Push or Pull? Performance-Pay, Incentives, and Information *

David Rietzke[†] and Yu Chen[‡]

April 4, 2019

Abstract

We study a principal-agent model wherein the agent is better informed of the prospects of the project, and the project requires both an observable and unobservable input. We characterize the optimal contracts, and explore the trade-offs between high and low-powered incentive schemes. We discuss the implications for push and pull programs used to encourage R&D activity, but our results are relevant in other contexts.

KEYWORDS: Pay for Performance, Moral Hazard, Adverse Selection, Observable Action, Principal-Agent Problem
JEL Classifications: D82, D86, O31

---

* A previous working paper was circulated under the title, "Push or pull? Grants, prizes, and information". We thank Stan Reynolds, Andreas Blume, Asaf Plan, John Wooders, Martin Dufwenberg, Rabah Amir, Derek Lemoine, Dakshina De Silva, Brian Roberson, Junichiro Ishida, Matthew Mitchell, Tim Flannery, Dominik Grafenhofer, Dawen Meng, and Bo Chen for helpful insights. We also thank participants at the Lancaster University Conference on Auctions, Competition, Regulation, and Public Policy, EARIE 2015, and the TILEC Conference on Competition, Standardization, and Innovation. Chen gratefully acknowledges financial support from the National Natural Science Foundation of China (Grant No.71673133) and the National Social Science Fund of China (Grant No.16BJL035).

† Department of Economics, Lancaster University,

‡ University of Graz, Austria. Yu Chen unexpectedly passed away in early 2019. He was a great friend and co-author and will be sincerely missed.

1 Introduction

To what extent should incentives be tied to performance? This question is relevant in many areas, including labor markets – where it relates to the debate on salaries vs. piece rates (see, e.g. Lazear, 1986, 2000) – and innovation incentives, where it pertains to the efficacy of "push" and "pull" programs (see, e.g., Kremer, 2002). Push programs, such as research grants or R&D tax credits, subsidize inputs; payments are not contingent on results. Pull programs, such as inducement prizes, or patent buyouts, tie rewards to output.

Adverse selection (AS) and moral hazard (MH) are inherent challenges in incentive provision. Given these problems, Kremer raises the concerns that push programs may finance projects unlikely to succeed, and provide weak incentives for unobservable inputs. Indeed, the MH literature stresses the importance of performance-pay; in the canonical model,[1] compensation must be tied to output to provide an incentive for effort. Yet low-powered incentives in which compensation is weakly, or not at all tied to performance, are commonly used. In this paper, we explore trade-offs between high and low-powered incentives in a model with AS and MH. We show that performance-pay may not be optimal for all types, but is always optimal for the highest types.

We consider a principal-agent model wherein a risk-neutral funder (he; the principal) motivates a risk-neutral researcher (she; the agent) to undertake an R&D project. The outcome depends on the researcher's private type, and two essential and complementary inputs – "investment" and "effort". Investment is contractible; effort is not.[2] If she succeeds, the researcher profits by marketing the technology, but this incentive is insufficient from the funder's perspective. To motivate greater R&D activity, the funder specifies a transfer independent of performance – a "grant" – and a payment for success – a "prize".

Our results reveal that the virtues of performance-pay depend on the relative strengths of AS and MH. In our model, a prize creates a strong incentive for effort, but generates costly rent for the researcher due to AS. A grant

[1] See, e.g., Grossman and Hart (1983) or Bolton and Dewatripont (2005) (Ch. 4).

[2] We use the terms "observable", "contractible", and "verifiable" interchangeably.

effectively limits rent, but creates only an indirect incentive for effort (by motivating investment). The virtue of the prize depends on the balance of these trade-offs. In some circumstances, the optimal prize is zero for a range of types. For high enough types, however, the prize is always strictly positive; moreover, when MH is more severe, the prize is strictly positive for all types.

We contribute to the contracting literature under AS/MH. In many models, output is the only verifiable signal available to the principal.[3] This renders performance-pay indispensable, as output-independent rewards will not affect marginal incentives. While it may be infeasible to monitor research effort, some inputs, such as large-scale capital investments, may be easier to verify. If so, then investment can be encouraged with rewards tied only to these expenditures. But a researcher's effort may be more productive when she has better equipment with which to work. Then, as long as there is *some* benefit to success, greater investment increases the marginal returns to effort. A similar intuition obtains in multitasking models (e.g. Hölmstrom and Milgrom, 1991; Meng and Tian, 2013). Given multiple complementary tasks, a stronger incentive on one task induces greater effort on others.

We also contribute to the literature on innovation incentives under MH, which has largely focused on pull programs;[4] few studies have examined the interactions between push and pull programs taking MH into account. Maurer and Scotchmer (2003) argue that repeated interaction between grantees and grantors resolves the MH problem. Our insights complement their's, as they are relevant in a static setting. Fu et al. (2012) show that grants may facilitate greater competition in a contest between researchers with asymmetric capital endowments. We abstract from competition to focus on the role of information.

[3]Studies close to this analysis include Lewis and Sappington (2000a,b) and Ollier and Thomas (2013). There are notable exceptions, which will be discussed.

[4]For instance, there is a large literature on optimal patent design; e.g., Gilbert and Shapiro (1990) Klemperer (1990), O'donoghue et al. (1998), Cornelli and Schankerman (1999), and Hopenhayn and Mitchell (2001). See Hall (2007) for a survey. And a literature on alternatives to intellectual property such as prizes or contracts; e.g., Wright (1983), Kremer (1998), Shavell and Van Ypersele (2001), Hopenhayn et al. (2006), Weyl and Tirole (2012), and Che et al. (2015). See Maurer and Scotchmer (2003) for an overview.

Many studies have explored trade-offs between high and low-powered incentives. Due to the "effort-substitution problem" Hölmstrom and Milgrom show the potential for "fixed-wage" contracts.[5] This fixed wage is independent of *any* signal received by the principal, while the grant in our model depends on investment, but is independent of performance. Baker (1992) shows that performance-pay may be muted if performance is weakly correlated with verifiable measures. Low-powered incentives may also arise as a means of risk-sharing (see, e.g., Prendergast, 1999); we abstract from risk-sharing as all parties are risk-neutral in our model. Allowing costly monitoring of effort, Prendergast (2002) shows that performance-pay may be more beneficial if the principal is uncertain of the "correct" action an agent should take.[6]

Also related, Zhao (2008) and Chen (2010, 2012) allow for partially observable actions, but abstract from AS. In a class of models following Laffont and Tirole (1986), the agent devotes unobservable effort directed at cost reduction, then chooses an observable output. But these models tend to involve "false moral hazard", which arises when there is a deterministic relationship between type, unobservable effort, and a contractible signal (see Laffont and Martimort, 2009, Ch.7). Meng and Tian (2013) study a multitasking model with AS and MH, and explain why the agent may be led to specialize on certain tasks. Finally, our framework relates closely to Laffont (1995), Lewis and Sappington (2000b) and Ollier and Thomas (2013). Differentiating our model is the presence of partially observable complementary actions, and the researcher's profit motivation, which both play a critical role in our model.

[5] The effort-substitution problem arises in multitasking models when efforts are substitutes. As a result, a stronger incentive on one task reduces effort on the other task.

[6] There is also a distinct literature that examines the question of whether it is optimal to monitor input or output if it is too costly to observe both (see, Maskin and Riley, 1985).

2 The Model

The Primitives

A researcher undertakes an R&D project whose outcome – success or failure – depends on her type, θ; investment, $x \in \mathbb{R}_+$; and effort, $y \in [0,1]$. Investment is contractible; effort is not. Success is verifiable at no cost. θ is a random variable with CDF, F, (smooth) PDF, f, and support $\Theta = [\underline{\theta}, \overline{\theta}] \subset (0,1]$. The researcher knows the true θ; the funder knows only its distribution. For each $\theta \in \Theta$, assume $f(\theta) > 0$, and $h'(\theta) < 0$, where $h(\theta) = \frac{1-F(\theta)}{f(\theta)}$ is the inverse hazard rate.

Given x, y, and θ, the probability of success is $\theta y \rho(x)$. The function, $\rho : \mathbb{R}_+ \to [0,1]$, is twice continuously with $\rho(0) = 0$, $\rho' > 0$ and $\rho'' < 0$. Notice that investment and effort are both essential for success, and are complements. If the researcher chooses (x, y), she incurs a cost $C(x, y) = x + \psi(y)$. For ease of exposition, we let $\psi(y) = \frac{c}{\alpha} y^\alpha$, where $\alpha > 1$, but our main results may be generalized (see Section 3.4). If she succeeds, the researcher earns $\pi > 0$, and the funder captures $W > 0$, otherwise both receive nothing. π may reflect the researcher's monetary profit in the product market or some intrinsic motivation. We will refer to π as the researcher's profit. W could be interpreted as the consumer surplus associated with the technology. Absent intervention, the payoff to a type-θ researcher who chooses (x, y), is $\Pi(x, y, \theta) = \theta y \rho(x)\pi - C(x, y)$. We focus on settings where π is "small", relative to W. This is quite natural in the context of R&D as the social value of an innovation often exceeds the value to the innovator (see, e.g., Hall et al., 2009). For simplicity, assume for all $y \in [0,1]$ and $\theta \in \Theta$,

$$\frac{\partial \Pi(x, y, \theta)}{\partial x}|_{x=0} = \theta y \rho'(0)\pi - 1 < 0, \tag{1}$$

which implies $\Pi(\cdot, y, \theta)$ is strictly decreasing and $\max_{x \geq 0, y \in [0,1]} \Pi(x, y, \theta) = 0$.[7]

[7] If $\max_{x \geq 0, y \in [0,1]} \Pi(x, y, \theta) > 0$ then the researcher's outside option is type-dependent, which could give rise to the phenomenon of "countervailing incentives" (see, e.g., Lewis and Sappington, 1989). In our model, so long as the funder induces the researcher to invest more than she would otherwise, this issue does not arise (see Rietzke and Chen, 2016).

Feasible Contracts and the Funder's Problem

The funder designs contracts to motivate R&D activity. A contract specifies a transfer independent of performance, $g \in \mathbb{R}$, a prize for success, $v \in \mathbb{R}_+$, and an investment, $x \in \mathbb{R}_+$. We interpret $g > 0$ as a grant, and $g < 0$ as an *entry fee*. Following Innes (1990) and Poblete and Spulber (2012), the funder is subject to a "free-disposal" constraint, requiring $v \geq 0$.[8] Without loss of generality, the researcher is eligible for any rewards only if she chooses the agreed-upon, x'. If she deviates from x' and chooses $x \neq x'$, her payoff is $\Pi(x, y, \theta) \leq \max_{x \geq 0, y \in [0,1]} \Pi(x, y, \theta) = 0$. So, as long as her payoff is non-negative when she chooses x', it is never optimal to deviate.

