Hypothesis Testing under Moral Hazard*

Silvia Martínez-Gorricho
University of Concepción†

Carlos Oyarzun
University of Queensland‡

November 25, 2018

Abstract

A decision maker commits to ex-post inefficient standards of evidence to reduce *low* type agents' incentives to improve their signal distribution (*information undermining*). We characterize when optimal standards are *harsh* (*soft*), i.e., require more (less) favorable evidence than ex-post statistically optimal tests to choose the alternative preferred by the agent. Optimal standards are always soft for low priors that the type is low and harsh for high priors. We characterize when optimal standards are *confirmative* and *conservative* (c.f., Li (2001)), i.e., when they are adjusted to favor and disfavor, respectively, the decision that would be made if only prior information was available. Under some conditions the decision maker deviations from ex-post optimality are always confirmative. Equilibria under information undermining are always Pareto dominated by decision making in absence of information undermining for high enough prior probabilities that the type is low. Commitment to ex-post inefficient standards yield equilibria that Pareto dominate the equilibria in absence of commitment if and only if standards are soft. Instead, if *high* type agents make efforts to improve their signal distribution (*information generation*), most of the above results reverse.

Keywords: Moral Hazard, Commitment, Ex-post inefficiency, Confirmativism, Conservatism, Standard of Evidence.

JEL Classification: C72; D82; D83.

*This manuscript is still in progress; for the most recent version, please write to c.oyarzun@uq.edu.au. We thank comments from seminar participants at ACE 2015 (QUT, Brisbane), AETW 2018 (ANU, Canberra), University of New South Wales, The University of Queensland, Universidad de Chile, and the joint Seminar Series of Universidad del Bío-bío and Universidad de Concepción. In particular, we thank Juan Carlos Carbajal, Eduardo Engel, Gabriele Gratton, Keiichi Kawai, Hongyi Li, Andrew McLennan, and Rodrigo Wagner for genuinely fruitful discussions. We are specially grateful to Heiko Gerlach, Jonas Hedlund, Li Hao, Takashi Kunimoto, Claudio Mezzetti, and Miguel Sánchez-Villalba, who have carefully read previous versions of this manuscript and provided constructive comments. James Butcher, from the Department of Psychology of the University of Minnesota, provided us several references on test-defenisveness in psychology. Financial support from the Spanish Ministry of Economics and Competitiveness (ECO 2012-34928) is gratefully acknowledged.

†Universidad de Concepción, Departamento de Economía, Victoria 471, 4070374, Concepción, Chile. E-mail: silviamg@ua.es

‡Corresponding author. School of Economics, University of Queensland, 608 Colin Clark Building, Brisbane, Queensland, 4072, Australia. E-mail: c.oyarzun@uq.edu.au. Tel.: +61-7-3365-6579.

1 Introduction

In the wake of the replicability crisis, the scientific community has come to the realisation that methodologies and practices followed for decades, or even centuries, are not immune to malign incentives and research misconduct in particular.[1] In the search of solutions, a number of alternatives have been put over the table; for instance, recently, the Ministry of Science and Technology and courts in China have moved towards hard penalties on scientific misconduct —most extreme views have even considered the death penalty.[2] There is a wide range of fields, however, where research misconduct, in practice, is impossible to be detected and even more difficult to be proven —thus, rendering potential punishment virtually ineffective.[3] In this paper, we will explore an alternative manner to disincentivize research misconduct by altering the design of the decision process that misconduct seeks to influence.

The logic of the problem is not exclusive to research misconduct. In civil litigation, tampering with evidence and obstruction of evidence are pervasive. E.g., Sanchirico (2004) points out that *"[A]ccording to many judges and practitioners... documents that should be produced in response to a discovery request are regularly shredded, altered, or suppressed."* Another example relates to *test-defensiveness* in psychological screening tests. For instance, Butcher (2002) writes *"When taking psychological tests at pre-employment, pilots who have personality problems and other mental health symptoms can respond in a way to 'mask' those problems."* [4] The principle analyzed in this paper can also be applied to alleviate tampering with evidence, test defensiveness, or other forms of evidence manipulation.

The examples above involve a decision maker who chooses between two actions, and an agent who prefers one of them over the other, regardless of her type. Before choosing an action, the decision maker observes evidence that is partially informative about whether the agent's type is the one granting the action she prefers: an editor decides whether a piece of research is worth to be published, a judge decides whether a defendant is innocent, and a manager decides whether a pilot applying for a job in a commercial airline is mentally fit. A low type agent, however, can incur in *information underminning*, i.e., make efforts to alter the evidence observed by the decision maker to improve the chances of obtaining a favorable decision, e.g., research misconduct, evidence tampering, or test-defensiveness. High type agents can engage in activities to affect the decision process as well. For instance, Butcher (1994) points out that high average performance of pilots in psychological tests could be explained by fit pilots attempts to display "overly favorable response patterns." Since high type agents and the

[1] Di Tillio et al. (2017) provide a historical account of the development of experimental methodologies.

[2] See STATNEWS June 23, 2017, https://www.statnews.com/2017/06/23/china-death-penalty-research-fraud/.

[3] We stress that this is not to say that punishment is completelly ineffective in all fields, but to acknowledge that within several of them, uncovering misconduct may not be practical (see, e.g., Fanelli (2009)).

[4] Airlines screening processes to select their pilots were subject to intensive scrutinity in 2015, in the aftermath of a Germanwings plane crash in the Alps, believed to be deliberetly caused by the pilot.

decision maker incentives are aligned, we refer to these agents' activities as *information generation*. Our analysis considers both types of *moral hazard*, information undermining and information generation.

In absence of concerns about collateral effects on the incentives of the agent, such decisions are analysed as standard statistical decision problems (see, e.g., Neyman and Pearson (1933), Karlin and Rubin (1956), DeGroot (2005)). In this paper, we analyze a simple model of *commitment* to standards of evidence, swayed away from optimal statistic decisions, in order to disincentivize information undermining and incentivize information generation.[5] We show that the direction of the deviations from ex-post optimality are determined by two factors: (i) whether information undermining or information generation effects are dominant, and (ii) whether prior beliefs that the agent's type is low are relatively high or relatively low. We also analyze the welfare consequences of information undermining and generation, and the welfare consequences of commitment to standards.

It is instructive to focus first on pure information undermining. Sections 2-7 are devoted to this case and Section 8 extends the model to consider both information undermining and information generation. For the pure information undermining model, we fully establish when the optimal standards of evidence of the decision maker are *harsh* and *soft*, i.e., require, respectively, more and less favorable evidence than statistically optimal tests to choose the alternative preferred by the agent. Perhaps counterintuitively, the optimal commitment for a decision maker facing information undermining is not necessarily setting harsh standards — indeed such a conjecture is always wrong within a non-trivial set of prior beliefs. In particular, a feature of distributions ordered according the monotone likelihood ratio property (MLRP) discovered by Milgrom (1981) allows us to show that standards are always soft for low priors that the type is low and harsh for high priors.

We introduce the pure information undermining and non-commitment framework in Section 2. We consider a manager (he) who decides whether to hire or reject a candidate (she) who is applying for a job. The candidate is fit or unfit, with an exogenously given prior probability. The manager's losses are given by the cost of hiring (rejecting) an unfit (fit) candidate. Before making a decision, the manager observes a signal of fitness whose distribution can be partially manipulated by unfit candidates if they exert a costly non-observable effort.[6] Fit candidates have a fixed signal distribution that dominates the distribution of unfit candidates in the sense of MLRP (although with less intensity if unfit candidates exert effort), and make no test-defensive effort (we drop this assumption in Section 8).[7] The candidate's payoff is given by a fixed payment that she only gets if she is hired, net of the cost of the effort she made.[8]

[5] For related examples, see Prendergast (1993), Gilbert and Klemperer (2000), Hart and Moore (2007), Matouschek (2004), and Laux (2008).

[6] The signal realization, however, is non-manipulable (it is neither concealable nor forgeable).

[7] For instance, consider a commercial airline screening pilots using psychological tests. An unfit candidate can get special online training programs that allow her to increase her chances to pass undetected through personality tests. A number of websites offer such training products.

[8] Our model is different from that in Spence (1973) signaling game where the candidate undertakes visible actions to signal competence. Here, the manager does not observe whether the candidate undertook actions to manipulate the distribution of hard

In Section 3, we introduce commitment: we study a two-stage dynamic game. First, the manager commits to a standard that is observed by the candidate, and second, having observed the standard, the candidate (if she is unfit) chooses her level of effort.[9] Then, the manager observes a signal realization from the corresponding distribution and hires the candidate only if the evidence meets the standard. The ex-post optimal standard balances the manager's concerns with both types of error: rejecting fit candidates and accepting unfit ones. The manager, however, is also concerned with his ability to indirectly control the precision of the signal by committing to a standard: Lemma 2 shows that effort by unfit candidates is a strategic complement (substitute) of the standard in the submodular (supermodular) region of F, and therefore the manager optimally deters effort by being soft (harsh).

In Section 4 we provide a simple characterization of whether the manager is harsh or soft in equilibrium, in terms of the *prior unfitness odds*, i.e., the prior probability that the candidate is unfit divided by the prior probability that the candidate is fit. Proposition 1 shows that, for all prior unfitness odds below (above) a certain cut-off, the manager is soft (harsh) in equilibrium —a consequence of equilibrium standards being located in the submodular (supermodular) region of F if the probability that the candidate is unfit is low (high).

The analysis above leads to qualitative results relating the direction of the distortion of the standards to the decisions that would be made in absence of evidence. If the prior expected cost of hiring is smaller (greater) than the expected cost of rejecting, a manager unable to observe the signal would decide to hire (reject) the candidate. Accordingly, we say that such a manager *leans towards hiring* (*rejection*). If, in equilibrium, a manager is soft when he leans towards hiring or harsh when he leans towards rejection, then we say that he is *confirmative*. Analogously, if, in equilibrium, a manager is harsh when he leans towards hiring or soft when he leans towards rejection, then we say that he is *conservative* (c.f., Li (2001)). Corollary 1 reveals that, if the prior unfitness odds' cut-off for soft and harsh standards favors hiring (rejection), a leaning-towards-rejection (hiring) manager is always confirmative, whereas a leaning-towards-hiring (rejection) manager is conservative when his beliefs are moderate and confirmative otherwise. Conservative standards may arise only for some intermediate beliefs, and at most in one of the directions. In contrast, confirmative standards always arise for prior beliefs favoring each choice. Proposition 2 characterizes the family of distributions such that, regardless of the cost function, the manager is always *uniformly confirmative* (i.e., the distortion of the standard is in the same direction as his prior beliefs whenever his standard is not ex-post efficient). Uniform conservatism, on the other hand, is ruled out. Section 6 contains a number of examples; it also shows that standards may vary discontinuously with prior beliefs: the manager may impose very different standards on candidates whose priors' differences are negligible.

evidence; he just observes the actual realization of this signal. Thus, our model involves moral hazard.

[9] In terms of interpretation, the manager does not need to literally commit to a standard. It is enough that he has a reputation, within the potential candidates, of applying such a standard.

In Section 7, we turn our attention to the welfare consequences of information undermining and commitment to standards. We first compare welfare in absence of information undermining with welfare in equilibrium of the dynamic game. The manager is essentially always worse-off under information undermining. Harsh standards can also fully overturn the benefits from information undermining to unfit candidates: Proposition 3 shows that, under mild assumptions, for high enough prior unfitness odds, the equilibrium of the dynamic game is Pareto dominated by the outcome without information undermining. On the other hand, comparing the equilibrium of the dynamic and static games (that is, the set up with and without commitment), Corollary 3 shows that commitment not only makes the manager better-off: the dynamic game Pareto dominates the equilibrium of the static game if (and only if) the manager is soft in equilibrium.

Finally, in Section 8, we allow for both information undermining and information generation. Here we assume that fit candidates solve a similar problem to that solved by unfit candidates, however fit candidates have a default advantage to face the test and a different cost structure. Under pure information generation, the main results of the pure information undermining model are reversed: (i) for all prior unfitness odds below (above) a certain cut-off, the manager is harsh (soft) in equilibrium, and (ii) a manager can be uniformly conservative but cannot be uniformly confirmative. Under both information undermining and information generation, the direction of the distortion of the standard is determined by the magnitude of the effects that the manager can induce over fit and unfit candidates. In particular, if fit candidates' initial advantage is large, their marginal benefits from effort decrease relatively fast and hence the strategic effects of the standard on these candidates is modest. As such the resulting standards distortions are qualitatively similar to those of the pure information undermining model. Similarly, if unfit candidates' marginal cost of making effort is large, the strategic effects of the standard on these candidates is modest. As such the resulting standards distortions are qualitatively similar to those of the pure information generation model.

Related literature. The analysis in this paper is complementary to Li (2001) theory of conservatism. Both papers highlight how the quality of information may be determined by the decision making process. Li (2001) studies decision making by a committee in which the level of precision of the aggregated information is determined by the private effort exerted by each committee member and candidates are not strategic. While Li (2001) focuses on the benefit of conservatism to mitigate free-riding from committee members, our emphasis is on the role of commitments to standards to discourage unfit candidates' information undermining.

