produce no visible output in the end, thereby making it very difficult to distinguish those who explore from those who slack off. This fact gives rise to a defining feature of our model where the high type may have no feasible means to fully separate from the low type at his own will, no matter how much cost he is willing to incur: if the high type settles on the safe project, the low type can surely pool with the high type by following the same path; if the high type takes on the risky project, the low type can still easily mimic (at an observational level) with high probability by simply slacking off. The bottom line is that the risky project has a small informational value because of its unpredictable and sporadic nature, which in turn gives rise to the signaling role of the safe project.

In a competitive environment where there is no direct return to minor innovations, the low type simply has no choice but to slack off. Given this, if the high type is to choose the risky project, the two types are observationally indistinguishable, which necessarily lowers the market belief about the agent's productivity. Moreover, as the high type achieves a major innovation sporadically over time and "exit" from the market, the market belief, conditional on no major innovation having occurred, monotonically declines and eventually reaches a level where supporting the risky project can no longer be justified, at which point the R&D market shuts down entirely. In an extreme case where the market holds a relatively pessimistic belief to begin with, the R&D market fails in a stark manner, and no R&D activities, either safe or risky, can take place at all.

This incentive structure changes rather drastically once minor innovations are rewarded via "subsidies" so that the low type has an incentive to take on the safe project with some positive probability.[4] When the subsidy is lucrative enough, the low type occasionally chooses the safe project and voluntarily separates from the high type by achieving a minor innovation. This self-sorting behavior is crucial, as it keeps the market belief high and within some tight bounds, which degenerate to a single point in the continuous-time limit. We derive the socially optimal subsidy scheme and show that the incidence of major innovations is enhanced substantially by rewarding minor innovations. Moreover, we also analyze a scheme which

---

[4]It should be noted that we use the term "subsidy" purely for expositional purposes and very broadly: it intends to capture any gains that accrue from each innovation above its intrinsic value. It includes not only financial supports and grants from public entities but also other more intangible benefits such as social recognition and reputation. Some of them are chosen consciously as policy variables whereas others should more appropriately be regarded as exogenous. Our main aim here is hence to investigate the consequences of artificially creating value for minor innovations regardless of its origin.

directly subsidizes major innovations and make a case for why rewarding minor innovations is superior to rewarding major innovations directly. This result draws clear contrast to the static counterpart where such a scheme can never be optimal.

Although the main contention of this paper is largely normative, our model also exhibits reputation dynamics that capture a norm peculiar to the academic profession. In academia, there is a pervasive belief that "no publications are better than a few mediocre publications," especially at an early stage of one's career. The presence of such a norm is an indication that researchers do indeed pursue different approaches and self-sort themselves into different categories. In our setup, under the optimally chosen subsidy, the high type chooses the risky project while the low type randomizes between the safe project and no effort. Note that, as we have emphasized earlier, the choice of the safe project results in an outcome that is informationally distinct and hence fully reveals the agent's type, which gives rise to the unique reputation dynamics of our model: along the equilibrium path, the reputation of an agent who has achieved a minor innovation is actually lower than that of an agent who has achieved nothing at all.[5]

Related literature: The paper that is most closely related to ours in spirit is Manso (2010) who analyzes a two-period moral-hazard model where the agent decides not only how much effort to supply (work or shirk) but also how much risk to take (explore or exploit) in each period. In this extended setup, he argues that it is essential to tolerate, or even reward, early failure in order to encourage exploratory activities.[6] Our analysis shares the same motivation and builds on a similar (trinary-choice) framework that incorporates a choice between exploration and exploitation. The main difference is that Manso (2010) assumes only one type of success where exploration in his context is about finding the true success probability. He also builds on a bandit problem with a risky arm whose payoff is unknown to the agent whereas ours is cast in the framework of a signaling model in which the agent is privately informed about his innate productivity. Given these differences, our focus is placed more on the principal's learning and ways in which to regulate this dynamic learning process. In addition, we also explicitly consider an optimal stopping problem in a infinite-horizon model – in particular the continuous-time limit of the equilibrium – which allows us to illustrate the optimal extent of tolerance for early failure in further detail.

Our analysis also contributes to the diverse literature on dynamic signaing. In our model,

---

[5]This phenomenon is to be distinguished from the case of counter-signaling where very high types refrain from signaling. In models of counter-signaling, it is typically the case that only medium types engage in costly signaling while high and low types pool on no signaling. See Feltovich et al. (2002), Araujo et al. (2007) and Chung and Eso (2013) for various examples of counter-signaling.

[6]Also, see Ederer and Manso (2013) who provide experimental evidence that the combination of tolerance for early failure and reward for long-term success is quite effective in motivating innovation.

low-type agents reveal their productivity gradually by achieving minor innovations, and the resultant dynamics are hence governed by their indifference condition. In this regard, the paper is closely related to a strand of literature which examines the consequences of dynamic signaling via exit decisions (of firms/sellers) in competitive environments (Hörner, 2002; Bar-Isaac, 2003; Daley and Green, 2012; Gul and Pesendorfer, 2012; Atkeson et al., 2015).[7] The paper is in particular closely related to Atkeson et al. (2015) who consider an environment where firms make unobservable investments upon entry and consumers gradually learn their types over time. With the slow diffusion of information, a mild version of the lemons problem persists where low-type firms enter the market and pool with high-type firms. They then show that in a competitive market environment with spot prices, entry regulation can complement the role of reputation and improve welfare. Two differences from our work are worth noting here. First, their regulatory scheme aims at correcting the inefficiency at the entry level whereas our scheme works through the regulation of exit behavior, reflecting the fact that entry regulation is typically not viable in industries of our interest. Second, in their model, firms choose whether to continue or not once they decide their types upon entry. In contrast, there is no explicit choice of exit in our model: low-type agents effectively "exit" from the market whenever they adopt a strategy which fully reveals their type, but what choice of action constitutes the exit option, if any, depends also on the high type's strategy.

From the market's point of view, the problem we analyze can be seen as a variant of the canonical two-armed bandit problem with one safe arm (terminating the project) and one risky arm (continuing the project). There are now increasingly many works, often called strategic experimentation, where a group of individuals, rather than a single individual, face a bandit problem (Bolton and Harris, 1999; Bergemann and Välimäki, 2000; Keller et al., 2005; Klein and Rady, 2011; Bonatti and Hörner, 2011). In a broad sense, our model also falls into this strand of literature in that the experimentation process involves decisions of multiple individuals. A key departure is that we consider an environment where experiments are "intermediated" by an informed party, i.e., the agent, who may or may not behave as the market wishes.[8]

---

[7]Along this line, Murto and Välimäki (2011) analyze a setting where players are *ex ante* symmetrically informed but gradually learn their types and exit when they become sufficiently pessimistic. Chen and Ishida (2015b) provide a model of hierarchical experimentation where the less confident principal gradually terminates the project over time.

[8]Chen and Ishida (2015a) consider an optimal stopping problem in a similar framework where the informed agent chooses either to work or to shirk while the principal learns his type through a sequence of outcomes. The major difference from this work is that the agent in the current setup faces not only a choice between working and shirking but also that between exploration and exploitation. In a different vein, Guo (2016) analyzes a model of "delegated experimentation" in which a principal delegates experimentation to an (uninformed) agent and solves for the optimal delegation rule.

2 Model

Environment: Consider an infinite-horizon dynamic signaling model in which an agent (e.g., researcher/entreprenuer) seeks to obtain external resources from a market of potential employers (e.g., university/investor) to undertake R&D activities. The agent's ability is either high ($a = h$) or low ($a = l$) where the prior probability that the agent is of the high type is $p_0 \in (0, 1)$. The ability type is the agent's private information which cannot be observed directly by the principal.

Timing: For most part, we describe and analyze the model in discrete time to clarify its timing structure although we will eventually focus on the continuous-time limit for clarity and tractability. More precisely, the agent and the market take actions at times $0, \Delta, 2\Delta, ...$, where we generally assume that Δ is some arbitrarily small number. The timing of events in each period proceeds as below:

1. The market offers a flow of funding, which we refer to as the investment flow, based on its belief about the agent's ability.