By the Revelation principle, it suffices to consider direct mechanisms. The funder commits to a menu, $\{v(\theta), g(\theta), x(\theta)\}_{\theta \in \Theta}$, where $v : \Theta \to \mathbb{R}_+$ is a prize schedule, $g : \Theta \to \mathbb{R}$ is a grant/fee schedule, and $x : \Theta \to \mathbb{R}_+$ is an investment schedule. Note that, where it does not cause confusion, we will liberally abuse notation sometimes let $v \in \mathbb{R}_+, g \in \mathbb{R}$, and $x \in \mathbb{R}_+$ denote particular prize, grant, and investment amounts. The researcher observes the menu, and if she participates, reports her type, $\hat{\theta}$, forming the contract, $\{v(\hat{\theta}), g(\hat{\theta}), x(\hat{\theta})\}$. The researcher then chooses her inputs, the project's outcome is realized, and transfers are made accordingly. If the researcher does not participate, both parties earn zero. We restrict attention to continuous, piecewise-differentiable prize, grant/fee, and investment schedules, but it will be shown (in the proof of Proposition 4) that this is without loss of generality.

Given (v, g, x), the researcher's effort, $y^*(x, v, \theta)$, solves $\max_{y \in [0,1]} \{\theta y \rho(x)(v + \pi) - C(x, y) + g\}$. At an interior solution, $y^*(\cdot)$ is unique-valued and given by, $y^*(x, v, \theta) = \left(\frac{\theta}{c}\rho(x)(v + \pi)\right)^{\beta}$. Let $\beta = \frac{1}{\alpha - 1}$; note that the elasticity of the researcher's effort with respect to $v + \pi$ is constant, and equal to β: $\frac{\partial y^*}{\partial(v+\pi)} \frac{v+\pi}{y^*} = \beta$. The payoff to a type-$\theta$ who reports $\hat{\theta}$ is,

$$u(\hat{\theta}|\theta) = \theta y^*(x(\hat{\theta}), v(\hat{\theta}), \theta)\rho(x(\hat{\theta}))(v(\hat{\theta}) + \pi) - C(x(\hat{\theta}), y^*(x(\hat{\theta}), v(\hat{\theta}), \theta)) + g(\hat{\theta}).$$

[8]This constraint ensures that a successful researcher cannot benefit by shrouding her success from the funder. This may be possible if a researcher who claims failure cannot be disproved by the funder.

Let $S(x, y, \theta) = \theta y \rho(x)(W + \pi) - C(x, y)$ denote total surplus and, abusing notation, let $u(\theta) \equiv u(\theta|\theta)$ and $y^*(\theta) \equiv y^*(x(\theta), v(\theta), \theta)$. Under truthful reporting, the funder's payoff is,

$$\phi = \int_{\underline{\theta}}^{\overline{\theta}} \left(\theta y^*(\theta)\rho(x(\theta))\,[W - v(\theta)] - g(\theta)\right) f(\theta)d\theta = \int_{\underline{\theta}}^{\overline{\theta}} \left[S(x(\theta), y^*(\theta), \theta) - u(\theta)\right] f(\theta)d\theta.$$

The second expression follows from the first by replacing $g(\theta)$ by $u(\theta)$ at each θ. The funder's payoff can be interpreted as expected consumer surplus, less the expected cost of funding, or expected total surplus, less the researcher's expected rent.[9]

Individual rationality (IR) requires $u(\cdot) \geq 0$; incentive compatibility (IC) requires for all $\theta, \hat{\theta} \in \Theta$, $u(\theta) \geq u(\hat{\theta}|\theta)$. As we have restricted attention to piece-wise differentiable mechanisms, by standard techniques, IC is satisfied if and only if at all points of differentiability:[10]

$$u'(\theta) = \frac{\partial u(\hat{\theta}|\theta)}{\partial \theta}\Big|_{\hat{\theta}=\theta} = y^*(\theta)\rho(x(\theta))(v(\theta) + \pi), \tag{IC-F}$$

and

$$\frac{d}{d\theta}\rho(x(\theta))(v(\theta) + \pi) \geq 0. \tag{IC-S}$$

Since (IC-F) implies $u' \geq 0$, IR is satisfied for all types so long as $u(\underline{\theta}) \geq 0$; since ϕ is strictly decreasing in u, this constraint binds in equilibrium. Using (IC-F) and setting $u(\underline{\theta}) = 0$: $\int_{\underline{\theta}}^{\overline{\theta}} u(\theta)f(\theta)d\theta = \int_{\underline{\theta}}^{\overline{\theta}} y^*(\theta)\rho(x(\theta))(v(\theta) + \pi)h(\theta)f(\theta)d\theta$. Let $\mathcal{I}(x, y, v, \theta) = h(\theta)y\rho(x)(v + \pi)$ denote the researcher's *virtual information rent.* The funder's problem may then be expressed,

[9] Our results extend to an environment where the principal also values the researcher's profit, but there is a social cost to raising funds, as in Laffont and Tirole (1986).

[10] See, e.g. Laffont and Tirole (1993) pp. 64 and 121.

$$\max_{x,v} \int_{\underline{\theta}}^{\overline{\theta}} \{S(x(\theta), y^*(\theta), \theta) - \mathcal{I}(x(\theta), y^*(\theta), v(\theta), \theta)\} f(\theta) d\theta$$

$$\text{s.t. } x(\cdot) \geq 0, \text{ (IC-S), and } v(\cdot) \geq 0. \quad \text{[P]}$$

Let $J(x, v, \theta)$ denote the integrand above. For each θ, we assume $\max_{x,v\geq 0} J(x, v, \theta) > 0$, which holds if W is sufficiently large,[11] and implies $\max_{x,y\geq 0} S(x, y, \theta) > 0$. Also assume for some $\hat{x} > 0$, $x > [<]\hat{x}$ implies $\rho''(x)\rho(x) + \beta\rho'^2(x) < [>]0$, which ensures J is strictly concave in x for $x > \hat{x}$ at the optimal v (for example, if $\rho(x) = 1 - \exp(-x)$ and $\beta = 1$ then $\hat{x} = \log(1 + \beta)$). Finally, we will assume throughout that the parameters are such that $y^*(\cdot) < 1$ at the solution to [P] (this holds if, for instance, $c > W + \pi$).

3 Results

Before characterizing the optimal contracts with AS and MH, we study three benchmarks: complete information, pure MH, and pure AS. We use uppercase letters (X, V, G, etc.) to denote optimal solutions. In the environments with MH, we let $Y(\theta) \equiv y^*(X(\theta), V(\theta), \theta)$.

3.1 First-Best: Complete Information

With complete information, the funder observes the true θ, as well as x and y. The funder's problem can be written $\max_{x,y,u\geq 0} \{S(x, y, \theta) - u\}$. Our first result characterizes the solution to this problem. We let $X_{FB}(\theta)$ and $Y_{FB}(\theta)$ denote the first-best investment and effort levels (respectively).

Proposition 1. *With complete information, the optimal means of funding is any combination of $V(\theta) \geq 0$ and $G(\theta)$ satisfying, $\theta\rho(X_{FB}(\theta))(V(\theta) + \pi) - C(X_{FB}(\theta), Y_{FB}(\theta)) + G(\theta) = 0$. $X_{FB}(\theta)$ and $Y_{FB}(\theta)$ are unique and satisfy,*

[11] Relaxing this assumption may yield an interval of low types who invest nothing (and receive no rewards from the funder), but will not change the qualitative conclusions of our analysis for the range of types investing a positive amount.

$\theta Y_{FB}(\theta)\rho'(X_{FB}(\theta))(W + \pi) = 1$ and $Y_{FB}(\theta) = (\theta\rho(X_{FB}(\theta))(W + \pi))^{\beta}$. Finally, $X'_{FB}(\theta) > 0$ and $Y'_{FB}(\theta) > 0$.

Proposition 1 shows that with complete information, the optimal means of funding may take the form of a pure prize ($V(\theta) > 0$ and $G(\theta) = 0$), a pure grant ($V(\theta) = 0$ and $G(\theta) = 0$) or some combination of the two. Indeed, with complete information and risk-neutrality, a grant of value, $g > 0$ is equivalent to a prize, v, with expected value $\theta y\rho(x)v = g$. Finally, note that the first-best investment/effort levels are increasing in θ.

3.2 Pure Moral Hazard

With pure MH, y is unobservable by the funder, but he observes θ and x. Given θ, the funder's problem may be expressed, $\max_{x,v,u\geq 0}\{S(x, y^*(x, v, \theta), \theta) - u\}$. The next result characterizes the optimal scheme with pure MH.

Proposition 2. *With pure MH, the optimal means of funding is a prize, $V(\theta) = W$ and an entry fee $G(\theta) < 0$. Moreover, $X(\theta) = X_{FB}(\theta)$, $Y(\theta) = Y_{FB}(\theta)$, and $G(\theta) = -S(X_{FB}(\theta), Y_{FB}(\theta), \theta)$.*

With pure MH, the optimal funding scheme takes the form of a "franchise contract" in which the agent is made a full residual claimant, and the funder extracts the researcher's rent through an entry fee.

Although the researcher's effort, $y^*(x, v, \theta) = \left(\frac{\theta}{c}\rho(x)(v + \pi)\right)^{\beta}$, is completely independent of g, it is important to emphasize that a positive grant *can* induce greater effort. Note that y^* is strictly increasing in both x and v. As investment is observable, the funder may condition the grant on this variable and elicit greater investment. Greater investment increases the marginal returns to effort and increases effort. A prize also induces greater investment (and thus operates along this indirect channel), but it also operates along a direct channel, since it is only received in the event of success.

3.3 Pure Adverse Selection

With pure AS, both x and y are observable by the funder, but θ is observed only by the researcher. The funder's problem is given by [P] (see Section 2), except the funder also chooses y, since effort is contractible. Our next result characterizes the optimal funding scheme under pure AS.

Proposition 3. *With pure AS, the optimal means of funding is a pure grant for all types. Moreover, (i)* $C(X(\theta), Y(\theta)) > G(\theta)$, $\frac{d}{d\theta}[C(X(\theta), Y(\theta)) - G(\theta)] > 0$*; and (ii)* $X(\theta) \leq X_{FB}(\theta)$, $Y(\theta) \leq Y_{FB}(\theta)$*, with equality only at* $\overline{\theta}$*.*

Proposition 3 shows that a pure-grant is optimal for all types in the model with pure AS. To understand why the optimal prize is zero, consider two types – a high type and low type. If the high type imitates the low type, the high type is more likely to succeed and receive the prize than the low type would be; therefore, the expected value of a prize offered to the low type is greater for a higher type. To prevent under-reporting, the high type must be compensated through an information rent. The grant, on the other hand, is received independently of outcome, so its value is the same for both types. Thus, for a given investment/effort schedule, prizes are a more expensive and inferior means of funding with pure AS.