This paper also relates to the literature on Bayesian persuasion by analyzing the optimal commitment that the receiver of information (the manager) can adopt in order to simultaneously minimize the expected losses from mistaken hiring and rejection decisions, and discourage the sender (the candidate) from undermining the quality of information.[10] As in Gentzkow and Kamenica (2011), the sender (the candidate) chooses a signal distribution which generates a distribution of posteriors that is Bayes plausible and whose realization is observed by the receiver. But our framework differs from theirs in several aspects: (i) in our setting, there is asymmetric information: the sender is perfectly informed about the state of the world; (ii) our model exhibits moral hazard: the receiver does not observe the signal distribution chosen by the sender, only its realization; (iii) the sender is restricted to choose among a parametrized subset of signal distributions, and it is costly for her to choose a signal distribution that is probabilistically more persuasive; and (iv) the receiver can affect the incentives of the sender who attempts to persuade him by his commitment to a standard.

Our paper contributes to a quickly growing literature on research practices and economic incentives (see, e.g., Di Tillio et al. (2016), Di Tillio et al. (2017), Chassang et al. (2012), Tetenov (2016), and references therein). In particular, the analysis in Di Tillio et al. (2016) and Di Tillio et al. (2017) focuses on the persuasion bias from scientists, considering explicitly the probabilistic structure that derives from sample selection. In contrast, our analysis abstracts from the specifics of information manipulation and rather focuses on the role of commitment to desincentivize (incentivize) information undermining (generation) in a moral hazard problem. Thus the insights provided in our paper are complementary to their results.

Finally, a plausible interpretation of effort in our setup is that it corresponds to test-gaming by the agent. In that respect, our setup shares some features with recent work by Frankel and Kartik (2017) (see also Fischer and Verrecchia (2000) and Ederer et al. (2017)). In Frankel and Kartik (2017), agents aim to maximize the expected value of the market about their quality, which in turn is determined directly by their "natural" type and indirectly by their "gaming" ability, a second dimension of their type. They provide conditions so that higher stakes yield equilibria that are less informative about the natural type. While both their work and ours exhibit moral hazard, our focus is on the optimal manner to tailor commitment to an evidence standard in order to encourage (discourage) effort from the high (low) type.[11,12]

[10] Our work can be embodied into two main areas of research on persuasion: a first area examines optimal responses to hard information (Fishman and Hagerty (1990), Glazer and Rubinstein (2004), Alonso and Matouschek (2008)) and the second area focuses on Bayesian persuasion through control of the receiver's information environment (Rayo and Segal (2010), Kolotilin (2015), Hedlund (2017), Perez and Prady (2012)). Also related is the work by Kolotilin et al. (2013), who show how commitment of the receiver to a restricted set of actions can improve information transmission in a Crawford and Sobel (1982) cheap talk setup.

[11] Ederer et al. (2017), on the other hand, analyze how "opaque" contracts help a principal to incentivise an agent with private information about her cost function from exerting effort in two tasks. In particular, "opaque" contracts help the principal to induce more balanced efforts, which is valuable if the preferences of the principal exhibit high complementarity between tasks. In contrast, in our paper standards distortions from ex-post optimality either increase or decrease the probability of acceptance for both agents in order to discourage low type's effort and encourage high type's effort, aiming to minimize the decision maker's probability of making wrong decisions.

[12] Our paper also relates, although to a lesser extent, to the literature on optimal evidentiary (legal) standards to induce adequate behavior in the presence of a diversity of precautionary and enforcement schemes (see, e.g., Demougin and Fluet (2008), Ganuza et al. (2012), Gerlach (2013), Kaplow (2011), Sanchirico (2004), Sanchirico (2010)). The role of evidentiary standards on the informativeness of hard evidence studied in this paper seems to have received little attention in the literature. Indeed, the model of optimal evidentiary standards analyzed by Rubinfeld and Sappington (1987) implicitly incorporates this role, and Stephenson (2008) analyzes the informative role of standards in the research effort of agencies seeking for court approvals, but their work has no counterpart to the characterizations of harsh versus soft standards, or confirmativism versus conservatism, provided here.

2 The Model

A *manager* decides whether to *hire* or *reject* a *candidate* for a position, and he prefers to hire the candidate if she is *fit* and to reject her if she is *unfit*. The manager, however, cannot observe the candidate's fitness. Each candidate is fit with a prior probability strictly between 0 and 1, and unfit otherwise. Prior probabilities are derived from observable information of the candidate on factors such as age, gender, education, etc. The prior unfitness odds

$$\kappa := \frac{\text{prior probability of unfit}}{\text{prior probability of fit}} \in (0, \infty)$$

will play an important role in the analysis. In the sequel we refer to κ simply as the *prior*.

For simplicity, the manager's weights on losses due to hiring unfit candidates and rejecting fit candidates are normalized to 1.[13] If the manager were to make his decision based on prior information only, then: (i) if $\kappa = 1$, the manager's expected loss from hiring and rejection would be exactly the same, and hence, he would be indifferent, and (ii) if $\kappa < (>)1$, the manager's expected loss from hiring would be smaller (greater) than the expected loss from rejection and thus, in absence of further evidence, the manager would choose hiring (rejection).

Definition 1 *The manager* leans towards hiring (rejection) *if* $\kappa < (>)1$.

2.1 Evidence

The manager runs a test to obtain further evidence on the candidates' fitness. The result of the test is the realization of a signal $z \in [0, 1]$. The cumulative distribution and density functions of fit candidates' signal are denoted by $F(\cdot, \overline{\theta})$ and $f(\cdot, \overline{\theta})$, respectively, where $\overline{\theta}$ is a parameter. The corresponding functions for unfit candidates who do not exert any effort to (probabilistically) improve the outcome of the test are $F(\cdot, \underline{\theta})$ and $f(\cdot, \underline{\theta})$, respectively, with $\underline{\theta} < \overline{\theta}$. The "parameter" $\theta \in \left[\underline{\theta}, \overline{\theta}\right] =: \Theta$ denotes a costly *effort* that may be exerted by unfit candidates in order to probabilistically improve the result of the test; the interior of Θ is denoted by Θ°. Throughout Sections 2-7 we assume that fit candidates make no effort; in Section 8 we drop this assumption. Thus, the domain of F is $D := [0, 1] \times \Theta$ and its interior is denoted by D°. We assume that $F(\cdot, \theta)$ is strictly increasing for all $\theta \in \Theta$ and F is well behaved in the following sense:

[13] The effects of asymmetric weights on each type of losses are isomorphous to those of asymmetric priors: an increase in the relative weight of accepting unfit candidates over the weight of rejecting fit candidates has the same effect as an increase in κ. Indeed, under symmetric weights, κ coincides with what Di Tillio et al. (2016) call the "acceptance hurdle." To be concrete, let $\gamma \in (0, 1)$ be the prior probability that the candidate is unfit and $\lambda \in (0, \infty)$ be the ratio of the loss associated to hiring an unfit candidate divided by the loss associated to rejecting a fit candidate. Then, the acceptance hurdle is given by $\kappa^{ah} := \frac{\gamma}{1-\gamma}\lambda$ and a manager who makes decisions according to his prior beliefs only accepts (rejects) the candidate if $\kappa^{ah} < (>)1$. Thus, working with $\lambda = 1$ is without loss of generality: any change in λ can be captured by changing κ. Throughout the paper $\lambda = 1$, thus changes in κ only reflect changes in γ, but the reader can also interpret changes in κ as changes in the ratio of the losses associated to wrong hiring or rejection.

Assumption F.1 *The distribution F is thrice continuously differentiable on $(0,1)\times\Theta$, $f(s,\theta)>0$ for all $s\in(0,1)$, and $F(0,\theta)=0$ and $F(1,\theta)=1$, for all $\theta\in\Theta$.*

Our next assumption on F is that θ indexes a family of signal distributions that satisfies the strict Monotone Likelihood Ratio Property (MLRP):

Assumption F.2 *For all $\theta,\theta'\in\Theta$, $\theta'>\theta$ implies*

$$\frac{1}{f(z,\theta')}\frac{\partial f(z,\theta')}{\partial z}>\frac{1}{f(z,\theta)}\frac{\partial f(z,\theta)}{\partial z}, \tag{1}$$

for all $z\in(0,1)$.

Hence, if $\theta'>\theta$, then $\frac{f(\cdot,\theta')}{f(\cdot,\theta)}$ is strictly increasing. The strict MLRP implies that $F(\cdot,\theta')$ strictly first-order stochastically dominates (FOSD) $F(\cdot,\theta)$ for $\theta'>\theta$; that is, $F(z,\theta)>F(z,\theta')$ for all $z\in(0,1)$, and, hence, $F(z,\cdot)$ is decreasing for all $z\in(0,1)$.

We assume that the marginal return to effort is weakly decreasing.

Assumption F.3 $\frac{\partial^2 F(z,\theta)}{\partial\theta^2}\geq 0$ *for all $(z,\theta)\in(0,1)\times\Theta$.*

Assumptions F.2 and F.3 imply that the derivative of $F(z,\cdot)$ with respect to θ is strictly negative: for all $z\in(0,1)$ and $\theta<\overline{\theta}$,

$$h(z,\theta):=\frac{\partial F(z,\theta)}{\partial\theta}<0.\text{[14]} \tag{2}$$

The sign of $\frac{\partial f(z,\theta)}{\partial\theta}$ plays a critical role in our analysis and has a simple interpretation: on those regions of the domain where it is negative (positive), i.e., where F is submodular (supermodular), the marginal effect of effort on the density function of the signal is negative (positive). Whether F is submodular or supermmodular at any point of its domain depending on whether z is small or large. In particular, let

$$m(z,\theta):=\frac{1}{f(z,\theta)}\frac{\partial f(z,\theta)}{\partial\theta}$$

for all $(z,\theta)\in D^\circ$. Proposition 5 in Milgrom (1981) implies that $m(\cdot,\theta)$ is increasing, and hence, $\frac{\partial m(\cdot,\theta)}{\partial z}\geq 0$ for all $\theta\in\Theta^\circ$.[15] At any (z,θ) such that $m(z,\theta)=0$, $\frac{\partial m(z,\theta)}{\partial z}\geq 0$ is equivalent to $\frac{\partial^2 f(z,\theta)}{\partial\theta\partial z}\geq 0$. In order to avoid non-substantial technicalities, we assume that the later inequality is strict at all such points:

[14] If $h(z,\theta)=0$ for some $z\in(0,1)$ and $\theta<\overline{\theta}$, then Assumption F.3 implies that $h(z,\theta')\geq 0$ for all $\theta'>\theta$, and hence, $F(z,\theta')\geq F(z,\theta)$, contradicting strict first-order stochastic dominance.

[15] By Assumptions F.1 and F.2, we have that:

$$\frac{\partial m(z,\theta)}{\partial z}=\frac{\partial^2 f(z,\theta)}{\partial z\partial\theta}\frac{1}{f(z,\theta)}-\frac{\partial f(z,\theta)}{\partial z}\frac{1}{f(z,\theta)}m(z,\theta)=\frac{\partial\left(\frac{\partial f(z,\theta)}{\partial z}\frac{1}{f(z,\theta)}\right)}{\partial\theta}\geq 0$$

for all $(z,\theta)\in D^\circ$.

Assumption F.4 *For all* $(z,\theta) \in D^\circ$, $m(z,\theta) = 0$ *implies* $\frac{\partial^2 f(z,\theta)}{\partial\theta\partial z} > 0$.

If we were to relax this assumption and that the inequality at (2) is strict, the exposition would be slightly more involved, but all qualitative aspects of our results would remain essentially unchanged. Assumption F.4 affords that there exists a function $\widetilde{s} : \Theta^\circ \to (0,1)$ whose graph separates the parts of D° where F is submodular and supermodular.

Remark 1 *Assume F.1-F.4. For all* $\theta < \overline{\theta}$ *there exists* $\widetilde{s}(\theta) \in (0,1)$ *such that*

$$\frac{\partial f(z,\theta)}{\partial\theta} \begin{cases} < 0 & \text{if } z < \widetilde{s}(\theta) \\ = 0 & \text{if } z = \widetilde{s}(\theta) \\ > 0 & \text{if } z > \widetilde{s}(\theta). \end{cases} \tag{3}$$

All proofs are provided in the appendix, with a few exceptions provided in the text or simply omitted.

2.2 Ex-post Optimal Standard

The manager is risk neutral and minimizes expected losses. The strict MLRP assumption guarantees that the manager's best response to any effort $\theta \in \Theta$ exerted by unfit candidates is an acceptance standard, that is, a "threshold" strategy (s) such that the manager hires the candidate if $z \geq s$ and rejects her if $z < s$ for some $s \in [0,1]$. For an arbitrary standard s and unfit candidates' effort θ, the probabilities of wrongful rejection and wrongful hiring are, respectively, $F(s,\overline{\theta})$ times the prior probability that the candidate is fit and $(1 - F(s,\theta))$ times the prior probability that the candidate is unfit. Thus, the trade-off faced by the manager is that a higher standard s decreases the probability of wrongful hiring, but increases the probability of wrongful rejection. The expected loss to the manager, as a function of an arbitrary hiring standard s and effort θ is an affine transformation of[16]

$$V(s,\theta) = F(s,\overline{\theta}) - \kappa F(s,\theta). \tag{4}$$

for all $(s,\theta) \in D$. It will often be convenient to indicate explicitly the dependence of V on the parameter κ, and then we will write $V(s,\theta;\kappa)$ instead of $V(s,\theta)$.