2. The agent chooses which project to work on from a set of feasible alternatives.

3. The outcome is realized, depending on the agent's ability and project choice.

4. The market belief is updated given the realized outcome.

Investment flows: At the beginning of each period, the market provides the investment flow upfront to the agent. Let w_t denote the investment flow (per unit of time) offered at time t, which the agent can either invest in his project or divert for his private benefits. The investment flow is necessary for the agent to undertake costly R&D activities: to implement a project, the investment flow must be at least as large as its implementation cost (to be discussed below). We assume that the market is perfectly competitive and simply offers w_t which equals the agent's expected productivity for the period $[t, t+\Delta)$. As such, throughout the analysis, we simply take the market as a nonstrategic offer process which computes the agent's expected productivity and matches the investment flow.

Project choices: Upon receiving w_t, the agent privately chooses which project to work on from a set of alternatives $X := \{R, S, N\}$. There are three possible choices for the agent: risky project R (exploration of new ideas), safe project S (exploitation of existing ideas) and no effort N (shirking). Let $x_t \in X$ denote the agent's project choice at time t. Different project choices lead to different outcomes in that the risky project may potentially generate major innovation while the safe project can constantly generate minor innovation.

Innovation outcomes: Given the agent's ability and project choice, the innovation outcome $y_t \in Y := \{R, S, \varnothing\}$ is realized and publicly observed. There are three potential outcomes, where $y_t = R$ denotes major innovation, $y_t = S$ minor innovation, and $y_t = \varnothing$ failure (no innovation). Two properties characterize the underlying R&D process of our setting. First, the risky project is highly ability-intensive in that only the high type can successfully complete it. Second, the risky project succeeds abruptly at random points in time while the safe project constantly yields some output with no uncertainty. Formally, we capture these properties by the following specification of production technologies:

- If $x_t = R$, the agent succeeds and generates one unit of major innovation ($y_t = R$) with probability $g^a \Delta$ and fails ($y_t = \varnothing$) with probability $1 - g^a \Delta$ where $g^h = g > g^l = 0$.[9]

- If $x_t = S$, the agent succeeds with probability one and generates Δ units of minor innovation ($y_t = S$) regardless of the ability type.

- If $x_t = N$, the agent fails ($y_t = \varnothing$) with probability one regardless of the ability type.

$x_t \backslash a$	h	l
R	$y_t = R$ with prob. $g\Delta$ $y_t = \varnothing$ with prob. $1 - g\Delta$	$y_t = \varnothing$ with prob. one
S	$y_t = S$ with prob. one	$y_t = S$ with prob. one
N	$y_t = \varnothing$ with prob. one	$y_t = \varnothing$ with prob. one

Table 1: innovation outcomes

Implementation costs: The cost of implementing a project varies over time, depending on the past innovation outcomes. We in particular assume that only the first success is costly to achieve – an assumption that is made to capture that there is a long-term gain from achieving an innovation.[10] More precisely, let c_t^x denote the cost of implementing x (per unit of time) at time t. We initially assume that $c_0^x = c^x$ where $c^R > c^S \geq c^N = 0$. Once the agent succeeds in x at time t, however, the game enters the post-innovation stage in the next period where the agent incurs no implementation cost (as long as he works on the same project), i.e., $c_s^x = 0$ for all $s > t$. As noted above, $w_t \geq c_t^x$ is required to implement x at time t.

---

[9]We use $g\Delta$ to approximate $1 - e^{-g\Delta}$ for an arbitrarily small Δ throughout the analysis. The same approximation applies to the discount factor which we introduce below.

[10]For a more practical interpretation, in the case of R&D intensive firms, the post-innovation stage corresponds to the phase where an innovating firm appropriates rent from its newly invented product. In academia, the post-innovation stage reflects the idea that first success is always the most difficult one to achieve, e.g., because one must identify a right topic and a right approach to the topic.

Payoffs: Each innovation carries an intrinsic benefit which we refer to as the social value. Let η^y denote the (per unit) social value of innovation outcome y where $\eta^R > \eta^S \geq \eta^\varnothing \geq 0$. Here, we consider an environment where the market value of an innovation may differ from its social value because an innovation may be "subsidized," as we discuss later. The market value is given by the sum of its social value η^y and a subsidy $b_t^y \geq 0$ and is hence denoted by $\pi_t^y = \eta^y + b_t^y$. Given b_t^y, the principal's payoff in $[t, t + \Delta)$ is

$$
u_t^P = \begin{cases} \pi_t^R - w_t \Delta & \text{if } y_t = R, \\ (\pi_t^S - w_t)\Delta & \text{if } y_t = S, \\ -w_t \Delta & \text{if } y_t = \varnothing, \end{cases}
$$

whereas the agent's payoff is

$$
u_t^A = (w_t - c_t^x)\Delta.
$$

Note that the agent does not derive any direct benefit from producing an innovation per se; his benefit comes indirectly as a success boosts his reputation and raises the subsequent investment flows. Both players discount future payoffs by a common discount factor $1 - r\Delta$.

Simplifying assumptions: Throughout the analysis, we make two assumptions to highlight our main contention.

Assumption 1 $\eta^R = v > \eta^S = 0$.

Assumption 2 $c^R = c > c^S = 0$.

Assumption 1 states that minor innovations are intrinsically valueless, which we assume strictly to emphasize that our main conclusion (that it is optimal to subsidize minor innovations) holds even when the safe project is socially inefficient. Assumption 2 is also innocuous as the level of c^S becomes totally irrelevant in the continuous-time limit: all of our subsequent results hold for any $c^S \in (0, c^R)$ as $\Delta \to 0$ (see Appendix B for more detail). In the interest of reducing the number of parameters and simplifying exposition, therefore, we analyze the case where $c^S = 0$, so that we can summarize the cost structure entirely by a single variable $c^R = c$.

3 A static benchmark

We begin our analysis with a static version of the model to illustrate the nature of our setting as well as to provide a preliminary discussion for our subsequent analysis. In this example, we suppose that the agent simply receives a post-innovation rent π^y at the end of the game

when the outcome is y. For now, the post-innovation rents are taken as exogenous while they will be endogenized in the following sections when we analyze the dynamic model.

The incentive compatibility constraints are all straightforward in this static setting with no dynamic reputation concerns. First, it is clear that the low type never chooses R, given that he can never succeed in it. The low type thus chooses S if $\pi^S > 0$ and N otherwise. In contrast, the high type chooses R if

$$g\pi^R - c \geq \max\{\pi^S, 0\} = \pi^S.$$

Assume $gv > c$, so that R is a socially efficient choice. Since minor innovations are intrinsically valueless, the first-best allocation is such that the high type chooses R while the low type chooses N. Below, we examine when this allocation can be implemented.

First, consider the case with no subsidies, i.e., $\pi^y = \eta^y$ for all $y \in Y$. Since $\pi^S = 0$ by assumption, no one has an incentive to choose S. This means that the low type always chooses N while the high type chooses R if it is feasible, i.e., $w \geq c$, under the maintained assumption. If the low type chooses N and the high type R, perfect competition forces the market to offer

$$w = p_0 gv.$$

If p_0 is large enough to satisfy

$$p_0 gv \geq c,$$

the first-best allocation can be supported as an equilibrium outcome. If this condition fails to hold, on the other hand, the high type is unable to choose R, and a version of the lemons problem surfaces. The only equilibrium in this case is such that the market offers $w = 0$ and both types have no choice but to choose N.

Given the scarcity of major innovations, it is reasonable to assume that p_0 is relatively small and hence $c > p_0 gv$, in which case no valuable innovations would be generated under perfect competition. Even in this case, the equilibrium allocation can be substantially improved by appropriately subsidizing innovations. Since the ultimate goal here is to induce major innovations, it is perhaps natural to consider a scheme that subsidizes b for each major innovation, i.e., $\pi^R = v + b$. This scheme indeed works and realizes the first-best project choice if b is set large enough to satisfy

$$p_0 g(v + b) \geq c.$$

Note that this scheme has no detrimental incentive effect as it keeps all the incentive conditions intact for both types.

Although it appears a bit awkward, one may also consider a scheme that subsidizes minor innovations to raise the investment flow above the requisite level c. Now suppose that b is subsidized for each minor innovation, i.e., $\pi^S = \eta^S + b = b$. Under this scheme, it is now strictly optimal for the low type to choose S, and the investment flow is obtained as

$$w = p_0 gv + (1 - p_0)b.$$

The high type then chooses R if

$$gv - c \geq b \quad \text{and} \quad w = p_0 gv + (1 - p_0)b \geq c,$$

provided that there exists b that can satisfy both conditions.