The optimal investment/effort schedules are distorted below the first-best and balance the trade-off between rent extraction and efficiency. Although this efficiency/rent extraction trade-off is standard in AS models,[12] we note the role played by π and free-disposal in our model. If we relax free-disposal, then the funder appropriates all of the researcher's rent by setting $v(\cdot) = -\pi$ and $g(\cdot) = C(x(\cdot)), y(\cdot))$ (see Lewis and Sappington, 2000b). But under free-disposal, the researcher must capture at least π in the event of success, which leaves an inappropriable rent. To prevent low types from overstating their ability, the optimal grant does not fully reimburse costs ($C(X, Y) > G$), and higher types must internalize a greater cost ($\frac{d}{d\theta}[C(X, Y) - G] > 0$).

[12] See, e.g., Maskin and Riley (1984)

3.4 Mixed Case: Adverse Selection and Moral Hazard

We now study the case of AS and MH. For a given θ, and a fixed y, we let $X_{FB|y}(\theta)$ denote the investment that maximizes total surplus: $X_{FB|y}(\theta) = \arg\max_{x\geq 0} S(x, y, \theta)$. $X_{FB|y}(\theta)$ is the unique solution to $\theta y \rho'(X_{FB|y}(\theta))(W + \pi) = 1$. We call $X_{FB|y}$ the first-best investment, given effort; we write $X_{FB|Y}(\theta)$ when $y = Y(\theta)$. The next result describes properties of the optimal funding scheme in the model with AS and MH. Before stating the result, define θ_v as follows:

$$\theta_v = \max\{\underline{\theta}, \theta | \beta\theta W = h(\theta)\pi(1 + \beta)\}.$$

Note that since $\beta\overline{\theta}W > h(\overline{\theta})\pi(1 + \beta) = 0$, $\theta_v < \overline{\theta}$.

Proposition 4. *In the model with AS and MH,*

(i) If $\theta_v > \underline{\theta}$, then for $\theta \in [\underline{\theta}, \theta_v]$, $V(\theta) = 0$, $G(\theta) > 0$, $G'(\theta) > 0$; moreover, $X(\theta) \leq X_{FB|Y}(\theta)$, with equality only at θ_v.

(ii) For $\theta \in (\theta_v, \overline{\theta}]$, $V(\theta) > 0$, $V'(\theta) > 0$, and $X(\theta) = X_{FB|Y}(\theta)$. Moreover, there exists $\theta_g < \overline{\theta}$ such that $\theta > \theta_g$ implies $G(\theta) < 0$ and $G'(\theta) < 0$.[13]

(iii) For $\theta \in [\underline{\theta}, \overline{\theta}]$, $X'(\theta) > 0$, $Y'(\theta) > 0$, $X(\theta) > G(\theta)$, and $X'(\theta) > G'(\theta)$. Moreover, $X(\theta) \leq X_{FB}(\theta)$ and $Y(\theta) \leq Y_{FB}(\theta)$, with equality only at $\overline{\theta}$.

Proposition 4 shows that the optimal prize may be zero for a range of low types, but is always strictly positive for high enough types. When the prize is zero, effort is incentivized indirectly through the grant; higher types receive large grants, but internalize a greater investment cost (i.e. $\frac{d}{d\theta}[X(\theta) - G(\theta)] > 0$). When the prize is positive, higher types receive larger prizes. The following corollary provides a necessary and sufficient condition under which the interval of types receiving pure-grant funding is non-empty.

Corollary 1. *The interval of types that receive pure grant funding is non-empty (i.e., $\theta_v > \underline{\theta}$) if and only if $\beta\underline{\theta}W < h(\underline{\theta})\pi(1 + \beta)$.*

[13]if $\theta_v = \underline{\theta}$, then Property (ii) holds for all $\theta \in [\underline{\theta}, \overline{\theta}]$

In order to incentivize higher levels of effort, the funder must either encourage greater investment or use a prize. Either way, greater information rent is generated for the researcher. All else equal, when the funder is more concerned about limiting information rent due to AS, he induces lower levels of effort. For reasons explained below, pure grant funding is more likely to be optimal when the funder wishes to induce lower levels of effort.

To elucidate the trade-offs in the model with AS/MH, let us consider an auxiliary problem. Suppose the funder wants to induce effort, $\overline{y} \in (0, 1)$, at some θ. The funder's instantaneous payoff at θ when $y^*(x, v, \theta) = \overline{y}$ is, $J = [S(x, \overline{y}, \theta) - \mathcal{I}(x, \overline{y}, v, \theta)] f(\theta)$. The funder chooses (x, v) to maximize J, subject to $y^*(x, v, \theta) = \overline{y}$. In this problem, the optimal (x, v) depends on two considerations: First, is the relative effectiveness of x and v to elicit effort, while limiting $\mathcal{I}$. Second, is the impact of x and v on S.

For simplicity, set $\beta = c = 1$. To understand the first consideration, note that the combinations of (x, v) that induce $\overline{y}$ satisfy, $y^*(x, v, \theta) = \theta\rho(x)(v + \pi) = \overline{y}$, or $\rho(x)(v + \pi) = \frac{\overline{y}}{\theta}$. Then $\mathcal{I}(x, \overline{y}, v, \theta) = h(\theta)\frac{\overline{y}^2}{\theta}$, which is independent from (x, v). Thus, any (x, v) that induces $\overline{y}$ yields the same virtual information cost, $\mathcal{I}$. However, note that investment contributes directly to total surplus, while the prize does not. In particular, $S(\cdot, \overline{y}, \theta)$ is strictly increasing [decreasing] for $x < [>] X_{FB|\overline{y}}$, while S is independent from v.

Jointly, these two considerations imply that it is optimal for the funder to induce a maximum investment of $x = X_{FB|\overline{y}}$, and provide any residual incentive necessary for effort through the prize.[14] If $y^*(\tilde{x}, 0, \theta) = \overline{y}$ for some $\tilde{x} \leq X_{FB|\overline{y}}$, then the funder sets $x = \tilde{x}$ and $v = 0$; in this case investment/effort are incentivized through the grant. If $y^*(X_{FB|\overline{y}}, 0, \theta) < \overline{y}$, the funder sets $x = X_{FB|\overline{y}}$, and specifies $v > 0$ to satisfy $y^*(X_{FB|\overline{y}}, v, \theta) = \overline{y}$. When $\overline{y}$ is

[14] Multiplicative separability between x, y, and θ simplifies the first consideration. For more general technologies, x and v may differ in terms of their ability to raise effort while limiting $\mathcal{I}$, and the optimal (x, v) will depend on the interplay between the two considerations in a more intricate way. In particular, the stronger is the complementarity between x and y, the cheaper it will be (in terms of limiting $\mathcal{I}$) to induce effort through x. In general, the nature of the interactions between the considerations can be quite complex and may vary over the distribution of types. The primary technical complication is determining if and when (IC-S) binds.

chosen optimally this scheme is precisely what is described in Proposition 4. A similar argument holds for any effort-cost function, ψ.

The optimal level of effort induced by the funder will depend on the trade-off between efficiency and rent extraction. Note that, in the auxiliary problem above, the information rent captured by the researcher is positively related to the effort level, $\overline{y}$. All else equal, when AS is more severe, the funder is more concerned about limiting the researcher's rent, and induces a lower level of effort. For lower levels of effort, pure grant funding is more likely to be optimal. When MH is more of a concern, larger prizes are used to encourage greater effort. θ_v can be interpreted as capturing the relative severity of AS and MH. Specifically, when AS is more severe relative to MH, θ_v is greater. Proposition 5 in Section 4 formalizes this intuition.

Proposition 4 also shows that the prize is strictly positive and increasing for $\theta \in (\theta_v, \overline{\theta}]$. Moreover, for θ high enough, the scheme resembles a franchise contract that emerged under pure MH (see Section 3.2). To understand these features, first note that the marginal benefit of effort (to the funder) is strictly increasing in θ. Larger prizes are therefore more attractive for higher types, as they can elicit greater effort. Second, as discussed in Section 3.3, a prize offered to some type, θ', generates rent for all higher types. But for θ' close to $\overline{\theta}$, the funder is less concerned about limiting the rent of a higher type (since the researcher is very unlikely to be of such a type), and the issue of limiting rent due to AS (which tilts against prizes) vanishes, while the MH problem (which tilts in favor or prizes) does not. Still, the following example demonstrates that θ_v can be arbitrarily close to $\overline{\theta}$.

Example 1. *Let* $\theta \sim U\left[\frac{2}{3}, 1\right]$*,* $\rho(x) = 1 - \exp(-x)$*,* $c = 5$*,* $\pi = 1$*,* $W = 4$*. Then for any* $\alpha > 8$ *(*$\beta < \frac{1}{7}$*),* $X(\theta) > 0$ *and* $\theta_v = \frac{\alpha\pi}{W+\alpha\pi} = \frac{\alpha}{4+\alpha}$*. Furthermore,* $\theta_v \to \overline{\theta} = 1$ *as* $\alpha \to \infty$ *(*$\beta \to 0$*).*

Example 1 shows that the interval of types that receive a prize can be arbitrarily small. Intuitively, when $\alpha \to \infty$ ($\beta \to 0$), $y^*(\cdot)$ becomes less sensitive to the prize. Reducing the prize has little impact on effort, and AS becomes the dominant information problem. We also note that if, due to a limitation on the researcher's time/energy, the upper-bound on effort binds for some $\tilde{\theta} < \theta_v$ (that is, $Y(\tilde{\theta}) = \left(\frac{\tilde{\theta}}{c}\rho(X(\tilde{\theta}))\pi\right)^{\beta} = 1$), then for $\theta > \tilde{\theta}$, MH imposes a non-binding constraint on the funder, and pure grant funding will be optimal for all types.[15]

In the next section, we explore more completely the circumstances under which performance-pay is or is not utilized, but before proceeding, we first mention the role of free-disposal and then comment on the particular functional form of ψ we have studied. If $\theta_v = \underline{\theta}$, the free-disposal constraint never binds and has no impact on our results. If $\theta_v > \underline{\theta}$ and we relax the free-disposal constraint, then it can be shown that for $\theta < \theta_v$, $G(\theta) > 0$, $X(\theta) = X_{FB|Y}(\theta)$, and $-\pi < V(\theta) = \frac{\theta\beta W - h(\theta)\pi(1+\beta)}{\theta\beta + h(\theta)(1+\beta)} < 0$. This scheme could be interpreted as one in which the funder purchases an equity stake in the researcher's project.