Consider the function $s^* : \Theta \times (0,\infty) \to [0,1]$ mapping pairs of efforts and priors (θ,κ) to the corresponding ex-post optimal standard for the manager. Define the *likelihood ratio function* $g : D \to \mathbb{R} \cup \{\infty\}$ with

$$g(s,\theta) := \frac{f(s,\overline{\theta})}{f(s,\theta)}$$

[16]The expected loss to the manager is $(V + \kappa)$ times the prior probability that the candidate is fit.

for all $(s,\theta) \in (0,1) \times \Theta$, and $g(0,\theta) := \lim_{s\to 0} g(s,\theta)$ and $g(1,\theta) := \lim_{s\to 1} g(s,\theta)$ for all $\theta \in \Theta$. By the strict MLRP, $g(\cdot,\theta)$ is strictly increasing for all $\theta < \overline{\theta}$. By Assumption F.1, sign $\left\{\frac{\partial V(s,\theta)}{\partial s}\right\}$ = sign $\{g(s,\theta) - \kappa\}$ for all $s \in (0,1)$, thus the optimal standard is

$$s^*(\theta,\kappa) = \begin{cases} 0 & \text{if } 0 < \kappa \leq g(0,\theta) \\ s^*_{\theta,\kappa} & \text{if } g(0,\theta) < \kappa < g(1,\theta) \\ 1 & \text{if } g(1,\theta) \leq \kappa, \end{cases} \tag{5}$$

where $s^*_{\theta,\kappa}$ is defined by

$$g(s^*_{\theta,\kappa},\theta) \equiv \kappa \tag{6}$$

for all $\kappa \in (g(0,\theta), g(1,\theta))$ and $\theta < \overline{\theta}$. Since $g(\cdot,\theta)$ is strictly increasing, equations (5) and (6) reveal that $s^*(\theta,\cdot)$ is weakly increasing for all $\theta < \overline{\theta}$.

A side note of independent interest is that $s^*(\cdot,\kappa)$ can be decreasing or increasing at $(\theta,\kappa) \in \Theta \times (0,\infty)$ depending on whether $(s^*(\theta,\kappa),\theta)$ is located in the submodular or supermodular region of F.[17]

Remark 2 *Assume F.1-F.2. The standard is a strategic substitute (complement) of the effort exerted by unfit candidates at* $(\theta,\kappa) \in \Theta \times (0,\infty)$ *(i.e.,* $\frac{\partial s^*(\theta,\kappa)}{\partial \theta} < (>)0$*) if and only if* $s^*(\theta,\kappa) \in (0,1)$ *and* $(s^*(\theta,\kappa),\theta)$ *is located in the submodular (supermodular) region of* F*.*

Intuitively, a higher effort θ exerted by unfit candidates lead the manager to decrease or increase the standard, depending on whether the marginal effect of effort on the density f is negative or positive, respectively. In the former (latter) scenario, a higher effort makes unfit candidates accumulate probability mass slower (faster) around the standard, decreasing (increasing) the manager's marginal benefit of increasing the standard.

2.3 Information undermining

Unfit candidates can make their signal distribution more similar to that of fit candidates by increasing the magnitude of θ. Their action set is Θ and the cost of changing their default signal distribution is given by the cost function $C : \Theta \to [0,\infty)$.

Assumption C. 1 *The cost function* C *is twice continuously differentiable, and satisfies* $C(\underline{\theta}) = C'(\underline{\theta}) = 0$ *and* $C''(\theta) > 0$ *for all* $\theta \in \Theta^\circ$*. Furthermore,* $C'(\overline{\theta}) > -h(s,\overline{\theta})$ *for all* $s \in (0,1)$*.*

[17] We say that $s^*(\cdot,\kappa)$ is strictly decreasing (increasing) at (θ,κ) if there is an interval $I_{\theta,\kappa} \subset \Theta$ containing θ and such that $\theta'' > \theta' \in I_{\theta,\kappa}$ implies $s^*(\theta'',\kappa) < (>) s^*(\theta',\kappa)$.

We assume that unfit candidates are risk neutral and their opportunity cost of not being hired is normalized to 1. Thus, for any standard $s \in [0,1]$ set by the manager, unfit candidates' expected loss is given by

$$U(s,\theta) = F(s,\theta) + C(\theta) \tag{7}$$

for all $\theta \in \Theta$.

Given a standard s, the optimal level of effort, denoted by $\theta^*(s)$, is a minimizer of $U(s,\cdot)$ within Θ. Since $F(0,\theta) = 0$ and $F(1,\theta) = 1$ for all $\theta \in \Theta$, $\theta^*(0) = \theta^*(1) = \underline{\theta}$. That is, effort is redundant if all candidates are hired or all candidates are rejected. Given a standard $s \in (0,1)$, an interior solution $\theta^*(s)$ satisfies

$$C'(\theta^*(s)) = -h(s,\theta^*(s)), \tag{8}$$

where we recall that $-h(s,\theta) := -\frac{\partial F(s,\theta)}{\partial \theta}$ is the rate of change of the unfit candidates' probability of being hired with respect to effort, for a given standard s. Furthermore, by Assumption C.1, the minimizer $\theta^*(s)$ always satisfies (8) and $\theta^*(s) \in \Theta^\circ$ for all $s \in (0,1)$.

Whether effort is a strategic complement or substitute of the standard at (s,θ) is determined by whether (s,θ) is located in the submodular or supermodular region of F: the Implicit Function Theorem and (8) imply

$$\frac{d\theta^*(s)}{ds} = -\frac{\partial f(s,\theta^*(s))}{\partial \theta}\left(C''(\theta^*(s)) + \frac{\partial h(s,\theta^*(s))}{\partial \theta}\right)^{-1} \tag{9}$$

for all $s \in (0,1)$. Since $C''(\theta^*(s)) > 0$ and $h(s,\cdot)$ is assumed to be weakly increasing for all $s \in (0,1)$, the sign of the effect of the standard on the optimal level of effort exerted by unfit candidates is determined by the sign of $\frac{\partial f(s,\theta^*(s))}{\partial \theta}$, i.e., by whether F is sub or supermodular at $(s,\theta^*(s))$. Sumarizing, we have the following:

Remark 3 *Assume F.1, F.3, and C.1. Unfit candidates' effort is a strategic complement (substitute) of the manager's standard at $s \in [0,1]$ (i.e., $\frac{d\theta^*(s)}{ds} > (<)0$) if and only if $(s,\theta^*(s))$ is located in a submodular (supermodular) region of F.*

We now provide the first of the three lemmas leading to Proposition 1, one of the key results of the paper.

Lemma 1 *Assume F.1-F.4 and C.1. There exists $\hat{s} \in (0,1)$ such that*

$$\frac{\partial f(s,\theta^*(s))}{\partial \theta} \begin{cases} < 0 & \text{if } 0 \leq s < \hat{s} \\ = 0 & \text{if } s = \hat{s} \\ > 0 & \text{if } \hat{s} < s \leq 1. \end{cases} \tag{10}$$

We call this property *single modularity-switch*: there exists a cut-off $\hat{s}$ for the standard, that we call the *modularity-switch point*, such that, (i) for standards smaller than $\hat{s}$, F is submodular at $(s, \theta^*(s))$, and (ii) for standards greater than $\hat{s}$, F is supermodular at $(s, \theta^*(s))$. The proof, provided in the appendix, relies heavily on the fact that we can separate the parts of the domain of F where it is submodular from the parts where it is supermodular, using the function $\widetilde{s}$ defined in Remark 1. Lemma 1 and (9) reveal that θ^* is strictly increasing over $[0, \hat{s}]$ and strictly decreasing over $[\hat{s}, 1]$. The left panel of Figure 1 shows θ^* for Example 1 (to be discussed below).

We end this section pointing out that, provided that F.1-F.3 and C.1 hold, $g(\cdot, \theta^*(\cdot))$ is strictly increasing:

$$\frac{dg(s,\theta^*(s))}{ds} = \frac{\partial g(s,\theta^*(s))}{\partial s} - \frac{g(s,\theta^*(s))}{f(s,\theta^*(s))}\frac{\partial f(s,\theta^*(s))}{\partial \theta}\frac{d\theta^*(s)}{ds} > 0, \tag{11}$$

for all $s \in (0, 1)$. The first term on the right-hand side is strictly positive, due to the strict MLRP, and the second one is nonnegative by condition (9).

2.4 The Static Game

In the sequel, the manager and candidates' preferences, including κ, F, and C, are common knowledge. We start our analysis considering a *static game*, denoted by Γ_1, with the following timing: (i) first, Nature choses the candidate's type (fit or unfit) and reveals it to the candidate; (ii) the manager chooses a hiring standard $s \in [0, 1]$, knowing neither the candidate's type nor the effort she exerted if she were unfit; and the candidate (if she is unfit) chooses her level of effort $\theta \in \Theta$ without knowing the manager's hiring standard; (iii) Nature choses a signal realization $z \in [0, 1]$ according to $F(\cdot, \theta)$; and (iv) the manager hires the candidate if and only if $z \geq s$, and rejects her otherwise.

Since fit candidates are non-strategic, Γ_1 can be collapsed to a simultaneous move game between the manager and unfit candidates choosing their respective actions in stage (ii), and such that their expected payoffs (where expectations are taken with respect to Nature moves) are given by (4) and (7), respectively. A *Pure Strategy Nash Equilibrium* of Γ_1 is a pair $(s_{NE}, \theta_{NE}) \in D$, with $s_{NE} = s^*(\theta_{NE}, \kappa)$ and $\theta_{NE} = \theta^*(s_{NE})$, and hence, satisfying equations (5) and (8). If $g(0, \underline{\theta}) < \kappa < g(1, \underline{\theta})$, then $s_{NE} \in (0, 1)$ and hence,

$$g(s_{NE}, \theta^*(s_{NE})) \quad = \quad \kappa. \tag{12}$$

On the other hand, if $\kappa \leq g(0, \underline{\theta})$, then $(s_{NE}, \theta_{NE}) = (0, \underline{\theta})$, and if $\kappa \geq g(1, \underline{\theta})$, then $(s_{NE}, \theta_{NE}) = (1, \underline{\theta})$. The appendix contains details on the uniqueness of the Nash equilibrium of Γ_1 and related comparative statics (see Proposition 9). The standard set by the manager in the Nash equilibrium of Γ_1 is ex-post efficient. Hence it serves as a reference point for the ex-post inefficient standards that arise in the dynamic game to be anlyzed in the subsequent sections.

3 Commitment to Standards: the Dynamic Game

In this section we analyze how the manager's commitment to a standard affects the equilibrium outcome of the game. By committing to a hiring standard, the manager can affect unfit candidates' incentives to undermine information.

We formalize the role of commitments to a standard by analyzing a dynamic game, denoted by Γ_2. The timing of Γ_2 is identical to that of Γ_1, except that stage (ii) is broken in two sub-stages: in sub-stage (ii.1) the manager commits to a standard $s \in S$; and, in sub-stage (ii.2), having observed the standard chosen by the manager, the unfit candidate chooses effort θ. Unfit candidates' set of strategies, denoted by $\Theta^{[0,1]}$, is the set of functions $\theta : [0,1] \to [\underline{\theta}, \overline{\theta}]$. Given a function $\theta \in \Theta^{[0,1]}$, the manager's optimal choice of standard minimizes

$$V(s, \theta(s)) \equiv F(s, \overline{\theta}) - \kappa F(s, \theta(s)). \tag{13}$$

with respect to $s \in [0,1]$.

Since the manager commits to a standard that is observed by the candidate and fit candidates make no effort, Γ_2 can be collapsed to a dynamic game of two periods between the manager and unfit candidates: in period 1 the manager announces the standard to the candidate, and in period 2, the candidate (if she is unfit) chooses her effort. The expected payoffs to the manager and unfit candidates are given by (13) and (7), respectively. Our analysis focuses on *Subgame Perfect Equilibria in Pure Strategies*, i.e., duplets $(s^*_{SP}, \theta^*) \in [0,1] \times \Theta^{[0,1]}$ such that $V(s^*_{SP}, \theta^*(s^*_{SP})) \leq V(s, \theta^*(s))$ for all $s \in [0,1]$, and $U(s, \theta^*(s)) \leq U(s, \theta)$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$ and $s \in [0,1]$. Under Assumptions F.1, F.3, and C.1, such an equilibrium always exists; for details, see Remark 9 in the appendix.

Since standards play both a decision role and an incentive role in this game, equilibrium standards, in general, are not ex-post efficient. The total derivative of the manager's expected loss function with respect to the standard is

$$\begin{aligned} \frac{dV(s, \theta^*(s))}{ds} &= \frac{\partial V(s, \theta^*(s))}{\partial s} + \frac{\partial V(s, \theta^*(s))}{\partial \theta} \frac{d\theta^*(s)}{ds} && (14) \\ &= f(s, \overline{\theta}) - \kappa f(s, \theta^*(s)) - \kappa h(s, \theta^*(s)) \frac{d\theta^*(s)}{ds}, && (15) \end{aligned}$$

for all $s \in (0,1)$. As in the static game, in setting the standard, the manager takes into account its direct effect on the quality of the decision: a higher standard yields a higher chance of wrongful rejection and a lower chance of wrongful hiring —respectively, the first and second terms on the right hand side of equation (15). In the dynamic game, however, the manager also takes into account the strategic effect that arises through the candidate's incentives to exert effort —the third term on the right hand side of equation (15). This consideration leads to deviations from ex-post optimality. The following definition classifies deviations of the optimal standard under commitment from ex-post optimal standards.

Definition 2 *Let (s^*_{SP}, θ^*) be an equilibrium of Γ_2. A manager is* soft (harsh) *at (s^*_{SP}, θ^*) if $s^*_{SP} < (>) s^*(\theta^*(s^*_{SP}), \kappa)$.*

A harsh (soft) manager's standard is strictly higher (lower) than the ex-post optimal standard, given the effort exerted by unfit candidates: in the margin, the manager rejects (accepts) candidates whose expected loss, upon observing the signal, is strictly less (greater) than hiring (rejecting) her.

