2011

Subjective Performance and the Value of Blind Evaluation

Curtis R. Taylor, *Duke University*
Huseyin Yildirim, *Duke University*

Review of Economic Studies (2010) **01**, 1–37 0034-6527/10/00000001$02.00

Subjective Performance and the Value of Blind Evaluation[1]

CURTIS R. TAYLOR

Duke University

HUSEYIN YILDIRIM

Duke University

First version received November 2008; final version accepted June 2010 (Eds.)

The incentive and project selection effects of agent anonymity are investigated in a setting where an evaluator observes a subjective signal of project quality. Although the evaluator cannot commit *ex ante* to an acceptance criterion, she decides up front between *informed review*, where the agent's ability is directly observable, or *blind review*, where it is not. An ideal acceptance criterion balances the goals of incentive provision and project selection. Relative to this, informed review results in an excessively steep equilibrium acceptance policy: the standard applied to low-ability agents is too stringent and the standard applied to high ability agents is too lenient. Blind review, in which all types face the same standard, often provides better incentives, but it ignores valuable information for selecting projects. The evaluator prefers a policy of blind (resp. informed) review when the ability distribution puts more weight on high (resp. low) types, the agent's payoff from acceptance is high (resp. low), or the quality signal is precise (resp. imprecise). Applications discussed include the admissibility of character evidence in criminal trials, and academic refereeing.

"If Justice is pictured blindfold, it is because she judges causes, not men, and not because the prime faculty of an arbitrator is lack of discernment."

–Charles Wagner, *Justice*, 1905, p. 133.

1. INTRODUCTION

The settings in which an evaluator must rely only on her subjective impressions to judge output or performance are ubiquitous. In cultural environments individuals are asked to evaluate wine, food, art, poetry, movies, and music. In retail settings experts and panel participants review a vast array of consumer products. In criminal trials and lawsuits, juries are charged with weighing evidence, and in academia faculty evaluate exams, manuscripts, and grant proposals. Given that subjective evaluation is endemic to so many significant situations, it is important to understand what elements add or detract from its efficacy. A key question in this regard is whether or not the reviewer should be permitted to use supplemental information such as the applicant's identity and prior record in the current evaluation, i.e., should the reviewer be "informed" or "blind"?

1. The authors thank Andrea Prat, four anonymous referees, and numerous seminar participants for helpful comments. They also benefited from suggestions by Atila Abdulkadiroğlu, Peter Arcidiacono, Jeremy Burke, Yeon-Koo Che, Andrew Daughety, Preston McAfee, Marco Ottaviani, and Jennifer Reinganum. Sergiu Ungureanu provided valuable research assistance. The authors are responsible for any remaining errors.

At first glance, the answer to this question may seem obvious: in situations where an individual's current output or performance – and not his innate ability – is the object of evaluation, the review process should be blind whenever feasible in order to minimize bias.[1] Note, however, that not all "bias" is undesirable. Because evaluation is often noisy, an effective use of information may dictate that individuals with stronger track records face lower standards.

In fact, the mode of review, blind or informed, varies both across and within evaluation settings. Wine tasting, for example, is virtually always performed blindly.[2] Similarly, in classical music, Goldin and Rouse (2000) note that most major U.S. symphony orchestras adopted some form of blind auditioning for hiring new members in the 1970s and 80s. Likewise, nearly all licensing and competency examinations such as bar exams and medical board exams are scored blindly.[3]

There are numerous settings in which a mixture of blind and informed review procedures are employed. For instance, there is a long-standing controversy among legal scholars about when judges should allow juries to hear character evidence, what McCormick (1954, pp. 340–41) defines as "a generalized description of one's disposition, or of one's disposition in respect to a general trait such as honesty, temperance, or peacefulness." Prior to the advent of Internet search engines, there was a similar debate in academia about whether journal submissions should be reviewed blindly. Blank (1991) reports that among 38 well-known journals in chemistry, biology, physics, mathematics, history, psychology, political science, sociology, and anthropology, 11 used blind review, as did 16 of 38 major economics journals. In a more recent survey of 553 journals across 18 disciplines, Bachand and Sawallis (2003) find that 58% employed blind review.

There are also numerous settings in which informed review is the predominant mode of evaluation. For example, the identity of students is commonly known to the grader when evaluating course examinations and projects. Similarly, grant proposals typically contain not only a description of the project, but the academic track record of the primary investigator; i.e., his education, major publications, and previous grants.

Most extant studies on the effects of blind versus informed review (summarized in the next section) have been experimental or empirical. While revealing important insights, many of these investigations have presented conflicting evidence, making it difficult – in the absence of a coherent theory – to draw general conclusions or make consistent policy recommendations. In this paper we study a simple game-theoretic model that focuses on three common features of many review processes: (1) the *applicant* can improve the quality of his *project* by expending effort; (2) evaluation is a noisy process in which the reviewer observes only an imperfect subjective signal of quality; and (3) knowing the identity of the applicant would provide the reviewer with additional information about his ability to produce a high quality project.[4]

1. Of course, one reason to adopt blind review is to avoid taste-based discrimination such as racism or sexism (or charges thereof) by evaluators. While this is clearly an important consideration in many settings, the model presented here focuses on a type of statistical discrimination (Arrow, 1973) in which blind review forces an evaluator to ignore an agent's productivity.

2. See Taber (2005) for an account of the most famous blind tasting, the Judgment of Paris, held in 1976, which is accredited with putting California wines on the world stage.

3. Bar exams in the U.S. contain an essay section which involves subjectivity in grading. Similarly, the United States Medical Licensing Exam contains a Clinical Skills section on which medical students and graduates are evaluated according to several subjectively scored criteria.

4. Ottaviani and Wickelgren (2009) also investigate an evaluation setting but with quite different features; e.g., symmetric information and learning.

The applicant cares only about having his project accepted, while the evaluator is a Bayesian decision maker who weighs her payoffs from accepting good and bad projects.[5] The equilibrium of the model is examined under three regimes: *commitment* – which is an ideal benchmark setting where the quality signal is verifiable and the evaluator can credibly commit up front to an acceptance criterion, *informed review* – in which the evaluator observes the applicant's ability, and *blind review* – in which the applicant's ability is hidden. In all three cases the reviewer follows a simple equilibrium strategy: accept the project if and only if the quality signal is above a certain threshold, or *standard*.

Under informed review, the evaluator – not surprisingly – applies weak standards to high-ability applicants and tough standards to low-ability ones. In fact, these standards are *too* weak and *too* tough when compared with the ideal review process. In a sense, the benchmark process calls for a more "fair" standard across applicants, even though no direct preference for equity is assumed.

The reason the ideal review policy is flatter than the one implemented under informed review is that it is designed not only to select good projects, but also to provide incentives to produce them. Both weak and tough standards generate poor incentives, albeit for opposing reasons. The marginal return to effort is low to an agent who is either very likely to have his project accepted or very likely to have it rejected. The optimally designed acceptance policy thus creates better incentives for agents at both ends of the type distribution by raising the standards facing high-ability agents and lowering those facing low-ability ones. This policy, however, is not time-consistent. Once the applicant has invested effort in the project and submitted it for evaluation, the reviewer would prefer to renege and apply a steeper (informationally-efficient) acceptance policy. Hence, if the quality signal observed by the evaluator is not verifiable (e.g., because it is impossible or impractical to quantify), then it will not be possible for her to credibly implement the relatively flat ideal acceptance criterion. It may, however, be possible for her to commit to remain ignorant about the applicant's type and apply a completely flat standard; that is, to perform blind review.

Under blind review, the evaluator sets a uniform standard as if she were assessing an applicant of average ability. This policy provides good incentives for applicants at both ends of the type distribution, but blind review is also clearly suboptimal when compared with the ideal policy. Specifically, blind review does not allow the evaluator to use any information about applicant ability to mitigate noise in the review process.

Hence, both informed and blind review procedures are suboptimal, but for different reasons. On one hand, *ex post* project selection is better under informed review, on the other hand, *ex ante* incentives are often better under blind review. Thus, the evaluator's preference between review procedures will depend on the environment, especially on the distribution of ability in the applicant pool and the informational content of the quality signal. When the distribution of applicants contains a large proportion of high-ability agents, then assessing project quality is relatively less important than providing incentives, and the evaluator, therefore, prefers blind review. Conversely, when the applicant pool contains a large proportion of low-ability agents, then project selection is paramount and the evaluator prefers informed review. In a similar vein, when the signal on project quality is very precise, then observing the applicant's ability provides little additional information and blind review is optimal. On the other hand, when the

5. The model can be altered to allow the applicant to care about project quality so long as there is some residual incongruity between his payoffs and the evaluator's.

quality signal is very imprecise, then observing ability provides significant incremental information, and informed review is the preferred mode of evaluation.

The remainder of the paper is organized as follows. The relevant literature is reviewed in the next section. In Section 3 the basic model is presented. Sections 4, 5, and 6 contain the analysis of the commitment benchmark, informed review, and blind review settings respectively. In Section 7 the factors influencing the evaluators equilibrium choice of review policy are determined. We discuss two applications of the theory in Section 8. In subsection 8.1 we contribute to the debate on character evidence, arguing that our findings support its use in cases of blue-collar street crime such as robbery or assault but not in cases of white-collar corporate crime such as embezzlement or price fixing. In subsection 8.2 we offer an explanation of the seemingly paradoxical claim that it is optimal to use blind review to evaluate scholarly manuscripts and informed review to assess grant proposals. In Section 9 three generalizations of the basic model are analyzed: competition among evaluators (9.1), costly false rejections (9.2), and non-productive effort (9.3). Section 10 contains concluding remarks and a discussion of future work. The proofs of all propositions and lemmas are relegated to the Appendix.

2. RELATED LITERATURE

There is a large empirical and experimental literature on the impact of anonymity on the academic publication process, ably surveyed by Snodgrass (2006). In particular, papers by Blank (1991) in economics, Horrobin (1982) in modern languages, Link (1998) in medicine, Peters and Ceci (1982) in psychology, and Zuckerman and Merton (1971) in physics, found compelling evidence that informed review is likely to introduce status, gender, or geographical bias in evaluation of scholarly manuscripts.

In the 1970s and 80s, concern about gender-biased hiring caused most major U.S. symphony orchestras to adopt some form of blind auditioning. Goldin and Rouse (2000) estimate that the switch to blind auditions explains up to 25 percent of the increase in female orchestra musicians hired over the intervening years.

The theoretical literature on subjective performance evaluation is relatively small [e.g., Levin (2003) and MacLeod (2003)] and almost exclusively addresses contracting problems within an agency setting. This paper considers a complementary setting in which transfers between the parties are not allowed and the principal can decide to remain ignorant of the agent's ability. The perverse incentive effect associated with better information at the core of this paper is reminiscent of similar results found in career concern models, either in the form of reduced effort by the agent [e.g., Dewatripont *et al.* (1999), and Holmstrom (1999)], or in the form of concealing his private information [e.g., Morris (2001), and Prat (2005)]. Unlike in our static framework, the agent in these models cares about the principal's belief about his ability. In the same spirit, several papers such as Cremer (1995), Riordan (1990), and Sappington (1986) have highlighted the potential benefits of committing to an imperfect monitoring technology in a contracting environment. In a somewhat different context, Fryer and Loury (2005) also observe that restrictions on what information can be used for selection may have serious consequences for incentives. This study also contributes to the literature on discretion versus rules (e.g., Milgrom and Roberts (1988)), which emphasizes that commitment to even an imperfect institution is sometimes better than no commitment at all.

The paper also belongs to a growing literature on mechanism design without transfers, originating with the work by Holmstrom (1977 and 1984) in the case of full commitment, and by Crawford and Sobel (1982) in the case of no commitment.

More directly related to our investigation is the paper by Seidmann (2005). Seidmann introduces limited commitment and imperfectly verifiable messages to the Crawford and Sobel model to demonstrate that a "right to silence" may indirectly benefit the innocent by inducing the guilty to remain silent. Unlike these papers that consider hidden information, we explore an environment of hidden action and compare full commitment (the benchmark), no commitment (informed review), and commitment to an imperfect institution (blind review).

A final strand of related research is the literature on statistical discrimination, which recognizes the potential tension between fairness and efficiency. Papers by Norman (2003) and Persico (2002) demonstrate that a more fair treatment of different groups need not interfere with a socially efficient allocation of resources. Depending on the elasticity of each group's production function, to insist on a more equal treatment can also shift equilibrium production toward a more socially efficient level.

The paper most closely related to this one is Coate and Loury (1993). They study a model in which two identifiable groups that are *ex ante* identical invest in human capital. Employers receive noisy subjective signals regarding investment levels and decide who to hire. There are assumed to be multiple equilibria of the investment/evaluation game. Coate and Loury suppose that one group coordinates with employers on an equilibrium with a modest standard and higher investment, while the other group gets stuck in a Pareto inferior equilibrium with a high standard and low investment. In the setting investigated here, by contrast, agent ability is drawn from a continuum and represents real *ex ante* heterogeneity in productivity. Moreover, players are assumed to coordinate on the unique Pareto superior equilibrium. Coate and Loury argue that forcing employers to use the same standard across groups can correct inefficient coordination failure. The focus here is on a different but complementary question – when is it in the best interest of an evaluator to commit herself not to use fundamentally valuable information in the review process? The potential benefit of blind review in this context is not to break coordination failure, but to raise productivity at both ends of the ability spectrum by pooling incentives.

3. THE BASIC MODEL

There are two risk-neutral parties: an applicant (the agent) and an evaluator (the principal) who play a three-stage game. In the first stage, the principal commits to a review policy, which is either *informed* (she directly observes the agent's type) or *blind* (she does not observe the agent's type).

In the second stage, the agent, who knows the review policy and knows his own type θ, exerts effort $p \in [0,1]$ to prepare a *project* for review by the principal. The ultimate quality of the project is high ($q = h$) with probability p, or low ($q = l$) with probability $1 - p$.[6] The agent's effort cost is given by

$$C(p;\theta) = \frac{p^2}{2\theta},$$

6. The focus of this investigation is the tradeoff between provision of incentives and the efficient use of information. Although there are a number of ways of exploring this tradeoff, the setting considered here is probably the simplest. A version of the model with continuous quality yields similar results; see Taylor and Yildirim (2007).

where $\theta \in [\underline{\theta}, \overline{\theta}] \subset \mathbb{R}_+$.[7] Hence, θ is a measure of the agent's productivity and may represent either his innate ability or his experience.

In the final stage of the game the agent submits the project to the principal for evaluation. The principal does not observe p or q directly, but receives a subjective (i.e., non-verifiable) signal of quality, $\sigma \in [\underline{\sigma}, \overline{\sigma}]$. Based upon the outcome of this signal – and the agent's type if the review policy is informed – the principal decides whether to accept or reject the project.

The principal prefers to accept high-quality projects and to reject low-quality ones. In particular, her exogenous payoff from accepting a high-quality project is $v > 0$ and from accepting a low-quality one is $-c < 0$. Her payoff from rejecting a low-quality project is taken to be zero, which is sensible since she otherwise could get "something for nothing" in a degenerate equilibrium where she always rejects and the agent never exerts effort. The principal's cost from rejecting a high-quality project is also taken to be zero. This greatly simplifies the analysis and is reasonable in many settings. The more general case in which she also suffers a loss from a false rejection is, however, analyzed in subsection 9.2. It is notationally convenient to define the principal's cost benefit ratio from accepting a project by $r \equiv \frac{c}{v}$.

The agent prefers the project to be accepted, regardless of its underlying quality. Specifically, he receives an exogenous gross payoff of $u > 0$ if the principal accepts the project, and zero if she rejects it. No monetary transfers between the parties are permitted.

The agent's type, θ, is distributed according to the distribution function $G(\cdot)$, possessing density $g(\cdot)$ and finite mean $E[\theta]$. The signal, σ, is drawn from one of two distributions: $F_h(\cdot)$ (with density $f_h(\cdot)$) if project quality is high, or $F_l(\cdot)$ (with density $f_l(\cdot)$) if it is low. For analytical convenience, assume $f_q(\cdot)$ is bounded and twice differentiable. The likelihood ratio is defined by $L(\sigma) \equiv \frac{f_h(\sigma)}{f_l(\sigma)}$, and satisfies the following regularity conditions.

Assumption 1 (Signal Technology). *The Likelihood ratio satisfies:*

(i) $L'(\sigma) > 0$,
(ii) $L(\underline{\sigma}) = 0$ *and* $L(\overline{\sigma}) = \infty$,
(iii) $\lim_{\sigma \to \overline{\sigma}} L(\sigma)(1 - F_q(\sigma)) = \lambda_q$ *exists for* $q \in \{l, h\}$ *and* $\lambda_h > 0$.

Part (i) is the familiar monotone likelihood ratio property (MLRP) indicating that higher signals are associated with high project quality. Part (ii) says that the most extreme signals (which occur with probability zero) are perfectly informative. This ensures equilibrium existence. Part (iii) is a boundary condition used below to identify the set of agent types that exert zero effort in equilibrium. The requirement $\lambda_h > 0$ is necessary because all types of agent would otherwise exert zero effort under informed review. All aspects of the environment are common knowledge, and the solution concept is Pareto efficient Perfect Bayesian equilibrium (PBE).