The functional form of ψ we have studied implies a constant elasticity of effort with respect to the reward for success; however, the qualitative features of our main results hold for more general technologies. Consider any ψ with $\psi'(0) = \psi(0) = 0$, and $\psi''(y) > 0$ for $y > 0$. It can be shown that the elasticity, ϵ, of the researcher's effort with respect to $v + \pi$ depends only on $z = \theta\rho(x)(v + \pi)$. The qualitative nature of Proposition 4 (and Corollary 1) hold so long as $\epsilon'(z) \geq 0$ for all z. This condition is analogous to condition (12b) in Ollier and Thomas, and is sufficient to rule out a binding (IC-S) constraint.

4 Further Analysis and Discussion

4.1 Further Analysis: Comparative Statics

Comparative Statics

Our next result explores how θ_v depends on the parameters of the model. First, it will be useful to introduce a parameter that captures the severity of the AS problem. To this effect, let us parameterize the distribution of θ by $t \in \mathbb{R}$

[15] In Rietzke and Chen (2016), we illustrate a similar result in a binary-effort model.

such that, at each $\theta \in [\underline{\theta}, \overline{\theta})$, $\frac{\partial h(\theta;t)}{\partial t} > 0$. Greater t reflects a more severe AS problem in the following sense: Recall that higher types capture rent through their ability to mimic lower types. Fix $\theta' < \overline{\theta}$, and consider the event that $\theta > \theta'$, conditional on θ being in some neighborhood of θ'. When this event is more likely, the greater is the expected rent relinquished in this neighborhood (for a given (x, v)), and can we say the AS problem is more severe. This is precisely what is captured by t: a higher t increases the likelihood of this event at each $\theta' \in [\underline{\theta}, \overline{\theta})$.

Proposition 5. *Suppose $\theta_v > \underline{\theta}$. Then θ_v is strictly decreasing in β and W, and is strictly increasing in π and t.*

Proposition 5 reveals that pure grant funding is used for a wider range of types when: (1) the funder's value for the project is modest (i.e., W is not "too large"); (2) the researcher associates a higher value to the project (i.e., π is large); (3) the researcher's effort is less sensitive to the reward for success (i.e., β is small); and (4) the AS problem is more severe (i.e., t is large).

Point (1) holds since effort is less valuable to the funder when W is small, rendering prizes less attractive. Points (2)-(3) speak to the severity of the MH problem: When π is large, the researcher has a stronger natural incentive to exert effort. When β is small, the researcher's effort is less sensitive to the prize; therefore, even a large reduction in the prize results in a modest reduction in effort. Point (4) holds since AS raises the expected cost of performance-pay; the greater is this problem, the less attractive prizes become.

The complementarity between investment/effort is also important for the usefulness of the grant in our model. As a point of comparison, in multitasking models when efforts are substitutes, the principal reduces incentives on more-easily observed tasks to avoid crowding-out effort on less-easily observed tasks. By a similar logic, if investment/effort were substitutes in our model then greater investment reduces effort, rendering the grant an ineffective means of eliciting effort.

We now explore how the prize depends on the parameters of the model.

Proposition 6. *For $\theta > \theta_v$, $V(\theta)$ is strictly increasing in β and W, and is strictly decreasing in π and t.*

Proposition 6 shows that larger prizes are utilized under antithetical circumstances to those described in Proposition 5. The intuitions follow as the counterpoint to those results. Specifically, larger prizes are used when (i) W is large, (ii) the MH problem is more severe (π is small or β is large), or (iii) the AS problem is less severe (t is small). When W is large, the researcher's effort is more valuable to the funder, which makes prizes more attractive. When MH is more severe larger prizes are more attractive due to their ability to encourage greater effort. When AS is less severe, the information rent cost associated with higher levels of effort is lower. Higher levels of effort are therefore more attractive, and larger prizes are utilized.

4.2 Discussion

Grace and Kyle (2009) note that there is limited evidence on how push and pull incentives work together. Our results describe precisely how these incentives can work together to help resolve AS and MH problems. Our results suggest that hybrid incentive schemes, which combine push and pull elements, are particularly effective in dealing with these information problems. Our findings call for stronger push incentives when AS is more of a concern, and stronger pull incentives when MH is more of a concern. Finally, our results call for the use of matching grants, in which the grantee shares in the cost of the project, to prevent researchers from overstating the prospects of a project. In this section, we discuss the empirical relevance of our theoretical predictions, and discuss several interesting extensions/avenues for future research.

Pharmaceutical Innovation

A prominent example of push/pull incentives working synergistically in practice is the U.S. Orphan Drug Act. This policy includes both R&D tax credits (push) as well as priority review vouchers (pull), which lengthen the effective patent life of a new drug. Lichtenberg and Waldfogel (2003) present evidence that this policy was effective in spurring the development of orphan drugs. Still, there remains considerable debate among policy-makers and researchers regarding the optimal balance of push and pull incentives (see, e.g., Kremer and Glennerster, 2004). Renwick et al. (2016) note an imbalance in the structure of incentives currently available. The bulk of funding is in the form of push incentives and targets the early stages of drug discovery. Pull programs, to the extent that they are used, primarily target later development and commercialization stages.

The early stage of drug discovery involves searching for and selecting compounds that may be effective in treating a particular condition. It is a highly scientific endeavor and requires highly specialized knowledge. Later stages of development/commercialization involve clinical testing and marketing expenses (Sloan and Hsieh, 2007, Ch. 3). Generally speaking, the development phase, while consuming a significant fraction of total R&D costs, requires less specialized knowledge, particularly in the case of well-understood conditions. As Cockburn and Henderson (2001) write: "For example, the design of clinical trials in the case of a compound designed to treat a relatively well understood condition such as hypertension is relatively straightforward...". To the extent that drug discovery entails more specialized knowledge than drug development, the observed pattern of funding is consistent with our theoretical predictions.[16] Of course, other factors beyond the scope of our model may also influence contractual forms. For example, drug discovery is more likely to involve non-profit research institutions (including university researchers) and small bio-technology firms, while drug development is typically undertaken by large for-profit pharmaceutical firms. As such, capital constraints may present more of a barrier at the early discovery phase (later in this section we provide a more detailed discussion of capital constraints).

[16]Note that our model does not incorporate a multi-stage R&D process. However, one can interpret the pharmaceutical example as consisting of two distinct R&D projects – one involving the search for a viable compound, and the second involving safety and efficacy testing, each entailing a unique information environment.

Federal Procurement

U.S. government agencies, such as NASA and the Department of Defense (DOD), are often interested in procuring innovative technologies; our model can readily be interpreted in this context. In this environment, a pure-grant corresponds to a cost or cost-sharing contract;[17] a pure-prize corresponds to a firm fixed-price contract. Between these two extremes are a number of hybrid schemes such as cost-plus-fixed-fee completion contracts, which cover costs and award a fee, conditional on a final product being delivered.

As the Federal Acquisition Regulation (FAR) stipulates, contracting officers are encouraged to consider project complexity when designing contracts. It seems reasonable that in more technical, complex environments, contractors hold specialized knowledge and AS would be a first-order concern. All else equal, our results would therefore predict that cost-reimbursement contracts are more likely to emerge in more complex environments. Empirical evidence provided by Girth and Lopez (2018) is consistent with this prediction. The authors analyze government procurement contracts from 2000-2014 in three areas of varying complexity: housekeeping (least complex), construction (moderately complex), and space R&D (most complex). While about 95% of housekeeping contracts and 98% of construction contracts fell into the category of fixed price, only about 57% of the R&D contracts fell into this category. Almost 28% of R&D contracts fell into the cost-reimbursement category, while less than 1% of the housekeeping and construction contracts were of this type. Anecdotal evidence also suggests that push funding is more prominent in more complex environments. In guidance to DOD contracting officers, Grady (2016) writes: "For example, in a situation where great technical uncertainty exists, there is also great likelihood of cost uncertainty. These factors dictate the selection of a cost-reimbursement type of contract.".

But it is worth noting that more complex environments may also involve greater risk for contractors, and the design of contracts may be influenced by risk-sharing considerations. Disentangling whether it is risk-sharing or AS

[17] As Grady (2016) writes: "...the best efforts language in cost-type contracts allows for contractors to receive payment even if a final deliverable is not achieved."

that is driving contractual forms in this environment is an interesting question for future research.[18]

Cost-Sharing and Matching Grants

Once concern with push incentives is that they may pay for research unlikely to succeed. To deal with this issue, when the pure grant emerges in our model, higher types receive larger grants but must internalize a greater investment cost; this ensures that low types are unwilling to accept large grants.[19]

Cost-sharing programs are commonly used in practice. For instance, as part of its Horizon 2020 program, the European Commission awards grants to small and medium sized business that cover 70% of firms' costs.[20] As this program covers a fixed percentage of firm costs, firms seeking larger grants will necessarily internalize a greater cost. In the U.S., The Energy Policy Act of 2005 requires Department of Energy contractors to bear at least 20% of project costs for most R&D contracts, and at least 50% for demonstration projects (i.e., those projects closer to market). The Advanced Technology Program (ATP), which ran from 1991 - 2005, provided matching grants to industry for early-stage R&D. Small firms were required to cover their indirect and overhead costs, while large firms were required to cost-share at least 60% of total project costs (National Research Council, 2001). An assessment of the ATP by the National Research Council suggests that the cost-sharing served as a means of screening. The report states, "Because the development of new technologies is inherently risky, regular assessment is vital to ensure continued technical viability, with cost-sharing requirements acting as an effective safeguard." (National Research Council, 2001).

Link and Ruhm (2009) provide empirical evidence that is consistent with

[18] See, also, Bajari and Tadelis (2001), for an alternative theory linking project complexity to contractual forms, in a setting without AS.

[19] Maurer and Scotchmer (2003) point out that a matching grant can be an effective screening device with AS; our result also takes MH into account. Cost sharing has been advocated in other contexts for dealing with AS and MH (see, e.g., Laffont and Tirole, 1986).

[20] `http://ec.europa.eu/programmes/horizon2020/en/h2020-section/sme-instrument`

the use of matching grants as an effective screening tool. The authors analyze phase II grants made by the National Institute of Health's (NIH) Small Business Innovation Research (SBIR) award program between 1992 - 2001. Although the SBIR program does not require cost-sharing, roughly 50% of grant recipients contribute additional internal funding to the project. The study finds that grantees, who contribute their own resources to the project, are significantly more likely to bring their innovation to market. The authors conclude that "...our results suggest that if the NIH is interested in managing its SBIR program so as to increase the probability of commercialization, it might wish to consider conditioning receipt of a Phase II award on university involvement and own and/or internal business funding".

Although cost-sharing has been used by a number of U.S. federal agencies, it has sparked some controversy. At the National Science Foundation (NSF), for example, concern was raised that cost-sharing, combined with the competitive grant-awarding process, skewed funding to researchers with more access to capital, but less-valuable projects. A National Science Board review of cost-sharing policies at NSF noted that, "...cost sharing can foster unequal competitiveness among grantee institutions based on their ability and willingness to contribute cost sharing resources to NSF-sponsored projects." (National Science Board, 2009). As a result, the NSF no longer permits voluntary committed cost-sharing, and places firm restrictions on the use of mandatory cost-sharing.