The following lemma characterizes deviations from the ex-post optimal decision rule according to whether the signal distribution is submodular or supermodular in the region of the domain where the equilibrium is located.

Lemma 2 *Assume F.1-F.3 and C.1. Let (s^*_{SP}, θ^*) be an equilibrium of Γ_2 such that $s^*_{SP} \in (0, 1)$. The manager is soft (harsh) at (s^*_{SP}, θ^*) if and only if F is strictly submodular (strictly supermodular) at $(s^*_{SP}, \theta^*(s^*_{SP}))$.*

The proof of Lemma 2 is provided in the appendix, however, the essence of the argument can be grasped from the analysis of the *pseudo likelihood ratio* function, $v : D_v \to \mathbb{R} \cup \{\infty\}$, given by $v(s) := \frac{dF(s,\overline{\theta})}{ds} \left(\frac{dF(s,\theta^*(s))}{ds} \right)^{-1}$ for all $s \in D_v := \left\{ s \in (0,1) : \frac{dF(s,\theta^*(s))}{ds} \neq 0 \right\} \cup \{0, 1\}$. Notice that

$$v(s) = f(s, \overline{\theta}) \left(f(s, \theta^*(s)) + h(s, \theta^*(s)) \frac{d\theta^*(s)}{ds} \right)^{-1} \tag{16}$$

for all $s \in D_v$. From (15), if $s^*_{SP} \in (0, 1)$ is the equilibrium standard of Γ_2, then $v(s^*_{SP}) = \kappa$, and from (16) and (9), $v(s^*_{SP}) > (<) g(s^*_{SP}, \theta^*(s^*_{SP}))$ if $(s^*_{SP}, \theta^*(s^*_{SP}))$ located in the submodular (supermodular) region of the domain of F.[18] Hence, since $g(\cdot, \theta^*(s))$ is increasing, the ex-post optimal standard is higher (lower) than the one in equilibrium. The right panel of Figure 1 shows $v(\cdot)$ and $g(\cdot, \theta^*(\cdot))$ for the game described in Example 1 (to be discussed below, in Section 6). For any value of κ corresponding to a value of s in the horizontal axis such that $\kappa = v(s) > g(s, \theta^*(s))$, $g(s', \theta^*(s)) = \kappa$ for some $s' \in (s, 1]$; hence, the manager is soft for the prior κ. Remark 3 and Lemma 2 imply the following:

[18] Notice that if $\frac{dF(s,\theta^*(s))}{ds} \leq 0$ for some $s \in (0, 1)$, then from (15) we have $\frac{dV(s,\theta^*(s))}{ds} > 0$, and hence s cannot be an equilibrium standard. All examples of the paper have $v(s) > 0$ for all $s \in (0, 1)$. There are, however, games such that $v(s) < 0$ within some intervals. For instance, the game Γ_2 defined by the family of distributions $F(s, \theta) = \theta s^{10} + (1 - \theta)s$ and effort cost $C(\theta) = \frac{1}{2}\theta^2$ for all $s \in [0, 1]$ and $\theta \in \Theta = [0, 1]$, has $v(s) < 0$ for all $s \in (0.52, 0.65)$.

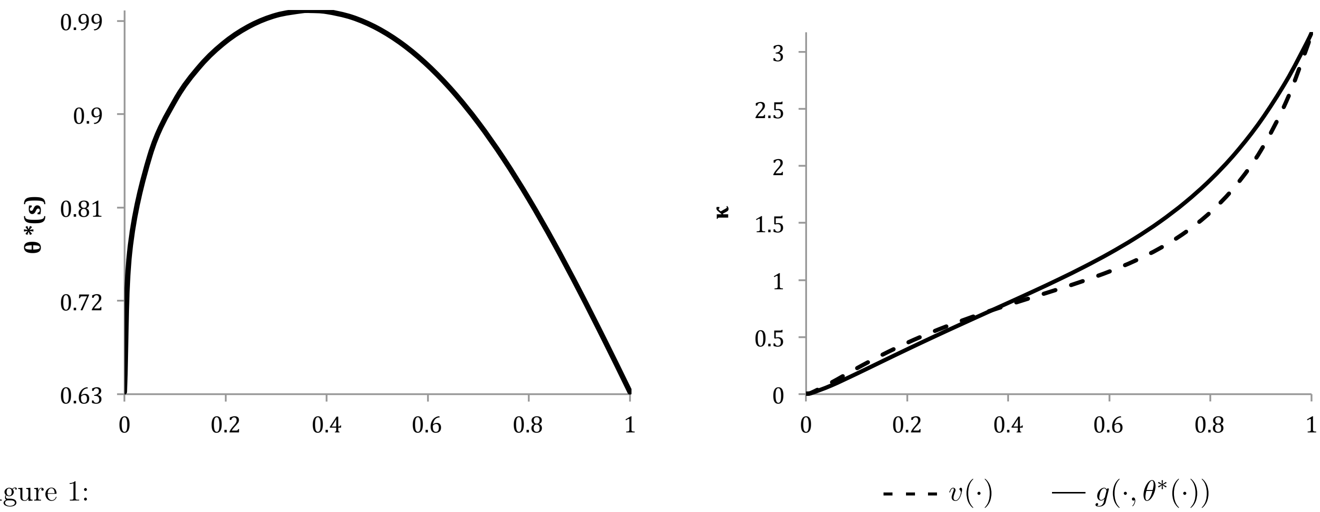

Figure 1:

Example 1. Left panel: best response of unfit candidate. Right panel: equilibrium standards in the games Γ_1 (solid line) and Γ_2 (dashed line) for each κ.

Remark 4 *Assume F.1-F.3 and C.1. Let* (s^*_{SP}, θ^*) *be an equilibrium of* Γ_2 *such that* $s^*_{SP} \in (0, 1)$*. The manager is soft (harsh) at* (s^*_{SP}, θ^*) *if and only if unfit candidates' effort is a strategic complement (substitute) of the standard at* $(s^*_{SP}, \theta^*(s^*_{SP}))$*.*

Intuitively, if higher standards lead to a higher (lower) effort exerted by unfit candidates, then, in order to discourage effort, the manager decreases (increases) his standard with respect to the ex-post optimal standard. These strategic effects generate ex-post inefficiencies in the actual statistical problem of hiring or rejecting the candidate. The optimal ex-post inefficiency balances this trade-off. Importantly for the analysis in the sequel, Lemmas 1-2 imply that the manager is soft (harsh) in equilibria (s^*_{SP}, θ^*) such that $s^*_{SP} < (>) \hat{s}$.

4 Standards of Evidence and Prior Beliefs

This section studies how equilibrium deviations from the ex-post optimal standard relate to the manager's prior beliefs about the candidate's fitness. In Subsection 4.1, we provide a lemma that reveals that equilibrium standards are increasing in the prior κ. In Subsection 4.2, we show that equilibrium standards are soft if the prior probability that the candidate is unfit is low and harsh if it is high.

4.1 Increasing Standards

We define the *classical statistical problem* as the problem solved by the manger in absence of information undermining. The expected losses to the manager in this problem are

$$V(s, \underline{\theta}) = F(s, \overline{\theta}) - \kappa F(s, \underline{\theta}),$$

for all $s \in [0,1]$. Let $\underline{\kappa} := g(0, \underline{\theta})$ (corresp. $\overline{\kappa} := g(1, \underline{\theta})$) be the largest (corresp. smallest) prior κ such that every candidate is accepted (corresp. rejected) in the classical statistical problem. The strict MLRP implies $\underline{\kappa} < 1 < \overline{\kappa}$.

As in the static game, we typically have a unique equilibrium in the dynamic game. Some dynamic games, however, have multiple equilibria for some prior κ. Let $\kappa \rightrightarrows \mathcal{S}^*(\kappa)$ be the correspondence mapping $\kappa \in (0, \infty)$ to the set of standards in pure strategy subgame perfect equilibrium. We say that $\mathcal{S}^*$ is *weakly* (*strictly*) *increasing* within a given real interval $I \subseteq (0, \infty)$ if $\kappa' > \kappa$, $s \in \mathcal{S}^*(\kappa)$ and $s' \in \mathcal{S}^*(\kappa')$ imply $s' \geq s$ $(s' > s)$, for all $\kappa, \kappa' \in I$.

Lemma 3 *Assume F.1-F.3 and C.1. The correspondence $\mathcal{S}^*$ is weakly increasing over $(0, \infty)$ and strictly increasing over $(\underline{\kappa}, \overline{\kappa})$. Further*

$$\mathcal{S}^*(\kappa) \begin{cases} = \{0\} & \text{if } 0 < \kappa \leq \underline{\kappa} \\ \subset (0,1) & \text{if } \underline{\kappa} < \kappa < \overline{\kappa} \\ = \{1\} & \text{if } \overline{\kappa} \leq \kappa. \end{cases}$$

The proof of Lemma 3 reveals that $\mathcal{S}^*$ being weakly increasing only hinges on $F(\cdot, \overline{\theta})$ being strictly increasing and not in any of the other assumptions. The increasingness of $\mathcal{S}^*$ will play an important role in the sequel.

Another implication of Lemma 3 is that the set of priors that make tests worthy are the same with our without information undermining. In the classical statistical problem, if $\kappa \in (\underline{\kappa}, \overline{\kappa})$ the manager's optimal standard is strictly between 0 and 1, and then we say that the problem is *test-worthy.* From Lemma 3, the set of priors for which running a test is worthy is the same, with or without information undermining.

4.2 The Soft-Harsh Pattern of Standards of Evidence

Our next result shows that in the range of priors for which we observe non-trivial tests in the classical problem, i.e., when the statistically optimal standard is interior, there is a cut-off $\widetilde{\kappa}$ such that, within that range, the manager is soft for priors below $\widetilde{\kappa}$ and harsh for priors above $\widetilde{\kappa}$. Outside that range the manager is ex-post efficient.

Proposition 1 *Assume F.1-F.4 and C.1. In any game* Γ_2*, there exists* $\widetilde{\kappa} \in (\underline{\kappa}, \overline{\kappa})$ *such that*

$$\text{the manager is} \begin{cases} \text{ex-post efficient} & \text{if } \kappa \in (0, \underline{\kappa}] \\ \text{soft} & \text{if } \kappa \in (\underline{\kappa}, \widetilde{\kappa}) \\ \text{harsh} & \text{if } \kappa \in (\widetilde{\kappa}, \overline{\kappa}) \\ \text{ex-post efficient} & \text{if } \kappa \in [\overline{\kappa}, \infty). \end{cases} \tag{17}$$

The proof of Proposition 1, provided in the appendix, builds upon Lemmas 1-3: by Lemma 3, relatively low (high) priors are associated with relatively low (high) standards in equilibrium. By Lemma 1, if the equilibrium standards are relatively low (high), then the equilibrium is located in the submodular (supermodular) region of F, and hence, by Lemma 2, the manager is soft (harsh).

If priors are extreme ($\kappa \leq \underline{\kappa}$ or $\kappa \geq \overline{\kappa}$), the equilibrium standards are ex-post efficient. If $\kappa = \widetilde{\kappa}$, the manager is ex-post efficient if and only if $\mathcal{S}^*(\widetilde{\kappa}) = \{\hat{s}\}$. Thus, the manager is ex-post efficient whenever the evidence from the test is not valuable as it is unable to overturn prior beliefs, i.e., if $\kappa \in (0, \underline{\kappa}] \cup [\overline{\kappa}, \infty)$; or, whenever the evidence is valuable but a commitment to a standard is not, i.e., if $\kappa = \widetilde{\kappa}$ and $\mathcal{S}^*(\widetilde{\kappa}) = \{\hat{s}\}$.

5 Confirmativism and Conservatism

We now study whether deviations from ex-post optimality of the standard under commitment sway the manager towards or against the alternative he would choose according to his prior beliefs, in absence of any evidence.

5.1 Confirmativism and Conservatism vs. Prior Beliefs

The definitions of confirmativism and conservatism (c.f., Li (2001)) are useful in the subsequent analysis:

Definition 3 *Let* (s^*_{SP}, θ^*) *be an equilibrium of* Γ_2*. The manager is* conservative *at* (s^*_{SP}, θ^*) *if he leans towards rejection (hiring) and he is soft (harsh) at* $(s^*_{SP}, \theta^*(s^*_{SP}))$*. The manager is* confirmative *at* (s^*_{SP}, θ^*) *if he leans towards rejection (hiring) and he is harsh (soft) at* $(s^*_{SP}, \theta^*(s^*_{SP}))$*.*

Proposition 1 allows us to characterize whether the manager is conservative or confirmative in equilibrium according to the priors. First, we need to extend the definitions of conservatism and confirmativism to the correspondence $\mathcal{S}^*$ in a natural way: we say that the manager is *conservative (confirmative, ex-post efficient) at* κ if he is conservative (confirmative, ex-post efficient) at $(s, \theta^*(s))$ for all $s \in \mathcal{S}^*(\kappa)$. The following corollary is implied by Proposition 1 and Definitions 1 and 3.