Example 1 (Signals). *The following signal technology will be used in illustrative examples below. The principal's subjective signal, σ, is drawn from one of the two*

7. The functional form assumed for the effort cost is analytically helpful but not critical for the qualitative nature of the results presented below.

triangular densities on [0,1]: $f_h(\sigma) = 2\sigma$ or $f_l(\sigma) = 2(1-\sigma)$. This implies $L(\sigma) = \frac{\sigma}{1-\sigma}$, $\lambda_h = 2$, and $\lambda_l = 0$ in agreement with Assumption 1.

4. THE COMMITMENT BENCHMARK

The fundamental problem facing the principal is an inability to commit. Because her evaluation results in a non-verifiable assessment, once the principal observes σ, she will accept the project if an only if she expects a positive payoff. While this is optimal *ex post* (after the agent has sunk effort), it is not generally desirable from an *ex ante* perspective. To highlight this, the benchmark case of the principal's optimal review policy with commitment is characterized in this section.

To begin, note that there is no scope for blind review in this context. With full power of commitment, the principal can choose to ignore information whenever it is advantageous. Next, it is straightforward to verify that MLRP implies the principal optimally uses a *standard* $s(\theta)$ when conducting a review. That is, she accepts the project of a type θ agent if and only if $\sigma \geq s(\theta)$.

Given an arbitrary standard s, a type θ agent will choose p so as to maximize his expected payoff

$$U(p, s; \theta) = u[p(1 - F_h(s)) + (1-p)(1 - F_l(s))] - \frac{p^2}{2\theta}, \tag{4.1}$$

subject to the downward and upward feasibility restrictions, $p \geq 0$ and $p \leq 1$. The first term in (4.1) is the agent's benefit from acceptance u times the probability the project is accepted whether it is good $p(1 - F_h(s))$ or bad $(1-p)(1 - F_l(s))$, while the second term is his cost of effort, $C(p, \theta)$.

Combining the first-order condition with the upward feasibility restriction yields the agent's reaction function:

$$P(s, \theta) = \min\{\theta u(F_l(s) - F_h(s)),\ 1\}. \tag{4.2}$$

If the feasibility restriction does not bind, then the agent's reaction function is "hump-shaped" in s. To see this, note first that the most extreme standards elicit no effort at all, $P(\underline{\sigma}, \theta) = P(\overline{\sigma}, \theta) = 0$. Next, define the *neutral signal* by $s^* \equiv L^{-1}(1)$. Then

$$P_s(s, \theta) = \theta u(1 - L(s)) f_l(s)$$

is positive for $s < s^*$ (effort is increasing in the standard), and negative for $s > s^*$ (effort is decreasing in the standard). Low standards elicit little effort because projects are rarely rejected and high standards elicit little effort because they are rarely accepted. The agent exerts maximal effort when facing the intermediate standard s^*. Intuitively, if $\sigma = s^*$, then the posterior on project quality is the same as the prior. Setting the neutral standard, thus, minimizes bias in the evaluation process which maximizes the agent's incentives. (See Figure 1.) If θ is sufficiently high, however, then the upward feasibility restriction will bind and the agent will exert effort of $P(s, \theta) = 1$ over an intermediate range of standards containing s^*; i.e., the "hump" of the reaction function will be truncated to a "plateau."

Inducing the agent to exert effort is only part of the principal's objective. An optimal review policy must both provide incentives for the agent and select high-quality projects as often as possible. Specifically, the principal will commit herself to a standard that

maximizes her expected payoff[8]

$$V(s,p) = vp(1 - F_h(s)) - c(1-p)(1 - F_l(s)), \tag{4.3}$$

subject to the agent's reaction function in (4.2).

The first term in the principal's objective is her benefit from accepting a good project, v, times the probability the project is good and the standard is achieved, $p(1 - F_h(s))$, and the second term is her cost from accepting a bad project, $-c$, times the probability the project is bad and the standard is achieved, $(1-p)(1 - F_l(s))$.

Substituting the agent's reaction function into (4.3) results in the function

$$V(s, P(s,\theta)) = vP(s,\theta)(1 - F_h(s)) - c(1 - P(s,\theta))(1 - F_l(s)). \tag{4.4}$$

This function is continuous, and hence achieves a maximum on the compact interval $[\underline{\sigma}, \overline{\sigma}]$. The following assumption ensures sufficiency of the first-order condition.[9]

Assumption 2 (Single-Peaked Preferences). *The function $V(s, P(s,\theta))$ is strictly quasi-concave in s whenever $P(s,\theta) < 1$.*

Ignoring the feasibility restrictions for the moment and differentiating (4.4) with respect to s yields the first-order condition

$$\underbrace{V_s(s, P(s,\theta))}_{\text{Selection Effect}} + \underbrace{V_p(s, P(s,\theta))P_s(s,\theta)}_{\text{Incentive Effect}} = 0. \tag{4.5}$$

Denote the solution to this equation by $s_0^C(\theta)$. This is the optimal standard the principal would announce for an agent whose type θ fell in the range $[\theta_-^C, \theta_+^C]$, where the feasibility restrictions do not bind.

Equation (4.5) highlights the tradeoff facing the principal: selection versus incentives. The first term in (4.5) represents the selection effect. As discussed in the next section, setting this term alone equal to zero results in the standard, $s^I(\theta)$, that accepts a project if and only if it has positive expected value to the principal. The second term in (4.5) represents the incentive effect. As discussed above, setting this term alone equal to zero results in the neutral standard, s^*, that maximizes the agent's effort. In general, it is not possible to set both terms to zero simultaneously. In other words, there is tension between the efficient use of information and motivating the agent.

Define the endpoints of the interval $[\theta_-^C, \theta_+^C]$ by

$$\theta_-^C \equiv \frac{r}{u(2\lambda_h + (r-1)\lambda_l)}, \tag{4.6}$$

and

$$\theta_+^C \equiv \min\{\theta \mid P(s_0^C(\theta), \theta) = 1\}. \tag{4.7}$$

8. There are two situations to consider. The principal might announce the review policy either *before* or *after* observing the agent's type. Because the expectation over θ of $V(s, P(s,\theta))$ is separable in θ, the optimal review policy in either case is found by maximizing this function with respect to s for each value of $\theta \in [\underline{\theta}, \overline{\theta}]$.

9. It is easy to verify that if $L(s) \leq 1$, Assumption 2 is automatically satisfied; and if $L(s) > 1$, it is satisfied whenever

$$-\frac{(L(s) + \frac{r-1}{2})(1 - F_l(s))L'(s)}{(L(s)-1)(L(s)+r)} + f_l(s) < 0.$$

This inequality turns out not to be too stringent, owing to the assumptions that $f_l(s)$ is bounded, and $\lim_{s\to\overline{\sigma}}[L(s)(1 - F_l(s))]$ exists.

The following result characterizes the benchmark solution.

Proposition 1 (Equilibrium under Commitment).

Principal: *The principal sets the standard*

$$s^C(\theta) \equiv \begin{cases} \overline{\sigma}, & \text{if } \theta < \theta_-^C \\ s_0^C(\theta), & \text{if } \theta \in [\theta_-^C, \theta_+^C] \\ \min\{s \mid P(s,\theta) = 1\}, & \text{if } \theta > \theta_+^C \end{cases}$$

Moreover, $s^C(\theta)$ is strictly decreasing for $\theta > \theta_-^C$, and $\lim_{\theta\to\infty} s^C(\theta) = \underline{\sigma}$.

Agent: *The agent chooses effort level*

$$p^C(\theta) \equiv \begin{cases} 0, & \text{if } \theta < \theta_-^C \\ P(s_0^C(\theta), \theta), & \text{if } \theta \in [\theta_-^C, \theta_+^C] \\ 1, & \text{if } \theta > \theta_+^C. \end{cases}$$

Moreover, $p^C(\theta)$ is continuous, and strictly increasing for $\theta \in (\theta_-^C, \theta_+^C)$.

The selection effect results in a negative relationship between an agent's ability and the standard set for him. Higher ability agents are more likely to produce good projects, so the principal accordingly lowers the standard confronting them. This response is, however, attenuated by the incentive effect. In order to elicit more effort, the principal commits to a profile of standards, $s^C(\theta)$, that is "too flat" to be ex post optimal. In other words, under $s^C(\theta)$ the principal may be forced to reject the project of a high ability agent or accept the project of a low ability one when she would prefer to do otherwise. This is most starkly illustrated for ability levels $\theta \geq \theta_+^C$. For these high-ability types, the upward feasibility restriction binds ($p^C(\theta) = 1$). Nevertheless, the standard facing such an agent, $s^C(\theta)$, is greater than the minimum standard, $\underline{\sigma}$. Hence, even though the principal *knows for sure* that the project is good, she commits herself to reject it with positive probability. Only by doing so can she induce the agent to exert effort in the first place.

At the other end of the spectrum are the low ability agents with $\theta \leq \theta_-^C$. These types of agents are effectively *pre-screened* in the sense that the principal commits never to accept their projects; i.e., $s^C(\theta) = \overline{\sigma}$. Consequently, these types exert no effort, so the downward feasibility restriction binds ($p^C(\theta) = 0$).

Finally, it is straightforward to verify that $s^* \in (s^C(\theta_+^C), \overline{\sigma})$. So, there exists a unique critical type

$$\theta^* \equiv \frac{r}{(1+r)u(F_l(s^*) - F_h(s^*))} \tag{4.8}$$

such that $s^C(\theta^*) = s^*$. For this one type of agent there is no conflict between selection and incentives. In particular, the standard, s^*, set for type θ^* both induces maximal effort and leads to an *ex post* optimal acceptance decision. For all other types, $\theta \neq \theta^*$, however, the benchmark solution $s^C(\theta)$ strikes a balance between selecting good projects *ex post* and providing incentives *ex ante*.

Example 2 (Commitment). *Suppose the signal technology of Example 1 and that $v = c = u = 1$.*[10] *From (4.2), the agent's reaction function is*

$$P(s,\theta) = \min\{2\theta s(1-s),\ 1\}.$$

10. For a fully parametric example, see Taylor and Yildirim (2007).

For $\theta < 2$ this is hump-shaped and attains a maximum at $s^* = \frac{1}{2}$. From (4.3), the principal's payoff is

$$V(s,p) = p(1-s^2) - (1-p)(1-s)^2.$$

Substituting $P(s,\theta)$ into this and maximizing yields the commitment solution

$$s^C(\theta) = \begin{cases} 1, & \text{if } \theta < \theta_-^C \\ \frac{1}{3} + \frac{1}{6\theta}, & \text{if } \theta \in [\theta_-^C, \theta_+^C] \\ \left(\frac{1}{2}\right)\left(1 - \sqrt{1 - \frac{2}{\theta}}\right), & \text{if } \theta > \theta_+^C, \end{cases}$$

where $\theta_-^C = \frac{1}{4}$ and $\theta_+^C = 1 + \frac{3\sqrt{2}}{4}$. As Proposition 1 indicates, $s^C(\theta)$ is decreasing for $\theta > \theta_-^C$. Low types of agent with $\theta \leq \theta_-^C$ are prescreened and induced to exert no effort, while high types with $\theta \geq \theta_+^C$ are induced to exert full effort. Finally, setting $s^C(\theta) = s^*$ and solving reveals that the critical type is $\theta^* = 1$.

5. INFORMED REVIEW

If the principal is unable to credibly commit to a standard, then she cannot act as a Stackelberg leader, maximizing her expected payoff subject to the agent's reaction function. (See the dashed iso-payoffs in Figure 1.) Instead, the agent's effort and principal's standard will be determined in a Cournot-Nash equilibrium.

If the principal has opted for informed review, then she observes the agent's type when choosing the standard. Hence, she maximizes her expected payoff, $V(s,p)$, given in (4.3), with respect to s, holding fixed p. The first-order condition is

$$V_s(s, P(s,\theta)) = -[vpL(s) - c(1-p)]f_l(s) = 0. \tag{5.9}$$

Rearranging this yields the principal's reaction function

$$S(p) = L^{-1}\left(r\frac{1-p}{p}\right). \tag{5.10}$$

Note that Assumption 1 implies: $S(0) = \overline{\sigma}$, $S(1) = \underline{\sigma}$, and $S'(p) = -\frac{r}{p^2 L'(s)} < 0$. This makes sense: if the principal believes the project is certainly bad ($p = 0$), then no signal realization will convince her to accept it. Similarly, if she believes the project is certainly good ($p = 1$), then no signal realization will deter her from accepting it. In general, the higher the principal believes p to be, the lower she sets the standard for acceptance.

Solving the agent and principal's reaction functions (4.2) and (5.10) results in the equilibrium standard and effort under informed review, $(s^I(\theta), p^I(\theta))$. (See Figure 1.) A degenerate equilibrium in which the principal never accepts the project ($s^I(\theta) = \overline{\sigma}$) and the agent never exerts effort ($p^I(\theta) = 0$) always exists. Indeed, for values of θ less than a cutoff θ_-^I this is the unique equilibrium, in which case the agent is prescreened. For higher values of θ, however, non-degenerate equilibria exist. In this case, the following observation, which obtains directly from the Envelope Theorem, implies that the set of equilibria are Pareto rankable.

Lemma 1. **(i)** *The principal's indirect payoff, $V(S(p),p)$, is increasing in p.*
(ii) *The agent's indirect payoff, $U(P(s,\theta),s;\theta)$, is decreasing in s.*

Because the principal's reaction function is downward-sloping, equilibria with lower standards (which the agent prefers) involve higher effort (which the principal prefers).

Hence, when multiple equilibria exist, the one with the lowest standard and highest effort is Pareto superior, and the players are presumed to coordinate on it.[11]

Momentarily ignoring the feasibility restrictions and substituting for p in (5.9) from (4.2) gives

$$V_s(s, P(s,\theta)) = -[vP(s,\theta)L(s) - c(1 - P(s,\theta))]f_l(s) = 0. \tag{5.11}$$

Define $s_0^I(\theta)$ to be the smallest root to this equation. Then $s_0^I(\theta)$ is the equilibrium standard when the feasibility restrictions on p do not bind.

Define the cutoff type by

$$\theta_-^I \equiv \frac{r}{u(\lambda_h - \lambda_l)}. \tag{5.12}$$

The following result characterizes the equilibrium under informed review.

Proposition 2 (Equilibrium under Informed Review).

Principal: *The principal sets the standard*

$$s^I(\theta) \equiv \begin{cases} \overline{\sigma}, & \text{if } \theta < \theta_-^I \\ s_0^I(\theta), & \text{if } \theta \geq \theta_-^I. \end{cases}$$

Moreover, $s^I(\theta)$ is strictly decreasing for $\theta > \theta_-^I$, and $\lim_{\theta\to\infty} s^I(\theta) = \underline{\sigma}$.

Agent: *The agent chooses effort level*

$$p^I(\theta) \equiv \begin{cases} 0, & \text{if } \theta < \theta_-^I \\ P(s_0^I(\theta), \theta), & \text{if } \theta \geq \theta_-^I. \end{cases}$$

Moreover, $p^I(\theta)$ is strictly increasing for $\theta > \theta_-^I$, and $\lim_{\theta\to\infty} p^I(\theta) = 1$.

As in the commitment benchmark, higher ability agents exert more effort in equilibrium and therefore face lower standards under informed review. Two differences from the commitment case are, however, readily apparent. First, at the low end of the type space, the range over which prescreening occurs is larger under informed review than under commitment ($\theta_-^I > \theta_-^C$). In other words, more types are induced to exert positive effort under commitment. Second, at the high end of the type space, the agent never exerts full effort under informed review while all types greater than θ_+^C do under commitment. Hence, at both extremes of the ability spectrum, the agent exerts less effort under informed review than under commitment. In fact, this is true for all types of agent as is stated in the following result.

Proposition 3 (Commitment vs. Informed Review). *The equilibrium profile of standards is flatter under commitment than under informed review and effort is higher. Specifically, suppose $\theta > \theta_-^C$ (else $s^C(\theta) = s^I(\theta) = \overline{\sigma}$), then*

$$s^C(\theta) \begin{cases} < s^I(\theta), & \text{if } \theta < \theta^* \\ = s^I(\theta), & \text{if } \theta = \theta^* \\ > s^I(\theta), & \text{if } \theta > \theta^*, \end{cases}$$

and $p^C(\theta) \geq p^I(\theta)$, with strict inequality if $\theta \neq \theta^*$.

11. The Pareto efficient equilibrium does not, however, maximize social surplus (i.e., the sum of the player's expected payoffs). In general, the principal sets too high a standard and the agent exerts too little effort in equilibrium because they do not account for the externalities their choices impose on the other player.

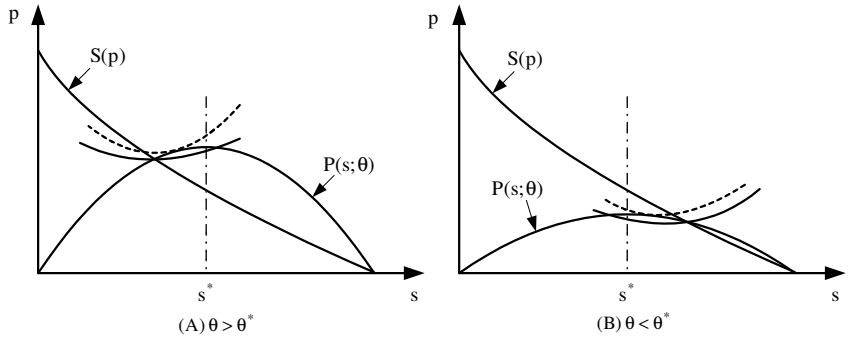

FIGURE 1
Evaluator's iso-payoffs and standard choices

The commitment standard, $s^C(\theta)$, strikes a balance between the goals of project selection and incentive provision, while the informed-review standard, $s^I(\theta)$, puts weight only on project selection. For high types, $\theta > \theta^*$, the incentive effect is positive; i.e., *raising* the standard induces more effort. Hence, commitment involves more *stringent* standards than informed review (panel A of Figure 1). On the other hand, for low types, $\theta < \theta^*$, the incentive effect is negative; i.e., *lowering* the standard induces more effort. Hence, commitment involves more *lenient* standards than informed review (panel B of Figure 1).