Capital Constraints

Some push programs provide upfront funding, which may be necessary if a researcher is capital-constrained. Yet as Scotchmer (2004, Ch. 8) points out, an appropriately designed pull program should be capable of attracting financing. Indeed, this is precisely the logic behind the "Pay for Success" model run by the U.S. Department of Labor.[21] Moreover, some push programs (e.g., R&D tax credits), do not provide funding upfront. It is therefore useful to explain why push incentives might emerge naturally under MH, rather than out of

[21] https://www.doleta.gov/workforce_innovation/success.cfm

necessity due to a capital constraint. Nevertheless, our results are relevant for understanding a related issue. Consider a researcher who has a strong incentive to devote her time/energy to a project (i.e. π is large and/or β is small), but is unwilling to raise the necessary capital (In our model, the unwillingness of the researcher to raise capital is captured by (1).). While a pull program *could* be used to encourage investment, in our model, pure-grant funding may be optimal for a wider range of types.

Still, it is worth commenting on how a capital constraint affects our results. In our model, the willingness of a higher type to internalize a greater investment cost acts as a credible signal of their ability. But, in an extreme case where the researcher has no access to capital, the grant must fully reimburse investment. Higher types then lose the ability to credibly signal their type, (IC-S) inevitably binds, and bunching may arise. To see why (IC-S) must bind, recall that (IC-S) requires that the product, $z(\cdot) = \rho(x(\cdot))(v(\cdot) + \pi)$ is non-decreasing. If $z(\cdot)$ were strictly increasing over some interval of types, then a researcher of *any* type in this interval would maximize her payoff by reporting the highest type, since she would not have to bear any of the cost of her investment, but would reap the benefit of a large prize/probability of success. Therefore, (IC-S) must bind and $z(\cdot)$ must be constant. If $v(\cdot) = 0$ over some interval, then $z'(\cdot) = 0$ implies $x'(\cdot) = 0$ and bunching will arise.[22]

5 Conclusion

We have characterized the optimal contracts in a setting with AS/MH and partially observable actions. In contrast to typical findings in MH models, we showed that performance-pay may not be optimal for all types, but is always utilized for the highest types. Our results are useful for understanding the basic trade-offs caused by AS/MH between push and pull programs used to encourage R&D activity, but our results are relevant in other contexts, e.g., worker compensation.

[22]The issue is analogous to the wealth constraint in Lewis and Sappington (2000b).

Appendix

For clarity, we omit the arguments of functions when there is no ambiguity.

Proof of Proposition 1

The funder's payoff is strictly decreasing in u, so optimality dictates $u = 0$. The first-best investment/effort levels solve, $\max_{x,y\geq 0} S(x, y, \theta)$. By assumption, $\max_{x,y\geq 0} S(x, y, \theta) > 0$ and $Y_{FB} < 1$, so the optimal investment/effort levels are interior and must satisfy the following first-order conditions:

$$S_x(X_{FB}, Y_{FB}, \theta) = \theta Y_{FB} \rho'(X_{FB})(W + \pi) = 1 \tag{2}$$

and,

$$S_y(X_{FB}, Y_{FB}, \theta) = \theta \rho(X_{FB})(W + \pi) = cY_{FB}^{\alpha-1} \tag{3}$$

Equation (3) implies $Y_{FB} = \left(\frac{\theta}{c}\rho(X_{FB})(W + \pi)\right)^{\beta}$, where $\beta = \frac{1}{\alpha-1}$. Combining equations (2) and (3) yields,

$$\frac{\theta^{\beta+1}}{c^{\beta}}\rho(X_{FB})^{\beta}\rho'(X_{FB})(W + \pi)^{\beta+1} = 1. \tag{4}$$

We now show that X_{FB} and Y_{FB} are unique. First, as $Y_{FB} = \left(\frac{\theta}{c}\rho(X_{FB})(W + \pi)\right)^{\beta}$, if X_{FB} is unique then Y_{FB} is also unique. The second-order necessary condition requires $\rho''(X_{FB})\rho(X_{FB}) + \rho'(X_{FB})^2 \leq 0$; equivalently, $X_{FB} \geq \hat{x}$. As, $\rho(\cdot)^{\beta}\rho'(\cdot)$ is strictly decreasing for $x > \hat{x}$, there is a unique solution to (4) satisfying $X_{FB} \geq \hat{x}$. Thus, X_{FB} and Y_{FB} are unique.

Next, we show that $X'_{FB}(\theta) > 0$ and $Y'_{FB}(\theta) > 0$. To do so, we will apply the Implicit Function Theorem to (4) and show that $X'_{FB} > 0$; it is then straightforward to confirm that $Y'_{FB} > 0$. Differentiating the LHS of (4) with respect to θ and simplifying yields,

$$X'_{FB} = -\frac{(1 + \beta)\rho(X_{FB})\rho'(X_{FB})}{\rho''(X_{FB})\rho(X_{FB}) + \beta\rho'(X_{FB})^2}. \tag{5}$$

The numerator on the right-hand side of (5) is strictly positive by assumption. The denominator is strictly negative so long as $X_{FB} > \hat{x}$; we will now show that this is the case. As the second-order necessary condition requires $X_{FB} \geq \hat{x}$, it suffices to show that $X_{FB} \neq \hat{x}$.

Consider sequentially optimizing S first over y then x. For a given x, the optimal choice of y, $\tilde{y}(x)$, satisfies the first-order condition $S_y(x, \tilde{y}(x), \theta) = 0$, yielding $\tilde{y}(x) = \left(\frac{\theta}{c}\rho(x)(W+\pi)\right)^{\beta}$. The first-best investment level can then be found by solving $\max_{x\geq 0} \tilde{S}(x,\theta)$, where $\tilde{S}(x,\theta) \equiv S(x,\tilde{y}(x),\theta) = \frac{1}{1+\beta}\frac{\theta^{1+\beta}}{c^{\beta}}\rho(x)^{1+\beta}(W+\pi)-x$. The necessary first-order condition is $\tilde{S}_x(X_{FB},\theta) = \frac{\theta^{1+\beta}}{c^{\beta}}\rho(X_{FB})^{\beta}\rho'(X_{FB})(W+\pi)^{1+\beta} - 1 = 0$.

We proceed by contradiction. Contrary to the claim, suppose $X_{FB} = \hat{x}$. Then $\tilde{S}_x(\hat{x},\theta) = 0$. Now see that $\tilde{S}_{xx}(x,\theta) = \frac{\theta^{1+\beta}}{c^{\beta}}(W+\pi)^{1+\beta}\rho(x)^{\beta-1}\left[\rho(x)\rho''(x) + \beta\rho'(x)^2\right] > 0$ for $x < \hat{x}$. We have now shown that $\tilde{S}_x(\hat{x},\theta) = 0$ and for all $x < \hat{x}$, $\tilde{S}_{xx}(x,\theta) > 0$. This means that for all $x < \hat{x}$, $\tilde{S}_x(x,\theta) < 0$. But since $\tilde{S}(0,\theta) = 0$, for any $x \leq \hat{x}$ it must be that $\tilde{S}(x,\theta) < 0$. In particular, $\tilde{S}(\hat{x},\theta) < 0 = \tilde{S}(0,\theta)$, which contradicts the hypothesis that $X_{FB} = \hat{x}$. Therefore, it must be that $X_{FB} > \hat{x}$, as claimed. □

Proof of Proposition 2

The funding scheme outlined in the Proposition gives the funder the first-best payoff of $\max_{x,y\geq 0} S(x,y,\theta)$; therefore, this is an optimal scheme. To see that it is the *unique* optimal funding scheme, note that when the researcher invests X_{FB}, then, $y^*(X_{FB}, v, \theta) = Y_{FB}$ if and only if $v = W$. It follows that any $v \neq W$ leads to strictly lower total surplus, and the funder's payoff is strictly less than $\max_{x,y\geq 0} S(x,y,\theta)$.

Proof of Proposition 3

We first ignore (IC-S); we then verify it is non-binding at the solution to the relaxed problem. As the funder's payoff is strictly decreasing in v, the optimal prize is zero. Setting $V = 0$, the point-wise first-order conditions are,

$$\theta Y\rho'(X)(W+\pi) - 1 - hY\rho'(X)\pi = 0, \tag{6}$$

and

$$Y = \left(\frac{\theta}{c}\rho(X)[W + \pi - h\pi]\right)^{\beta}. \tag{7}$$

Similar arguments to those in the proof of Proposition 1 reveal that X and Y are unique, and $X > \hat{x}$. Now, fix $\theta < \overline{\theta}$; we show $X(\theta) < X_{FB}(\theta)$ and $Y(\theta) < Y_{FB}(\theta)$. Combining (6) and (7):

$$\rho(X)^{\beta}\rho'(X)\frac{[\theta(W + \pi) - h\pi]^{\beta+1}}{c^{\beta}} = 1. \tag{8}$$

As $h(\theta) > 0$ for $\theta < \overline{\theta}$, (4) and (8) imply $\rho(X)^{\beta}\rho'(X) > \rho(X_{FB})^{\beta}\rho'(X_{FB})$. But since $\rho^{\beta}\rho'$ is strictly decreasing for $x > \hat{x}$, it follows that $X < X_{FB}$. Using (7), it is then straightforward to see that $Y(\theta) < Y_{FB}(\theta)$. Finally, as $h(\overline{\theta}) = 0$, (6) and (7) imply $X(\overline{\theta}) = X_{FB}(\overline{\theta})$ and $Y(\overline{\theta}) = Y_{FB}(\overline{\theta})$.

Next, we show that (IC-S) is satisfied. When $V = 0$, (IC-S) is satisfied if $X' \geq 0$. Applying the Implicit Function Theorem to (8):

$$\begin{aligned} &\rho(X)^{\beta-1}\left[\beta\rho'^2 + \rho''(X)\rho(X)\right](\theta(W + \pi) - h\pi)^{\beta+1} X' \\ &\qquad + \rho(X)^{\beta}\rho'(X)(\beta + 1)\left[\theta(W + \pi) - h\pi\right]^{\beta}\left[W + \pi - h'\pi\right] = 0. \end{aligned}$$

As $X > \hat{x}$, the term in square brackets on the first line above is strictly negative; moreover, the expression on the second line is strictly positive. It follows that $X' > 0$. It is then straightforward to confirm $Y' > 0$. This establishes part (ii).