Corollary 1 *Assume F.1-F.4 and C.1. For every game* Γ_2 *and* $\kappa \neq \widetilde{\kappa}$*, the manager is*

1. *confirmative at* κ *if and only if* $\underline{\kappa} < \kappa < \min\{1, \widetilde{\kappa}\}$ *or* $\max\{1, \widetilde{\kappa}\} < \kappa < \overline{\kappa}$*, and*

2. *conservative at* κ *if and only if* $\min\{1, \widetilde{\kappa}\} < \kappa < \max\{1, \widetilde{\kappa}\}$.

Corollary 1 reveals that commitments to standards typically involves a mix of confirmativism and conservatism when the manager faces candidates with different priors of being fit.

5.2 Uniform Confirmativism

This subsection analyzes conditions on the signal distribution F which lead to a non-conservative manager for all priors.

Definition 4 *A manager is* uniformly confirmative *(*conservative*) if and only if he is confirmative (conservative) for some priors and he is not conservative (confirmative) for any priors.*

The following result is a direct consequence of Corollary 1 (and hence, its proof is omitted). It provides a necessary and sufficient condition for uniform confirmativism and rules out uniform conservatism.

Corollary 2 *Assume F.1-F.4 and C.1. In any game* Γ_2*, (i) the manager is uniformly confirmative if and only if* $\widetilde{\kappa} = 1$ *and (ii) the manager cannot be uniformly conservative.*

There are families of distributions that lead to uniform confirmativism for all cost function satisfying Assumption C.1. An interesting case, because of its simplicity, is the *rotation distribution* (RD) family, defined by $[\underline{\theta}, \overline{\theta}] = [0, 1]$ and

$$f(z, \theta) = \theta f(z, 1) + (1 - \theta) f(z, 0)$$

for all $(z, \theta) \in (0, 1)^2$, such that $f(\cdot, 1)/f(\cdot, 0)$ is strictly increasing in z over $(0, 1)$. The modularity switch-point of any game Γ_2 with F in the RD family corresponds to the crossing point of $f(\cdot, 0)$ and $f(\cdot, 1)$. The manager is uniformly confirmative in any game Γ_2 with F in the RD family. In the sequel, we provide a characterization of uniform confirmativism, including the RD family as a particular case.

We say that a distribution F *has a neutral signal* $s^\star \in (0, 1)$ if $f(s^\star, \theta) = f(s^\star, \theta')$ for every θ, $\theta' \in \Theta$ (c.f., Milgrom (1981)). If F has a neutral signal $s^\star$, then the modularity-switch point of any game Γ_2 with distribution F is equal to the neutral signal, $\hat{s} = s^\star$. Having a neutral signal is a necessary condition on the family of distributions F to generate uniform confirmativism for all cost function.

Lemma 4 *Assume F.1-F.4. If the manager of a game* Γ_2 *with distribution* F *is uniformly confirmative for all cost function satisfying Assumption C.1, then* F *has a neutral signal.*

We now introduce a second property of distributions that plays a key role in the analysis of uniform confirmativism. For each $\theta \in \Theta^\circ$, we define implicitly the function $s \mapsto s_f(\cdot, \theta)$ by

$$h(s,\theta) = h(s_f(s,\theta),\theta)$$

with $s_f(s,\theta) \neq s$ for all $s \in [0,1] \setminus \{\widetilde{s}(\theta)\}$, and $s_f(\widetilde{s}(\theta),\theta) = \widetilde{s}(\theta)$. The continuity of h, $h(0,\theta) = h(1,\theta) = 0$, and $\frac{\partial h(s,\theta)}{\partial s} < (>) 0$ for all $s \in (0, \widetilde{s}(\theta))$ ($s \in (\widetilde{s}(\theta), 1)$) guarantee that s_f is well defined. We say that s and $s_f(s,\theta)$ are *matched* at θ, i.e., they induce the same value for $h(\cdot,\theta)$. If each standard is matched to the same standard at all θ, i.e., if $s_f(s,\theta)$ does not depend on θ, then we say that F satisfies the fixed matching property:

Definition 5 *A cumulative distribution F has the* fixed matching property *(FMP) if $h(s,\theta) = h(s',\theta)$ for some $\theta \in \Theta^\circ$ implies $h(s,\theta') = h(s',\theta')$ for all $\theta' \in \Theta^\circ$, for all $s, s' \in [0,1]$.*

If F has the FMP, then the fixed-matched standard of s is denoted by $s_f(s)$. If F has the FMP then (for details, refer to Claim 2 in the appendix): F has a neutral signal, $\theta^*(s) = \theta^*(s_f(s))$ for all $s \in [0,1]$ and every cost function, and $V(s, \theta^*(s); 1) = V(s_f(s), \theta^*(s_f(s)); 1)$, i.e.,

$$F(s,\overline{\theta}) - F(s,\theta^*(s)) = F(s_f(s),\overline{\theta}) - F(s_f(s),\theta^*(s_f(s))) \tag{18}$$

for all $s \in [0,1]$ and every cost function. In turn, (18) and Corollary 2 imply that FMP is a sufficient condition for uniform confirmativism.

Lemma 5 *Assume F.1-F.4. If F satisfies the FMP, then the manager is uniformly confirmative for all cost function satisfying Assumption C.1.*

Our next result reveals that FMP is not a necessary condition for F to generate uniform confirmativism for all cost function, and provides the conditions required on those parts of its domain violating FMP, in order to have uniform confirmativism.

Proposition 2 *Assume F.1-F.4. In every game Γ_2, the manager is uniformly confirmative for all cost function satisfying Assumption C.1 if and only if*

(i) F has a neutral signal $s^\star$, and

(ii)

$$F(s,\overline{\theta}) - F(s,\theta) < F(s_f(s,\theta),\overline{\theta}) - F(s_f(s,\theta),\theta) \tag{19}$$

implies

$$f(s,\overline{\theta}) \leq (\geq) f(s,\theta) - h(s,\theta)\frac{\partial f(s,\theta)}{\partial \theta}\left(\frac{\partial h(s,\theta)}{\partial \theta}\right)^{-1} \tag{20}$$

for all $s \in (0, s^\star)$ *(*$s \in (s^\star, 1)$*) and* $\theta \in \Theta^\circ$*.*

Condition (ii) in Proposition 2 guarantees that any pair (s, θ) satisfying (19) does not minimize the manager's expected loss when $\kappa = 1$; that is, it is not a critical point of $V(\cdot, \theta^*(\cdot); 1)$ for any cost function.

We end this section providing two simple properties on the functional form of F that directly yield uniform confirmativism, using the results above. The first property is separability of the off-diagonal terms of the Hessian of F: if there exist functions $\alpha : (0,1) \to \mathbb{R}$ and $\beta : \Theta^\circ \to \mathbb{R}$ such that $\frac{\partial f(z,\theta)}{\partial \theta} = \alpha(z)\beta(\theta)$ for all $z \in (0,1)$ and $\theta \in \Theta^\circ$, then F satisfies FMP —for details, see Claim 3 in the appendix. This family nests the RD family with, e.g., $\alpha(z) = f(z, 0) - f(z, 1)$ and $\beta(\theta) = -1$. The second property is rotational symmetry of the off-diagonal terms of the Hessian of F: if $\frac{\partial f(z,\theta)}{\partial \theta} = -\frac{\partial f(1-z,\theta)}{\partial \theta}$ for all $z \in (0, \frac{1}{2})$ and $\theta \in \Theta^\circ$, then F satisfies FMP, with $s_f(s, \theta) = 1 - s$ for all $s \in (0,1)$ and $\theta \in \Theta^\circ$.[19] In the next section we provide a number of examples.

6 Examples

In this section, we provide a number of examples of the results above. In games whose family of distributions F and cost function C yield a monotone pseudo likelihood ratio, computations are particularly simple. In Subsection 6.1 we analyze this case and provide an example. Subsection 6.2 provides two examples of uniform confirmativism. In absence of monotonicity of the pseudo likelihood ratio, standards may vary discontinuously with changes in beliefs. We consider such an example in Subsection 6.3.

6.1 Examples with Monotone Pseudo Likelihood Ratio

If the magnitude of the changes in effort in response to changes in the standard are relatively modest, the pseudo likelihood ratio function v is monotone —despite of the strategic term that makes it different from the likelihood ratio function $g(\cdot, \theta^*(\cdot))$. When v is monotone we say that the game Γ_2 satisfies the *monotone pseudo likelihood ratio* (MPLR) property. It is intuitive that as effort costs increase, unfit candidates become less responsive to changes in the standard. Below we provide sufficient conditions for high effort cost to yield MPLR. We first introduce two additional assumptions.

Assumption F.5 *All first, second, and third-order partial derivatives of* F *are continuous real functions defined over* D *and* $F(0, \theta) = 0$ *and* $F(1, \theta) = 1$*, for all* $\theta \in \Theta$*.*

Assumption F.5 strengthens Assumption F.1 by ensuring that the derivatives of F, up to the third order, are bounded.

[19] This follows directly from $h(s, \theta) = \int_0^s \frac{\partial h(z,\theta)}{\partial z} dz = -\int_{1-s}^1 \frac{\partial h(z,\theta)}{\partial z} dz = h(1 - s, \theta)$ for all $s \in (0,1)$ and $\theta \in \Theta^\circ$.

Our arguments below will often involve the analysis of the derivative of v. The existence of this derivative requires C to be thrice differentiable:

Assumption C. 2 *C is thrice continuously differentiable over Θ and $C''(\theta) > 0$ for all $\theta \in \Theta$.*

The following remark provide sufficient conditions for high effort costs to lead to MPLR.

Remark 5 *Assume F.2, F.3, F.5, C.1, C.2, and $f > 0$ (over D). Then, there exists $\overline{\lambda} > 0$ such that the game Γ_2 defined by F and the cost function $\lambda^{-1}C$ satisfies MPLR, for all $\lambda \in (0, \overline{\lambda})$.*

The following remark will be useful in the analysis of the examples satisfying MPLR:

Remark 6 *Suppose that Γ_2 satisfies MPLR. Then,*

$$\widetilde{\kappa} = \hat{\kappa} := g(\hat{s}, \theta^*(\hat{s}))$$

and therefore, Proposition 1, and Corollaries 1 and 2 hold, mutatis mutandis, replacing $\widetilde{\kappa}$ by $\hat{\kappa}$. Further, $\mathcal{S}^*(\hat{\kappa}) = \{\hat{s}\}$, and thus the manager is ex-post efficient at $\hat{\kappa}$.

Confirmative and conservative standards. The following example illustrates how the manager adopts confirmative or conservative standards depending on his prior beliefs.

Example 1 *Consider the family of distributions indexed by $\theta \in \Theta = [\underline{\theta}, \overline{\theta}]$ with $\underline{\theta} \in (0, \overline{\theta} - (e\overline{\theta})^{-1})$,*[20]

$$F(z, \theta) = z^{\theta}$$

for all $z \in [0, 1]$. Accordingly, the corresponding family of density functions is given by

$$f(z, \theta) = \theta z^{\theta - 1},$$

the likelihood ratio function is

$$g(z, \theta) = \frac{\overline{\theta}}{\theta} z^{\overline{\theta} - \theta},$$

and

$$h(z, \theta) = z^{\theta} \ln z$$

for all $z \in (0, 1)$ and $\theta \in \Theta$. $F(z, \theta)$ is strictly submodular (supermodular) in the region $z < (>) e^{-\frac{1}{\theta}}$ and $\frac{\partial h(z,\theta)}{\partial z} = 0$ at the locus

$$z = e^{-\frac{1}{\theta}}. \tag{21}$$

[20]This condition, jointly with the cost function of this example guarantee that Assumption C.1 holds. Notice that $\overline{\theta} - (e\overline{\theta})^{-1} > 0$ is equivalent to $\overline{\theta} > e^{-1/2}$.

The unfit candidate's disutility of effort is given by the function $C(\theta) = \frac{1}{2}(\theta - \underline{\theta})^2$ *for all* $\theta \in \Theta$*. Thus, given a standard* s*, she solves the problem*

$$\min_{\theta \in \Theta} U(s, \theta) = \min_{\theta \in \Theta} \left\{ s^{\theta} + \frac{1}{2}(\theta - \underline{\theta})^2 \right\}.$$

The unfit candidate's optimal level of effort θ *as a function of the manager's standard is the root of*

$$-s^{\theta^*(s)} \ln s = \theta^*(s) - \underline{\theta} \tag{22}$$

for all $s \in (0,1)$*. The only duplet* $(s, \theta) \in [0,1] \times \Theta$ *that satisfies both (21) and (22) (and* $z = s$*) is* $(\hat{s}, \theta^*(\hat{s}))$ *with* $\theta^*(\hat{s}) = \frac{1}{2}(\underline{\theta} + (\underline{\theta}^2 + 4e^{-1})^{\frac{1}{2}})$*; and* $\hat{s} = e^{\frac{-1}{\theta^*(\hat{s})}}$*, being the latter the modularity-switching point. Therefore,* $\hat{\kappa} \equiv g(\hat{s}, \theta^*(\hat{s})) = \frac{\bar{\theta}}{\theta^*(\hat{s})} e^{1 - \frac{\bar{\theta}}{\theta^*(\hat{s})}} < 1$.