Example 3 (Informed Review). *Suppose the signal technology of Example 1 and that $v = c = u = 1$. From (5.10), the principal's reaction function is*

$$S(p) = 1 - p.$$

Solving this and the agent's reaction function,

$$P(s, \theta) = \min\{2\theta s(1 - s),\ 1\},$$

yields the equilibrium standard under informed review,

$$s^I(\theta) = \begin{cases} 1, & \text{if } \theta < \theta^I_- \\ \frac{1}{2\theta}, & \text{if } \theta \geq \theta^I_-, \end{cases}$$

where $\theta^I_- = \frac{1}{2}$. Comparison with Example 2 reveals that the region of prescreening is larger under informed review than under commitment, $\frac{1}{2} > \frac{1}{4}$. Notice also that no finite type ever exerts full effort under informed review. For $\theta > \theta^I_-$ the equilibrium profile under informed review, $s^I(\theta)$, is steeper than the one under commitment, $s^C(\theta)$, and imposes higher standards for $\theta < 1$ and lower standards for $\theta > 1$. It is straightforward to check that effort is uniformly lower under informed review.

6. BLIND REVIEW

If the principal observes only a subjective signal of project quality, then she will not be able to commit to the relatively flat profile of standards $s^C(\theta)$. Nevertheless, in this case, it may be possible for her to commit to remain ignorant of the agent's type when performing an evaluation. That is, she may be able to implement a policy of blind review and impose the same completely flat standard s^B on all types of agent (see Figure 2). While blind review forces the principal to disregard information that is valuable for project selection, it can be an effective method for providing incentives. For instance, a blind review procedure with $s^B = s^*$ would raise the effort of all types relative to a policy of informed review. Of course, even under blind review the principal can only implement a standard, s^B, that is *ex post* optimal given the information she possesses. The question is whether it is ever advantageous for her to commit to possessing less information.

In order to investigate blind review, it is analytically convenient to rule out cases in which the upward feasibility restriction on effort binds. Hence, the following additional assumption is imposed below.[12]

Assumption 3 (Bounded Effort). *The highest ability agent never exerts full effort,*

$$\overline{\theta}u(F_l(s^*) - F_h(s^*)) < 1.$$

Because the principal does not observe θ, the equilibrium standard is a best response to the agent's expected effort:

$$s^B = L^{-1}\left(r\frac{1 - E[p^B(\theta)]}{E[p^B(\theta)]}\right). \tag{6.13}$$

Similarly, the agent's effort is a best response to the standard,

$$p^B(\theta) = \theta u(F_l(s^B) - F_h(s^B)). \tag{6.14}$$

Finally, taking the expectation of (6.14) over θ gives

$$E[p^B(\theta)] = E[\theta]u(F_l(s^B) - F_h(s^B)). \tag{6.15}$$

These three equations define the equilibrium standard and effort under blind review. Comparing the solution to (6.13), (6.14), and (6.15) to the solution to (4.2) and (5.10) yields the following characterization.

Proposition 4 (Equilibrium under Blind Review).

Principal: *The principal sets the standard equal to the one she would have set for the mean type of agent under informed review, $s^B = s^I(E[\theta])$.*
Agent: *The agent chooses effort level $p^B(\theta) = P(s^I(E[\theta]), \theta)$.*

While this result derives technically from linearity of the agent's reaction function in θ and linearity of the principal's objective in p, it, nevertheless, seems intuitive that – when ignorant of the agent's type – the principal would set a standard as if she faced the average type in the population (see Figure 2).

The following result provides a comparison between the standards and induced effort under informed and blind review.

12. A stronger assumption, which is easier to check, is simply $\overline{\theta}u \leq 1$.

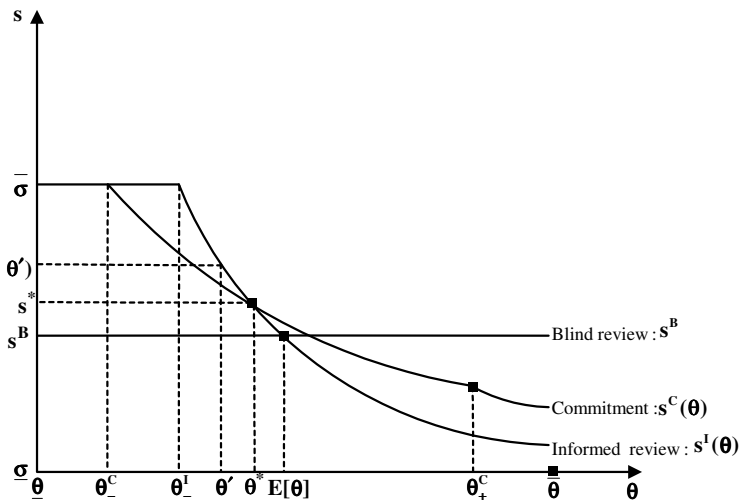

FIGURE 2
Equilibrium standards

Proposition 5 (Comparing Outcomes).

(i) *Agents with less than average ability face a lower standard under blind review than under informed review, and agents with greater than average ability face a higher standard:*

$$s^B \begin{cases} < s^I(\theta), & \text{if } \theta < E[\theta] \text{ and } E[\theta] > \theta^I_- \\ > s^I(\theta), & \text{if } \theta > E[\theta]. \end{cases}$$

(ii) *Expected effort is higher under blind review than under informed review if the mean type is sufficiently close to the critical type. Specifically,* $E[p^B(\theta)] - E[p^I(\theta)]$ *increases as* $E[\theta]$ *approaches* θ^**, and* $E[p^B(\theta)] - E[p^I(\theta)] > 0$ *if* $E[\theta] = \theta^*$.

Part (i) of Proposition 5 is a direct consequence of Propositions 2 and 4. Because the profile of standards under informed review is decreasing and because blind review is equivalent to an informed review over the mean type, the standard imposed under blind review is higher for types above the mean and lower for types below the mean than would be imposed under informed review. Combining this with the fact that agents always like lower standards (part (ii) of Lemma 1) reveals that an agent whose ability is above the mean prefers informed review, while one whose ability is below the mean prefers blind review. Hence, under a regime of blind review, agents of higher than average ability would attempt to identify themselves while agents of lower ability would not (see subsection 9.1 for more on this point).

To obtain intuition for part (ii) of Proposition 5 suppose first that $E[\theta] = \theta^*$. Then the standard under blind review is the one that maximizes the effort of all types, $s^B = s^*$. Clearly, no other review policy will elicit higher average effort than blind review in this

case. Next, suppose (as depicted in Figure 2) that $E[\theta]$ is slightly greater than θ^*.[13] Then $s^B = s^I(E[\theta])$ will be less than s^*. For types $\theta > E[\theta]$, blind review still imposes a higher standard than informed review, so these types would continue to exert more effort under blind review. Types in the interval $[\theta^*, E[\theta])$, however, would exert more effort under informed review because it calls for a higher standard, $s^I(\theta) \in (s^B, s^*]$. Of course, s^B is also too low for a neighborhood of types less than θ^*. However, there is a type $\theta' < \theta^*$ (shown in Figure 2) for whom the excessively low standard s^B would elicit the same effort as the excessively high one $s^I(\theta') > s^*$. For all types $\theta < \theta'$, blind review would induce strictly higher effort. In other words, if $E[\theta] \neq \theta^*$, then there is a band of types around θ^* who would exert more effort under informed review while the types outside this band would exert more effort under blind review. If $E[\theta]$ is distant from θ^*, then the band of types who work harder under informed review is large, and it is clearly the superior evaluation procedure because it provides both better selection and better incentives.

It is worth remarking on the reason blind review provides better incentives than informed review when $E[\theta]$ is close to θ^*. Under informed review, high-ability agents rest on their laurels, knowing that the principal will give them the benefit of the doubt. Low-ability agents also exert little effort under informed review, but for the opposite reason – they know that the principal will discriminate against them. Blind review pools high and low ability agents together and improves incentives at both ends of the type spectrum.

Example 4 (Blind Review). *Suppose the signal technology of Example 1 and that $v = c = u = 1$. Also suppose there are two possible types, $\theta \in \{\frac{1}{2}, \frac{3}{2}\}$, that are equally likely. Hence, $E[\theta] = 1$, which (as noted in Example 2) is also the critical type θ^*. The table below displays the equilibrium standards, efforts and error probabilities under informed and blind review.*

	INFORMED	BLIND
STANDARD: $\theta = \frac{1}{2}$	1	$\frac{1}{2}$
STANDARD: $\theta = \frac{3}{2}$	$\frac{1}{3}$	$\frac{1}{2}$
EFFORT: $\theta = \frac{1}{2}$	0	$\frac{1}{4}$
EFFORT: $\theta = \frac{3}{2}$	$\frac{2}{3}$	$\frac{3}{4}$
$\Pr\{\text{ACCEPT}\|q = 0\}$	$\frac{1}{9}$	$\frac{1}{4}$
$\Pr\{\text{REJECT}\|q = 1\}$	$\frac{1}{9}$	$\frac{1}{4}$

Because $E[\theta] = \theta^*$, the flat standard under blind review induces higher effort from both types of agent. Indeed, as noted in Example 3, $\theta^I_- = \frac{1}{2}$, so the low-ability agent exerts no effort under informed review. Although blind review provides better incentives, this comes at a cost. The probability of making a mistake (either accepting a bad project or rejecting a good one) is substantially higher under blind review.

7. THE EQUILIBRIUM REVIEW POLICY

Having characterized equilibrium under informed and blind review, we now turn to the principal's choice between these policies at the beginning of the game.

13. the discussion is analogous for $E[\theta] < \theta^*$.

7.1. *The Ability Distribution*

Define the principal's equilibrium payoff under informed review by $V^I(\theta) \equiv V(s^I(\theta), p^I(\theta))$. By proposition 4 the principal's expected equilibrium payoff under blind review is

$$\begin{aligned} E[V^B(\theta)] &= vE[p^B(\theta)](1 - F_h(s^B)) - c(1 - E[p^B(\theta)])(1 - F_l(s^B)) \\ &= vp^I(E[\theta])(1 - F_h(s^I(E[\theta]))) - c(1 - p^I(E[\theta]))(1 - F_l(s^I(E[\theta]))) \\ &= V^I(E[\theta]). \end{aligned}$$

Therefore, when choosing between review policies, the principal compares her expected payoff under informed review, $E[V^I(\theta)]$, with her expected payoff under blind review, $V^I(E[\theta])$. Evidently, if $V^I(\cdot)$ is convex or concave everywhere, then Jensen's inequality will suffice to rank the two payoffs irrespective of the type distribution. In general, however, $V^I(\cdot)$ is S-shaped, possessing both a convex and a concave region, as the following lemma records.

Lemma 2. *If θ^I_- is sufficiently low, then there exist two cutpoints, $\theta_L \leq \theta_H$, such that $V^I(\theta)$ is strictly convex for $\theta < \theta_L$ and strictly concave for $\theta > \theta_H$.*

The intuition behind Lemma 2 is that for very high types, effort is close to its maximum, and so is the principal's payoff. Thus, there are diminishing marginal returns to ability. For very low types, on the other hand, a rise in ability not only raises effort but significantly improves the probability of making a correct acceptance decision.

In light of Lemma 2, it is clear that the principal's choice between the two review procedures depends crucially on the distribution of types. If, for instance, all types receiving positive probability under $g(\cdot)$ are above θ_H (where $V^I(\cdot)$ is concave), then Jensen's inequality implies that blind review dominates informed review. On the other hand, if $g(\cdot)$ puts weight only on types less than θ_L (where $V^I(\cdot)$ is convex), then the principal prefers informed review. More generally, determination of which review procedure is optimal depends on whether high types or low types are more prevalent in the population, as is stated in the following result.

Proposition 6 (Extreme Types and the Agent's Payoff). *Suppose that the support of the ability distribution includes both the (low) region where $V^I(\cdot)$ is convex and the (high) region where it is concave; i.e., $\underline{\theta} < \theta_L \leq \theta_H < \overline{\theta}$. Then,*

(i) *there exist ϵ^B and $\epsilon^I > 0$ such that the principal prefers blind review if $G(\theta_H) < \epsilon^B$, and informed review if $G(\theta_L) > 1 - \epsilon^I$; and*

(ii) *there exist $0 < u_L \leq u_H < \infty$ such that the principal prefers blind review if $u > u_H$, and informed review if $u < u_L$.*

Part (i) of Proposition 6 indicates that incentives are more important than project selection when evaluating high-ability agents, so blind review is optimal. Project selection, however, becomes the dominant concern when evaluating low-ability agents, and informed review is, therefore, preferable in this case.

The second part of Proposition 6 states that, fixing the ability distribution, the principal is also more likely to prefer blind review, as the agent's payoff from acceptance, u, increases. Note from (4.2) that an increase in u is equivalent to an increase in θ.

Hence, agents with high rewards from acceptance will behave like those with high ability, in which case blind review is the principal's preferred mode of evaluation.

A question that naturally arises in light of Proposition 6 is whether blind review is the preferable evaluation process only when the ability distribution places sufficient weight on high values of θ. Mathematically this comes down to asking whether $V^I(\theta)$ is concave only at high ability levels. The following result establishes that this is not so.

Lemma 3. *The principal's equilibrium payoff function $V^I(\theta)$ is strictly concave at θ^*.*

Observe from (4.8) that θ^* does not depend on the ability distribution. In particular, it is increasing in the principals cost to benefit ratio r and decreasing in the agent's payoff from acceptance u and in the difference $F_l(s^*) - F_h(s^*)$. Hence, $V^I(\theta)$ may possess a concave region at virtually any point in the type space. It follows, therefore, that it is not necessary for the ability distribution to include a large fraction of high types in order for blind review to be optimal. This is formalized as follows.

Proposition 7 (Non-Extreme Types). *There exist $\epsilon > 0$ and $\Delta > 0$ such that if $|E[\theta] - \theta^*| < \epsilon$ and $|\theta - \theta^*| < \Delta$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$, then the principal prefers blind review.*

From part (ii) of Proposition 5, expected effort is higher under blind review than under informed review when the critical type is close to the mean. Moreover, when the support of the ability distribution is concentrated around θ^*, then observing the realization of θ is of little value to the principal, and blind review is, therefore, optimal.

Example 5 (Optimal Review). *Suppose the signal technology of Example 1 and that $v = c = u = 1$. From Example 3 it is straightforward to compute*

$$V^I(\theta) = \left(1 - \frac{1}{2\theta}\right)^2.$$

This is S-shaped, with an inflection point at $\theta_H = \theta_L = \frac{3}{4}$. In accordance with Proposition 6, blind review is optimal for any ability distribution with $\underline{\theta} \geq \frac{3}{4}$, and informed review is optimal for any distribution with $\overline{\theta} \leq \frac{3}{4}$. In accordance with Proposition 7, blind review is also optimal for any distribution sufficiently concentrated around the critical type $\theta^* = 1$. If, as in Example 4, there are two types, $\theta \in \{\frac{1}{2}, \frac{3}{2}\}$, that are equally likely, then the principal's expected equilibrium payoffs under informed and blind review are respectively $E[V^I(\theta)] = \frac{2}{9}$ and $V^I(E[\theta]) = \frac{1}{4}$.

7.2. *The Informativeness of the Signal*

An important element in determining the optimal review process is the *informativeness* of the signal observed by the evaluator. The informativeness of the signal varies across applications depending on such factors as the expertise of the evaluator and the stage at which the project is submitted for review. In what follows, we define a notion of informativeness based on MLRP along the lines suggested by Milgrom (1981), and then investigate its effect on the choice of the review policy.

Let $\alpha \in \mathbb{R}_+$ be a parameter measuring the informativeness of the signal, and write the probability distribution and corresponding density functions respectively as $F_q(\sigma; \alpha)$

and $f_q(\sigma;\alpha)$, for $q = h, l$. Intuitively, a signal technology is more informative if it is more likely to generate a high signal when project quality is high, and a low signal when project quality is low. This is formalized as follows.

Definition 1 (Informativeness). Signal technology $\{f_h(\sigma;\alpha_1), f_l(\sigma;\alpha_1)\}$ is said to be more informative than $\{f_h(\sigma;\alpha_0), f_l(\sigma;\alpha_0)\}$ if $\frac{f_h(\sigma;\alpha_1)}{f_h(\sigma;\alpha_0)}$ increases and $\frac{f_l(\sigma;\alpha_1)}{f_l(\sigma;\alpha_0)}$ decreases in σ whenever $\alpha_1 > \alpha_0$.

That is, in addition to Assumption 1, we also impose MLRP separately on both $f_h(\sigma;\alpha)$ and $f_l(\sigma;\alpha)$.