Does it ever make sense to subsidize minor innovations so as to induce major innovations? In this static benchmark, the answer is a resounding no because that would distort both types' incentives. First, under the scheme that subsidizes minor innovations, the low type is necessarily induced to work on the inefficient safe project (although this efficiency loss is assumed to be arbitrarily small throughout the analysis).[11] Second, this scheme also has a distortionary incentive effect for the high type: it makes the safe project more attractive, giving the high type a stronger incentive to choose S instead of R. In particular, if

$$gv > (2 - p_0)c,$$

there is no subsidy level b that can induce the high type to choose R.

This argument suggests that there is no efficiency rationale to subsidize minor innovations in this static environment, as any compensation practice which rewards minor innovations above their social value can never be welfare-enhancing. As we will argue below, however, this conclusion is reversed once dynamic incentives are explicitly taken into account: a scheme that subsidizes minor innovations exhibits several nice properties while minimizing the distortionary incentive effect for the high type; in fact, the distortion disappears altogether in the continuous-time limit.

4   On the failure of the R&D market

We now examine a fully dynamic version of the model where the agent needs to be incentivized continuously over time. In this section, we analyze the model with no subsidies, i.e., $\pi_t^y = \eta^y$, $y \in Y$, for all t, to show once again that the R&D market can break down rather starkly under some feasible conditions. The underlying reasoning is basically the same as in the static case but is now augmented with Bayesian belief updating.

---

[11]Throughout the analysis, we assume that only the first success is costly to achieve. Since the agent is guaranteed to succeed in S at a first try, the total cost of working on S is $c^S \Delta$. As such, the cost converges to zero in the continuous-time limit for any c^S.

4.1 Equilibrium investment flows

In the dynamic setting, the market observes a sequence of outcomes and continuously updates its belief at each instance. Formally, let $h_t := (y_0, y_\Delta, ..., y_{t-\Delta}) \in H_t$ denote the history of the game at time t with $h_0 = \varnothing$, and $H := \bigcup_{n=0}^{\infty} H_{n\Delta}$ the set of all histories. We denote by $p_t = p(h_t)$ the market belief at time t, which is defined as the probability that the agent is of the high type conditional on h_t.

Given the current belief, the market offers the investment flow w_t for the period $[t, t+\Delta)$. Under perfect competition, the market is forced to offer the amount that equals the agent's expected productivity. Letting $\sigma_t^a(x)$ be the behavior strategy of type a, the equilibrium investment flow is obtained as

$$w_t = p_t\big(\sigma_t^h(R)g\pi_t^R + \sigma_t^h(S)\pi_t^S\big) + (1 - p_t)\big(\sigma_t^l(R)g\pi_t^R + \sigma_t^l(S)\pi_t^S\big).$$

taking π_t^R and π_t^S as given.

4.2 Continuation equilibria with no reputation concerns

The amount of information revealed by each innovation outcome depends on the strategies of both types. In some cases, an innovation outcome can fully reveal the agent's type, i.e., either $p_t = 0$ or $p_t = 1$, after which the agent no longer has any reputation concerns. With no reputation concerns, the agent's problem becomes essentially static, and the continuation equilibria in these contingencies can be straightforwardly characterized as below.

- If $p_t = 0$, it is (at least weakly) optimal for the low type to choose S. We suppose that the low type always chooses S thereafter. Let V_t^S denote the continuation payoff in this contingency which emerges in equilibrium when the low type achieves a minor innovation.

- If $p_t = 1$, it is strictly optimal for the high type to choose R. We suppose that the high type always chooses R thereafter. Let V_t^R denote the continuation payoff in this contingency which emerges in equilibrium when the high type achieves a major innovation.

4.3 Project choice with reputation concerns

When $p_t \in (0, 1)$, the agent must take into account how his action affects his reputation. Define $p_{t+\Delta}^y(p)$, or simply $p_{t+\Delta}^y$, as the next-period belief when $p_t = p \in (0, 1)$ and $y_t = y$. We also denote by $U_t^a(p)$ the continuation payoff for type a as a function of the current

belief.[12] Then, given p_t and $w_t \geq c$, if the agent chooses R in period t, the expected payoff is given by

$$(w_t - c)\Delta + (1 - r\Delta)\big((1 - g^a\Delta)U_{t+\Delta}^a(p_{t+\Delta}^\varnothing) + g^a\Delta V_t^R\big). \tag{1}$$

Similarly, if the agent chooses S and N, the expected payoffs are

$$w_t\Delta + (1 - r\Delta)U_{t+\Delta}^a(p_{t+\Delta}^S), \tag{2}$$

and

$$w_t\Delta + (1 - r\Delta)U_{t+\Delta}^a(p_{t+\Delta}^\varnothing), \tag{3}$$

respectively.

Observe that since $g^l = 0$, N strictly dominates R for the low type, so that he would never choose R under any circumstance. Similarly, from (1) and (3), N strictly dominates R for the high type if

$$U_{t+\Delta}^h(p_{t+\Delta}^\varnothing) > \frac{(1 - r\Delta)gV_t^R - c}{(1 - r\Delta)g}. \tag{4}$$

This condition always holds in the limit if $c > gV_t^R \geq \frac{g^2v}{r}$, in which case there is no feasible way to let the high type take on R.[13] Since this is clearly uninteresting for the purpose of our study, we assume that the value of a major innovation is high enough for the high type to choose R when the continuation payoff is sufficiently small.

Assumption 3 $\frac{g^2v}{r} \geq c$.

4.4 Equilibrium with no subsidies

The essential aspect of the current setting is that the agent faces a resource constraint and needs to obtain external resources to carry out R&D activities. Because of this feature, we may have an equilibrium in which the principal offers $w_t = 0$ and the agent has no choice but to choose $x_t = N$, which we call trivial in what follows. Note that the trivial equilibrium is a form of pooling equilibrium which by itself provides no additional information to the market. As long as the trivial equilibrium prevails, therefore, the market belief stays constant over time, which bounds the investment flow from above. In particular, when the initial prior p_0 is not high enough to begin with, the investment flow is perpetually kept below the requisite level and only the trivial equilibrium can emerge as a consequence.

---

[12]In principle, the continuation payoff may depend not only on the current belief but also on the entire history of the game. In what follows, however, we only look at cases where the continuation payoff depends only on the current belief for a given t.

[13]When $p_t = 1$, the high type always chooses R and hence $w_t = gv$ when no subsidies are provided. This implies that $V_t^R \geq \frac{gv}{r}$

Proposition 1 *If $c > p_t g v$, the only continuation equilibrium is the repetition of the trivial equilibrium, i.e., $w_s = 0$ and both types choose $x_s = N$ for all $s \geq t$.*

PROOF: When only major innovations generate value as in this setup, the equilibrium investment flow w_t cannot exceed $p_t g v$. This implies that $c > p_t g v \geq w_0$, so that the investment flow cannot be high enough to support R. Given this, the agent can choose either S or N, but in either case, his expected productivity is zero by assumption. The principal thus offers $w_t = 0$, and the agent has no choice but to choose N as a consequence. Since no additional information is revealed in the trivial equilibrium, we have $p_{t+\Delta} = p_t$. As this argument applies to all the subsequent periods, the only possible equilibrium is the repetition of the trivial equilibrium where $w_t = s$ and both types choose $x_s = N$ for all $s \geq t$.∎

The R&D market failure is most severe when the initial prior belief is not high enough to support R from the beginning, i.e., $c > p_0 g v$. Even when the condition does not hold at time 0, however, the problem can still persist. Suppose that $p_0 g v \geq c$ and it is also optimal for the high type to choose R at time 0. Given this, the belief in the next period is given by

$$p_\Delta = \frac{(1 - g\Delta)p_0}{1 - p_0 + (1 - g\Delta)p_0},$$

which is evidently lower than p_0. As can be seen from this, the belief monotonically declines over time as long as the high type chooses R (while the low type chooses N), and any R&D activities are bound to disappear sooner or later.

This simple argument suggests that the R&D market can break down rather starkly due to a version of the lemons problem when only major innovations are valuable and hence rewarded. One way to solve this dilemma is then to artificially create value for innovations, be it major or minor, via "subsidies" so as to raise the agent's expected market value. In the next section, we in particular argue that a simple subsidy scheme – namely the one that exclusively subsidizes minor innovations – can alter the equilibrium allocation and correct this market failure to a considerable extent. To emphasize this point, we focus on a situation where the initial prior is not high enough to support R at the outset, so that the R&D market would break down in the most severe manner without any intervention.