A sufficient condition for MPLR is $\bar{\theta} > (\underline{\theta}^2 + 4e^{-1})^{\frac{1}{2}}$*. Under this condition, from Remark 6 and Corollary 1, the manager is confirmative if and only if* $\kappa \in (0, \hat{\kappa}) \cup \left(1, \frac{\bar{\theta}}{\underline{\theta}}\right)$*, conservative if and only if* $\kappa \in (\hat{\kappa}, 1)$*, and ex-post efficient if and only if* $\kappa = \hat{\kappa}$ *or* $\kappa \in \left[\frac{\bar{\theta}}{\underline{\theta}}, \infty\right)$*.*

The unfit candidate's equilibrium best-response and the manager's equilibrium standards for each κ *and both* Γ_1 *and* Γ_2 *are illustrated in Figure 1 for* $\Theta = [1 - e^{-1}, 2]$*.*

6.2 Examples of Uniform Cofirmativism

Several specifications satisfy uniform confirmativism. The family of distributions in the following example is in the RD family and, from Claim 3 in the Appendix, it satisfies FMP. Hence, by Lemma 5, this family of distributions leads to uniform confirmativism for all cost function.

Example 2 *Consider the cumulative distribution functions indexed by* $\theta \in \Theta = [0,1]$*,*

$$F(z, \theta) = \theta z^2 + (1 - \theta) z$$

for all $z \in [0,1]$*. Accordingly, the corresponding family of density functions is given by*

$$f(z, \theta) = \theta 2z + (1 - \theta),$$

that is, $f(\cdot, \theta)$ *is a convex combination of the densities* $2z$ *and the density of the standard uniform distribution, with weights* θ *and* $1 - \theta$*, respectively. The likelihood ratio function is*

$$g(z, \theta) = \frac{2z}{\theta 2z + (1 - \theta)},$$

and

$$h(z,\theta) = z(z-1)$$

for all $z \in [0,1]$ *and* $\theta \in \Theta$. $F(z,\theta)$ *is submodular (supermodular) in the region* $z < (>)\frac{1}{2}$ *and* $\frac{\partial f(z,\theta)}{\partial \theta} = 0$ *at* $z = \frac{1}{2}$. *Thus,* $\hat{s} = \frac{1}{2}$ *is the modularity-switch point.*

The unfit candidate's disutility of effort is given by the function $C(\theta) = \frac{1}{2}\theta^2$ *for all* $\theta \in \Theta$*; thus, she solves the problem*

$$\min_{\theta \in \Theta} U(s,\theta) = \min_{\theta \in \Theta} \left\{ \theta s^2 + (1-\theta)s + \frac{1}{2}\theta^2 \right\}.$$

The unfit candidate's optimal level of effort θ *as a function of the manager's standard is*

$$\theta^*(s) = s(1-s) \tag{23}$$

for all $s \in [0,1]$.

Since F *is in the RD family, Claim 3 and Lemma 5 imply that the manager is uniformly confirmative. Alternatively, direct computations reveal that* $dv(s)/ds > 0$ *for all* $s \in (0,1)$*; thus this game satisfies the MPLR property, and hence, Remark 6 implies* $\widetilde{\kappa} = \hat{\kappa} = g(\hat{s}, \theta^*(\hat{s})) = 1$, *and Corollary 2 implies that the manager is uniformly confirmative.*

Now we use Claim 3 to provide an example of uniform confirmativism with a family of distributions that is not in the RD family.

Example 3 *Consider the family of distribution functions indexed by* $\theta \in \Theta = [0,1]$,

$$F(z,\theta) = z^3 + (1-\theta^\alpha)z(1-z^2)$$

for all $z \in [0,1]$ *and* $\alpha \in (0,1)$. *From Claim 3,* F *satisfies the FMP. The modularity-switch point is* $\hat{s} = \frac{1}{\sqrt{3}}$.

The unfit candidate's effort disutility is given by $C(\theta) = \frac{1}{2}\theta^2$, *for all* $\theta \in [0,1]$, *thus he solves the problem*

$$\min_{\theta \in \Theta} U(s,\theta) = \min_{\theta \in \Theta} \left\{ s^3 + (1-\theta^\alpha)s(1-s^2) + \frac{1}{2}\theta^2 \right\}.$$

Therefore, her optimal effort, as a function of the manager's standard, is

$$\theta^*(s) = \left(\alpha s(1-s^2)\right)^{\frac{1}{2-\alpha}}$$

for all $s \in [0,1]$. *From Lemma 5, the manager is uniformly confirmative. In particular, from Corollary 1, he is* *confirmative in equilibrium if and only if* $\kappa \in (0,3) \setminus \{1\}$ *and ex-post efficient otherwise.*[21]

6.3 Non-monotone Pseudo Likelihood Ratio and Discontinuous Standards of Evidence

If MPLR does not hold, the standard in equilibrium can vary discontinuously with changes in κ: at critical points of $V(\cdot, \theta^*(\cdot))$ located in intervals where v is decreasing, the manager's expected loss is maximized. This discontinuity implies that candidates with arbitrarily similar priors may be subject to very different standards.

Remark 7 *Assume F.1-F.3 and C.1. Suppose that* v *is strictly decreasing over some interval* $(\underline{s}, \overline{s})$*, with* $0 < \underline{s} < \overline{s} < 1$*. Then, there exists a prior* $\kappa \in (0, \infty)$ *and* $\delta > 0$ *such that* $\kappa' < \kappa < \kappa''$*,* $s' \in \mathcal{S}^*(\kappa')$*, and* $s'' \in \mathcal{S}^*(\kappa'')$ *imply* $s'' - s' > \delta$*.*

The proof, provided in the appendix, uses indirect arguments. It is instructive, however, considering a simple case for which we here provide a constructive argument: suppose $v(s) > 0$ for all $s \in (0,1)$ and has two critical points: a unique interior local minimum $\underline{v} > \underline{\kappa}$ and a unique interior local maximum $\overline{v} < \overline{\kappa}$. Figure 2 provides an example satisfying these conditions. For any $\kappa \in [\underline{v}, \overline{v}]$, there are two local minimizers of $V(\cdot, \theta^*(\cdot); \kappa)$, corresponding to the two outer roots of $v(s) = \kappa$ in the interval $[\underline{v}, \overline{v}]$ of the range of v. We denote these roots by $s_1(\kappa)$ and $s_2(\kappa)$, with $s_1(\kappa) < s_2(\kappa)$, and observe that $dv(s_1(\kappa))/ds,\ dv(s_2(\kappa))/ds > 0$.

Let the inner root of $v(s) = \kappa$ in the interval $[\underline{v}, \overline{v}]$ of the range be denoted by $s_3(\kappa)$ and notice that $dv(s_3(\kappa))/ds < 0$. The difference between the expected loss to the manager between choosing $s_2(\kappa)$ and $s_1(\kappa)$ is

$$\begin{aligned} V(s_2(\kappa), \theta^*(s_2(\kappa))) - V(s_1(\kappa), \theta^*(s_1(\kappa))) \quad = \quad & \int_{s_1(\kappa)}^{s_3(\kappa)} (v(s) - \kappa) \left(f(s, \theta^*(s)) + h(s, \theta^*(s)) \frac{d\theta^*(s)}{ds} \right) ds \\ & - \int_{s_3(\kappa)}^{s_2(\kappa)} (\kappa - v(s)) \left(f(s, \theta^*(s)) + h(s, \theta^*(s)) \frac{d\theta^*(s)}{ds} \right) ds \end{aligned}$$

for all $\kappa \in [\underline{v}, \overline{v}]$. For $\kappa = \underline{v}$, $s_2(\kappa) = s_3(\kappa)$, so the first integral is strictly positive whereas the second integral is zero; hence $V(s_2(\kappa), \theta^*(s_2(\kappa))) > V(s_1(\kappa), \theta^*(s_1(\kappa)))$ and $\mathcal{S}^*(\underline{v}) = \{s_1(\kappa)\}$. Analogously, for $\kappa = \overline{v}$, $s_1(\kappa) = s_3(\kappa)$, so the second integral is strictly positive whereas the first integral is zero; hence $V(s_2(\kappa), \theta^*(s_2(\kappa))) < V(s_1(\kappa), \theta^*(s_1(\kappa)))$ and $\mathcal{S}^*(\overline{v}) = \{s_2(\kappa)\}$. As κ increases within $[\underline{v}, \overline{v}]$, $s_1(\kappa)$ increases, $s_3(\kappa)$ decreases and the integrand of the first integral decreases, so the first integral (monotonically) decreases. At the same time, $s_2(\kappa)$ increases, and the integrand of the second integral increases, so the second integral (monotonically) increases.

[21] The parameter α does not affect the range of values of κ yielding confirmativism, yet it does affect the magnitude of the deviation from ex-post optimality.

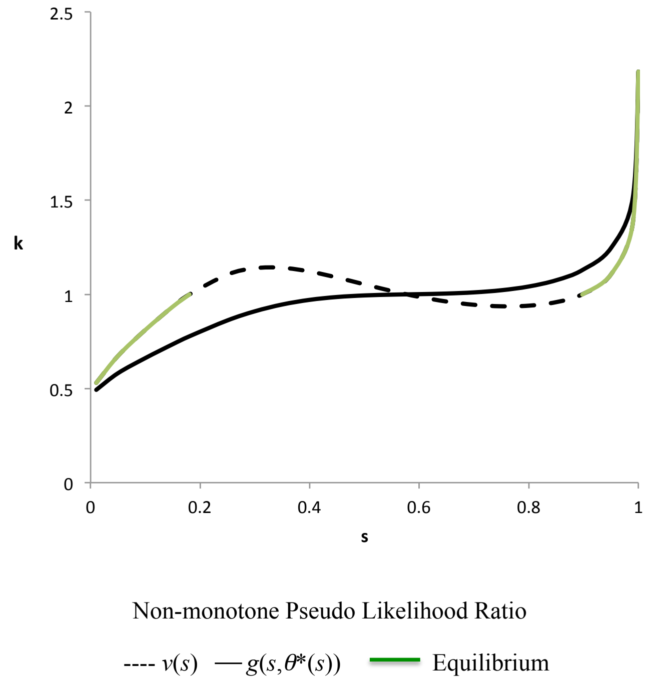

Figure 2: **Discontinuous standards.** Equilibrium standards for each prior κ in the dynamic game Γ_2 from Example 4 (solid green line) and the static game Γ_1 corresponding to that specification of F and C (solid black line).

Thus, there exists a unique value of $\kappa \in [\underline{v}, \overline{v}]$, denoted by κ°, such that

$$\mathcal{S}^*(\kappa) = \begin{cases} \{s_1(\kappa)\} & \text{if } \underline{v} \leq \kappa < \kappa^\circ \\ \{s_1(\kappa), s_2(\kappa)\} & \text{if } \kappa = \kappa^\circ \\ \{s_2(\kappa)\} & \text{if } \kappa^\circ < \kappa \leq \overline{v}. \end{cases}$$

In the following example, MPLR is not satisfied. Along with illustrating discontinuous standards, the example shows that MPLR is not a necessary condition for uniform confirmativism.

Example 4 *Consider the same family of distribution functions from Example 3 but with the constraint* $\overline{\theta} < \left(\frac{2}{2-\alpha}\right)^{\frac{1}{\alpha}} \left(\frac{2\alpha}{3\sqrt{3}}\right)^{\frac{1}{2-\alpha}} < 1$ *in the parameter constellation. The candidate's best response function is the same as in Example 3, because* $\theta^*(\cdot)$ *does not depend on* $\overline{\theta}$*. The modularity-switch point is* $\hat{s} = 1/\sqrt{3}$ *and* $\hat{\kappa} = 1$ *as before, but now MPLR does not hold. The function* v *is positive and non-monotone, having two interior local extreme: a local maximum at* $s < \hat{s}$ *and a local minimum at* $s > \hat{s}$*, as illustrated in Figure 2 for* $\alpha = 1/2$*. Since* $dv(\hat{s})/ds < 0$*, the modularity-switch point* $\hat{s}$ *is no longer an optimal standard for any value of* κ*. As a consequence, the manager's rejection standard varies discontinuously with* κ*. Routine computations yield* $\underline{\kappa} = 1 - \overline{\theta}^\alpha$ *and* $\overline{\kappa} = 1 + 2\overline{\theta}^\alpha$*. The equilibrium standard increases in* κ *(from Lemma 3) and varies discontinuously with* κ *at* $\tilde{\kappa} = 1$*, where* *the optimal standard jumps from $s_1(1)$ to $s_2(1)$.*[22] *Thus, the manager is soft if and only if $\kappa \in (1-\overline{\theta}^{\alpha}, 1)$ and harsh if and only if $\kappa \in (1, 1+2\overline{\theta}^{\alpha})$. Consequently, the manager is uniformly confirmative. In particular, he is confirmative in equilibrium if and only if $\kappa \in (1-\overline{\theta}^{\alpha}, 1+2\overline{\theta}^{\alpha}) \setminus \{1\}$ and ex-post efficient otherwise.*

7 Welfare Analysis

In this section we analyze the impact of information undermining on the welfare of all players, namely the manager, fit candidates, and unfit candidates. First, in Subsection 7.1 we compare directly welfare in equilibrium of the dynamic game with welfare in absence of information undermining. In Subsection 7.2, we analyze how the manager's commitment ability affects welfare in equilibrium.

7.1 Welfare Consequences of Information undermining

Information undermining always makes the manager worse off when the problem is test-worthy. On the other hand, fit candidates are worse-off under high priors leading to equilibrium standards that are higher than in absence of information undermining. Within the highest of those priors, unfit candidates are also worse-off, and hence, the equilibrium of Γ_2 to be Pareto dominated by the outcome in absence of information undermining.