We now establish part (i). (IR), $V = 0$ and expression (1) jointly imply $G > 0$. Next, $U(\underline{\theta}) = V(\underline{\theta}) = 0$ means, $X(\underline{\theta}) + \psi(Y(\underline{\theta})) - G(\underline{\theta}) = \underline{\theta}\rho(X(\underline{\theta}))\pi > 0$. Hence, $X(\underline{\theta}) + \psi(Y(\underline{\theta})) > G(\underline{\theta})$. Then, using the definition of $U(\theta)$, (IC-F) can be written:

$$X' + \psi'(Y)Y' - G' = \theta\pi\left[Y\rho'(X)X' + Y'\rho(X)\right] > 0. \tag{9}$$

We have shown $X(\underline{\theta}) + \psi(Y(\underline{\theta})) > G(\underline{\theta})$ and $X' + \psi'(Y)Y' > G'$, which means,

$X + \psi(Y) > G$. Finally, we show $G' > 0$. Fix θ; (1) and (6) imply $\theta W > h\pi$. Re-writing (7), $\theta\rho(X)\pi - \psi'(Y) = -\rho(X)(\theta W - h\pi) < 0$, and re-writing (9), $X'[\theta Y \rho'(X)\pi - 1] + Y'[\theta\rho(X)\pi - \psi'(Y)] + G' = 0$. Since $X', Y' > 0$, Assumption (1) and $\theta\rho(X)\pi - \psi'(Y) < 0$ imply $G' > 0$. □

Proof of Proposition 4

We will first relax the problem [P] by ignoring (IC-S); we will then show that it is satisfied at the solution to the relaxed problem.

The relaxed problem amounts to pointwise maximization of J, (where J is the integrand of the problem [P]) subject to, $x \geq 0$ and $v \geq 0$. Plugging in $y^* = \left(\frac{\theta}{c}\rho(x)(v+\pi)\right)^{\beta}$ into J, we obtain

$$J(x, v, \theta) = \left(\frac{\theta}{c}\right)^{\beta} (v+\pi)^{\beta} \rho(x)^{1+\beta} \left[\theta(W+\pi) - (v+\pi)\left(\theta\frac{\beta}{1+\beta} + h\right)\right] - x.$$

By assumption, the optimal investment level is interior ($X > 0$); therefore, the associated first-order conditions/complementary slackness conditions are:

$$J_x = (1+\beta)\left(\frac{\theta}{c}\right)^{\beta} (V+\pi)^{\beta} \rho(X)^{\beta} \rho'(X) \left[\theta(W+\pi) - (V+\pi)\left(\theta\frac{\beta}{1+\beta} + h\right)\right] = 1, \tag{10}$$

$$J_v = \left(\frac{\theta}{c}\right)^{\beta} \rho(X)^{1+\beta} (V+\pi)^{\beta-1} \left[\theta\beta(W+\pi) - (V+\pi)\big(\theta\beta + h(1+\beta)\big)\right] \leq 0, \text{ and} \tag{11}$$

$$V J_v = 0; \quad V \geq 0.$$

We divide the remainder of the proof into four parts. In part I, we prove part (i) of the proposition. In part II, we show that for $\theta > \theta_v$, $V > 0$, $V' > 0$ and $X = X_{FB|Y}$. In part III we show that for all θ, $X' > 0$, $Y' > 0$, $X \leq X_{FB}$, and $Y \leq Y_{FB}$ (with equality only at $\overline{\theta}$). Note that once we show $X' > 0$ and, when $V > 0$, $V' > 0$, we will have shown that $\frac{d}{d\theta}\rho(X(\theta))(V(\theta) + \pi) > 0$, which ensures that the solution to the relaxed problem satisfies (IC-S). In part IV, we show that for all θ, $G' < X'$ and $G < X$; we also show $G, G' < 0$ for θ sufficiently close to $\overline{\theta}$.

Part I: For $\theta \in [\underline{\theta}, \theta_v]$: $V = 0$, $G, G' > 0$ and $X \leq X_{FB|Y}$

In this section, we prove part (i) of the proposition. We assume $\theta_v > \underline{\theta}$, which means $\theta_v \beta W - h(\theta_v)(1+\beta)\pi = 0$. We first show $V = 0$ and $G > 0$ on $[\underline{\theta}, \theta_v]$. Let $k(\theta) \equiv \theta\beta W - h(\theta)(1+\beta)\pi > 0$, and note that $k(\theta_v) = 0$ and $k'(\theta) = \beta W - h'(\theta)(1+\beta)\pi > 0$. Now fix $\theta \in [\underline{\theta}, \theta_v]$. (11), together with the complementary slackness conditions, imply that $V > 0$ if and only if $k > 0$. But since $k(\theta_v) = 0$ and $k' > 0$, this means $k \leq 0$ and hence, $V = 0$. Then, as IR requires $u(\theta) = \theta\rho(X)Y\pi - X - \psi(Y) + G \geq 0$, and (1) implies $\theta\rho(X)Y\pi - X - \psi(Y) < 0$, we must have $G > 0$.

Next, we show $G' > 0$ on $[\underline{\theta}, \theta_v]$. (IC-F) together with IR ($u(\underline{\theta}) = 0$) imply that, at the optimal (X, V), $u(\theta) = \int_{\underline{\theta}}^{\theta} \rho(X(t))Y(t)(V(t)+\pi)$. By definition of u,

$$G(\theta) = \int_{\underline{\theta}}^{\theta} \rho(X(t))Y(t)(V(t)+\pi)dt + X(\theta) + \psi(Y(\theta)) - \theta\rho(X(\theta))Y(\theta)(V(\theta)+\pi). \tag{12}$$

Noting that $V = 0$ for $\theta \leq \theta_v$, $G'(\theta) = [1 - \theta\rho'(X)\pi]X'$. By (1), $G' > 0$ if $X' > 0$. We will show that $X' > 0$ in part III.

Finally, we show $X < X_{FB|Y}$ for $\theta < \theta_v$ and $X(\theta_v) = X_{FB|Y}(\theta_v)$. Fix $\theta \in [\underline{\theta}, \theta_v)$. $X_{FB|Y}$ satisfies $S_x(X_{FB|Y}, Y, \theta) = \theta Y \rho'(X_{FB|Y})(W+\pi) - 1 = 0$. By strict concavity of ρ, $X < X_{FB|Y}$ if $S_x(X, Y, \theta) > 0$. Using (10) and plugging in $V = 0$ and $Y = \left(\frac{\theta}{c}\rho(X)\pi\right)^{\beta}$, we obtain $J_x(X, 0, \theta) = \theta Y \rho'(X)(W+\pi) - 1 + Y\rho'(X)k(\theta) = 0$; equivalently, $J_x(X, 0, \theta) = S_x(X, Y, \theta) + Y\rho'(X)k(\theta) = 0$. $k(\theta) < 0$ for $\theta < \theta_v$, implies $S_x(X, Y, \theta) > 0$; thus $X < X_{FB|Y}$. But $k(\theta_v) = 0$ implies $S_x(X(\theta_v), Y(\theta_v), \theta_v) = 0$; so $X(\theta_v) = X_{FB|Y}(\theta_v)$.

Part II: For $\theta \in (\theta_v, \overline{\theta}]$: $V, V' > 0$ and $X = X_{FB|Y}$

We first show $V, V' > 0$ on $\theta \in (\theta_v, \overline{\theta}]$; we then show $X = X_{FB|Y}$. In part I we showed $V > 0$ if and only if $k = \theta\beta W - h(\theta)(1+\beta)\pi > 0$. By definition, $k(\theta_v) \geq 0$; since $k' > 0$, $k > 0$ for $\theta > \theta_v$, and hence, $V > 0$ for $\theta > \theta_v$. When $V > 0$, (11), together with the complementary slackness conditions, imply $V = \frac{\theta\beta W - h(1+\beta)\pi}{\theta\beta + h(1+\beta)}$. It follows that $V' = \frac{\beta(1+\beta)(W+\pi)(h(\theta)-h'(\theta))}{\theta\beta+h(\theta)(1+\beta)} > 0$, where the inequality holds since $h' < 0$. Plugging $Y = \left(\frac{\theta}{c}\rho(X)(V+\pi)\right)^{\beta}$ and $V = \frac{\theta\beta W - h(1+\beta)\pi}{\theta\beta + h(1+\beta)}$ into (10) yields $J_x = \theta Y \rho'(X)(W+\pi) - 1 = 0$; hence, $X = X_{FB|Y}$.

Part III: For all $\theta \in [\underline{\theta}, \overline{\theta}]$, $X' > 0$, $Y' > 0$, $X \leq X_{FB}$, and $Y \leq Y_{FB}$

We first show $X' > 0$ for all θ. To do so, we will apply the Implicit Function Theorem. For each θ, X satisfies $J_x(X, V, \theta) = 0$. Since J_x and V are differentiable, then so long as $J_{xx}(X, V, \theta) \neq 0$, the Implicit Function Theorem implies X is differentiable and $X' = -\frac{J_{xv}(X,V,\theta)V' + J_{x\theta}(X,V,\theta)}{J_{xx}(X,V,\theta)}$. We will now show that $J_{xx}(X, V, \theta) < 0$ for all θ. For an arbitrary x, see that,

$$J_{xx}(x, V, \theta) = \left(\frac{\theta}{c}\right)^{\beta} (V+\pi)^{\beta}(1+\beta)\rho(x)^{\beta-1}\left[\rho(x)\rho''(x) + \beta\rho'(x)^2\right]\left[\theta(W+\pi) - (V+\pi)\left(\theta\frac{\beta}{1+\beta} + h\right)\right].$$

(10) implies that $\left[\theta(W+\pi) - (V+\pi)\left(\theta\frac{\beta}{1+\beta} + h\right)\right] > 0$, and so $J_{xx}(x, V, \theta) \leq (<)0$ if and only if $\rho(x)\rho''(x) + \beta\rho'(x)^2 \leq (<)0$; equivalently, $X \geq (>)\hat{x}$. We claim that $X > \hat{x}$. To establish the claim, first note that at each θ, a second-order necessary condition is, $J_{xx}(X, V, \theta) \leq 0$; equivalently, as seen above, $X \geq \hat{x}$. To establish the claim, it therefore suffices to show $X \neq \hat{x}$. We proceed by contradiction; suppose, at some θ, $X = \hat{x}$. Then, $J_x(\hat{x}, V, \theta) = 0$. But for $x < \hat{x}$, $J_{xx}(x, V, \theta) > 0$, which means $J_x(x, 0, \theta) < 0$ for all $x < \hat{x}$. Since $J(0, V, \theta) = 0$, this means $J(x, V, \theta) < 0$ for any $x \leq \hat{x}$. In particular, $J(\hat{x}, V, \theta) < 0 = J(0, V, \theta)$, which contradicts the optimality of $\hat{x}$. Therefore, $X > \hat{x}$ and hence $J_{xx}(X, V, \theta) < 0$.