Assumption 4 $g v > c > p_0 g v$.

5   Subsidizing minor innovations

We are now ready to evaluate the consequences of rewarding minor innovations above their social value in the environment described above. To make our contention, however, we take

an approach from the reverse direction: instead of directly evaluating the welfare impact of rewarding minor innovations, we first identify a subsidy scheme that can implement the second-best allocation, and then show that one of those schemes is a simple scheme that subsidizes minor innovations at a constant rate.[14] From this result, we argue that the reward compression, which appears to overcompensate minor innovations, is not so harmful after all and can even be welfare-improving under some plausible conditions.

5.1 The social Planner's problem

We consider a social planner who can commit to a subsidy scheme at the outset of the game. There is a transfer (or taxation) cost in that the social planner incurs a cost $\theta > 0$ for each dollar transferred. The social welfare in this context is then defined as the discounted sum of values generated from R (hereafter, simply the value of innovation) net of the discounted sum of transfer costs (the transfer cost). A feasible subsidy scheme is a pair (b^S, b^R) where $b^y : H \to \mathbb{R}_+$ maps a current history to a real-valued transfer for each $y = S, R$.

The social planner's objective is to device a subsidy scheme which maximizes the social welfare. To this end, it is important to note that although the whole set of histories can be quite large, some of the histories are actually irrelevant for the purpose of our study and can be excluded from consideration without loss of generality. This fact allows us to reduce the domain and simplify the problem to a large extent. Below, we first comment on the relevant domain of the social planner's problem which we consider in the subsequent analysis.

First, it is often convenient to divide our game into two distinct stages – before and after a major innovation is achieved – as the problems in these stages are quite different and in fact independent from each other. As we have seen above, once a major innovation is achieved, the agent's type is fully revealed, and the high type chooses R indefinitely even without any subsidies. This indicates that there is no feasible way to improve the social welfare via subsidies once the game enters into this post-innovation stage. For this reason, when deriving an optimal scheme, we suppose that no subsidies are provided once a major innovation is achieved and hence $V_t^R = V^R := \frac{gv}{r}$ for all t.[15]

---

[14]Note that we do not consider direct revelation mechanisms because our goal here is not to identify a class of optimal mechanisms *per se*, but rather to evaluate the efficiency properties of a particular compensation practice. This draws contrast to Halac et al. (2016) who analyze optimal contracts for experimentation with hidden information about the agent's ability type.

[15]Of course, there is a possibility that rewarding subsequent major innovations further enhances welfare because a higher post-innovation rent V^R can provide a stronger incentive for the high type to choose R (for $h_t \in \overline{H}_t$), thereby allowing the social planner to extend the contract duration. However, given that our aim is to make an argument for why it makes sense to reward minor innovations, this possibility is an issue of independent interest: any subsidy in the post-innovation stage can affect the equilibrium allocation only through its effect on V^R; as such, V^R is a sufficient statistic of whatever happens in this stage. Since our argument holds for any V^R that satisfies the maintained assumptions, if it is ever optimal to provide

Second, since the expected productivity of an agent who has achieved no major innovation gradually declines over time, it must become optimal at some point not to support R any more. This implies that we can focus on schemes which terminate and provide no subsidies after some finite time.[16] Let T denote such a termination date (which we also interchangeably refer to as the contract duration) after which no subsidies are offered, i.e., the agent must achieve a major innovation until time T or otherwise he will receive a continuation payoff of zero after that.

Given these restrictions, we can now define

$$\overline{H}_t := \{h_t : y_s \neq R \text{ for all } s < t\}.$$

as the set of histories in which no major innovation has yet been achieved, and

$$\overline{H}^T := \bigcup_{n=0}^{\frac{T}{\Delta}} \overline{H}_{n\Delta}.$$

as the set of all such histories up to time T. Then, the above restrictions are formally equivalent to $b^S(h) = b^R(h) = 0$ for all $h \notin \overline{H}^T$, and hence $\overline{H}^T$ constitutes the relevant domain of our problem. We say that a subsidy scheme is *optimal* if it implements the second-best allocation, i.e., the one that maximizes the social welfare among all feasible schemes on the relevant domain.

5.2 The high type's optimal strategy

For any $h_t \in \overline{H}_t$, given $w_t \geq c$, the high type chooses R over N only if

$$(w_t - c)\Delta + (1 - r\Delta)\big(g\Delta V^R + (1 - g\Delta)U^h_{t+\Delta}(p^\varnothing_{t+\Delta})\big) \geq w_t\Delta + (1 - r\Delta)U^h_{t+\Delta}(p^\varnothing_{t+\Delta}),$$

which is simplified to

$$(1 - r\Delta)g\big(V^R - U^h_{t+\Delta}(p^\varnothing_{t+\Delta})\big) \geq c.$$

By rearranging and taking the limit, we obtain

$$\frac{gv}{r} - \frac{c}{g} \geq U^h_{t+\Delta}(p^\varnothing_{t+\Delta}), \tag{5}$$

additional incentives by subsidizing subsequent major innovations to raise V^R, the social planner can do so while maintaining the elements of the optimal scheme we will discuss below. Alternatively, our analysis can be seen as analyzing a situation where V^R is already optimally set, possibly by subsidies to subsequent major innovations.

[16]Although it is in principle possible to design an optimal scheme which provides subsidies indefinitely over time, we can always rewrite this into a scheme which terminates at some termination date since only thing that matters is the continuation payoff. It is hence without loss of generality to focus on this class of schemes.

where we often refer to the left-hand side of (5) as the static gain from working on R. The key to this dynamic incentive structure is that a higher continuation payoff diminishes the current incentive to take on R, implying that w_t should be kept as small as possible while satisfying the resource constraint when necessary.

Suppose for now that there is some exogenously given termination date T and moreover that the high type always chooses R and $w_t = c$ for any history. Under these conditions, the continuation payoff at any time $t < T$ converges to

$$\int_0^{T-t} \frac{g^2 v}{r} e^{-(r+g)s} ds = gv\Big(\frac{1}{r} - \frac{1}{r+g}\Big)\big(1 - e^{-(r+g)(T-t)}\big).$$

From this, (5) can be written as

$$\frac{gv}{r} - \frac{c}{g} > gv\Big(\frac{1}{r} - \frac{1}{r+g}\Big)\big(1 - e^{-(r+g)(T-t)}\big). \tag{6}$$

Among other things, the high type's current incentive to work on R depends crucially on $T - t$, i.e., the remaining time the agent has until the termination date. If $c > \frac{g^2 v}{r+g}$, we can define $\overline{T}$ such that

$$\frac{gv}{r} - \frac{c}{g} = gv\Big(\frac{1}{r} - \frac{1}{r+g}\Big)\big(1 - e^{-(r+g)\overline{T}}\big), \tag{7}$$

from which we obtain

$$\overline{T} = \frac{1}{r+g} \ln \frac{g^3 v}{r\big((r+g)c - g^2 v\big)},$$

Since the high type's incentive to work on R is maximized when $w_t = c$ as long as he fails, this is the upperbound of the contract duration in which the high type can be induced to work on R for all $t \in [0, T]$. If $\frac{g^2 v}{r+g} \geq c$, on the other hand, (6) can be satisfied for any remaining time, and we let $\overline{T} = \infty$ in this case. The upper bound plays a crucial role in shaping the optimal scheme because there exists no equilibrium in which the high type chooses R for all $t \in [0, T]$ if T is set to be larger than $\overline{T}$. The optimal design of the subsidy scheme thus eventually comes down to how much to subsidize each innovation and for how long.

5.3 Optimal indirect scheme

With a termination date and time discounting, we can show that the second-best allocation should have effort spread out evenly over the entire contract duration. In fact, in the second-best allocation to be implemented, the high type chooses R with some constant probability $k \in (0, 1]$ for all $t \in [0, T]$; the implemented equilibrium is hence characterized thoroughly by a pair (k, T). For clarity, we say that an equilibrium is *full-funding* if $k = 1$ and *partial-funding* if $k < 1$.