Manager. The manager in the pure information undermining model can only be worse-off than in the classical statistical problem:

Lemma 6 *Assume F.1-F.3 and C.1. The manager's expected loss in any subgame perfect equilibrium of Γ_2 is strictly greater than in the classical statistical problem in every test-worthy problem.*

Proof. For all $\kappa \in (\underline{\kappa}, \overline{\kappa})$,

$$F(s^*(\underline{\theta}, \kappa), \overline{\theta}) - \kappa F(s^*(\underline{\theta}, \kappa), \underline{\theta}) \leq F(s, \overline{\theta}) - \kappa F(s, \underline{\theta}) \leq F(s, \overline{\theta}) - \kappa F(s, \theta^*(s))$$

for all $s \in [0,1]$, where the first inequality is strict except for $s = s^*(\underline{\theta}, \kappa)$; and the second inequality is strict except for $s \in \{0,1\}$, by strict MLRP and the fact that $\theta^*(s) > \underline{\theta}$ for all $s \in (0,1)$. ■

Candidates. Fit candidates are better off in the subgame perfect equilibrium of the game Γ_2 if and only if the manager chooses a lower standard in equilibrium than in the classical statistical problem: $s^*_{SP} < s^*(\underline{\theta}, \kappa)$. If fit candidates are better-off under information undermining, then unfit candidates are better-off as well as the possibility of making effort improve the prospects of unfit candidates. In absence of information undermining, the expected loss to unfit candidates is $F(s^*(\underline{\theta}, \kappa), \underline{\theta})$, and, if $s^*_{SP} < s^*(\underline{\theta}, \kappa)$, then $F(s^*(\underline{\theta}, \kappa), \underline{\theta}) > F(s^*_{SP}, \underline{\theta}) + 0 \geq$

[22] Both $s_1(1)$ and $s_2(1)$ satisfy the equality : $s(1-s^2) = \frac{1}{\alpha}\left(\left(\frac{2-\alpha}{2}\right)^{\frac{1}{\alpha}} \overline{\theta}\right)^{2-\alpha}$.

$F(s^*_{SP}, \theta^*(s^*_{SP})) + C(\theta^*(s^*_{SP}))$. Indeed, even if fit candidates are worse-off under information undermining (due to a higher standard) unfit candidates might be better-off if their gains from exerting effort are greater than the increase in losses associated to higher effort costs, i.e., if

$$C(\theta^*(s^*_{SP})) < F(s^*(\underline{\theta}, \kappa), \underline{\theta}) - F(s^*_{SP}, \theta^*(s^*_{SP})).$$

We thus have the following observation:

Remark 8 *Assume F.1-F.3 and C.1. An equilibrium (s^*_{SP}, θ^*) of Γ_2 is Pareto dominated by the outcome of the classical statistical problem if and only if the problem is test-worthy and $F(s^*(\underline{\theta}, \kappa), \underline{\theta}) - F(s^*_{SP}, \theta^*(s^*_{SP})) \leq C(\theta^*(s^*_{SP}))$.*

The main result of this section provides sufficient conditions yielding a range of prior beliefs such that the equilibrium of the dynamic game is dominated by the outcome of the classical statistical problem.

Proposition 3 *Assume F.1-F.4 and C.1. (i) There exists $\breve{\kappa} > \underline{\kappa}$ such that both fit and unfit candidates are better-off in the equilibrium of Γ_2 than in the classical statistical problem for all $\kappa \in (\underline{\kappa}, \breve{\kappa})$, and (ii) if F.5 and C.2 hold, and $f(1, \underline{\theta}) > 0$, then there exists $\check{\breve{\kappa}} < \overline{\kappa}$ such that the outcome of the classical statistical problem Pareto Dominates the subgame perfect equilibrium of Γ_2 for all $\kappa \in (\check{\breve{\kappa}}, \overline{\kappa})$.*[23]

The following example illustrates the previous findings.

Example 5 *We revisit Example 2. In this example,*

$$v(s) = \frac{2s}{1 + 2s(1-s)(2s-1)},$$

$g(s, \underline{\theta}) = 2s$, and $F(s, \underline{\theta}) = s$, for all $s \in (0, 1)$. Thus, we have $F(s^*(\underline{\theta}, \kappa), \underline{\theta}) = \frac{\kappa}{2}$ for $\kappa \in [0, 2]$. Also,

$$F(s, \theta^*(s)) + C(\theta^*(s)) = s(1 - s(1-s)^2) + \frac{1}{2}s^2(1-s)^2 = s - \frac{1}{2}s^2(1-s)^2,$$

for all $s \in [0, 1]$.

On the other hand, if (s^*_{SP}, θ^*) is an equilibrium with $s^*_{SP} \in (0, 1)$, then $v(s^*_{SP}) = \kappa$ and hence

$$s^*_{SP}(1 - \kappa(1 - s^*_{SP})(2s^*_{SP} - 1)) = \frac{\kappa}{2}.$$

Therefore

$$F(s^*(\underline{\theta}, \kappa), \underline{\theta}) - F(s^*_{SP}, \theta^*(s^*_{SP})) \leq C(\theta^*(s^*_{SP}))$$

[23] The proof reveals that the equilibrium of Γ_2 is indeed unique for all $\kappa < \breve{\kappa}$ and all $\kappa > \check{\breve{\kappa}}$.

if and only if $\frac{s^*_{SP}(1-s^*_{SP})}{2s^*_{SP}-1} \in (0, 2\kappa]$*, which is equivalent to* $s^*_{SP} \in [0.56, 1)$ *or* $\kappa \in [1.06, 2)$*. The equilibrium of* Γ_2 *is Pareto dominated by the outcome of the Classical Statistical Problem if and only if* $\kappa \in [1.06, 2)$*.*

7.2 Ex-post Inefficient Standards and Welfare

In this section we analyze how the welfare of all players is affected by the ability of the manager to commit to a hiring standard. We thus compare equilibrium payoffs in the dynamic and static games. Recall that Proposition 1 states that the manager is soft for all prior beliefs such that $\kappa \in (\underline{\kappa}, \widetilde{\kappa})$ and harsh for all prior beliefs such that $\kappa \in (\widetilde{\kappa}, \overline{\kappa})$. Candidates a priori relatively likely to be fit ($\kappa \in (\underline{\kappa}, \widetilde{\kappa})$) are better-off by the manager's commitment because it leads to a downward distortion on the standard and, in the case of unfit candidates, less effort is exerted in equilibrium. In contrast, candidates a priori relatively unlikely to be fit ($\kappa \in (\widetilde{\kappa}, \overline{\kappa})$) are worse-off by the manager's commitment because of the increase in the probability of being rejected, due to the upward distortion on the standard. In the case of unfit candidates, this effect dominates the decrease in costs due to the lower exerted effort. Candidates with high enough prior probability of being fit (unfit), i.e., $\kappa \leq \underline{\kappa}$ ($\kappa \geq \overline{\kappa}$), are automatically hired (rejected) in both the static and dynamic games, regardless the signal observed in the test. Hence, candidates with such prior probabilities do not benefit from commitment. Summarizing, we have the following corollary:

Corollary 3 *Assume F.1-F.3 and C.1. A subgame perfect equilibrium of* Γ_2 *Pareto dominates the equilibrium of* Γ_1 *if and only if the manager is soft. Both fit and unfit candidates are worse-off in* Γ_2 *than in* Γ_1 *if and only if the manager is harsh.*

Proof. The result is a direct consequence of the Envelope Theorem. ■

8 Strategic Fit Candidates

In this section we analyze a third game in which the manager commits to an evidence standard, and then both fit and unfit candidates can exert effort to affect the distribution of the signal. The resulting variation of the model yields a signalling game. In contrast to most signalling games, candidates' actions (i.e., the effort exerted) are not observed by the manager. These actions, however, probabilistically increase the value of the signal observed by the manager.

8.1 Framework

Fit candidates are endowed with $\underline{\theta}_q \in \Theta^\circ$. Given a family of distributions F, $\underline{\theta}_q$ provides a measure of the intrinsic advantage of fit candidates in the test: if neither fit nor unfit candidates exert effort, then their corresponding signal distributions are $F(\cdot, \underline{\theta}_q)$ and $F(\cdot, \underline{\theta})$, respectively.[24]

The cost function of unfit candidates in this section is denoted by C_u —and hence in the sequel Assumption C.1 is interpreted as an assumption on C_u. The cost function of fit candidates is denoted by $C_q : D_q \to \mathbb{R}$, with $D_q = \{(\theta; \underline{\theta}_q) \in \Theta^2 : \theta \geq \underline{\theta}_q\}$ where for each $(\theta; \underline{\theta}_q) \in D_q$, θ represents effort and $\underline{\theta}_q$ is the initial advantage of fit candidates.

Assumption C. 3 *C_q is twice continuously differentiable, with $C_q(\underline{\theta}_q; \underline{\theta}_q) = C_q'(\underline{\theta}_q; \underline{\theta}_q) = 0$, $C_q'(\theta; \underline{\theta}_q) < C_u'(\theta)$, and $C_q''(\theta; \underline{\theta}_q) > 0$, for all $\theta \in \left(\underline{\theta}_q, \overline{\theta}\right)$ and $\underline{\theta}_q \in \Theta$.*

Assumption F.6 *$h(s, \overline{\theta}) = 0$ for all $s \in (0, 1)$.*

Assumption F.6 implies that $\overline{\theta}$ cannot be an optimal choice for fit candidates, as $C_q'(\overline{\theta}; \underline{\theta}_q) > -h(s, \overline{\theta})$ for all $s \in (0, 1)$. The resulting game is denoted by Γ_3.

Analysis of the model. Let θ_q^* and θ_u^* be the best response function of fit and unfit candidates, respectively. Assumption C.3 implies that $\theta_q^*(s) > \theta_u^*(s)$ for all $s \in [0, 1]$.

Since the manager commits to an evidence standard, the objective function of his minimization problem is an affine transformation of $\mathcal{V} : [0, 1] \to \mathbb{R}$, with

$$\mathcal{V}(s) := F(s, \theta_q^*(s)) - \kappa F(s, \theta_u^*(s)) \tag{24}$$

for all $s \in [0, 1]$. Let $g(s, \theta_q^*(s), \theta_u^*(s)) := \frac{f(s, \theta_q^*(s))}{f(s, \theta_u^*(s))}$ be the likelihood ratio function for all $s \in (0, 1)$,[25]

$$D_i(s) := \frac{dF(s, \theta_i^*(s))}{ds} = f(s, \theta_i^*(s)) + h(s, \theta_i^*(s)) \frac{d\theta_i^*(s)}{ds}$$

for all $s \in [0, 1]$ and $i = q, u$, and define the pseudo likelihood ratio function $w : D_w \to \mathbb{R} \cup \{\infty\}$, with $w(s) := \frac{D_q(s)}{D_u(s)}$ for all $s \in D_w := \{s \in (0, 1) : D_u(s) \neq 0\} \cup \{0, 1\}$. A *Subgame Perfect Equilibrium in Pure Strategies* of Γ_3 is a triplet $(s_{SP}^*, \theta_q^*, \theta_u^*)$ where s_{SP}^* is a minimizer of (24). As before, $\mathcal{S}^*$ denotes the correspondence mapping the priors κ to the set of standards in equilibrium, i.e., those that minimize (24) for each κ.

Counterparts to Lemmas 1-2. The single modularity-switch property holds in Γ_3 for the best response function of both fit and unfit candidates. The counterpart to Lemma 1 is:

[24] The case $\underline{\theta}_q = \overline{\theta}$ corresponds to the setup of the previous sections.

[25] In contrast to the function $g(\cdot, \theta^*(\cdot))$ defined in the analysis of the model with non-strategic fit candidates, $g(\cdot, \theta_q^*(\cdot), \theta_u^*(\cdot))$ may not be monotone.

Lemma 7 *Assume F.1-F.4, F.6, C.1, and C.3. For any game* Γ_3 *and* $i = u, q$*, there exists* $\hat{s}_i \in (0, 1)$ *such that*

$$\frac{\partial f(s, \theta_i^*(s))}{\partial \theta} \begin{cases} < 0 & \text{if } 0 \leq s < \hat{s}_i \\ = 0 & \text{if } s = \hat{s}_i \\ > 0 & \text{if } \hat{s}_i < s \leq 1. \end{cases} \tag{25}$$

The proof is analogous to the proof of Lemma 1 and it is omitted.