Since $J_{xx}(X,V,\theta) < 0$, $X' > 0$ if $J_{xv}(X,V,\theta)V' + J_{x\theta}(X,V,\theta) > 0$. We first show $J_{xv}(X,V,\theta)V' = 0$ for all θ. We then show $J_{x\theta}(X,V,\theta) > 0$. See that,

$$J_{xv}(X,V,\theta) = (1+\beta)\left(\frac{\theta}{c}\right)^{\beta}(V+\pi)^{\beta-1}\rho(X)^{\beta}\rho'(X)\big[\beta\theta(W+\pi) - (V+\pi)\left(\theta\beta + (1+\beta)h\right)\big].$$

As shown in part II, when $V > 0$, $V = \frac{\theta\beta W - h\pi(1+\beta)}{\theta\beta+(1+\beta)h}$; plugging this into the expression above yields $J_{xv}(X,V,\theta) = 0$. If $\theta_v > \underline{\theta}$, then plugging θ_v and $V(\theta_v) = 0$ into the expression above yields, $J_{xv}(X(\theta_v), V(\theta_v), \theta_v) = 0$. Thus, for $\theta \geq \theta_v$, $J_{xv}(X,V,\theta) = 0$. Then since $V' = 0$ for $\theta \in [\underline{\theta}, \theta_v)$, $J_{xv}(X,V,\theta)V' = 0$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$. Next, we show $J_{x\theta}(X,V,\theta) > 0$. See that,

$$\begin{aligned} J_{x\theta}(X,V,\theta) = {} & \beta(1+\beta)\frac{\theta^{\beta-1}}{c^{\beta}}(V+\pi)^{\beta}\rho(X)^{\beta}\rho'(X)\left[\theta(W+\pi) - (V+\pi)\left(\theta\frac{\beta}{1+\beta} + h\right)\right] \\ & + (1+\beta)\left(\frac{\theta}{c}\right)^{\beta}(V+\pi)^{\beta}\rho(X)^{\beta}\rho'(X)\left[W + \pi - (V+\pi)\frac{\beta}{1+\beta} - (V+\pi)h'\right]. \end{aligned}$$

The term in square brackets in the first line is strictly positive by (10). The term in square brackets on the second line is strictly positive since $h' < 0$, and $V = \max\left\{0, \frac{\theta\beta W - h(1+\beta)\pi}{\theta\beta + h(1+\beta)}\right\} \leq W$, which means $W + \pi - (V+\pi)\frac{\beta}{1+\beta} > 0$. Hence, $J_{x\theta}(X,V,\theta) > 0$. This establishes that $X' > 0$. As $Y = \left(\frac{\theta}{c}\rho(X)(V+\pi)\right)^{\beta}$, since $X' > 0$ and $V' \geq 0$, clearly $Y' > 0$. Additionally, note that $X' > 0$ and $V' \geq 0$ imply that the solution to the relaxed problem satisfies (IC-S).

We will now show that $X(\theta) \leq X_{FB}(\theta)$ and $Y(\theta) \leq Y_{FB}(\theta)$, with equality only at $\overline{\theta}$. Fix $\theta < \overline{\theta}$. Since $\rho^{\beta}\rho'$ is strictly decreasing for all $x > \hat{x}$ and since $X > \hat{x}$ (as shown previously in this part) and $X_{FB} > \hat{x}$ (as shown in the proof of Proposition 1), $X < X_{FB}$ if and only if $\rho(X)^{\beta}\rho'(X) > \rho(X_{FB})^{\beta}\rho'(X_{FB})$. Using equations (4) and (10), it holds that $\rho(X)^{\beta}\rho'(X) > \rho(X_{FB})^{\beta}\rho'(X_{FB})$ if and only if, $\Gamma(V) > 0$, where

$$\Gamma(v) = \frac{\theta^{1+\beta}}{c^\beta}(W+\pi)^{1+\beta} - (1+\beta)\left(\frac{\theta}{c}\right)^\beta (v+\pi)^\beta \left[\theta(W+\pi) - (v+\pi)\left(\theta\frac{\beta}{1+\beta} + h\right)\right].$$

First, suppose $V > 0$. Plugging $V = \frac{\theta\beta W - h(1+\beta)\pi}{\theta\beta + h(1+\beta)}$ into Γ,

$$\begin{aligned}
\Gamma(V) &= \frac{\theta^{1+\beta}}{c^\beta}(W+\pi)^{1+\beta} \\
&\quad -(1+\beta)\left(\frac{\theta}{c}\right)^\beta \left[\frac{\theta\beta(W+\pi)}{\theta\beta + h(1+\beta)}\right]^\beta \left[\theta(W+\pi) - \left(\frac{\theta\beta(W+\pi)}{\theta\beta + h(1+\beta)}\right)\left(\frac{\theta\beta + h(1+\beta)}{1+\beta}\right)\right] \\
&= \frac{\theta^{1+\beta}}{c^\beta}(W+\pi)^{1+\beta} - \frac{\theta^{1+\beta}}{c^\beta}\left[\frac{\theta\beta}{\theta\beta + h(1+\beta)}\right]^\beta (W+\pi)^{1+\beta} \\
&> \frac{\theta^{1+\beta}}{c^\beta}(W+\pi)^{1+\beta} - \frac{\theta^{1+\beta}}{c^\beta}(W+\pi)^{1+\beta} = 0.
\end{aligned}$$

The strict inequality holds since $\left[\frac{\theta\beta}{\theta\beta + h(1+\beta)}\right]^\beta < 1$ for $\theta < \overline{\theta}$. Thus, for any $\theta \in (\theta_v, \overline{\theta})$, $X(\theta) < X_{FB}(\theta)$. Next, we show $X < X_{FB}$ when $V = 0$. Suppose $\theta_v > \underline{\theta}$ and fix $\theta \in [\underline{\theta}, \theta_v]$.

$$\begin{aligned}
\Gamma(V) &= \frac{\theta^{1+\beta}}{c^\beta}(W+\pi)^{1+\beta} - (1+\beta)\left(\frac{\theta}{c}\right)^\beta \pi^\beta \left[\theta(W+\pi) - \pi\left(\theta\frac{\beta}{1+\beta} + h\right)\right] \\
&= \frac{\theta^{1+\beta}}{c^\beta}(W+\pi)^{1+\beta} - \left(\frac{\theta}{c}\right)^\beta \pi^\beta \left[\theta(W+\pi) + \theta\beta W - h\pi(1+\beta)\right] \\
&\geq \frac{\theta^{1+\beta}}{c^\beta}(W+\pi)^{1+\beta} - \frac{\theta^{1+\beta}}{c^\beta}\pi^\beta(W+\pi) > 0
\end{aligned}$$

The weak inequality holds since $\theta\beta W - h\pi(1+\beta) \leq 0$ for $\theta \in [\underline{\theta}, \theta_v]$. The strict inequality holds since $\pi^\beta(W+\pi) < (W+\pi)^{1+\beta}$. We have now shown that $X < X_{FB}$ for all $\theta < \overline{\theta}$. To see that $X(\overline{\theta}) = X_{FB}(\overline{\theta})$, note that $V(\overline{\theta}) = W$ and $h(\overline{\theta}) = 0$; plugging these into (10), $X(\overline{\theta})$ satisfies, $\frac{\overline{\theta}^{1+\beta}}{c^\beta}\rho(X(\overline{\theta}))^\beta \rho'(X(\overline{\theta}))(W+\pi)^{1+\beta} = 1$. From (4), it follows that $X(\overline{\theta}) = X_{FB}(\overline{\theta})$.

Finally, we show $Y \leq Y_{FB}$, with equality only at $\overline{\theta}$. For $\theta < \overline{\theta}$, $X < X_{FB}$

and $V < W$; thus, $Y_{FB} = \left(\frac{\theta}{c}\rho(X_{FB})(W+\pi)\right)^{\beta} > \left(\frac{\theta}{c}\rho(X)(V+\pi)\right)^{\beta} = Y$. But since $X(\overline{\theta}) = X_{FB}(\overline{\theta})$ and $V(\overline{\theta}) = W$, $Y(\overline{\theta}) = Y_{FB}(\overline{\theta})$.

Part IV: $X' > G'$, $X > G$, and $G, G' < 0$ for $\theta \in [\theta_g, \overline{\theta}]$, where $\theta_g < \overline{\theta}$

We first show $X' > G'$ and $X > G$. By (12), $G' = [1 - \theta Y\rho'(X)(V+\pi)]\, X' - \theta\rho(X)YV'$, which means $G' - X' = -\theta Y\rho'(X)(V+\pi)X' - \theta\rho(X)YV' < 0$; the inequality holds since $X' > 0$ and $V' \geq 0$. Thus, $G' < X'$. Now note that, $G(\underline{\theta}) - X(\underline{\theta}) = \psi(Y(\underline{\theta})) - \underline{\theta}Y(\underline{\theta})\rho(X(\underline{\theta}))(V(\underline{\theta})+\pi)$. Plugging in $Y(\underline{\theta}) = \left(\frac{\underline{\theta}}{c}\rho(X(\underline{\theta}))(V(\underline{\theta})+\pi)\right)^{\beta}$: $G(\underline{\theta}) - X(\underline{\theta}) = -\frac{1}{1+\beta}\left(\frac{\underline{\theta}}{c}\rho(X(\underline{\theta}))(V(\underline{\theta})+\pi)\right)^{1+\beta} < 0$. $G(\underline{\theta}) < X(\underline{\theta})$ and $G' < X'$ imply $G < X$ for all θ.

We now show $G < 0$ for θ sufficiently close to $\overline{\theta}$. To do so, we will show $G(\overline{\theta}) < 0$. Plugging $X(\overline{\theta}) = X_{FB}(\overline{\theta})$, $Y(\overline{\theta}) = Y_{FB}(\overline{\theta})$, and $V(\overline{\theta}) = W$ into (12) yields, $G(\overline{\theta}) = \int_{\underline{\theta}}^{\overline{\theta}} \rho(X(t))Y(t)(V(t)+\pi)dt - S^*(\overline{\theta})$, where $S^*(\theta) = S(X_{FB}(\theta), Y_{FB}(\theta), \theta)$. For $\theta < \overline{\theta}$, $X(\theta) < X_{FB}(\theta)$, $Y(\theta) < Y_{FB}(\overline{\theta})$, and $V(\theta) < W$, which means $G(\overline{\theta}) < \int_{\underline{\theta}}^{\overline{\theta}} \rho(X_{FB}(t))Y_{FB}(t)(W+\pi)dt - S^*(\overline{\theta})$. By the Envelope Theorem, $S^{*\prime}(\theta) = \rho(X_{FB}(\theta))Y_{FB}(\theta)(W+\pi)$, and hence, $S^*(\overline{\theta}) = S^*(\underline{\theta}) + \int_{\underline{\theta}}^{\overline{\theta}} \rho(X_{FB}(t))Y_{FB}(t)(W+\pi)dt$. Therefore, $G(\overline{\theta}) < \int_{\underline{\theta}}^{\overline{\theta}} \rho(X_{FB}(t))Y_{FB}(t)(W+\pi)dt - S^*(\overline{\theta}) = S^*(\overline{\theta}) - S^*(\underline{\theta}) - S^*(\overline{\theta}) = -S^*(\underline{\theta}) < 0$. Since $G(\overline{\theta}) < 0$, by continuity, there exists some $\theta_{g1} < \overline{\theta}$, such that $\theta > \theta_{g1}$ implies $G(\theta) < 0$.