To illustrate how to pin down the optimal contract duration, consider a full-funding equilibrium. In this equilibrium, conditional on no major innovation having occurred, the (limit) probability that the agent is of the high type is given by $\overline{p}_t := \frac{p_0 e^{-gt}}{1-p_0+p_0 e^{-gt}}$ for $t \leq T$.[17] If the contract duration is extended for one more period at time T, the agent generates a major innovation with probability $\overline{p}_T g\Delta$, and the expected value generated is $\overline{p}_T g\Delta(v + (1 - r\Delta)V^R)$.[18] In the meanwhile, the market needs to pay $c\Delta$ to the agent, so that the amount of subsidy needed is $c\Delta - \overline{p}_T gv$. This cost and benefit must be balanced at the optimal contract duration, i.e.,

$$\overline{p}_T g\big(v + (1 - r\Delta)V^R\big) = \theta(c - \overline{p}_T gv).$$

whose continuous-time limit is given by

$$gv\Big(1 + \theta + \frac{g}{r}\Big)\frac{p_0 e^{-gT}}{1 - p_0 + p_0 e^{-gT}} = \theta c. \tag{8}$$

Let $\widehat{T}$ denote the solution to (8) which is given by

$$\widehat{T} := \frac{1}{g}\left(\ln\frac{p_0}{1 - p_0} + \ln\frac{gv(1 + \theta + \frac{g}{r}) - \theta c}{\theta c}\right),$$

which is the break-even point under the presumption that the high type always chooses R.

Note that if θ is too large, it is clearly optimal to provide zero subsidies, and no social surplus is generated under the maintained assumptions.[19] As this only results in a trivial solution, we assume that θ is sufficiently small, so that $\widehat{T} > 0$ and the socially optimal allocation involves strictly positive subsidies.

Assumption 5 $p_0 gv > \frac{\theta c}{1+\theta+\frac{g}{v}}$.

For expositional purposes, we say that a subsidy scheme is *history-independent* if $b^y(h) = b^y$ for all $h \in \overline{H}^T$. The next proposition is the main result of this paper, which states that the second-best allocation can be implemented by a simple history-independent scheme that exclusively subsidizes minor innovations at a constant rate.

Proposition 2 *Let $\tilde{k}$ be a solution to $\int_0^{\frac{\widehat{T}}{\tilde{k}}} \frac{\tilde{k}g^2 v}{r}e^{-(r+\tilde{k}g)t}dt = \frac{gv}{r} - \frac{c}{g}$. Then, the following history-independent scheme is optimal: for all $h \in \overline{H}^{T^*}$,*

$$b^R(h) = 0, \ b^S(h) = k^*c, \tag{9}$$

---

[17]Note that $\overline{p}_t$ is to be distinguished from the belief p_t since $\overline{H}_t$ contains a history where the agent has achieved a minor innovation, in which case $p_t = 0$.

[18]After the termination date, only the agent who has achieved a major innovation can continue to work, and the continuation value is precisely V^R.

[19]As we will see below, if Assumption 5 fails to hold, it is optimal to set $T = 0$. Given no subsidies, the agent has no choice but to choose N for all t as we have seen above.

where

- *if $\overline{T} \geq \widehat{T}$, $k^* = 1$ and $T^* = \widehat{T}$;*

- *if $\widehat{T} > \overline{T}$, $k^* = \tilde{k}$ and $T^* = \frac{\widehat{T}}{k}$.*

In the implemented equilibrium, the high type chooses R with probability k^* for all $t \in [0, T^*]$.

PROOF: See Appendix A.

In words, the optimal scheme subsidizes k^*c for each minor innovation until time T^*, given that no major innovation has been achieved. For expositional purposes, we refer to this as an indirect scheme as it attempts to induce major innovations indirectly by subsidizing minor innovations. The optimal indirect scheme implements a full-funding equilibrium in which the high type is induced to choose R for all $t \in [0, T^*]$ if $\overline{T} > \widehat{T}$, i.e.,

$$\frac{1}{r+g} \ln \frac{g^3 v}{r\big((r+g)c - g^2 v\big)} \geq \frac{1}{g}\left(\ln \frac{p_0}{1-p_0} + \ln \frac{gv(1 + \theta + \frac{g}{r}) - \theta c}{\theta c} \right). \tag{10}$$

A key determinant of k^* is the marginal transfer cost θ. Under the maintained assumptions, there exists some $\overline{\theta}$ such that (10) holds for $\theta \geq \overline{\theta}$.

It is worth pointing out that the equilibrium implemented by this optimal scheme exhibits unique reputation dynamics: on the equilibrium path, the market belief declines to zero as soon as the agent succeeds in S whereas it is strictly positive when he has achieved nothing. That is, throughout the contract duration, the agent actually loses reputation by achieving a minor innovation, which may capture a pervasive view in academia that "no publications are better than a few mediocre publications" at an early stage of one's career. This rather peculiar property of the reputation dynamics is a reflection of the fact that exploration of new ideas is inherently time-consuming and requires persistent effort that pays off only randomly over time. Most notably on this point, March (1991) notes that "returns from exploration are systematically less certain, more remote in time, and organizationally more distant from the locus of action and adaptation," and "the search for new ideas, markets, or relations has less certain outcomes, longer time horizons, and more diffuse effects than does further development of existing ones." The time-consuming and uncertain nature of exploratory activities is a crucial driving force which determines the structure of innovations, both in our model and in reality.

5.4 Discussion: how and why the indirect scheme works

Proposition 2 states that the second-best allocation can be implemented by a simple stationary scheme where the subsidy rate is held constant for $t \in [0, T]$. The reason why this simple

scheme works is because it induces self-sorting among low-type agents to sustain the market belief constant over time. The key mechanism at work is that the low type is held indifferent between S and N throughout the contract duration: if the high type chooses R, the low type can separate from the high type by choosing S, in which case he chooses S and receives $b^S(h) = c$ until time T; if he instead chooses N, he partially pools with the high type and receives $w_t = c$. On the equilibrium path, therefore, the low type randomizes between S and N to keep the expected productivity invariably at c for any history. This means that at any $t \in [0, T^*)$, we must have

$$p_t g v + (1 - p_t) \sigma_t^l(S) c = c \Leftrightarrow \sigma_t^l(S) = \frac{c - p_t g v}{(1 - p_t) c},$$

taking p_t as given. The belief in the next period is then obtained as

$$p_{t+\Delta} = \frac{p_t(1 - g\Delta)}{(1 - p_t)\frac{p_t(gv-c)}{(1-p_t)c} + p_t(1 - g\Delta)} = \frac{(1 - g\Delta)c}{g(v - c\Delta)},$$

which is constant over time and converges to $\frac{c}{gv}$ as $\Delta \to 0$.

An important virtue of the proposed scheme is that it has no distortionary incentive effect for the high type, which draws clear contrast with the static counterpart where rewarding minor innovations necessarily distorts the high type's incentives. The key difference is that the agent needs to be incentivized continuously over time in the dynamic setting. In the second-best allocation, the investment flow must equal c whenever the high type is supposed to choose R, implying that the agent can secure at least a payoff of c in this contingency. As is common in this type of dynamic setting, the presence of this dynamic rent inevitably weakens the current incentive to work hard because even if the agent fails today, he can still have another chance tomorrow.

The proposed scheme is designed to exploit this property, using this dynamic rent to induce the low type to occasionally choose S and "exit" from the market. To see this, observe that the high type chooses R over S if

$$(w_t - c)\Delta + (1 - r\Delta)\big(g\Delta V^R + (1 - g\Delta)U_{t+\Delta}^h(p_{t+\Delta}^\varnothing)\big) \geq w_t\Delta + (1 - r\Delta)U_{t+\Delta}^h(p_{t+\Delta}^S).$$

Given that (5) holds, this condition is necessarily satisfied if $U_{t+\Delta}^h(p_{t+\Delta}^\varnothing) \geq U_{t+\Delta}^h(p_{t+\Delta}^S)$, which suggests that rewarding minor innovations is entirely harmless as long as the continuation payoff of choosing S does not exceed that of choosing N. Now suppose that the high type deviates and chooses S. In this contingency, the market falsely updates its belief to $p_t = 0$, but still offers $w_t = c$ under the proposed scheme, thereby giving him exactly the same continuation payoff after the deviation, i.e., $U_{t+\Delta}^h(p_{t+\Delta}^\varnothing) = U_{t+\Delta}^h(p_{t+\Delta}^S)$. As such, the incentive to choose S is no stronger than the incentive to choose N, and there is hence no incentive distortion from subsidizing minor innovations.