Next, normalize α so that as $\alpha \to 0$, the signal technology becomes completely uninformative, namely $F_l(\sigma;\alpha) - F_h(\sigma;\alpha) \to 0$ for all σ; and as $\alpha \to \infty$, it becomes completely informative, namely $F_l(\sigma;\alpha) \to 1$ for $\sigma \neq \underline{\sigma}$, and $F_h(\sigma;\alpha) \to 0$ for $\sigma \neq \overline{\sigma}$. In particular, when the signal technology is completely informative, the principal observes only the highest signal, $\overline{\sigma}$, if project quality is high, and only the lowest signal, $\underline{\sigma}$, if it is low. Finally, in order to compare different signal technologies, it is necessary to impose an assumption on the likelihood ratio, $L(\sigma;\alpha) \equiv \frac{f_h(\sigma;\alpha)}{f_l(\sigma;\alpha)}$.

Assumption 4 (The Neutral Signal). *$L(s^*;\alpha) = 1$, for all $\alpha \in (0,\infty)$.*

In other words, s^* continues to be the unique neutral signal, independent of α. This assumption implies that, as α increases, $L(\sigma;\alpha)$ rotates counter-clockwise around the point $(s^*, 1)$ such that, as $\alpha \to \infty$, $L(\sigma;\alpha) \to 0$ if $\sigma < s^*$ and $L(\sigma;\alpha) \to \infty$ if $\sigma > s^*$. (see Lemma A1 in the Appendix).

A simple parametric extension of the signal technology presented in Example 1 satisfies Definition 1 and Assumption 4. Namely, it is straightforward to check that the family of signal technologies with $f_l(\sigma;\alpha) = (1+\alpha)(1-\sigma)^\alpha$ and $f_h(\sigma;\alpha) = (1+\alpha)\sigma^\alpha$, for $\sigma \in [0,1]$ and $\alpha \geq 0$, works.

The first result of this subsection reveals that a more informative signal elicits greater effort from the agent.

Lemma 4. *Suppose $p^I(\theta;\alpha) > 0$.*

(i) *$p^I(\theta;\alpha)$ strictly increases in α.*
(ii) *As $\alpha \to 0$, $p^I(\theta;\alpha) \to 0$ and $s^I(\theta;\alpha) \to \overline{\sigma}$ for all θ.*
(iii) *As $\alpha \to \infty$, $p^I(\theta,\alpha) \to \min\{\theta u, 1\}$ and*

$$s^I(\theta;\alpha) \to \begin{cases} s^*, & \text{if } \theta u < 1 \\ \underline{\sigma}, & \text{if } \theta u \geq 1. \end{cases}$$

A more informative signal provides a tighter measure of the agent's performance and improves incentives. In the extreme cases, the agent exerts no effort when the signal is completely uninformative, and the maximum effort when it is completely informative. An implication of this observation is that the principal is indifferent between blind and informed review, in either of the extreme cases ($\alpha = 0$ or $\alpha = \infty$). For signal technologies near – but not equal to – the extremes, however, the principal has a strict preference for either blind or informed review, as is stated in the following key result.

Proposition 8 (Informativeness). *Suppose $\overline{\theta}u < 1$. Then,*

(i) *When the signal technology is either completely uninformative or completely informative, the principal is indifferent between blind and informed review. In particular,* $\lim_{\alpha\to 0} E[V^I(\theta;\alpha)] = \lim_{\alpha\to 0} E[V^B(\theta;\alpha)] = 0$, *and* $\lim_{\alpha\to\infty} E[V^I(\theta;\alpha)] = \lim_{\alpha\to\infty} E[V^B(\theta;\alpha)] = E[\theta]vu$.

(ii) *For a sufficiently uninformative signal technology, the evaluator prefers informed review. That is, for a small* $\alpha > 0$, $E[V^I(\theta;\alpha)] > E[V^B(\theta;\alpha)]$.

(iii) *For a sufficiently informative signal technology, the evaluator prefers blind review. That is, for a large* $\alpha < \infty$, $E[V^I(\theta;\alpha)] < E[V^B(\theta;\alpha)]$.

When the signal is imprecise, the principal opts to use the information contained in the agent's type to improve her selection decision; i.e, she uses informed review. On the other hand, when the signal is very precise, the principal eschews the information contained in the agent's type in order to provide better incentives; i.e., she uses blind review.

8. APPLICATIONS

The analysis presented in the preceding section can be applied to address questions regarding the equilibrium review policy in a variety of real-world settings. This is illustrated below in two examples: admission of character evidence in trials and academic refereeing. There are, of course, aspects of these settings not captured by our model. For instance, in the case of a criminal trial, the principal (jurist) might also be concerned with the future trajectory of the agent (defendant), not just his current conduct. In the case of academic refereeing, the choice of review process also influences the evaluator's incentives to perform a thorough evaluation. These important caveats notwithstanding, we believe that the analysis performed above sheds important light on these as well as other real-world evaluation settings.

8.1. *Character Evidence in Trials*

There is a long-standing debate among legal scholars about whether character evidence – i.e., the past conduct (good or bad) of the defendant – should be admissible at trial. Those who oppose the use of character evidence argue: (1) an individual's past behavior is too weak a predictor of his current act; (2) the Jury may be prone to "cognitive error," giving too much weight to dispositional evidence; (3) Jurors might convict someone solely based on his bad character or his prior wrongdoings; (4) the court's resources should not be used on examining the details of someone's past; (5) banning character evidence forces parties to seek the best evidence for the current case.

Legal scholars who support the use of character evidence argue that the defendant's past acts may be a good predictor of the current act, because several influential studies in social psychology maintain the stability of criminal behavior associated with certain personality features, such as lack of self-control or low-empathy. Moreover, supporters of character evidence say that it may be of use in tailoring incentives to individual characteristics. Character evidence would be of use in determining defendants' criminal propensity, so that punishments could be appropriately adjusted.

The intellectual schism in the legal community over the admissibility of character evidence appears to extend even to the U.S. Federal Rules of Evidence.[14] For instance,

14. See `http://www.law.cornell.edu/rules/fre/rules.htm`.

opposing its use, Rule 404(b) states, "Evidence of other crimes, wrongs, or acts is not admissible to prove the character of a person in order to show action in conformity therewith." Conversely, supporting its use, Rule 406 states, "Evidence of the habit of a person or of the routine practice of an organization, whether corroborated or not and regardless of the presence of eyewitnesses, is relevant to prove that the conduct of the person or organization on a particular occasion was in conformity with the habit or routine practice." In general, courts allow character evidence if its probative value exceeds its prejudicial effect.[15]

In the context of the model presented above, one can think of the principal as a jurist and the agent as a defendant in a trial. Initially (before the trial), the defendant exerts some effort p to avoid criminal behavior. The cost of this effort depends on the defendant's characteristics such as his education, age, and employment status, as summarized by θ. Project acceptance corresponds to acquittal and rejection to conviction. If character evidence is admissible, then the jurist performs an informed review, and if it is banned, then she performs a blind review.

The close correlation between an individual's socioeconomic characteristics and the type of crime he is most likely to commit is well documented.[16] Specifically, blue-collar or street crimes, such as robbery and assault, tend to be committed by individuals who are relatively young, uneducated, and poor (low values of θ), while white-collar or corporate crimes, such as embezzlement or price fixing, tend to be committed by individuals who are relatively old, educated, and wealthy (high values of θ). In this light, Proposition 6 suggests that character evidence should be permitted in cases of street crime and eschewed in cases of white-collar crime. Interestingly, this analysis indicates that the overriding concern in cases of blue-collar crime is making a correct decision about whether to acquit or convict, while the predominant concern in cases of white-collar crime is deterrence.

8.2. *Journal Submissions and Grant Proposals*

As a second application of the theory, note that Bachand and Sawallis (2003) find that 58% of 553 journals across 18 academic disciplines use blind review of manuscripts; i.e., the referee is not informed of the author's identity.[17] Grant proposals, however, appear to always be evaluated non-blindly. Indeed, in many if not most cases (e.g., NSF, NIH, Russell Sage Foundation, Alfred P. Sloan Foundation) principal investigators are required to submit a *Curriculum Vitae* to be reviewed as part of their proposal.

A major difference between a journal submission and a grant proposal is the stage at which the research is evaluated. Grant proposals, by their nature, are speculative and preliminary. They contain few, if any, concrete findings. Rather, they are hypothetical in scope and generally provide only a conceptual plan, or road map, for how the proposed research is to be performed. Journal submissions, by contrast, are intended to disseminate

15. For an excellent discussion of the alternative views on character evidence, see Sanchirico (2001). Also see Lippke (2008), and, for the British perspective, Redmayne (2002).

16. On the age distribution of crime, see Steffensmeier *et al.* (1989). On the link between economic inequality and violent crime see Blau and Blau (1982), and on the correlation between education and crime see Ehrlich (1975), and Lochner and Moretti (2004).

17. In disciplines such as economics, where working papers are customarily uploaded to the web, commitment to blind review is not possible because a referee can ascertain the author's identity by searching for a few key words. Perhaps this is why over 80% of the top-50 economics journals no longer attempt to practice blind review. Many disciplines, however, do not post manuscripts on the Internet, making blind review still a feasible option.

actual research findings. They are supposed to contain substantial analysis and explicit results.

It stands to reason, therefore, that the signal an evaluator receives about the ultimate quality of the research is much less informative when reviewing a grant proposal than when reviewing a journal submission. Proposition 8 accords neatly with this interpretation. Grant proposals, which are submitted at a very early stage of the project, should be subjected to informed review, while manuscripts, which are submitted at a more mature stage, should be reviewed blindly. In other words, the primary concern in grant review should be the selection of good projects, while the dominant concern in manuscript review should be the provision of incentives for authors to write high-quality papers.

9. GENERALIZATIONS AND EXTENSIONS

In this section, the basic model is extended in three dimensions to highlight the robustness of the results obtained above and glean some important additional insights.

9.1. *Competing Evaluators and Informed Review Bias*

In practice there are often multiple evaluators (e.g., schools, companies, and academic journals) that compete for high-quality applications. In this subsection, the basic model is extended to show how competition among evaluators impacts the equilibrium mode of review.

Suppose there are two *ex ante* symmetric evaluators, $i = 1, 2$, who simultaneously and publicly announce their review policies, $\tau_i \in \{I, B\}$. Upon observing τ_1 and τ_2, each agent then exerts effort and applies to one evaluator. To parameterize the degree of competition, one of three possible situations is assumed to obtain. With probability $1-\phi$ an agent is *unattached* (i.e., he is free to apply to either evaluator); with probability $\frac{\phi}{2}$ he is attached to evaluator 1; and with probability $\frac{\phi}{2}$ he is attached to evaluator 2. Attachments are independent across agents and over types. For simplicity, also assume that re-applications are not feasible, and, in case of indifference, an unattached agent selects between the evaluators with equal probability.

Let $\pi_i^{\tau_1,\tau_2}$ be evaluator i's expected payoff in the subgame with review policies τ_1 and τ_2. Note that, if $\tau_1 = \tau_2 = I$, then *ex ante* each agent is equally likely to apply to either evaluator, resulting in equal payoffs, $\pi_i^{I,I} = \frac{1}{2}E[V^I(\theta)]$. Similarly, if $\tau_1 = \tau_2 = B$, then a straightforward argument shows that, in equilibrium, both evaluators adopt the same standard tailored to the population mean, $E[\theta]$, yielding equal payoffs, $\pi_i^{B,B} = \frac{1}{2}V^I(E[\theta])$. The equilibrium characterization with different review policies is less obvious.

Suppose $\tau_1 = B$ and $\tau_2 = I$. Moreover, suppose, in equilibrium, the mean type that applies to evaluator 1 is m_1. This implies that an unattached type θ prefers evaluator 1 whenever $\theta < m_1$. Hence, the conditional mean ability for evaluator 1 is

$$M_1(m_1;\phi) \equiv \frac{(1-\frac{\phi}{2})\int_{\underline{\theta}}^{m_1}\theta\, dG(\theta) + \frac{\phi}{2}\int_{m_1}^{\overline{\theta}}\theta\, dG(\theta)}{(1-\frac{\phi}{2})G(m_1) + \frac{\phi}{2}(1-G(m_1))}.$$

In equilibrium, the conditional mean, $m_1 = \mu_1(\phi)$, must solve

$$M_1(m_1;\phi) - m_1 = 0. \tag{9.16}$$

Lemma 5. *There exists a unique solution, $\mu_1(\phi)$ to (9.16). The function $\mu_1(\phi)$ is strictly increasing and has boundary values $\mu_1(0) = \underline{\theta}$ and $\mu_1(1) = E[\theta]$.*

Lemma 5 says that as each agent becomes less likely to be attached, fewer high types apply to evaluator 1 for blind review. This causes it to raise its standard and further discourages applications by high unattached types in the remaining pool. In particular, in the absence of attached types ($\phi = 0$), complete unraveling occurs and all agents apply to evaluator 2 for informed review.

In light of Lemma 5, equilibrium payoffs for evaluators 1 and 2 in the subgame with $\tau_1 = B$ and $\tau_2 = I$ are given, respectively, by

$$\pi_1^{B,I}(\phi) = \left[\left(1 - \frac{\phi}{2}\right) G(\mu_1(\phi)) + \frac{\phi}{2}(1 - G(\mu_1(\phi)))\right] V^I(\mu_1(\phi))$$

and

$$\pi_2^{B,I}(\phi) = \frac{\phi}{2} \int_{\underline{\theta}}^{\mu_1(\phi)} V^I(\theta)\, dG(\theta) + \left(1 - \frac{\phi}{2}\right) \int_{\mu_1(\phi)}^{\overline{\theta}} V^I(\theta)\, dG(\theta).$$

The following lemma characterizes these payoffs.

Lemma 6. *In the unique equilibrium with $\tau_1 = B$ and $\tau_2 = I$, the evaluator's payoffs have the following properties:*

(i) *$\pi_1^{B,I}(\phi)$ is strictly increasing, and $\pi_1^{B,I}(0) = 0$ and $\pi_1^{B,I}(1) = \frac{1}{2}V^I(E[\theta])$.*
(ii) *$\pi_2^{B,I}(\phi)$ is strictly decreasing, and $\pi_2^{B,I}(0) = E[V^I(\theta)]$ and $\pi_2^{B,I}(1) = \frac{1}{2}E[V^I(\theta)]$.*

In other words, the evaluator who uses blind review is better off when there are more attached types, because they have a direct positive effect on her payoff as well as a positive indirect effect through attracting high unattached types, who, by Lemma 5, anticipate a lower standard. By the same token, the evaluator who uses informed review is worse off when there are more attached types.

Having characterized the evaluators' payoffs in each subgame, the equilibrium review policies can now be determined.

Proposition 9 (Equilibrium with Competing Evaluators). *Suppose $\frac{1}{2}V^I(E[\theta]) < E[V^I(\theta)] < V^I(E[\theta])$. Then, there exist two cutpoints, $0 < \phi^* \leq \phi^{**} < 1$, such that for $\phi < \phi^*$, the unique equilibrium has $\tau_1 = \tau_2 = I$, whereas for $\phi > \phi^{**}$, the unique equilibrium has $\tau_1 = \tau_2 = B$. For $\phi \in [\phi^*, \phi^{**}]$, both symmetric and asymmetric review policies may occur in equilibrium.*

The message of Proposition 9 is that competition among evaluators to attract high quality applications is likely to lead them to adopt informed review, even when each would individually prefer blind review.

9.2. *Two Types of Error*

Up to now, it has been assumed that the evaluator suffers a loss only from a false acceptance. In some settings, however, she may also suffer a loss from a false rejection. For instance, misjudging a potentially good musician is probably as costly for the performance of a symphony orchestra as hiring a potentially bad one. In Taylor and Yildirim (2007) it

is shown that accounting for both types of error does not qualitatively change the main results derived from the basic model, especially those pertaining to the comparison of informed and blind review. Thus, in this subsection only the novel insights associated with this generalization are highlighted.

Suppose, in addition to the loss $-c < 0$ from a false acceptance, the principal also incurs a loss $-\overline{c} < 0$ from a false rejection. While this does not alter the agent's payoff in (4.1), the expected loss from rejecting a good project needs to be subtracted from the principal's payoff in (4.3):

$$\overline{V}(s,p) = vp(1 - F_h(s)) - c(1-p)(1 - F_l(s)) - \overline{c}pF_h(s). \tag{9.17}$$

Maximizing (9.17) with respect to s, the principal's reaction function is

$$\overline{S}(p) = L^{-1}\left(\overline{r}\frac{1-p}{p}\right), \tag{9.18}$$

where $\overline{r} \equiv \frac{c}{v+\overline{c}}$. Because $L' > 0$, (9.18) implies that the principal is more likely to accept a project as her loss from a false rejection increases. Applying the Envelope Theorem, the principal's indirect payoff satisfies

$$\frac{d}{dp}\overline{V}(\overline{S}(p),p) = v(1 - F_h(\overline{S}(p))) + c(1 - F_l(\overline{S}(p))) - \overline{c}F_h(\overline{S}(p)),$$

which is clearly positive if $\overline{S}(p)$ is close to $\underline{\sigma}$, and negative if $\overline{S}(p)$ is close to $\overline{\sigma}$. Hence, in contrast to part (i) of Lemma 1, the principal's indirect payoff does not monotonically increase in p. In fact, since $\overline{S}(0) = \overline{\sigma}$ and $\overline{S}'(p) < 0$, her indirect payoff strictly decreases in p whenever p is small, because such a project is very likely to be rejected, and thus exposes the principal to risk of false rejection.[18] In fact, a sufficiently small p may result in a negative payoff for the evaluator, $\overline{V}(\overline{S}(p),p) < 0$, which never occurs in the base model. The evaluator would, of course, avoid a negative payoff if she could commit to prescreening those agents who are unlikely to exert a high enough effort. Intuitively, in the absence of commitment by the principal, some low-ability agents will take advantage of her fear of false rejections by submitting projects with very low values of p.