Finally, we show $G'(\theta) < 0$ for θ sufficiently high. As shown previously in this section, $G'(\theta) = [1 - \theta Y(\theta)\rho'(X(\theta))(V(\theta)+\pi)]\, X'(\theta) - \theta Y(\theta)\rho(X(\theta))V'(\theta)$. Note that $1 - \overline{\theta}Y(\overline{\theta})\rho'(X(\overline{\theta}))(V(\overline{\theta})+\pi) = 1 - \overline{\theta}Y_{FB}(\overline{\theta})\rho'(X_{FB}(\overline{\theta}))(W+\pi) = 0$. Since X' is finite, and $1 - \theta Y(\theta)\rho'(X(\theta))(V(\theta)+\pi)$ is continuous, the term, $X'(\theta)[1 - \theta Y(\theta)\rho'(X(\theta))(V(\theta)+\pi)]$, will be arbitrarily close to zero for θ sufficiently close to $\overline{\theta}$. Moreover, for any $\theta \in (\theta_v, \overline{\theta}]$, the term, $\theta Y(\theta)\rho(X(\theta))V'(\theta)$, is strictly positive, and bounded away from zero. Therefore, there must exist some $\theta_{g2} < \overline{\theta}$ such that $\theta > \theta_{g2}$ implies $G'(\theta) < 0$. Letting $\theta_g = \max\{\theta_{g1}, \theta_{g2}\}$, for $\theta > \theta_g$, $G(\theta) < 0$ and $G'(\theta) < 0$.

Proof of Proposition 5

Fix $\theta < \overline{\theta}$ and let $L(\theta; t) = \frac{\theta}{h(\theta;t)}$. Note that $\frac{\partial L}{\partial \theta} > 0$ and $\frac{\partial L}{\partial t} < 0$. By definition, $L(\theta_v; t) = \frac{(1+\beta)\pi}{\beta W}$. Applying the Implicit Function Theorem, it is straightforward to establish the properties of θ_v given in the Proposition. □

Proof of Proposition 6

By Proposition 4, when $V(\theta) > 0$, $V(\theta) = \frac{\beta\theta W - h(\theta;t)\pi(1+\beta)}{\theta\beta + h(\theta;t)(1+\beta)}$. Differentiating $V(\theta)$ with respect to β, W, π, and t the comparative statics given in the Proposition follow. □

References

Bajari, P. and S. Tadelis (2001). Incentives versus transaction costs: a theory of procurement contracts. *RAND Journal of Economics 32*(3), 387–407.

Baker, G. P. (1992). Incentive Contracts and Performance Measurement. *Journal of Political Economy*, 598 – 614.

Bolton, P. and M. Dewatripont (2005). *Contract Theory*. MIT press.

Che, Y.-K., E. Iossa, and P. Rey (2015). Prizes Versus Contracts as Incentives for Innovation. CEIS Working Paper No. 358.

Chen, B. (2010). All-or-Nothing Monitoring: Comment. *The American Economic Review 100*(1), 625–627.

Chen, B. (2012). All-or-Nothing Payments. *Journal of Mathematical Economics 48*(3), 133–142.

Cockburn, I. M. and R. M. Henderson (2001). Scale and scope in drug development: unpacking the advantages of size in pharmaceutical research. *Journal of Health Economics 20*, 1033–1057.

Cornelli, F. and M. Schankerman (1999). Patent Renewals and R&D Incentives. *The RAND Journal of Economics 30*, 197–213.

Fu, Q., J. Lu, and Y. Lu (2012). Incentivizing R&D: Prize or Subsidies? *International Journal of Industrial Organization 30*(1), 67–79.

Gilbert, R. and C. Shapiro (1990). Optimal Patent Length and Breadth. *The RAND Journal of Economics 21*, 106 – 112.

Girth, A. M. and L. E. Lopez (2018). Contract Design, Complexity, and Incentives: Evidence From U.S. Federal Agencies. *The American Review of Public Administration 62*, 1–13.

Grace, C. and M. Kyle (2009). Comparative Advantages of Push and Pull Incentives for Technology Development: Lessons for Neglected Disease Technology Development. *Global Forum Update on Research for Health 6*, 147–151.

Grady, C. M. (2016). *Guidance on Using Incentive and Other Contract Types*. Washington DC: U.S. Department of Defense.

Grossman, S. J. and O. D. Hart (1983). An Analysis of the Principal-Agent Problem. *Econometrica*, 7–45.

Hall, B. H. (2007). Patents and Patent Policy. *Oxford Review of Economic Policy 23*(4), 568–587.

Hall, B. H., J. Mairesse, and P. Mohnen (2009). Measuring the Returns to R&D. National Bureau of Economic Research w15622.

Hölmstrom, B. and P. Milgrom (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law 7*, 24–52.

Hopenhayn, H., G. Llobet, and M. Mitchell (2006). Rewarding Sequential Innovators: Prizes, Patents, and Buyouts. *Journal of Political Economy 114*(6), 1041–1068.

Hopenhayn, H. A. and M. F. Mitchell (2001). Innovation Variety and Patent Breadth. *The RAND Journal of Economics 32*, 152 – 166.

Innes, R. D. (1990). Limited Liability and Incentive Contracting with Ex-Ante Action Choices. *Journal of Economic Theory 52*(1), 45–67.

Klemperer, P. (1990). How Broad Should the Scope of Patent Protection Be? *The RAND Journal of Economics 21*, 113–130.

Kremer, M. (1998). Patent Buyouts: A Mechanism for Encouraging Innovation. *The Quarterly Journal of Economics 113*(4), 1137–1167.

Kremer, M. (2002). Pharmaceuticals and the Developing World. *The Journal of Economic Perspectives 16*(4), 67–90.

Kremer, M. and R. Glennerster (2004). *Strong Medicine: Creating Incentives for Pharmaceutical Research on Neglected Diseases*. Princeton University Press.

Laffont, J. J. (1995). Regulation, Moral Hazard and Insurance of Environmental Risks. *Journal of Public Economics 58*(3), 319–336.

Laffont, J. J. and D. Martimort (2009). *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press.

Laffont, J. J. and J. Tirole (1986). Using Cost Observation to Regulate Firms. *The Journal of Political Economy*, 614–641.

Laffont, J. J. and J. Tirole (1993). *A Theory of Incentives in Procurement and Regulation*. Cambridge, Mass: MIT Press.

Lazear, E. P. (1986). Salaries and Piece Rates. *Journal of Business*, 405–431.

Lazear, E. P. (2000). The Power of Incentives. *American Economic Review 90*(2), 410–414.

Lewis, T. R. and D. Sappington (1989). Countervailing Incentives in Agency Problems. *Journal of Economic Theory 49*(2), 294–313.

Lewis, T. R. and D. E. Sappington (2000a). Contracting With Wealth-Constrained Agents. *International Economic Review 41*(3), 743–767.

Lewis, T. R. and D. E. Sappington (2000b). Motivating Wealth-Constrained Actors. *American Economic Review 90*(4), 944–960.

Lichtenberg, F. R. and J. Waldfogel (2003). Does Misery Love Company - Evidence from Pharmaceutical Markets before and after the Orphan Drug Act. *National Bureau of Economic Research*, 1–24.

Link, A. N. and C. J. Ruhm (2009, June). Bringing science to market: commercializing from NIH SBIR awards. *Economics of Innovation and New Technology 18*(4), 381–402.

Maskin, E. and J. Riley (1984). Monopoly with incomplete information. *The RAND Journal of Economics 15*, 171–196.

Maskin, E. and J. Riley (1985). Input Versus Output Incentive Schemes. *Journal of Public Economics 28*, 1–23.

Maurer, S. M. and S. Scotchmer (2003). Procuring Knowledge. National Bureau of Economic Research w9903.

Meng, D. and G. Tian (2013). Multi-task Incentive Contract and Performance Measurement with Multidimensional Types. *Games and Economic Behavior 77*(1), 1–45.

National Research Council (2001). The advanced technology program: Assessing outcomes. Washington DC: The National Academies Press. https://doi.org/10.17226/10145.

National Science Board (2009). Investing in the future: NSF cost sharing policies for a robust federal research enterprise.

O'donoghue, T., S. Scotchmer, and J.-F. Thisse (1998). Patent Breadth, Patent Life, and the Pace of Technological Progress. *Journal of Economics & Management Strategy 7*(1), 1–32.

Ollier, S. and L. Thomas (2013). Ex Post Participation Constraint in a Principal-Agent Model with Adverse Selection and Moral Hazard. *Journal of Economic Theory 148*(6), 2383–2403.

Poblete, J. and D. Spulber (2012). The Form of Incentive Contracts: Agency with Moral hazard, Risk Neutrality, and Limited liability. *The RAND Journal of Economics 43*(2), 215–234.

Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of economic literature 37*(1), 7–63.

Prendergast, C. (2002). The Tenuous Trade-off between Risk and Incentives. *Journal of Political Economy 110*(5), 1071–1102.

Renwick, M. J., V. Simpkin, and E. Mossialos (2016). *Targeting innovation in antibiotic drug discovery and development: The need for a One Health–One Europe–One World Framework*. European Observatory on Health Systems and Policies, Copenhagen (Denmark).

Rietzke, D. and Y. Chen (2016). Push or Pull? Performance Pay, Incentives, and Information. Lancaster University, Department of Economics Working Paper Series, No. 2016/015.

Scotchmer, S. (2004). *Innovation and Incentives*. MIT press.

Shavell, S. and T. Van Ypersele (2001). Rewards Versus Intellectual Property Rights. *Journal of Law and Economics 44*, 525 – 547.

Sloan, F. A. and C.-R. Hsieh (2007). *Pharmaceutical innovation: incentives, competition, and cost-benefit analysis in international perspective*. Cambridge University Press.

Weyl, E. G. and J. Tirole (2012). Market Power Screens Willingness-to-Pay. *Quarterly Journal of Economics 127*, 1971 – 2003.

Wright, B.-D. (1983). The Economics of Invention Incentives: Patents, Prizes, and Research Contracts. *The American Economic Review* (73), 691 – 707.

Zhao, R. R. (2008). All-or-Nothing Monitoring. *The American Economic Review 98*(4), 1619–1628.