19

5.5   Subsidizing major innovations?

In the current setting, there actually exist other subsidy schemes that can equivalently implement the second-best allocation. In fact, it is possible to construct an optimal scheme that exclusively subsidies major innovations as in the static benchmark; for expositional purposes, we call this a direct scheme as it subsidizes major innovations directly to induce major innovations. Here, we briefly discuss this possibility and build on this argument to make a case for why rewarding minor innovations is superior to rewarding major innovations directly.

Consider an alternative scheme in which $b^S(h) = 0$ for all $h \in \overline{H}^T$. In this hypothetical environment where only major innovations are valuable and hence rewarded, the low type clearly has no incentive to choose S. Since no minor innovations are observed on the equilibrium path under this scheme, we define

$$\underline{H}_t := \{h_t : y_s = \varnothing \text{ for all } s < t\} \subset \overline{H}_t, \ \underline{H}^T := \bigcup_{n=0}^{\frac{T}{\Delta}} \underline{H}_{n\Delta}.$$

Given some history $h_t \in \underline{H}_t$ and the corresponding belief $p_t = p(h_t)$, the investment flow at time t is then obtained as

$$w_t = p_t g\bigl(v + b^R(h_t)\bigr),$$

which must be at least as large as c to induce R from the high type. Since the low type chooses N for all t, the market belief p_t must decline monotonically over time. This immediately suggests that no history-independent scheme that exclusively subsidizes major innovations can replicate the second-best allocation, as it necessarily implies $w_t > c$ for $t \in [0, T)$ if it is to induce R for all $t \in [0, T]$. To keep the investment flow at c, the subsidy rate must then increase over time to offset a decrease in the market belief.

Proposition 3 *The following direct scheme is optimal:*

$$b^S(h) = 0 \text{ for } h \in \overline{H}^{T^*}, \ b^R(h) = \begin{cases} \frac{c}{p(h)g} - v & \text{for } h \in \underline{H}^{T^*}, \\ 0 & \text{for } h \notin \underline{H}^{T^*}, \end{cases}$$

where T^* is determined as in Proposition 2.

PROOF: Under the proposed scheme, the low type earns zero payoff regardless of his action and hence chooses N with probability one. The expected productivity of the agent is then invariably c for any given p_t. Now consider the same offer process as in Proposition 2 which offers $w_t = c$ with probability k and $w_t = 0$ with the remaining probability. If k and T are set appropriately, the high type then chooses R whenever $w_t = c$. Moreover, since the

amount of subsidy is exactly the same as under the optimal indirect scheme, this indeed implements the second-best allocation. ∎

The optimal direct scheme generates major innovations at the same rate while requiring the same amount of subsidy as does the optimal indirect scheme. Both of these schemes hence implement the second-best allocation and are equivalent in terms of social welfare. We argue, however, that there are reasons to believe that the indirect scheme is superior to the direct scheme in a practical sense. Below, we raise three major reasons for why it makes more sense to reward minor innovations to induce major innovations.

Simplicity: A major advantage of the indirect scheme is that it is history-independent and hence substantially easier to implement. In contrast, in the optimal direct scheme, the subsidy rate depends on the market belief and must increase over time at a rate which depends on the minute details of the underlying environment. We argue that this virtue of simplicity is particularly important when we attempt to regulate a large and diverse community, such as academia, where it is prohibitively costly to tailor an elaborated scheme for each individual and the class of enforceable schemes can be quite limited.

Manipulability: The more elaborated an incentive scheme becomes, the easier it is to manipulate the scheme, as illustrated most notably by Holmström and Milgrom (1987). This reasoning also applies to our direct scheme: the subsidy rate that monotonically increases over time can give rise to an incentive to game the system if outside parties cannot observe when a major innovation is achieved. To see this possibility in more detail, suppose that the market (which should now be interpreted as an individual employer) is now also a strategic player, and further that the arrival of a major innovation is observed only privately by the market and the agent but not by the social planner. Upon the arrival of a major innovation, the market can then choose when to "publicize" it to maximize the joint payoff of the market and the agent. Suppose that a major innovation is achieved at some time t. If it is disclosed immediately, the joint continuation payoff is $b_t^R + v + V^R$; if disclosed at the next instance $t + \Delta$, it is $(1 - r\Delta)(b_{t+\Delta}^R + v + V^R)$. It is then optimal for the principal to wait if

$$\frac{b_{t+\Delta}^R - b_t^R}{\Delta} \geq r(b_{t+\Delta}^R + v + V^R).$$

Letting $\Delta \to 0$, we obtain

$$\dot{b}_t^R = -\frac{c}{p_t^2 g}\dot{p}_t = \frac{(1-p_0)c}{p_0 e^{-gt}} \geq r(b_t^R + v + V^R),$$

which holds if r is small relative to g. Note that no such incentive arises under the indirect scheme since the value of each major innovation stays constant over time.

21

Externalities: We have thus far assumed that minor innovations are intrinsically valueless to make our conclusion more emphatic. In all likelihood, however, it is a little too extreme to assume that those minor innovations offer absolutely zero value to society. This is particularly true when they yield some external effects. For instance, it is reasonable to argue that even when minor innovations per se are useless, they may still raise the overall productivity due to knowledge spillovers: after all, a situation where a sizable group of researchers work hard at generating something new is more conducive to innovations than to a situation where the vast majority simply slack off and do nothing. If minor innovations provide positive externalities even slightly, the indirect scheme becomes strictly superior to the direct scheme that can induce effort only from a limited fraction of researchers.

Before we conclude, it should be emphasized that the aim of our analysis is to provide a stylized description of when and under what conditions rewarding minor innovations entails less incentive distortion, and we make some strong assumptions to make this point clearer. Our theoretical results should thus be regarded as more of a theoretical benchmark, for they hold only under a set of stringent conditions. In particular, when we consider a more realistic case with $c^S > 0$, the assumption that the total cost of implementing S, as captured by $c^S \Delta$, vanishes to zero in the limit plays a crucial role for our indirect scheme to be (approximately) optimal.[20] Needless to say, this is a very strong assumption which is made strictly for expositional clarity rather than for realism.

In contrast, when $c^S \Delta$ is bounded away from zero, the indirect scheme entails an additional social cost as it induces the low type to waste resources on the inefficient safe project. As we have argued, however, our argument still holds true in a qualitative sense even in this case: rewarding minor innovations induces self-sorting among low-type agents, thereby sustaining the market belief within some tight bounds, while its distortionary incentive effort becomes less attenuated when the agent needs to be incentivized continuously over time. In this more realistic scenario, this efficiency loss must be traded off against the aforementioned benefits; obviously, the benefits dominate the cost when $c^S \Delta$ is sufficiently small.

6 Conclusion

This paper illustrates how an R&D market can fail due to a version of the lemons problem and how optimally R&D activities can be subsidized to overcome this problem. We argue that

---

[20]Under the assumption that only the first success is costly, the total cost the agent needs to incur is $c^S \Delta$ even if he works on S for all t. Even when we suppose $c^S > 0$, the optimal indirect scheme we identified can approximately implement the second-best allocation as $\Delta \to 0$. See Appendix B for more detail.

there are substantial efficiency gains from rewarding minor innovations above their market value, and further that it is superior to rewarding major innovations directly even when minor innovations *per se* are intrinsically valueless. Rewarding minor innovations induces self-sorting among less productive agents which is essential in ensuring more time and resources for more productive ones to try out riskier but potentially more valuable approaches. Our argument implies that the current compensation practice, which appears to overcompensate minor innovations, is not necessarily welfare-reducing and can even be conducive to major innovations when dynamic incentives are explicitly taken into account.

As a final note, we consider a pure signaling context where the agent knows his own type precisely throughout the analysis. A possible extension of this setting is therefore to incorporate the agent's learning about his own type. Although the extension clearly involves technical issues of independent interest, the same economic principle should apply in this case as well: less confident agents eventually choose the safe project and drop out when the expected return of doing so is sufficiently high; to induce this self-sorting, minor innovations must be rewarded sufficiently to compensate for the loss of reputation. The essential element of our story is hence that the agent has some private information, regardless of whether his information is perfect or not. Nevertheless, since both elements of learning and signaling must be present in reality, it is of some interest to consider the possibility of imperfect information in this type of environment.

Appendix A: proof of Proposition 2

To characterize the optimal scheme, we first take the contract duration T as given and, for each T, identify subsidy schemes that maximize the value of innovation, which we refer to as surplus-maximizing for clarity. Among all those surplus-maximizing schemes, we then identify the one that entails the minimum transfer cost. Since there is at least one such scheme for each T, this gives us a set of candidate schemes across all T. Identifying this set is an important intermediate step because the optimal scheme by definition must exist within this class.