The discussion thus far suggests two novel insights that are confirmed in Proposition 10: first, the evaluator may receive a negative equilibrium payoff from some intermediate type agents, and second, there may be too little equilibrium prescreening compared with the commitment benchmark.

Proposition 10 (Two Types of Error). *Suppose both types of error are costly to the evaluator, i.e., $c, \overline{c} > 0$. Then, in equilibrium,*

(i) *under informed review there exist two types, $\overline{\theta}^I_- < \overline{\theta}^I_r$, such that*

$$\overline{V}(\theta)\begin{cases} = 0 & \text{if } \theta \le \overline{\theta}^I_- \text{ or } \theta = \overline{\theta}^I_r \\ < 0 & \text{if } \overline{\theta}^I_- < \theta < \overline{\theta}^I_r \\ > 0 & \text{if } \theta > \overline{\theta}^I_r, \end{cases}$$

(ii) *if $\overline{c}$ is sufficiently large, then, compared with commitment, there is less prescreening under informed review, namely, $\overline{\theta}^I_- < \overline{\theta}^C_-$.*

18. The nonmonotonicity of the evaluator's indirect payoff implies that equilibria under informed review are not always Pareto rankable. For consistency, we continue to assume that the players coordinate on the equilibrium with the highest effort and lowest standard.

Hence, when the evaluator is sufficiently concerned about a false rejection, there may be too little prescreening under informed review as opposed to too much prescreening as observed in the base model. The reason is as suggested above: when the evaluator fears rejecting a good project, she cannot credibly discourage some intermediate type agents from submitting projects in equilibrium, even though they are unlikely to produce high quality.

9.3. *Non-Productive Effort*

In some settings, applicants can engage in both productive activities, which enhance their current performance, and influence activities, which may enhance the signal received by the evaluator but do not enhance the actual quality of the project. For instance, an individual submitting a manuscript for publication may spend significant time and energy on formatting and layout, rather than on improving content. To capture this possibility, extend the base model by supposing that an agent can exert two kinds of effort: a productive effort, p, and an unproductive effort, e. In particular, even when the true quality of the project is low, which happens with probability $1-p$, the evaluator still receives the signal from F_h with probability e. In this case, the principal's expected payoff is given by

$$V(s,p,e) = p(1-F_h(s))v - e(1-p)(1-F_h(s))c - (1-e)(1-p)(1-F_l(s))c.$$

Differentiating with respect to s yields

$$V_s(s,p,e) = [(-p+(1-p)er)L(s) + (1-p)(1-e)r]vf_l(s).$$

Note that if $-p+(1-p)er \geq 0$, then $S(p,e) = \overline{\sigma}$. Suppose $-p+(1-p)er < 0$. Then, the evaluator's reaction function is

$$L(S(p,e)) = \frac{(1-p)r-(1-p)er}{p-(1-p)er}.$$

For $e=0$ this yields the base-model reaction function.

As for the agent, assume that ability affects only the cost of exerting productive effort. That is, all types of agent are equally capable of influence activity.[19] In particular, a type θ agent's cost of exerting p and e is $\frac{1}{2}\left(\frac{p}{\sqrt{\theta}}+\frac{e}{\sqrt{t}}\right)^2$, where $t>0$ is a constant. Hence his expected utility is

$$U(p,e,s;\theta) = u[(1-(1-p)(1-e))(1-F_h(s))+(1-p)(1-e)(1-F_l(s))] - \frac{1}{2}\left(\frac{p}{\sqrt{\theta}}+\frac{e}{\sqrt{t}}\right)^2.$$

This also yields the base model when $e=0$.

Lemma 7. *For any $s \in [\underline{\sigma}, \overline{\sigma}]$, the best response of a type θ agent is as follows: for $\theta \leq t$, $p=0$ and $e = \min\{tu(F_l(s)-F_h(s)),\ 1\}$, and for $\theta \geq t$, $e=0$ and $p = \min\{\theta u(F_l(s)-F_h(s)),\ 1\}$.*

The agent's payoff function is submodular in the two kinds of effort ($U_{pe}<0$), so the marginal return to productive effort is highest when non-productive effort is zero and

19. Technically, what matters is that influence ability not be perfectly correlated with productive ability.

vice versa. If $\theta < t$, then non-productive effort costs the agent less than productive effort, and if $\theta > t$, then the reverse is true.

Clearly, under informed review the principal will prescreen all types $\theta < t$; i.e., $s = \overline{\sigma}$ and $p = e = 0$. For $\theta \geq t$, the equilibrium is as characterized in Proposition 2, namely $s = s^I(\theta)$ and $p = p^I(\theta)$. The equilibrium under blind review is more difficult to characterize, mainly because the equilibrium standard s^B is a function of t. In general, blind review should perform relatively poorly compared with informed review when non-productive effort is possible. Under informed review, low-ability agents who specialize in non-productive effort can be excised from the applicant pool, while under blind review they cannot. The following result reveals that when there is a significant fraction of agents who exert non-productive effort, then blind review is infeasible.

Proposition 11 (Rent Seeking). *Suppose $\overline{\theta}u \leq 1$. There exists $\underline{t} < \overline{\theta}$ such that if $t \geq \underline{t}$, then blind review results in the degenerate equilibrium: $s^B = \overline{\sigma}$ and $p^B = e^B = 0$. Hence, if $t \in [\underline{t}, \overline{\theta})$, then informed review is optimal.*

This result says that if the fraction of "spoilers" in the population is high enough ($G(t) \geq G(\underline{t})$), then the principal would reject all projects under blind review. (As the proof of the proposition makes clear, $\underline{t}$ is decreasing in r, so that the fraction of spoilers can be small if the cost to benefit ratio is high.) However, so long as some fraction of agents prefer to exert productive effort ($G(t) < 1$), the principal receives a positive payoff under informed review because she only evaluates submissions from agents with $\theta \geq t$. Hence, the ability of agents to exert non-productive effort mitigates against the adoption of blind review.

10. CONCLUSION

The issue at the heart of this paper concerns the tradeoff between the effective use of information and the provision of incentives, in a setting where commitment to a review standard is infeasible. It was shown in this context that when the evaluator observes the innate ability of the applicant, the equilibrium review policy is unduly biased – the standards facing high ability applicants are too weak and those facing low-ability applicants are too tough. While this policy uses information optimally *ex post*, it provides poor incentives *ex ante*. In particular, if the evaluator could commit to a review procedure, then she would implement a flatter (less biased) one. Commitment to such a policy is, however, often impractical or impossible because of the subjective nature of performance measures: the taste of a fine wine, the skill of a classical musician, the preponderance of evidence. Although it is not possible to commit to a highly tailored acceptance procedure in such environments, it is often possible to commit to remain ignorant of the identity (and hence the ability) of the applicant; that is, to perform blind review.

The uniform standard implemented under blind review often provides good incentives for agents at both ends of the ability distribution, but it sacrifices information at the project selection stage. Hence, whether the evaluator prefers blind or informed review depends critically on whether incentives or project selection is more important to her. Blind review was shown to be the preferred mode of evaluation if the applicant pool contains a large proportion of high ability agents, the applicant's stakes from acceptance are relatively high, the subjective signal of project quality is fairly precise, wrong decisions are relatively less costly, or there is limited competition among evaluators.

Two applications were presented. In the context of criminal trials, the theory presented here suggests that character evidence should be admissible in cases of street crime but not white-collar crime. It was also argued that research projects in early stages of development (e.g., grant proposals) should be subjected to informed review, while projects in later stages (e.g., manuscripts) should be evaluated blindly.

There are a number of intriguing issues that remain to be addressed in future work. For instance, it would be interesting to investigate the effects of more general non-monetary incentive devices. In academia, for example, there is a gradation of rewards associated with acceptance of a manuscript (lead article, long or short paper, best paper prize, and so on) that can be used to induce individuals to exert more effort. Similarly, it would also be interesting to study the role of peer effects in a setting where an applicant's benefit from acceptance, u, is determined endogenously. Specifically, the status or prestige enjoyed by an agent whose project is accepted might well depend on the average quality of other projects accepted by the evaluator. Another intriguing avenue for future research would be to consider the possibility of psychological bias on the part of the evaluator and the role of blind review in mitigating prejudice. Finally, it would be instructive to study a dynamic version of the model in which the current ability of an agent depends on his history of prior projects. In particular, it would be useful to explore the interaction between the review policies of evaluators and the career trajectories of agents in such a setting. While these and other topics for further research appear fruitful, it seems likely that the fundamental message of this paper will remain intact. Fairness is not the only reason to level the playing field – often it is also the best thing the evaluator can do.

APPENDIX

This appendix contains the proofs of all lemmas and propositions presented in the text, as well as the statement and proofs of two technical lemmas.

Proof of Proposition 1. Note first that if a type θ agent is subject to s sufficiently close to $\overline{\sigma}$, then $P(s,\theta) < 1$, and hence $V(s, P(s,\theta))$ is strictly quasi-concave in s by Assumption 2. This means $s^C(\theta) = \overline{\sigma}$ if and only if $\lim_{s\to\overline{\sigma}} \frac{d}{ds}V(s, P(s,\theta)) > 0$. Differentiating (4.4) and noting $\lim_{s\to\overline{\sigma}} P(s,\theta) = 0$ and $\lim_{s\to\overline{\sigma}}[L(s)P(s,\theta)] = u\theta(\lambda_h - \lambda_l)$ by Assumption 1, it follows $s^C(\theta) = \overline{\sigma}$ and thus $p^C(\theta) = 0$ if and only if $\theta < \frac{r}{u[2\lambda_h + (r-1)\lambda_l]} = \theta^C_-$. Second, suppose $P(s,\theta) > 0$. If we ignore the constraint $P(s,\theta) \leq 1$, then by Assumption 2, $s^C(\theta) = s^C_0(\theta)$ is the unique solution to $\frac{d}{ds}V(s^C_0(\theta), P(s^C_0(\theta),\theta)) = 0$, or equivalently to (4.5). If $P(s^C_0(\theta),\theta) > 1$, then $s^C(\theta) \neq s^C_0(\theta)$; rather $s^C(\theta) = \min\{s | P(s,\theta) = 1\}$, because for $P(s,\theta) = 1$, (4.4) reduces to $V(s, P(s,\theta)) = v(1 - F_h(s^I))$, which is strictly decreasing in s. By construction, $s^C(\theta)$ is continuous for $\theta \neq \theta^C_-$. Since $\lim_{s\to\overline{\sigma}} V(s,p) = 0$ for any $p \in [0,1]$, $s^C(\theta)$ is also continuous at $\theta = \theta^C_-$. This implies that $p^C(\theta) = P(s^C(\theta),\theta)$ is continuous for all θ.

To prove that $s^C(\theta)$ is strictly decreasing for $\theta > \theta^C_-$, observe that for $\theta \in (\theta^C_-, \theta^C_+)$, $\frac{d}{d\theta}s^C(\theta) =^{sign} V_p(\cdot)P_{s\theta}(\cdot) + V_{sp}(\cdot)P_\theta(\cdot)$ by Assumption 2. Since $P_\theta(\cdot) = \frac{P(\cdot)}{\theta}$, and $V_p(\cdot)P_{s\theta}(\cdot) = \frac{V_p(\cdot)P_s(\cdot)}{\theta} = -\frac{V_s(\cdot)}{\theta}$ (where the last equality is due to the FOC), the r.h.s. simplifies to $V_p(\cdot)P_{s\theta}(\cdot) + V_{sp}(\cdot)P_\theta(\cdot) = \frac{1}{\theta}[V_{sp}(\cdot)P(\cdot) - V_s(\cdot)]$. From (4.3), $V_s(s,p) = -pf_h(s)v + (1-p)f_l(s)c$ and $V_{sp}(s,p) = -[f_h(s)v + cf_l(s)]$, which imply $V_{sp}(\cdot)P(\cdot) - V_s(\cdot) = -f_l(s)c < 0$. Hence, $\frac{d}{d\theta}s^C(\theta) < 0$ for $\theta \in (\theta^C_-, \theta^C_+)$. For $\theta > \theta^C_+$, we have $P(s^C(\theta),\theta) = 1$, and thus $\frac{d}{d\theta}s^C(\theta) = -\frac{P_\theta(\cdot)}{P_s(\cdot)} < 0$, where $P_s(\cdot) > 0$ because $s^C(\theta) < s^* = L^{-1}(1)$. For $\theta = \theta^C_+$, $s^C(\theta)$ is not differentiable, but it is clearly strictly decreasing. As a result, $s^C(\theta)$ is strictly decreasing for all $\theta > \theta^C_-$. Given that $s^C(\theta)$

is strictly decreasing and $V(s, P(s,\theta))$ is strictly quasi-concave by Assumption 2 for $\theta \in (\theta_-^C, \theta_+^C)$, it easily follows that $p^C(\theta)$ is strictly increasing for $\theta \in (\theta_-^C, \theta_+^C)$.

Finally, suppose, to the contrary, that $\lim_{\theta\to\infty} s^C(\theta) = \underline{s}^C > \underline{\sigma}$. Then, there would be a sufficiently large $\widehat{\theta} < \infty$ such that $p^C(\widehat{\theta}) = 1$ and $s^C(\widehat{\theta}) > \underline{\sigma}$. But in this case, the principal could strictly improve her payoff, $v(1 - F_h(s))$, by setting a standard, $(s^C(\widehat{\theta}) + \underline{\sigma})/2$. Hence, $\lim_{\theta\to\infty} s^C(\theta) = \underline{\sigma}$. ‖

Proof of Lemma 1. Since, by definition, $V_s(S(p), p) = U_p(P(s,\theta), s; \theta) = 0$, (4.1) and (4.3) imply that $\frac{d}{ds}U(P(s,\theta), s; \theta) = \frac{\partial}{\partial s}U(P(s,\theta), s; \theta) < 0$ and $\frac{d}{dp}V(S(p), p) = \frac{\partial}{\partial p}V(S(p), p) > 0$. ‖

Proof of Proposition 2. To prove the existence of the cutoff type θ_-^I, observe that $V(s,p)$ is strictly quasi-concave in s because $L'(s) > 0$. This means $s^I(\theta) = \overline{\sigma}$ if and only if $\lim_{s\to\overline{\sigma}} V_s(s, P(s,\theta)) > 0$. Since $\lim_{s\to\overline{\sigma}} P(s,\theta) = 0$ and $\lim_{s\to\overline{\sigma}} L(s)P(s,\theta) = u\theta(\lambda_h - \lambda_l)$ by Assumption 1, we have $s^I(\theta) = \overline{\sigma}$ and thus $p^I(\theta) = 0$ if and only if $\theta < \frac{r}{u(\lambda_h - \lambda_l)} = \theta_-^I$. Next, note that $p^I(\theta) < 1$ for all $\theta < \infty$; otherwise, if $p^I(\widehat{\theta}) = 1$ for some $\widehat{\theta} < \infty$, then $s^I(\widehat{\theta}) = S(1) = \underline{\sigma}$, to which the agent would reply by $p^I(\widehat{\theta}) = 0$, contradicting $p^I(\widehat{\theta}) = 1$. Note also that $p^I(\theta) < 1$ implies that $s^I(\theta) > \underline{\sigma}$ for all $\theta < \infty$.

Now, we prove that $s^I(\theta)$ is strictly decreasing for $\theta > \theta_-^I$. Recall from the text that $s^I(\theta) = s_0^I(\theta)$, which is the smallest root to (5.11), or equivalently to $\widehat{S}^{-1}(s,r) - P(s,\theta) = 0$, where $\widehat{S}^{-1}(s,r) \equiv S^{-1}(L^{-1}(r\frac{1-p}{p}))$. Differentiating the last equation with respect to θ, we find

$$\frac{\partial}{\partial\theta} s_0^I(\theta) = \frac{P_\theta(s_0^I(\theta), \theta)}{\widehat{S}_s^{-1}(s_0^I(\theta), r) - P_s(s_0^I(\theta), \theta)}.$$

Clearly, $P_\theta(s_0^I(\theta), \theta) = u[F_l(s_0^I(\theta)) - F_h(s_0^I(\theta))] > 0$ because $p^I(\theta) = P(s_0^I(\theta), \theta) < 1$. To show that $\widehat{S}_s^{-1}(s_0^I(\theta), r) - P_s(s_0^I(\theta), \theta) < 0$, note that $\widehat{S}_s^{-1}(s_0^I(\theta), r) < 0$ by our equilibrium selection, and thus $\widehat{S}_s^{-1}(s_0^I(\theta), r) - P_s(s_0^I(\theta), \theta) < 0$ whenever $P_s(s_0^I(\theta), \theta) \geq 0$. Suppose that $P_s(s_0^I(\theta), \theta) < 0$ and, to the contrary, that $\widehat{S}_s^{-1}(s_0^I(\theta), r) - P_s(s_0^I(\theta), \theta) \geq 0$. If $\widehat{S}_s^{-1}(s_0^I(\theta), r) - P_s(s_0^I(\theta), \theta) = 0$ for some θ, then by continuity, there would be a sufficiently small $\epsilon > 0$ such that $s_0^I(\theta) = s_0^I(\theta + \epsilon)$. But, from (4.2), this would imply $p^I(\theta) < p^I(\theta + \epsilon)$, and thus $s_0^I(\theta) \neq s_0^I(\theta + \epsilon)$, yielding a contradiction. Hence, $\widehat{S}_s^{-1}(s_0^I(\theta), r) - P_s(s_0^I(\theta), \theta) > 0$. But then, by continuity, we would have $\widehat{S}_s^{-1}(s_0^I(\theta) - \epsilon, r) - P_s(s_0^I(\theta) - \epsilon, \theta) < 0$ for some $\epsilon > 0$. Given $\widehat{S}^{-1}(\underline{\sigma}, r) - P(\underline{\sigma}, \theta) > 0$, this would generate another equilibrium with a greater effort, contradicting our equilibrium selection. Thus, for $\theta > \theta_-^I$ and thus $s_0^I(\theta) \neq \overline{\sigma}$, it must be that $\widehat{S}_s^{-1}(s_0^I(\theta), r) - P_s(s_0^I(\theta), \theta) < 0$. Hence, $\frac{\partial}{\partial\theta} s_0^I(\theta) < 0$.

Given that $\frac{\partial}{\partial\theta} s_0^I(\theta) < 0$, and from (5.10) that $p^I(\theta) = \frac{r}{r + L(s_0^I(\theta))}$, it immediately follows

$$\frac{\partial}{\partial\theta} p^I(\theta) = -\frac{rL'(\cdot)}{[r + L(\cdot)]^2} \frac{\partial}{\partial\theta} s_0^I(\theta) > 0.$$

Finally, to prove the limits as $\theta \to \infty$, suppose $\lim_{\theta\to\infty} p^I(\theta) \to \overline{p}^I < 1$. From (5.10), this would imply $\lim_{\theta\to\infty} s_0^I(\theta) = \underline{\sigma}^I > \underline{\sigma}$. But then, since $F_l(\underline{\sigma}^I) - F_h(\underline{\sigma}^I) > 0$, it must be that $\overline{p}^I = 1$ by (4.2), yielding a contradiction. Hence, $\overline{p}^I = 1$. As a result, $\lim_{\theta\to\infty} s^I(\theta) = \underline{\sigma}$. ‖

Proof of Proposition 3. Suppose $\theta > \theta_-^C$. Since $\theta_-^C < \theta_-^I$, clearly $s^C(\theta) < s^I(\theta) = \overline{\sigma}$ for $\theta \in (\theta_-^C, \theta_-^I)$. Next, note that $\theta_+^C > \theta^*$: otherwise, if $\theta_+^C \leq \theta^*$, then there would be some finite $\theta \geq \theta_+^C$ for which $p^I(\theta) = 1$, contradicting Proposition 2.

Suppose $\theta < \theta^*$, but, on the contrary, $s^C(\theta) = s_0^C(\theta) \geq s^I(\theta)$. Since $s^I(\theta) > s^I(\theta^*)$, it follows that $V_s(s^I(\theta), p^I(\theta)) = 0$ and $P_s(s^I(\theta), \theta) < 0$, which imply $\frac{d}{ds}V(s, P(s,\theta))\big|_{s=s^I(\theta)} < 0$, and by Assumption 2 reveal $s^C(\theta) < s^I(\theta)$ – a contradiction. Thus, $s^C(\theta) < s^I(\theta)$ for $\theta < \theta^*$. Given $\theta_+^C > \theta^*$, a similar line of argument shows that $s^C(\theta^*) = s^I(\theta^*)$, and $s^C(\theta) > s^I(\theta)$ for $\theta^* < \theta \leq \theta_+^C$. Finally, suppose $\theta > \theta_+^C$. Then, $s^C(\theta) = \min\{s | P(s,\theta) = 1\}$ by Proposition 1. If, on the contrary, $s^C(\theta) \leq s^I(\theta)$, then since $\theta > \theta_+^C$, we would have $1 = P(s^C(\theta), \theta) \leq P(s^I(\theta), \theta)$ because $P_s(s,\theta) > 0$ for $\theta > \theta_+^C$. But this would imply $p^I(\theta) = 1$, contradicting Proposition 1. Hence, $s^C(\theta) > s^I(\theta)$ for $\theta > \theta_+^C$. Finally, to prove $p^C(\theta) \geq p^I(\theta)$ with strict inequality if $\theta \neq \theta^*$, simply recall $P_s(s,\theta) =^{sign} s^* - s$ whenever $P(s,\theta) < 1$. ‖

Proof of Proposition 4. The claim directly follows from (6.13), (6.14), and (6.15). ‖

Proof of Proposition 5. Since $s^I(\theta)$ is strictly decreasing for $\theta > \theta_-^I$, and $s^B = s^I(E[\theta])$, part (i) immediately follows. To prove part (ii), note that $F_l(s) - F_h(s)$ is strictly quasi-concave in s, achieving its unique maximum at $s = s^*$. Since $s^I(E[\theta])$ is strictly decreasing, this implies that $F_l(s^B) - F_h(s^B) = F_l(s^I(E[\theta])) - F_h(s^I(E[\theta]))$ is strictly quasi-concave in $E[\theta]$, admitting its maximum at $E[\theta] = \theta^*$. Thus, for a given type θ, $p^B(\theta) - p^I(\theta)$ strictly increases in $E[\theta]$ as $E[\theta]$ approaches θ^*, and so does $E[p^B(\theta)] - E[p^I(\theta)]$. Moreover, if $E[\theta] = \theta^*$ so that $s^B = s^*$, then $p^B(\theta) - p^I(\theta) > 0$ for all $\theta \neq \theta^*$ and $p^B(\theta) - p^I(\theta) = 0$ for $\theta = \theta^*$, revealing that $E[p^B(\theta)] - E[p^I(\theta)] > 0$ at $E[\theta] = \theta^*$. ‖

Proof of Lemma 2. Consider $\theta > \theta_-^I$ so that $p^I > 0$. We first prove the following claim.

Claim. $\frac{p^{I''}}{p^{I'}s^{I'}} \to \infty$, as $\theta \to \infty$.

Proof. Since $p^I < 1$ by Proposition 2, $p^I = \theta u(F_l(s^I) - F_h(s^I))$. Differentiating with respect to θ, we obtain

$$p^{I'} = u[(F_l(s^I) - F_h(s^I)) + \theta(f_l(s^I) - f_h(s^I))s^{I'}]$$

and

$$p^{I''} = u[(f_l(s^I) - f_h(s^I))(2s^{I'} + \theta s^{I''}) + (f_l'(s^I) - f_h'(s^I))\theta(s^{I'})^2].$$

Again, by Proposition 2, we know $p^I \to 1$ and $s^I \to \underline{\sigma}$ as $\theta \to \infty$. Suppose $\theta s^{I'} \to a < 0$. Then, $p^{I'} \to u[f_l(\underline{\sigma}) - f_h(\underline{\sigma})]a < 0$ because $f_l(\underline{\sigma}) - f_h(\underline{\sigma}) > 0$, which contradicts $p^{I'} > 0$. Hence, $\theta s^{I'} \to 0$, which implies $s^{I'} \to 0$. Next, by definition $\frac{p^{I'}}{\theta s^{I'}} = u[\frac{F_l(s^I) - F_h(s^I)}{\theta s^{I'}} + f_l(s^I) - f_h(s^I)]$, and by L'Hopital's rule, $\lim_{\theta\to\infty} \frac{F_l(s^I) - F_h(s^I)}{\theta s^{I'}} = \lim_{\theta\to\infty} \frac{(f_l(s^I) - f_h(s^I))s^{I'}}{s^{I'} + \theta s^{I''}}$. This means for a sufficiently large θ, $\frac{p^{I'}}{\theta s^{I'}} \approx u[f_l(\underline{\sigma}) - f_h(\underline{\sigma})]\frac{2s^{I'} + \theta s^{I''}}{s^{I'} + \theta s^{I''}}$. Since $\frac{p^{I'}}{\theta s^{I'}} < 0$, $f_l(\underline{\sigma}) - f_h(\underline{\sigma}) > 0$ and $s^{I'} < 0$, we must have $s^{I'} + \theta s^{I''} > 0$ and $2s^{I'} + \theta s^{I''} < 0$. Now, observe that since $s^{I'} \to 0$ and $s^{I''} > 0$, it follows $s^{I''} \to 0$. Using L'Hopital's rule, $\frac{s^{I'}}{s^{I''}} \approx \frac{s^{I''}}{s^{I'''}}$ for a sufficiently large θ, implying that $s^{I'''} < 0$. A similar limit argument shows that $\frac{F_l(s^I) - F_h(s^I)}{s^{I'}} \approx \frac{[f_l(\underline{\sigma}) - f_h(\underline{\sigma})]s^{I'}}{s^{I''}} \approx \frac{[f_l'(\underline{\sigma}) - f_h'(\underline{\sigma})](s^{I'})^2 + [f_l(\underline{\sigma}) - f_h(\underline{\sigma})]s^{I''}}{s^{I'''}}$. Given $s^{I'''} < 0$, the numerator of the last ratio must be positive, or, equivalently $\frac{f_l'(\underline{\sigma}) - f_h'(\underline{\sigma})}{f_l(\underline{\sigma}) - f_h(\underline{\sigma})} > -\frac{s^{I''}}{(s^{I'})^2}$. Next, consider the following ratio:

$$\frac{p^{I''}}{p^{I'}s^{I'}} = \frac{u[(f_l(s^I) - f_h(s^I))(2s^{I'} + \theta s^{I''}) + (f_l'(s^I) - f_h'(s^I))\theta(s^{I'})^2]}{p^{I'}s^{I'}}.$$

Dividing the right side by $\theta(s^{I\prime})^2$, we obtain

$$\frac{p^{I\prime\prime}}{p^{I\prime}s^{I\prime}} = \frac{(f_l(s^I) - f_h(s^I))\frac{2s^{I\prime}+\theta s^{I\prime\prime}}{\theta(s^{I\prime})^2} + f_l'(s^I) - f_h'(s^I)}{\frac{p^{I\prime}}{\theta s^{I\prime}}}u.$$

Thus, for a large θ

$$\begin{aligned}\frac{p^{I\prime\prime}}{p^{I\prime}s^{I\prime}} &\approx \frac{[f_l(\underline{\sigma}) - f_h(\underline{\sigma})]\frac{2s^{I\prime}+\theta s^{I\prime\prime}}{\theta(s^{I\prime})^2} + f_l'(\underline{\sigma}) - f_h'(\underline{\sigma})}{[f_l(\underline{\sigma})f_h(\underline{\sigma})]\frac{2s^{I\prime}+\theta s^{I\prime\prime}}{s^{I\prime}+\theta s^{I\prime\prime}}}u \\ &= \frac{s^{I\prime}+\theta s^{I\prime\prime}}{\theta(s^{I\prime})^2}u + \frac{[f_l'(\underline{\sigma}) - f_h'(\underline{\sigma})]}{[f_l(\underline{\sigma}) - f_h(\underline{\sigma})]}\frac{s^{I\prime}+\theta s^{I\prime\prime}}{2s^{I\prime}+\theta s^{I\prime\prime}}u \\ &> \frac{s^{I\prime}+\theta s^{I\prime\prime}}{\theta(s^{I\prime})^2}u - \frac{s^{I\prime\prime}}{(s^{I\prime})^2}\frac{s^{I\prime}+\theta s^{I\prime\prime}}{2s^{I\prime}+\theta s^{I\prime\prime}}u \\ &= \frac{s^{I\prime}+\theta s^{I\prime\prime}}{2s^{I\prime}+\theta s^{I\prime\prime}}\frac{2}{\theta s^{I\prime}}u = +\infty,\end{aligned}$$

where the last line follows because $p^{I\prime} = -\frac{(p^I)^2L'}{r}s^{I\prime}$ and $L'(\underline{\sigma}) < \infty$, revealing $\frac{p^{I\prime}}{\theta s^{I\prime}} = -\frac{(p^I)^2L'}{\theta r} \to 0^-$ as $\theta \to \infty$. Furthermore, given $\frac{p^{I\prime}}{\theta s^{I\prime}} \approx u[f_l(\underline{\sigma}) - f_h(\underline{\sigma})]\frac{2s^{I\prime}+\theta s^{I\prime\prime}}{s^{I\prime}+\theta s^{I\prime\prime}}$, we have $\frac{2s^{I\prime}+\theta s^{I\prime\prime}}{s^{I\prime}+\theta s^{I\prime\prime}} \to -\infty$, which, together with $\frac{2}{\theta s^{I\prime}} \to -\infty$, proves the claim. ‖

Now, in general

$$V^{I\prime\prime} = (V_{ss}s^{I\prime} + V_{sp}p^{I\prime})s^{I\prime} + V_s s^{I\prime\prime} + (V_{ps}s^{I\prime} + V_{pp}p^{I\prime})p^{I\prime} + V_p p^{I\prime\prime}.$$

Consider the right side of this expression. The first term is zero by straightforward (but tedious) algebra. Moreover, $V_s = 0$ by the principal's FOC and $V_{pp} = 0$ by linearity. Hence,

$$V^{I\prime\prime} = V_{ps}s^{I\prime}p^{I\prime} + V_p p^{I\prime\prime}, \tag{A1}$$

where $V_{ps} = -(f_h(s^I)v + f_l(s^I)c)$ and $V_p = (1 - F_h(s^I))v + (1 - F_l(s^I))c$.

If $\theta \to \theta^I_-$, then $s^I \to \overline{\sigma}$, which implies $V_{ps} \to -f_h(\overline{\sigma})v - f_l(\overline{\sigma})c < 0$ and $V_p \to 0$. Hence, $V^{I\prime\prime} > 0$. Finally, for $\theta \to \infty$, we have $s^I \to \underline{\sigma}$, $V_{ps} \to -f_h(\underline{\sigma})v - cf_l(\underline{\sigma}) < 0$ and $V_p \to v + c > 0$. Moreover, since, by the earlier Claim, $\frac{p^{I\prime\prime}}{p^{I\prime}s^{I\prime}} \to +\infty$, the second term in (A1) dominates, and given $p^{I\prime\prime} < 0$, implies $V^{I\prime\prime} < 0$.

Overall, this establishes the existence of two cutpoints $\theta^I_- < \theta_L \leq \theta_H < \infty$ such that $V^I(\theta)$ is strictly convex for $\theta < \theta_L$ and strictly concave for $\theta > \theta_H$. ‖

Proof of Proposition 6. Suppose $\underline{\theta} < \theta_L \leq \theta_H < \overline{\theta}$, where θ_L and θ_H are the two cutpoints defined in Lemma 2, and by definition, they are independent of the type distribution. Since $V^I(\theta)$ is strictly concave for $\theta > \theta_H$, Jensen's inequality implies

$$\int_{\theta_H}^{\overline{\theta}} V^I(\theta)\frac{dG(\theta)}{1 - G(\theta_H)} < V^I\left(\int_{\theta_H}^{\overline{\theta}} \theta\frac{dG(\theta)}{1 - G(\theta_H)}\right),$$

or equivalently

$$\int_{\theta_H}^{\overline{\theta}} V^I(\theta)dG(\theta) < (1 - G(\theta_H))V^I\left(\int_{\theta_H}^{\overline{\theta}} \theta\frac{dG(\theta)}{1 - G(\theta_H)}\right).$$

Then, by adding the term $\int_{\underline{\theta}}^{\theta_H} V^I(\theta)dG(\theta)$ to both sides,

$$E[V^I(\theta)] < (1-G(\theta_H))V^I\left(\frac{\mu(G(\theta_H)) - \int_{\underline{\theta}}^{\theta_H}\theta dG(\theta)}{1-G(\theta_H)}\right) + \int_{\underline{\theta}}^{\theta_H} V^I(\theta)dG(\theta)$$
$$\leq (1-G_H)V^I\left(\frac{\mu(G_H) - \underline{\theta}G_H}{1-G_H}\right) + V^I(\theta_H)\,G_H \equiv \Phi(G_H),$$

where $G(\theta_H) \equiv G_H$ and $\mu(G_H)$ is the mean type given G_H. For $G_H = 0$, note that $E[V^I(\theta)] < V^I(\mu(0)) = \Phi(0)$ because $V^I(\theta)$ is strictly concave for $\theta > \theta_H$ by hypothesis, and $\mu(0) > \theta_H$. Given that $\Phi(G_H)$ is continuous, there must exist some $\epsilon^B > 0$ such that $E[V^I(\theta)] < V^I(\mu(G_H))$ whenever $G_H < \epsilon^B$. A similar line of argument shows that there must also exist some $\epsilon^I > 0$ such that $E[V^I(\theta)] > V^I(\mu(G_L))$ whenever $G_L \equiv G(\theta_L) > 1-\epsilon^I$, completing the proof of part (i).