First, if $\overline{T} > T$, it is possible to construct a full-funding equilibrium. To do this, the expected productivity of the agent must be c for any history in $\overline{H}^T$. Assumption 4 ($c > p_0 g v$) then implies that the low type must be induced to separate from the high type by choosing S with some probability. Now consider a scheme with $k = 1$. Under this scheme, the low type always receives c until time T regardless of his strategy: even if he chooses S and separates from the high type, he continues to receive $w_t = b_t^S = c$. The low type is therefore indifferent between N and S for the entire contract duration under the proposed scheme, and $w_t = c$

can be sustained by having the low type randomize between N and S. More precisely, we must have

$$p_t g v + (1 - p_t)\sigma_t^l(S)c = c \iff \sigma_t^l(S) = \frac{c - p_t g v}{(1 - p_t)c},$$

taking p_t as given. Given that $w_t = c$ and $\frac{g^2 v}{r+g} \geq c$, the high type prefers R over N as we have already seen. It is also easy to see that $u_{t+\Delta}^h(p_{t+\Delta}^\emptyset) > u_{t+\Delta}^h(p_{t+\Delta}^S)$, suggesting that it is optimal for the high type to work on R as long as $b^S(h) = c$ for $h \in \overline{H}^T$.

The situation becomes more complicated when the contract duration exceeds the upperbound $\overline{T}$, in which case it is no longer feasible to implement a full-funding equilibrium: even if $w_t \geq c$ for all $t \in [0, T]$, the high type has no incentive to choose R early on when the contract is still far from termination. Since the social planner wants to induce R earlier than later due to time discounting, a solution in this case is to provide only partial funding in order to keep the continuation payoff low enough. More precisely, it is essential to keep the continuation payoff weakly below $\frac{gv}{r} - \frac{c}{g}$ throughout the contract duration. The following statement derives the optimal scheme when the termination date exceeds the upperbound, which implements a partial-funding equilibrium where the high type is induced to work on R only for some fraction of time.

Lemma 1 *Let $\hat{k}_T$ be a solution to $\int_0^T \frac{\hat{k}_T g^2 v}{r} e^{-(r + \hat{k}_T g)t} dt = \frac{gv}{r} - \frac{c}{g}$. Then, for a given $T > \overline{T}$, the following scheme minimizes the transfer cost among all surplus-maximizing schemes: for all $h \in \overline{H}^T$,*

$$b^R(h) = 0, \ b^S(h) = \hat{k}_T c. \tag{11}$$

PROOF: The proposed scheme implements an equilibrium in which the high type chooses R with probability $\hat{k}_T$ over the whole contract duration. For the sake of the argument, we think of the market as a non-strategic offer process which gives $w_t = c$ with probability $\hat{k}_T$ and $w_t = 0$ with the remaining probability.[21] Now suppose that the high type chooses R whenever $w_t \geq c$. Given the offer process, his continuation payoff in the limit is then given by

$$U_t^h(p_t) = \int_0^{T-t} \frac{\hat{k}_T g^2 v}{r} e^{-(r + \hat{k}_T g)s} ds < \frac{gv}{r} - \frac{c}{g},$$

---

[21] We can think of this as an equilibrium outcome of some strategic bidding process when we consider a competitive market of potential employers. To see this, note that the current investment flow has no incentive effect in this setting. The only difference it makes is when the investment flow is so low that the resource constraint binds, in which case some employer may induce the high type to choose R by offering $w_t \geq c$. However, the expected productivity would be c in this equilibrium, so that employers are indifferent between offering $w_t = c$ and $w_t = 0$ and hence have no strict incentive to deviate.

for any $t > 0$. Therefore, by (5), the high type will choose R over N whenever $w_t \geq c$. Note also that if the high type ever deviates by choosing S, he will simply receive $\hat{k}_T c$ (by working on S), so the continuation payoff converges to the level when he repeatedly chooses N until period T. As this yields a strictly lower payoff, the high type chooses R whenever $w_t \geq c$ and N otherwise.

The low type's optimal strategy is effectively the same with a minor modification. Again, the low type is indifferent between pooling with the high type and separating from the high type (by choosing S): in any contingency, the expected per-period payoff for the low type is $\hat{k}_T c$. The only difference is that we now have some time intervals with $w_t = 0$, in which case the low type chooses N to pool with the high type. On the other hand, when $w_t = c$, the low type adopts a similar mixed strategy to keep the expected productivity of the agent at c, i.e.,

$$p_t g v + (1 - p_t)\sigma_t^l(S)\hat{k}_T c = c \iff \sigma_t^l(S) = \frac{c\Delta - p_t g v}{(1 - p_t)\hat{k}_T c},$$

taking p_t as given. This shows that the proposed subsidy scheme, along with the offer process, implements a partial-funding equilibrium in which the high type chooses R with probability $\hat{k}_T$.

Given this, we now show that the implemented partial-funding equilibrium is surplus-maximizing for a given $T > \overline{T}$. To see this, note that since the high type appropriates a constant fraction $\frac{\frac{gv}{r}}{v + \frac{gv}{r}} = \frac{g}{r+g}$ of the value of innovation, the problem is equivalent to maximizing the high type's expected payoff in period 0. Given $T > \overline{T}$, once the continuation payoff goes above $\frac{gv}{r} - \frac{c}{g}$, the high type would not choose R and no additional surplus could be generated. This means that even if the continuation payoff goes above $\frac{gv}{r} - \frac{c}{g}$, it does not contribute to increasing the value of innovation; as such, the continuation payoff must be bounded by $\frac{gv}{r} - \frac{c}{g}$. Since $\int_0^T \frac{\hat{k}_T g^2 v}{r} e^{-(r+\hat{k}_T g)t} dt = \frac{gv}{r} - \frac{c}{g}$ by definition, the implemented equilibrium maximizes the value of innovation.

Finally, we need to show that the proposed scheme minimizes the transfer cost among all surplus-maximizing schemes. To see this, consider an alternative equilibrium in which the high type chooses R with probability κ_t at time t as long as $h_t \in \overline{H}_t$. It follow from the above argument that any surplus-maximizing scheme must satisfy

$$\int_0^T \kappa_t \frac{g^2 v}{r} e^{-(r+\kappa_t g)t} dt = \frac{gv}{r} - \frac{c}{g}$$

$$\Rightarrow \int_0^T \kappa_t e^{-(r+\kappa_t g)t} dt = \int_0^T \hat{k}_T e^{-(r+\hat{k}_T g)t} dt.$$

Since $e^{-\kappa_t t} < e^{-\hat{k}_T t}$ if and only if $\kappa_t > \hat{k}_T$, we have

$$\int_0^T \kappa_t e^{-rt} dt \geq \int_0^T \hat{k}_T e^{-rt} dt. \tag{12}$$

In the equilibrium constructed above with $\hat{k}_T$, the low type's expected payoff is $\int_0^T \hat{k}_T c e^{-rt} dt$; in the alternative equilibrium, the low type's expected payoff is at least $\int_0^T \kappa_t c e^{-rt} dt$, which is larger than $\int_0^T \hat{k}_T c e^{-rt} dt$ by (12). In contrast, these two equilibria yields the same value of innovation and hence the high type's payoff, meaning that the transfer cost is higher in the alternative equilibrium. Since this holds for any nonconstant sequence of κ_t, this proves that the proposed scheme minimizes the transfer cost among all surplus-maximizing scheme.∎

Given the set of candidate schemes, we now simply need to find the optimal contract duration T. The argument is essentially the same as above: the social planner sets T at a point where the marginal cost equals the marginal benefit. It follows from (8) that in the full-funding equilibrium, the belief reaches this point at time $\widehat{T}$. Clearly, if $\overline{T} \geq \widehat{T}$, it is optimal to set $k^* = 1$ and $T^* = \widehat{T}$. If not, the high type cannot be induced to work on R for all $t \in [0, \widehat{T}]$, and it is hence optimal to set $k^* < 1$. Note that if the high type chooses R with probability k, the belief then reaches this point at time $\frac{\widehat{T}}{k}$. The optimal scheme is thus given by $k^* = \tilde{k}$ and $T^* = \frac{\widehat{T}}{\tilde{k}}$ where $\tilde{k}$ solves

$$\int_0^{\frac{\widehat{T}}{\tilde{k}}} \frac{\tilde{k} g^2 v}{r} e^{-(r+\tilde{k}g)t} dt = gv\Big(\frac{1}{r} - \frac{1}{r+\tilde{k}g}\Big)\Big(1 - e^{-(r+\tilde{g})\frac{\widehat{T}}{\tilde{k}}}\Big) = \frac{gv}{r} - \frac{c}{g}.$$

∎

Appendix B: the case with $c^S > 0$

In this appendix, we argue that it is innocuous to assume $c^S = 0$ by showing that as $\Delta \to 0$, there exists a sequence of equilibria converging to the equilibrium in Proposition 2 for any $c^S > 0$. For the appendix, we focus on the case where $\overline{T} \geq \widehat{T}$, so that it is optimal to implement a full-funding equilibrium. The argument can be easily extended to the case of a partial-funding equilibrium.