To prove the second part, fix a type distribution $G(\theta)$, and let $\widehat{\theta} \equiv u\theta \in [u\underline{\theta}, u\overline{\theta}]$ so that $\widehat{G}_u(\widehat{\theta}) = G(\frac{\theta}{u})$. Clearly, $\widehat{G}_u(\theta_H) = G(\frac{\theta_H}{u}) \to 0$ as $u \to \infty$, which implies that there is some $u_H > 0$ such that $\widehat{G}_u(\theta_H) < \epsilon^B$ for all $u > u_H$. Applying the result from part (i), we then have $E[V^I(\theta)] < V^I(E[\theta])$ for all $u > u_H$. A parallel line of argument reveals that there is some $u_L > 0$ such that $\widehat{G}_u(\theta_L) > 1-\epsilon^I$ for all $u < u_L$, and thus $E[V^I(\theta)] > V^I(E[\theta])$. ‖

Proof of Lemma 3. Recall that $f_h(s^*) = f_l(s^*)$. Moreover, $L'(s^*) = \frac{f_h'(s^*)-f_l'(s^*)}{f_l(s^*)} > 0$ by MLRP. From (A1)

$$V^{I\prime\prime} = V_{ps}s^{I\prime}p^{I\prime} + V_p p^{I\prime\prime}.$$

We wish to show that this is negative at $\theta = \theta^*$.

To obtain an expression for V_{ps}, invoke Young's Theorem and differentiate the principal's FOC with respect to p to get

$$V_{ps} = V_{sp} = -vp^I f_l(s^I)\frac{r}{(p^I)^2}.$$

To derive an expression for $s^{I\prime}(\theta^*)$, consider the matrix of second derivatives obtained by differentiating the FOCs of the agent and principal, respectively.

$$\begin{bmatrix} U_{pp} & U_{ps} \\ V_{sp} & V_{ss} \end{bmatrix} = \begin{bmatrix} -1 & \theta u(f_l(s^I) - f_h(s^I)) \\ -vp^I f_l(s^I)\frac{r}{(p^I)^2} & -vp^I f_l(s^I)L'(s^I) \end{bmatrix}.$$

Hence, differentiating the system of FOCs with respect to θ yields

$$\begin{bmatrix} -1 & \theta u(f_l(s^I) - f_h(s^I)) \\ -vp^I f_l(s^I)\frac{r}{(p^I)^2} & -vp^I f_l(s^I)L'(s^I) \end{bmatrix}\begin{bmatrix} p^{I\prime} \\ s^{I\prime} \end{bmatrix} = \begin{bmatrix} -u(F_l(s^I) - F_h(s^I)) \\ 0 \end{bmatrix}.$$

Applying Cramer's Rule and evaluating the result at $\theta = \theta^*$ reveals

$$s^{I\prime}(\theta^*) = -r\frac{u(F_l(s^*) - F_h(s^*))}{(p^I(\theta^*))^2 L'(s^*)}.$$

To derive expressions for $p^{I\prime}(\theta^*)$ and $p^{I\prime\prime}(\theta^*)$, differentiate $p^I(\theta) = \theta u[F_l(s^I(\theta)) - F_h(s^I(\theta))]$ and evaluate at $\theta = \theta^*$ to get

$$p^{I\prime}(\theta^*) = u(F_l(s^*) - F_h(s^*))$$

and

$$p^{I\prime\prime}(\theta^*) = -\theta^* u f_l(s^*)L'(s^*)(s^{I\prime}(\theta^*))^2,$$

where we use the facts that $f_h(s^*) = f_l(s^*)$ and $L'(s^*) = \frac{f_h'(s^*) - f_l'(s^*)}{f_l(s^*)}$.

Finally, noting $V_p = (1 - F_h(s^I))v + (1 - F_l(s^I))c$, and putting the pieces together give

$$V^{I\prime\prime}(\theta^*) = vp^*(\theta^*)f_l(\theta^*)L'(s^*)(s^{I\prime}(\theta^*))^2 - [(1 - F_h(s^*))v + (1 - F_l(s^*))c]\theta^* u f_l(s^*)L'(s^*)(s^{I\prime}(\theta^*))^2.$$

Factoring the common term $f_l(s^*)L'(s^*)(s^{I\prime}(\theta^*))^2 > 0$ out of the right side reveals that $V^{I\prime\prime}(\theta^*) < 0$ iff

$$vp^I(\theta^*) - [(1 - F_h(s^*))v + (1 - F_l(s^*))c]u\theta^* < 0,$$

iff

$$v\theta^* u[F_l(s^*) - F_h(s^*)] - [(1 - F_h(s^*))v + (1 - F_l(s^*))c]\theta^* u < 0,$$

iff

$$F_l(s^*) - F_h(s^*) - [(1 - F_h(s^*)) + (1 - F_l(s^*))r] < 0,$$

iff

$$-(1 - F_l(s^*))(1 + r) < 0,$$

which is evidently true. ‖

Proof of Proposition 7. By Lemma 3, $V^I(\theta)$ is strictly concave at θ^*. Hence, there exists $\Delta > 0$ such that

$$V^I(\theta) < V^I(\theta^*) + V^{I\prime}(\theta^*)(\theta - \theta^*)$$

for all $\theta \in (\theta^* - \Delta, \theta^* + \Delta)$ and $\theta \neq \theta^*$. Suppose $\underline{\theta} > \theta^* - \Delta$ and $\overline{\theta} < \theta^* + \Delta$. Then, taking expectations on each side of the above inequality yields

$$E[V^I(\theta)] < V^I(\theta^*) + V^{I\prime}(\theta^*)(E[\theta] - \theta^*). \tag{A2}$$

If $E[\theta] = \theta^*$, then this yields directly $E[V^I(\theta)] < V^I(E[\theta])$. Suppose $E[\theta] > \theta^*$. Then, setting $\epsilon = \frac{V^I(E[\theta]) - V^I(\theta^*)}{V^{I\prime}(\theta^*)} > 0$ and letting $E[\theta] - \theta^* < \epsilon$, we obtain

$$\begin{aligned} V^I(E[\theta]) &= V^I(\theta^*) + V^{I\prime}(\theta^*)\epsilon \\ &> V^I(\theta^*) + V^{I\prime}(\theta^*)(E[\theta] - \theta^*) \\ &> E[V^I(\theta)]. \end{aligned}$$

Finally, suppose $E[\theta] < \theta^*$. By way of contradiction, assume $E[V^I(\theta)] - V^I(E[\theta]) > 0$. Note that the value of θ^* does not affect this difference because it is independent of type distribution. Hence, by continuity of $V^I(\theta)$ and the fact that $V^{I\prime}(\theta)$ is bounded (both easily verified), for any θ^* sufficiently close to $E[\theta]$, we have

$$V^I(\theta^*) - V^I(E[\theta]) + V^{I\prime}(\theta^*)(E[\theta] - \theta^*) < E[V^I(\theta)] - V^I(E[\theta]).$$

But then, $V^I(\theta^*) + V^{I\prime}(\theta^*)(E[\theta] - \theta^*) < E[V^I(\theta)]$, contradicting (A2). Hence, $E[V^I(\theta)] - V^I(E[\theta]) \leq 0$. ‖

Lemma A1.

(i) *$F_h(\sigma;\alpha)$ decreases in α for $\sigma \neq \overline{\sigma}$; and $F_l(\sigma;\alpha)$ increases in α for $\sigma \neq \underline{\sigma}$.*
(ii) *For $\alpha_1 > \alpha_0$, $L(\sigma;\alpha_1) - L(\sigma;\alpha_0) =^{sign} \sigma - s^*$.*
(iii) *As $\alpha \to \infty$, we have $L(\sigma;\alpha) \to 0$ if $\sigma < s^*$, and $L(\sigma;\alpha) \to \infty$ if $\sigma > s^*$.*

PROOF. Take any $\alpha_1 > \alpha_0$. By Definition 1, $f_h(\sigma;\alpha)$ satisfies MLRP, which implies FOSD or $F_h(\sigma;\alpha_1) < F_h(\sigma;\alpha_0)$ for all $\sigma \neq \overline{\sigma}$. Using a parallel argument for $f^l(\sigma;\alpha)$, it also follows that $F_l(\sigma;\alpha_1) > F_l(\sigma;\alpha_0)$ for all $\sigma \neq \underline{\sigma}$. To prove part (ii), note first that $L(s^*;\alpha_1) - L(s^*;\alpha_0) = 0$ or equivalently $\frac{L(s^*;\alpha_1)}{L(s^*;\alpha_0)} = 1$ by Assumption 4. Next, since $\frac{L(\sigma;\alpha_1)}{L(\sigma;\alpha_0)} \equiv \frac{f_h(\sigma;\alpha_1)}{f_h(\sigma;\alpha_0)}\frac{f_l(\sigma;\alpha_0)}{f_l(\sigma;\alpha_1)}$, $\frac{L(\sigma;\alpha_1)}{L(\sigma;\alpha_0)}$ increases in σ. Suppose $\sigma < s^*$, and, to the contrary, $L(\sigma;\alpha_1) - L(\sigma;\alpha_0) \geq 0$ or $\frac{L(\sigma;\alpha_1)}{L(\sigma;\alpha_0)} \geq 1$. But then, $\frac{L(s^*;\alpha_1)}{L(s^*;\alpha_0)} > 1$ – a contradiction. Hence, $L(\sigma;\alpha_1) - L(\sigma;\alpha_0) < 0$. Now, suppose $\sigma > s^*$, and, to the contrary, $L(\sigma;\alpha_1) - L(\sigma;\alpha_0) \leq 0$ or $\frac{L(\sigma;\alpha_1)}{L(\sigma;\alpha_0)} \leq 1$. But then, $\frac{L(s^*;\alpha_1)}{L(s^*;\alpha_0)} < 1$ – a contradiction. Hence, $L(\sigma;\alpha_1) - L(\sigma;\alpha_0) > 0$. Finally, note that Assumption 4 implies that $L(\sigma;\alpha) = \psi(\sigma)^{h(\alpha)}$ for some ψ and h such that $\psi(\underline{\sigma}) = 0$, $\psi(s^*) = 1$, $\psi(\overline{\sigma}) = \infty$, $\psi' > 0$, and that $h(0) = 0$, $h(\infty) = \infty$, $h' > 0$. From here, if $\sigma < s^*$, then $\psi(\sigma) < 1$, and thus $L(\sigma;\alpha) \to 0$ as $\alpha \to \infty$. And if $\sigma > s^*$, then $\psi(\sigma) > 1$, and thus $L(\sigma;\alpha) \to \infty$ as $\alpha \to \infty$. ‖

PROOF OF LEMMA 4. Suppose $p^I(\theta;\alpha) > 0$. In an informed equilibrium, $p^I(\theta;\alpha) < 1$ by Proposition 2, which means $p^I(\theta,\alpha) = \theta u[F_l(s^I(\theta,\alpha);\alpha) - F_h(s^I(\theta,\alpha);\alpha)]$. Differentiating with respect to α, we obtain

$$p^I_\alpha(\theta;\alpha) = \theta u[F_{l,\alpha}(\cdot) - F_{h,\alpha}(\cdot) + (f_l(\cdot) - f_h(\cdot))s^I_\alpha(\theta;\alpha)].$$

Since $L(s^I(\theta,\alpha),\alpha) = \frac{1-p^I(\theta,\alpha)}{p^I(\theta,\alpha)}r$, we also obtain

$$s^I_\alpha(\theta,\alpha) = -\frac{L_\alpha(\cdot) + \frac{1}{(p^I(\cdot))^2}rp^I_\alpha(\theta,\alpha)}{L_s(\cdot)}.$$

Substituting for $s^I_\alpha(\theta,\alpha)$ and arranging terms yields

$$p^I_\alpha(\theta,\alpha) = \theta u\frac{(F_{l,\alpha}(\cdot) - F_{h,\alpha}(\cdot))L_s(\cdot) + f_l(\cdot)(L(s^I) - 1)L_\alpha(\cdot)}{L_s(\cdot) + \theta u(f_l(\cdot) - f_h(\cdot))\frac{r}{(p^I(\cdot))^2}}.$$

Notice that $(L(s^I) - 1)L_\alpha(\cdot) > 0$ for $s^I \neq s^*$, because $L(s^I) - 1 =^{sign} L_\alpha(s^I)$ by Lemma A1. Hence, the numerator is strictly positive. The denominator is also positive because of the stability of the equilibrium (i.e., the principal's reaction function crosses the agent's from above). Hence, $p^I_\alpha(\theta;\alpha) > 0$. To prove part (ii), note that as $\alpha \to 0$, we have $F_l(\sigma;\alpha) - F_h(\sigma;\alpha) \to 0$ for all σ, and thus $p^I(\theta;\alpha) \to 0$ and $s^I(\theta;\alpha) \to \overline{\sigma}$. Finally, since $p^I(\theta,\alpha)$ is increasing and bounded, $p^I(\theta,\alpha) \to \overline{p}$ as $\alpha \to \infty$ for some $\overline{p}$. Suppose $\overline{p} < \theta u < 1$. Then, $\lim_{\alpha\to\infty} s^I(\theta;\alpha) \neq \underline{\sigma}, \overline{\sigma}$, which means $F_l(s^I(\theta;\alpha),\alpha) \to 1$ and $F_h(s^I(\theta;\alpha),\alpha) \to 0$ as $\alpha \to \infty$, revealing $\overline{p} = \theta u$ – a contradiction. Hence, $\overline{p} = \theta u$. Moreover, since $\frac{1-\overline{p}}{\overline{p}}r = \frac{1-\theta u}{\theta u}r \neq 0, \infty$, Lemma A1 implies that $\lim_{\alpha\to\infty} s^I(\theta;\alpha) = s^*$ whenever $\theta u < 1$. A similar line of argument proves the case when $\theta u \geq 1$. ‖

PROOF OF PROPOSITION 8 Suppose $\overline{\theta}u < 1$. Part (i) follows directly from Lemma 4. To prove part (ii), note from (A1) that

$$V^I_{\theta\theta}(\theta;\alpha) = V_{ps}p^{I\prime}s^{I\prime} + V_p p^{I\prime\prime},$$

where

$$V_{ps} = -f_l(\sigma;\alpha)[L(\sigma;\alpha)v + c],$$
$$V_p = [1 - F_h(\sigma;\alpha)]v + [1 - F_l(\sigma;\alpha)]c,$$

and

$$p^{I\prime} = u[F_l(\sigma;\alpha) - F_h(\sigma;\alpha) + \theta(f_l(\sigma;\alpha) - f_h(\sigma;\alpha))s^{I\prime}],$$
$$p^{I\prime\prime} = u[(f_l(\sigma;\alpha) - f_h(\sigma;\alpha))(2s^{I\prime} + \theta s^{I\prime\prime}) + (f_l'(\sigma;\alpha) - f_h'(\sigma;\alpha))\theta(s^{I\prime})^2].$$

As $\alpha \to 0$, we have $p^I(\theta;\alpha) \to 0$ and $s^I(\theta;\alpha) \to \overline{\sigma}$ by Lemma 4, and thus $V_p \to 0$, $V_{ps} \to -f_h(\overline{\sigma})v$, and $p^{I\prime} \to -\theta u f_h(\overline{\sigma})s^{I\prime}$, implying that $V^I_{\theta\theta}(\theta;\alpha) \to \theta uv(f_h(\overline{\sigma}))^2(s^{I\prime})^2 > 0$. Thus, $V^I(\theta;\alpha)$ is strictly convex for a sufficiently small $\alpha > 0$, and $E[V^I(\theta;\alpha)] > E[V^B(\theta;\alpha)]$ by Jensen's inequality.

As $\alpha \to \infty$, we have $p^I(\theta;\alpha) \to u\theta$ and $s^I(\theta;\alpha) \to s^*$ by Lemma 4, which imply that $F_h(s^I(\theta;\alpha);\alpha) \to 0$, $F_l(s^I(\theta;\alpha);\alpha) \to 1$, $f_l(s^I(\theta;\alpha);\alpha) - f_h(s^I(\theta;\alpha);\alpha) \to 0$, $p^{I\prime} \to u$, and $V_p \to v$. In addition, as $\alpha \to \infty$, we also have $p^{I\prime\prime} \to u(f_l'(s^*;\infty) - f_h'(s^*;\infty))\theta(s^{I\prime})^2 < 0$ because $f_l'(s^*;\alpha) - f_h'(s^*;\alpha) = -L'(s^*;\alpha)f_l(s^*;\alpha) < 0$ for any α. Similarly, as $\alpha \to \infty$, $V_{ps} \to 0$ because $L(s^*;\infty) = 1$ and $f_l(s^*;\infty) = 0$. Overall, as $\alpha \to \infty$, $V^I_{\theta\theta}(\theta;\alpha) \to vu\theta(f_l'(s^*;\infty) - f_h'(s^*;\infty))(s^{I\prime})^2 < 0$. This means $V^I(\theta;\alpha)$ is strictly concave for a sufficiently large $\alpha < \infty$, and $E[V^I(\theta;\alpha)] < E[V^B(\theta;\alpha)]$ by Jensen's inequality. ‖

Proof of Lemma 5. Integrating by parts and arranging terms, (9.16) simplifies to:

$$H(m_1;\phi) \equiv 2(1-\phi)\int_{\underline{\theta}}^{m_1} G(\theta)d\theta + \phi m_1 - \phi E[\theta] = 0. \tag{A3}$$

Note that $H(m_1;\phi)$ is strictly increasing in m_1 and that $H(\underline{\theta};\phi) \le 0$ and $H(E[\theta];\phi) > 0$. Thus, there is a unique solution $m_1 = \mu_1(\phi)$ to (A3), and its boundary values are $\mu_1(0) = \underline{\theta}$ and $\mu_1(1) = E[\theta]$. Furthermore, $\mu_1(\phi)$ is strictly increasing because $H_\phi(m_1;\phi) = -2\int_{\underline{\theta}}^{m_1} G(\theta)d\theta + m_1 - E[\theta] < 0$ for $m_1 \le E[\theta]$. ‖

Proof of Lemma 6. Recall that

$$\pi_1^{B,I}(\phi) = \left[\left(1 - \frac{\phi}{2}\right)G(\mu_1(\phi)) + \frac{\phi}{2}(1 - G(\mu_1(\phi)))\right]V^I(\mu_1(\phi))$$

and

$$\pi_2^{B,I}(\phi) = \frac{\phi}{2}\int_{\underline{\theta}}^{\mu_1(\phi)} V^I(\theta)dG(\theta) + (1 - \frac{\phi}{2})\int_{\mu_1(\phi)}^{\overline{\theta}} V^I(\theta)dG(\theta).$$

From Lemma 5, we have $\pi_1^{B,I}(0) = 0$ and $\pi_1^{B,I}(1) = \frac{1}{2}V^I(E[\theta])$. Next, we differentiate $\pi_1^{B,I}(\phi)$ to obtain

$$\frac{d}{d\phi}\pi_1^{B,I}(\phi) = \left[(1-\phi)G(\mu_1) + \frac{\phi}{2}\right]\left[\frac{d}{d\theta}V^I(\mu_1)\right]\mu_1' + \left[-G(\mu_1) + (1-\phi)g(\mu_1)\mu_1' + \frac{1}{2}\right]V^I(\mu_1) > 0,$$

where $-G(\mu_1) + (1-\phi)g(\mu_1)\mu_1' + \frac{1}{2} \ge 0$ follows from (9.16).