For some given $\Delta > 0$, we first show how to select $b^S(h) = b$ to implement a full-funding equilibrium. Given $c^S > 0$, the low type (who has never chosen S before) has no incentive to choose S. Suppose that the last period in which the low type chooses S with positive

probability is $T - n\Delta$. In this case, we must have

$$\sum_{i=1}^{n}(1 - r\Delta)^i(b - w_{T-(n-i)\Delta}) = c^S. \tag{13}$$

In a period earlier, the low type also chooses S with positive probability, which means that

$$\sum_{i=1}^{n}(1 - r\Delta)^i(b - w_{T-(n+1-i)\Delta}) = c^S. \tag{14}$$

It follows from these that

$$b - w_{T-n\Delta} = \frac{r\Delta}{1 - r\Delta}\sum_{i=1}^{n}(1 - r\Delta)^i(b - w_{T-(n-i)\Delta}) = \frac{r\Delta}{1 - r\Delta}c^S.$$

Similarly,

$$b - w_t = \frac{r\Delta}{1 - r\Delta}c^S,$$

for $t = \Delta, 2\Delta, \cdots, T - (n+1)\Delta$.

This means that the agent's expected productivity must be $b - \frac{r\Delta}{1-r\Delta}c^S$ for $t = \Delta, 2\Delta, ..., T - n\Delta$. We can then derive a belief p such that

$$pgv + (1 - p)gb = b - \frac{r\Delta}{1 - r\Delta}c^S. \tag{15}$$

The left-hand side is the expected productivity when the low type chooses S with probability $g\Delta$ in every period (until $T - n\Delta$). Given this strategy, the belief stays constant over time for $t \in [\Delta, T - n\Delta]$. At $t = 0$, as p_0 is given and in general $p_0 < p$, the low type needs to choose S with a higher probability to satisfy $p_\Delta = p$ (and hence $w_0 > b - \frac{r\Delta}{1-r\Delta}c^S$. On the other hand, from period $T - (n-1)\Delta$ on, the low type only chooses N. The belief thus p_t goes down until period T, and so does

$$w_t = p_t gv. \tag{16}$$

To induce R from the high type, it is necessary to have $w_t \geq c$ for all $t \in [0, T]$. Since the belief declines monotonically for $t \in [T - (n-1)\Delta, T]$, we let $w_T = c$. Note that for $t \in [T - (n-1)\Delta, T]$, the belief follows

$$p_{t+\Delta} = \frac{p_t(1 - g\Delta)}{1 - p_t g\Delta}.$$

This completely characterizes the dynamic paths of p_t and w_t, from which we can find b and n to satisfy (13) and (15), which we denote $b(\Delta)$ and $n(\Delta)$ respectively.

We now show that the sequence of equilibria constructed above converges to the equilibrium in Proposition 2 as $\Delta \to 0$, i.e., $\lim_{\Delta \to 0} b(\Delta) = b$ and $\lim_{\Delta \to 0} n(\Delta)\Delta = 0$. Suppose on the contrary that $\lim_{\Delta \to 0} n(\Delta)\Delta > 0$. Then, since $w_t < b$ is strictly decreasing for $t \in [T - n(\Delta)\Delta, T]$, the left-hand side of (14) is strictly positive in the limit while the right-hand side converges to zero. As this is a contradiction, $\lim_{\Delta \to 0} n(\Delta) = 0$. It then follows from this that given $w_T = c$, $\lim_{\Delta \to 0} w_{T-n\Delta} = c$, and both w_t and b converge to c by (15).

References

Araujo, A., Gottlieb, D., and Moreira, H., 2007, A Model of Mixed Signals with Applications to Countersignaling, *RAND Journal of Economics*, 38, 1020-43.

Atkeson, A., Hellwig, C., and Ordonez, G., 2015, Optimal Regulation in the Presence of Reputation Concerns, *Quarterly Journal of Economics*, 130, 415-64.

Bar-Isaac, H., 2003, Reputation and Survival: Learning in a Dynamic Signalling Model, *Review of Economic Studies*, 70, 231-51.

Bergemann, D. and Hege, U., 1998, Venture Capital Financing, Learning, and Moral Hazard, *Journal of Banking and Finance*, 22, 703-35.

Bergemann, D. and Hege, U., 2005, The Financing of Innovation: Learning and Stopping, *RAND Journal of Economics*, 36, 719-52.

Bergemann, D. and Välimäki, J., 2000, Experimentation in Markets, *Review of Economic Studies*, 67, 213-34.

Bolton, P. and Harris, C., 1999, Strategic Experimentation, *Econometrica*, 67, 349-74.

Bonatti, A. and Hörner, J., 2011, Collaborating, *American Economic Review*, 101, 632-63.

Chen, C.-H. and Ishida, J., 2015a, Dynamic Performance Evaluation with Deadlines: The Role of Commitment, revised version of ISER Discussion Paper No. 919.

Chen, C.-H. and Ishida, J., 2015b, Hierarchical Experimentation, ISER Discussion Paper No. 949.

Chung, K.-S. and Eso, P., 2013, Persuasion and Learning by Countersignaling, *Economics Letters*, 121, 487-91.

Daley, B. and Green, B., 2012, Waiting for News in the Market for Lemmons, *Econometrica*, 80, 1433-1504.

Ederer, F. and Manso, G., 2013, Is Pay for Performance Detrimental to Innovation? *Management Science*, 59, 1496-1513.

Feltovich, N., Harbaugh, R., and To, T., 2002, Too Cool for School? Signaling and Countersignaling, *RAND Journal of Economics*, 33, 630-49.

Gul, F. and Pesendorfer, W., 2012, The War of Information, *Review of Economic Studies*, 79, 707-34.

Guo, Y., 2016, Dynamic Delegation of Experimentation, *American Economic Review*, 106, 1969-2008.

Halac, M., Kartik, N., and Liu, Q., 2016, Optimal Contracts for Experimentation, *Reiew of Economic Studies*, 83, 1040-91.

Hamermesh, D.S., 2015, Citations in Economics: Measurement, Uses and Impacts, NBER Working Paper 21754.

Holmström, B., and Milrom, P., 1987, Aggregation and Linearity in the Provision of Intertemporal Incentives, *Econometrica*, 55, 303-28.

Hörner, J., 2002, Reputation and Competition, *American Economic Review*, 92, 644-63.

Hörner, J. and Samuelson, L., 2013, Incentives for Experimenting Agents, *RAND Journal of Economics*, 44, 632-63.

Keller G., Rady, S., and Cripps, M., 2005, Strategic Experimentation with Exponential Bandits, *Econometrica*, 73, 39-68.

Klein, N.A. and Rady, S., 2011, Negatively Correlated Bandits, *Review of Economic Studies*, 78, 693-732.

March, J.G., 1991, Exploration and Exploitation in Organizational Learning, *Organizational Science*, 2, 71-87.

Manso, G., 2011, Motivating Innovation, *Journal of Finance*, 66, 1823-69.

Murto, P. and Välimäki, J., 2011, Learning and Information Aggregation in an Exit Game, *Review of Economic Studies*, 78, 1426-61.

Price, D.J. de Solla, 1965, Networks of Scientific Papers, *Science*, 149, 510-15.

Redner, S., 1998, How Popular is Your Paper? An Empirical Analysis of the Citation Distribution, *European Physical Journal B*, 4, 131-34.

Wallace, M.L., Lariviere, V., and Gingras, Y., Modeling a Century of Citation Distributions, *Journal of Informetrics*, 3, 296-303.