Again, from Lemma 5, we have $\pi_2^{B,I}(0) = E[V^I(\theta)]$ and $\pi_2^{B,I}(1) = \frac{1}{2}E[V^I(\theta)]$. Differentiating $\pi_2^{B,I}(\phi)$ yields

$$\frac{d}{d\phi}\pi_2^{B,I}(\phi) = \frac{1}{2}\left[\int_{\underline{\theta}}^{\mu_1(\phi)} V^I(\theta)dG(\theta) - \int_{\mu_1(\phi)}^{\overline{\theta}} V^I(\theta)dG(\theta)\right] - (1-\phi)V^I(\mu_1)g(\mu_1)\mu_1' < 0. \ ‖$$

Proof of Proposition 9. Suppose $\frac{1}{2}V^I(E[\theta]) < E[V^I(\theta)] < V^I(E[\theta])$. Since $0 < \frac{1}{2}E[V^I(\theta)] < \frac{1}{2}V^I(E[\theta])$, by Lemma 4 there is $\phi_1 \in (0,1)$ such that $\pi_1^{B,I}(\phi) < \frac{1}{2}E[V^I(\theta)]$ if and only if $\phi < \phi_1$. Moreover, since $\frac{1}{2}E[V^I(\theta)] < \frac{1}{2}V^I(E[\theta]) < E[V^I(\theta)]$, by Lemma

4 there is also $\phi_2 \in (0,1)$ such that $\pi_2^{B,I}(\phi) > \frac{1}{2}V^I(E[\theta])$ if and only if $\phi < \phi_2$. Define $\phi^* = \min\{\phi_1, \phi_2\}$ and $\phi^{**} = \max\{\phi_1, \phi_2\}$. Clearly, $0 < \phi^* \leq \phi^{**} < 1$. If $\phi < \phi^*$, then $\pi_1^{B,I}(\phi) < \frac{1}{2}E[V^I(\theta)]$ and $\pi_2^{B,I}(\phi) > \frac{1}{2}V^I(E[\theta])$, which imply that choosing informed review is a strictly dominant strategy for each evaluator, and thus the unique equilibrium has $\tau_1 = \tau_2 = I$. A similar argument shows that if $\phi > \phi^{**}$, then choosing blind review is a strictly dominant strategy for each evaluator, and thus the unique equilibrium has $\tau_1 = \tau_2 = B$. Suppose $\phi^* \neq \phi^{**}$ and $\phi \in (\phi^*, \phi^{**})$. If $\phi^* = \phi_2$ and $\phi^{**} = \phi_1$, then there are exactly two equilibria: $\tau_1 = \tau_2 = I$ and $\tau_1 = \tau_2 = B$. If, on the other hand, $\phi^* = \phi_1$ and $\phi^{**} = \phi_2$, then there are also two equilibria: $\tau_i = B$ and $\tau_j = I$ for $i, j = 1, 2$ and $i \neq j$. ‖

Proof of Proposition 10. Suppose $c > 0$ and $\overline{c} > 0$. To prove part (i), observe that $\overline{V}_{ps}(s,p) < 0$, $\overline{V}_p(\underline{\sigma}, p) = v + c > 0$, and $\overline{V}_p(\overline{\sigma}, p) = -\overline{c} < 0$. Thus, there is a unique $\sigma_d \in (\underline{\sigma}, \overline{\sigma})$ such that $\overline{V}_p(s,p) =^{sign} \sigma_d - s$. In equilibrium under informed review, the exact argument used in the proof of Proposition 2 reveals that $\overline{s}^I(\theta) = \overline{\sigma}$ (and $\overline{p}^I(\theta) = 0$) if and only if $\theta < \frac{\overline{r}}{u(\lambda_h - \lambda_l)} = \overline{\theta}_-^I$. Thus, $\overline{V}^I(\theta) = 0$ for $\theta < \overline{\theta}_-^I$. The arguments in Proposition 2 also reveal $\frac{d}{d\theta}\overline{s}^I(\theta) < 0$ and $\frac{d}{d\theta}\overline{p}^I(\theta) > 0$ for $\theta \geq \overline{\theta}_-^I$. Let $\theta \geq \overline{\theta}_-^I$. Then, $\frac{d}{d\theta}\overline{V}^I(\theta) = \frac{d}{d\theta}\overline{V}(\overline{s}^I(\theta), \overline{p}^I(\theta)) = \overline{V}_p(\overline{s}^I(\theta), \overline{p}^I(\theta))\frac{d}{d\theta}\overline{p}^I(\theta)$. Since $\frac{d}{d\theta}\overline{p}^I(\theta) > 0$, this implies $\frac{d}{d\theta}\overline{V}^I(\theta) =^{sign} \sigma_d - \overline{s}^I(\theta)$, and since $\frac{d}{d\theta}\overline{s}^I(\theta) < 0$, it also implies that there is a unique $\overline{\theta}_d > \overline{\theta}_-^I$ such that $\frac{d}{d\theta}\overline{V}^I(\theta) < 0$ for $\theta < \overline{\theta}_d$ and $\frac{d}{d\theta}\overline{V}^I(\theta) > 0$ for $\theta > \overline{\theta}_d$. Clearly, $\overline{V}^I(\theta) < 0$ for $\theta \leq \overline{\theta}_d$. Given that $\frac{d}{d\theta}\overline{V}^I(\theta) > 0$ for $\theta > \overline{\theta}_d$ and $\overline{V}^I(\theta) > 0$ for a sufficiently large θ, there is a unique $\overline{\theta}_r^I > \overline{\theta}_d$ such that $\overline{V}^I(\overline{\theta}_r^I) = 0$ and thus $\overline{V}^I(\theta) < 0$ for $\theta \in (\overline{\theta}_d, \overline{\theta}_r^I)$. Overall, there exist two types $\overline{\theta}_-^I$ and $\overline{\theta}_r^I$ such that $\overline{V}^I(\theta) = 0$ if $\theta \leq \overline{\theta}_-^I$ or $\theta = \overline{\theta}_r^I$; $\overline{V}^I(\theta) < 0$ if $\theta \in (\overline{\theta}_-^I, \overline{\theta}_r^I)$; and $\overline{V}^I(\theta) > 0$ if $\theta > \overline{\theta}_r^I$.

To prove part (ii), note that since, under commitment, setting $\overline{s}^C(\theta) = \overline{\sigma}$ is always feasible for the principal, $\overline{V}^C(\theta) \geq 0$ for all θ. Note also that type $\theta = \overline{\theta}_-^I \equiv \frac{\overline{r}}{u(\lambda_h - \lambda_l)}$ is not prescreened under commitment if and only if $\overline{V}^C(\overline{\theta}_-^I) > 0$. Let $\theta(\epsilon) = \overline{\theta}_-^I + \epsilon$ for $\epsilon > 0$. By definition, $\overline{s}^I(\theta(\epsilon)) < \overline{\sigma}$, and $\theta(\epsilon) \to \epsilon$ as $\overline{c} \to \infty$. Suppose $\overline{s}^C(\theta(\epsilon)) < \overline{\sigma}$. Then, $V^C(\overline{s}^C(\theta(\epsilon)), p^C(\theta(\epsilon))) \approx -[1 - F^h(\overline{s}^C(\theta(\epsilon)))]c < 0$ for a small ϵ and large $\overline{c}$, which means the principal is better off setting $\overline{s}^C(\theta(\epsilon)) = \overline{\sigma}$. ‖

Proof of Lemma 7. If $s \in \{\underline{\sigma}, \overline{\sigma}\}$, then $p = e = 0$ is clearly optimal for the agent. Suppose $s \in (\underline{\sigma}, \overline{\sigma})$. The second-order necessary Hessian condition for an interior maximum is

$$-[u(F_l(s) - F_h(s))]^2 - 2\frac{u(F_l(s) - F_h(s))}{\sqrt{\theta t}} \geq 0,$$

which obviously fails. Hence, the agent's utility function is a saddle which is maximized as indicated in the statement of the lemma. ‖

Proof of Proposition 11. First we define the unique number $\underline{t} < \overline{\theta}$. To do this, consider the function

$$\rho(t) \equiv \int_t^{\overline{\theta}} \theta \, dG(\theta) - rtG(t).$$

Observe that: $\rho(\underline{\theta}) = E[\theta]$, $\rho(\overline{\theta}) = -r\overline{\theta}$ and $\rho'(t) = -(1+r)tg(t) - rG(t) < 0$. Hence, there exists a unique number $\underline{t} < \overline{\theta}$ satisfying $\rho(\underline{t}) = 0$.

Next, we characterize equilibrium under blind review in the presence of non-productive effort. Given that agents with $\theta < t$ exert only non-productive effort and

agents with $\theta \geq t$ exert only productive effort, the principal's reaction function is defined by

$$L(s^B) = \frac{\left(1 - e^B G(t) - \int_t^{\overline{\theta}} p^B(\theta)\, dG(\theta)\right) r}{\int_t^{\overline{\theta}} p^B(\theta)\, dG(\theta) - e^B G(t) r}.$$

Also from Lemma 7, $p^B(\theta) = \theta u[F_l(s^B) - F_h(s^B)]$ and $e^B = tu[F_l(s^B) - F_h(s^B)]$. Substituting these expressions into the principal's reaction function implicitly defines the standard under blind review

$$L(s^B) = \frac{\left(\frac{1}{u(F_l(s^B) - F_h(s^B))} - tG(t) - \int_t^{\overline{\theta}} \theta\, dG(\theta)\right) r}{\rho(t)}.$$

For $t < \underline{t}$, $\rho(t) > 0$ so the right side of this expression is clearly finite, which implies $s^B < \overline{\sigma}$. As $t \to \underline{t}$, $\rho(t) \to 0$, so the right side of the above expression approaches infinity, which implies $s^B = \overline{\sigma}$. $\|$

REFERENCES

ARROW, K. (1973), "The Theory of Discrimination", in O. Ashenfelter and A. Rees, (eds.) *Discrimination in Labor Markets* (Princeton, NJ: Princeton University Press) 3–33.

BACHAND, R. and SAWALLIS, P. (2003), "Accuracy in the Identification of Scholarly and Peer-Reviewed Journals and the Peer-Review Process Across Disciplines", *Serials Librarian*, **45**, 39–59.

BLANK, R. (1991), "The Effects of Double-Blind versus Single-Blind Refereeing: Experimental Evidence from the *American Economic Review*", *American Economic Review*, **81**, 1041–67.

BLAU, J. and BLAU, P. (1982), "The Cost of Inequality: Metropolitan Structure and Violent Crime", *American Sociological Review*, **47**, 114–129.

COATE, S. and LOURY, G. (1993), "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, **83**, 1220–40.

CREMER, J. (1995), "Arm's Length Relationships", *Quarterly Journal of Economics*, **110**, 275–95.

DEWATRIPONT, M., JEWITT, I. and TIROLE, J. (1999), "The Economics of Career Concerns, Part I: Comparing Information Structures", *Review of Economic Studies*, **66**, 183–98.

EHRLICH, I. (1975), "On the Relation between Education and Crime", in F. Thomas Juster, (eds.) *Education, Income, and Human Behavior* (Cambridge, MA: NBER 1975) 313–38.

FRYER, R. and LAURY, G. (2005), "Affirmative Action in Winner-Take-All Markets", *Journal of Economic Inequality*, **3**, 263–80.

GOLDIN, C. and ROUSE, C. (2000), "Orchestrating Impartiality: The Impact of 'Blind' Auditions on Female Musicians", *American Economic Review*, **90**, 715–41.

HOLMSTROM, B. (1999), "Managerial incentive problems: A dynamic perspective", *Review of Economic Studies*, **66**, 169–82.

HOLMSTROM, B. (1984), "On the Theory of Delegation", in M. Boyer and R. Kihlstrom, (eds.) *Bayesian Models in Economic Theory* (New York: North-Holland) 115–41.

HOLMSTROM, B. (1977), "On Incentives and Control in Organizations" (Ph.D. Thesis, Stanford University).

HORROBIN, D. (1982), "Peer Review: A Philosophically Faulty Concept Which is Proving Disastrous to Science", *Behavioral and Brain Sciences*, **5**, 217–8.

LEVIN, J. (2003), "Relational Incentive Contracts", *American Economic Review*, **93**, 835–57.

LINK, A. (1994), "US and Non-US Submission: An Analysis of Review Bias", *Journal of the American Medical Association*, **272**, 147–9.

LIPPKE, R. (2008), "Criminal Record, Character Evidence, and the Criminal Trial", *Legal Theory*, **14**, 167–91.

LOCHNER, L. and MORETTI, E. (2004), "The Effects of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports", *American Economic Review*, **94**, 155–89.

MACLEOD, B. (2003), "Optimal Contracting with Subjective Evaluation", *American Economic Review*, **93**, 216–40.

MCCORMICK, C. (1954), *Handbook of the Law of Evidence* (St. Paul, MN: West).

MILGROM, P. (1981), "Good News and Bad News: Representation Theorems and Applications", *Bell Journal of Economics*, 380–91.

MILGROM, P. and ROBERTS, J. (1988), "An Economic Approach to Influence Activities in Organizations", *American Journal of Sociology*, **94** (Supplement on Organizations and Institutions), S154–79.

MORRIS, S. (2001), "Political Correctness", *Journal of Political Economy*, **109**, 231–65.

NORMAN, P. (2003), "Statistical Discrimination and Efficiency", *Review of Economic Studies*, **70**, 615–27.

OTTAVIANI, M. and WICKELGREN, A. (2009), "Approval, Regulation and Learning, with Application to Timing of Merger Control" (Mimeo, Northwestern University, Kellogg School of Management).

PERSICO, N. (2002), "Racial Profiling, Fairness, and Effectiveness of Policing", *American Economic Review*, **92**, 1472–97.

PRAT, A. (2005), "The Wrong Kind of Transparency", *American Economic Review*, **95**, 862–77.

PETERS, D. and CECI, S. (1982), "Peer Review Practices of Psychological Journals: The Fate of Published Articles, Submitted Again", *Behavioral and Brain Sciences*, **5**, 187–252.

PRENDERGAST, C. (1999), "The Provision of Incentives in Firms", *Journal of Economic Literature*, **37**, 7–63.

REDMAYNE, M. (2002), "The Relevance of Bad Character", *Cambridge Law Journal*, **61**, 684–714.

RIORDAN, M. (1990), "What is Vertical Integration?", in M. Aoki, B. Gustafsoon, and O. Williamson, (eds.) *The Firm as a Nexus of Treaties* (London: European Sage).

SANCHIRICO, C. (2001), "Character Evidence and the Object of Trial", *Columbia Law Review*, **101**, 1227-1311.

SEIDMANN, D. (2005), "The effects of a right to silence", *Review of Economic Studies*, **72**, 593–614.

SNODGRASS, R. (2006), "Single- Versus Double-Blind Reviewing: An Analysis of the Literature", SIGMOD *Record*, **35**, 8–21.

STEFFENSMEIER, D., ANDERSEN, A. E., HARER, M. and STREIFEL, C. (1989), "Age and the Distribution of Crime", *American Journal of Sociology,* **94**, 803–31.

TABER, G. (2005), *Judgement of Paris: California vs France and the Historic 1976 Paris Tasting that Revolutionized Wine* (New York: Simon & Schuster).

TAYLOR, C. and YILDIRIM, H. (2007), "Subjective Performance and the Value of Blind Evaluation" (Mimeo, Duke University).

WAGNER, C. (1905), *Justice.* Translated by M.L. Hendee (New York: McClure, Phillips & Co).

ZUCKERMAN, H. and MERTON, R. (1971), "Patterns of Evaluation in Science: Institutionalization, Structure and Functions of the Referee System", *Minerva*, **9**, 66–100.