In equilibria such that $(s, \theta_q^*(s))$ and $(s, \theta_u^*(s))$ are located in the submodular and supermodular regions of F, respectively,[26] the efforts exerted by fit and unfit candidates are, respectively, strategic complement and strategic substitute of the standard. In such equilibria, the manager is harsh because his expected loss is decreasing in fit candidates' effort and increasing in unfit candidates' effort. An analogous argument yields that the manager is soft in equilibria such that $(s, \theta_q^*(s))$ and $(s, \theta_u^*(s))$ are located in the supermodular and submodular regions of F respectively.[27]

If both $(s, \theta_q^*(s))$ and $(s, \theta_u^*(s))$ are located in the submodular region of F, or both are located in the supermodular region of F, then the manager faces a trade-off between encouraging effort by fit candidates and discouraging effort by unfit candidates. In this case the magnitudes of the strategic effects of the standard on each type of candidate's effort play a critical role. In particular, we define the *strategic ratio*

$$r(s) := \frac{h(s, \theta_q^*(s)) \frac{d\theta_q^*(s)}{ds}}{h(s, \theta_u^*(s)) \frac{d\theta_u^*(s)}{ds}}$$

for all $s \in [0, 1] \setminus \{0, \hat{s}_u, 1\}$. We also define $r(0) = \lim_{s \to 0} r(s), r(\hat{s}_u) \lim_{s \to \hat{s}_u} r(s)$, and $r(1) = \lim_{s \to 1} r(s)$, respectively, whenever these limits exist. The counterpart to Lemma 2 is:

Lemma 8 *Assume F.1-F.3, F.6, C.1, and C.3. For any equilibrium* $(s_{SP}^*, \theta_q^*, \theta_u^*)$ *of* Γ_3 *with* $s_{SP}^* \in (0, 1)$*,*

(i) if $(s_{SP}^*, \theta_q^*(s_{SP}^*))$ *and* $(s_{SP}^*, \theta_u^*(s_{SP}^*))$ *are located in the supermodular (submodular) and submodular (supermodular) regions of F, i.e.,* $s_{SP}^* \in (\hat{s_q}, \hat{s}_u)$ $(s_{SP}^* \in (\hat{s_u}, \hat{s}_q))$*, respectively, then the manager is soft (harsh); and*

(ii) if both $(s_{SP}^*, \theta_q^*(s_{SP}^*))$ *and* $(s_{SP}^*, \theta_u^*(s_{SP}^*))$ *are located in the submodular region (i.e.,* $s_{SP}^* \in (0, \min\{\hat{s_q}, \hat{s}_u\})$*), then the manager is soft (harsh) if and only if* $r(s_{SP}^*) < (>)\kappa$*. Analogously, if both* $(s_{SP}^*, \theta_q^*(s_{SP}^*))$ *and* $(s_{SP}^*, \theta_u^*(s_{SP}^*))$ *are located in the supermodular region (i.e.,* $s_{SP}^* \in (\max\{\hat{s_q}, \hat{s}_u\}, 1)$*), then the manager is soft (harsh) if and only if* $r(s_{SP}^*) > (<)\kappa$*.*

Now we discuss the intuition of part (ii). If the ratio of the strategic effect of the standard on the signal

[26] For instance, if $\tilde{s}(\cdot)$ is strictly increasing, then this is the case for $s_{SP}^* \in (\hat{s}_u, \hat{s}_q)$.

[27] If $\tilde{s}(\cdot)$ is strictly decreasing, then this is the case for $s_{SP}^* \in (\hat{s}_q, \hat{s}_u)$.

distribution of fit candidates divided by the corresponding effect of the standard on the signal distribution of unfit candidates is smaller (greater) than the prior, then the later (former) effect dominates and the manager is relatively more concerned with the effect of his commitment on the effort exerted by unfit (fit) candidates. Since the effort exerted by unfit (fit) candidates is a strategic complement of the standard over the submodular region of F and a strategic substitute of the standard over the supermodular region of F, the manager optimally commits to soft (harsh) and harsh (soft) standards respectively in those regions.

Lemmas 7 and 8 imply that if the strategic concerns of the manager are dominated by the effect of changes in the standard on the effort of unfit candidates (i.e., when $r(s)$ is relative low), then, as the prior increases, a pattern of soft and then harsh standards arises —provided that the standard in equilibrium is increasing in the prior. In the next subsection, we provide conditions within the extended model for the manager's strategic concerns to be dominated by unfit candidates' efforts. On the other hand, if the strategic concerns of the manager are instead dominated by the effect of changes in the standard on the effort of fit candidates (i.e., when $r(s)$ is relative high), then a pattern of harsh and then soft standards arises as priors increase. In Subsection 8.3, we provide conditions for the manager's concerns to be dominated by the effort of fit candidates. In Subsection 8.4, we have a glance on other patterns of optimal standard commitments that may arise when neither type's effort is dominant.

8.2 Dominating Information Undermining in the Extended Model

Provided that the manager strategic concerns are dominated by the effect of the standard on the effort of unfit candidates, most qualitative aspects of the results we derived for the game Γ_2 arise here as well. For instance, if fit candidates have a very high intrinsic advantage, then they may benefit little from increasing their effort as the returns to effort vanish. In what follows, we analyze such an scenario.

If the initial advantage of fit candidates is large, then the thesis of Proposition 1 remains valid in the extended model: the manager is soft for relatively low prior probabilities that the candidate is unfit and harsh for relatively high prior probabilities that the candidate is unfit. If additionally F satisfies the FMP, then the manager is uniformly confirmative. Formally, let $\underline{\kappa}(\underline{\theta}_q) := \frac{f(0,\underline{\theta}_q)}{f(0,\underline{\theta})}$ and $\overline{\kappa}(\underline{\theta}_q) := \frac{f(1,\underline{\theta}_q)}{f(1,\underline{\theta})}$ for all $\underline{\theta}_q \in \Theta^\circ$. The analogous to Assumption C.2 in the setup of this section is:

Assumption C. 4 *$C_q(\cdot,\underline{\theta}_q)$ is thrice continuously differentiable and $C_q''(\theta,\underline{\theta}_q) > 0$ over $[\underline{\theta}_q,\overline{\theta}]$ for all $\underline{\theta}_q \in (\underline{\theta},\overline{\theta})$.*

We have the following result:

Proposition 4 *Assume F.2-F.6, C.1-C.4, and $f > 0$. Then, there exists $\underline{\theta}_q \in \Theta^\circ$ such that for all $\underline{\theta}_q \in \left(\underline{\theta}_q,\overline{\theta}\right)$, (i) there exists $\widetilde{\kappa}(\underline{\theta}_q) \in (\underline{\kappa}(\underline{\theta}_q),\overline{\kappa}(\underline{\theta}_q))$ such that (17) holds, mutatis mutandis, replacing $\widetilde{\kappa}$, $\underline{\kappa}$, and $\overline{\kappa}$, with $\widetilde{\kappa}(\underline{\theta}_q)$, $\underline{\kappa}(\underline{\theta}_q)$, and $\overline{\kappa}(\underline{\theta}_q)$, respectively; and (ii) if F satisfies FMP, then the manager is uniformly confirmative.*

The proof, provided in the appendix, spells out that the Lemmata 1-3 and the argument in the proof of Proposition 1 can be applied in the game Γ_3 for large enough $\underline{\theta}_q$. The same applies for the sufficiency of FMP for uniform confirmativism.

8.3 Information Generation

In this section, we consider the case in which the manager strategic concerns are dominated by the effect of the standard on the effort exerted by fit candidates. It is instructive to first analyze a pure information generation game, denoted by Γ_4, in which unfit candidates exert no effort to affect the signal distribution, i.e., $\theta_u^*(s) = \underline{\theta}$ for all $s \in [0,1]$. All other assumptions on Γ_3, laid out in Subsection 8.1, remain. Loosely speaking, Γ_4 is the game that we obtain from any game Γ_3 with cost function C_u, taking the limit as $C_u(\theta)$ goes to infinity for all $\theta > \underline{\theta}$.

The objective function of the manager's minimization problem in this setup is an affine transformation of:

$$\mathcal{W}(s, \theta_q^*(s)) := F(s, \theta_q^*(s)) - \kappa F(s, \underline{\theta}). \tag{26}$$

for all $s \in [0,1]$. A *Subgame Perfect Equilibrium in Pure Strategies* of Γ_4 is a duplet (s_{SP}^*, θ_q^*) where s_{SP}^* is a minimizer of (26). As before $\mathcal{S}^*$ maps priors to the set of minimizers of (26).

In this model, the likelihood ratio function, $\widetilde{g} : [0,1] \times [\underline{\theta}_q, \overline{\theta}] \to \mathbb{R} \cup \{\infty\}$, is given by

$$\tilde{g}(s, \theta) := \frac{f(s, \theta)}{f(s, \underline{\theta})}$$

for all $s \in [0,1]$ and $\theta \in [\underline{\theta}_q, \overline{\theta}]$. And the pseudo likelihood ratio function, $\widetilde{v} : [0,1] \to \mathbb{R} \cup \{\infty\}$, is given by

$$\tilde{v}(s) := \tilde{g}(s, \theta_q^*(s)) + \frac{h(s, \theta_q^*(s))}{f(s, \underline{\theta})} \frac{d\theta_q^*(s)}{ds}$$

for all $s \in [0,1]$.

In contrast to unfit candidates, fit candidates' effort increases the manager's expected payoff. It is thus intuitive that the analogous to Lemma 2, characterizing whether the manager is soft or harsh in a given equilibrium (s_{SP}^*, θ_q^*) according to whether F is submodular or supermodular at $(s_{SP}^*, \theta_q^*(s_{SP}^*))$, reverses the thesis of that result:

Lemma 9 *Assume F.1-F.3, F.6, and C.3. Let (s_{SP}^*, θ_q^*) be an equilibrium of Γ_4 such that $s_{SP}^* \in (0,1)$. The manager is harsh (soft) at (s_{SP}^*, θ_q^*) if and only if F is strictly submodular (strictly supermodular) at $(s_{SP}^*, \theta_q^*(s_{SP}^*))$.*

Further, observe that the thesis of Lemma 7 (the analogous to Lemma 1 in the benchmark model) still holds here for $i = q$. So does a version of Lemma 3 (see Lemma 13 in the appendix) stating that $\mathcal{S}^*$ is increasing and strictly increasing over $(\underline{\underline{\kappa}}, \overline{\overline{\kappa}})$, where

$$\underline{\underline{\kappa}} := \inf_{s\in(0,1)} \left\{ \frac{F(s,\theta_q^*(s))}{F(s,\underline{\theta})} \right\} \quad \text{and} \quad \overline{\overline{\kappa}} := \sup_{s\in(0,1)} \left\{ \frac{1-F(s,\theta_q^*(s))}{1-F(s,\underline{\theta})} \right\}.$$

Connecting Lemma 9 with these two observations allows us to obtain that, analogously to the results in the benchmark model, if only fit candidates exert effort, then the manager's standard is harsh (soft) for priors κ less (greater) than $\tilde{\tilde{\kappa}} := \tilde{g}(\hat{s}_q, \theta_q^*(\hat{s}_q))$.

Proposition 5 *Assume F.1-F.4, F.6, and C.3. For any game* Γ_4*,* $\underline{\underline{\kappa}} < \tilde{\tilde{\kappa}} < \overline{\overline{\kappa}}$ *and*

$$\text{the manager is} \begin{cases} \textit{ex-post efficient} & \textit{if } \kappa \in \big(0, \underline{\underline{\kappa}}\big) \\ \textit{harsh} & \textit{if } \kappa \in (\underline{\underline{\kappa}}, \tilde{\tilde{\kappa}}) \\ \textit{ex-post efficient} & \textit{if } \kappa = \tilde{\tilde{\kappa}} \\ \textit{soft} & \textit{if } \kappa \in (\tilde{\tilde{\kappa}}, \overline{\overline{\kappa}}) \\ \textit{ex-post efficient} & \textit{if } \kappa \in (\overline{\overline{\kappa}}, \infty). \end{cases} \tag{27}$$

The natural analogue to Corollary 1, characterising when the manager is confirmative and conservative in terms of the priors, is provided in the appendix (see Corollary 5).

Uniform Conservatism. The analogue to Corollary 2 stating when the manager is uniformly conservative is provided in the appendix (see Corollary 6). The characterization of the families of distributions such that the manager is uniformly conservative under information generation is slightly simpler than the characterization of the family of distributions such that the manager is uniformly confirmative in the case of information undermining:

Proposition 6 *Assume F.1-F.4, and F.6. In every game* Γ_4*, the manager is uniformly conservative for all initial advantage* $\underline{\theta}_q \in \Theta^\circ$ *and cost function* C_q *satisfying Assumption C.3 if and only if* F *has a neutral signal.*

Welfare Analysis under Information Generation. We define the *information-generation classical statistical problem* as the problem solved by the manager in absence of information generation (and information undermining) when fit candidates have an initial advantage $\underline{\theta}_q$. The manager's expected loss in this problem is

$$\mathcal{W}(s, \underline{\theta}_q) = F(s, \underline{\theta}_q) - \kappa F(s, \underline{\theta}),$$

for all $s \in [0, 1]$.

Proposition 7 *Assume F.1-F.4, F.6 and C.3. (i) If F.5 and C.4 hold, then, there exists* $\widetilde{\kappa} > \underline{\underline{\kappa}}$ *such that both fit and unfit candidates are worse-off in the equilibrium of* Γ_4 *than in the information-generation classical statistical problem for all* $\kappa \in (\underline{\underline{\kappa}}, \widetilde{\kappa})$*, and (ii) there exists* $\widetilde{\widetilde{\kappa}} < \overline{\overline{\kappa}}$ *such that the subgame perfect equilibrium of* Γ_4 *Pareto Dominates the outcome of the classical statistical problem for all* $\kappa \in (\widetilde{\widetilde{\kappa}}, \overline{\overline{\kappa}})$.[28]

Finally, we define the static game $\widetilde{\Gamma}_1$, between the manager and the fit candidate, which is analogous to the game Γ_1, between the manager and the unfit candidate. Therefore, in $\widetilde{\Gamma}_1$, the expected loss to the manager and the expected loss to fit candidates are given, respectively, by

$$\mathcal{W}(s, \theta_q) := F(s, \theta_q) - \kappa F(s, \underline{\theta}) \quad \text{and} \quad U_q(s, \theta_q) := F(s, \theta_q) + C_q(\theta_q) \tag{28}$$

for all $s \in [0, 1]$ and $\theta_q \in [\underline{\theta}_q, \overline{\theta}]$.

Corollary 4 *Assume F.1-F.3, F.6, and C.3. A subgame perfect equilibrium of* Γ_4 *Pareto dominates the equilibrium of* $\widetilde{\Gamma}_1$ *if and only if the manager is soft. Both fit and unfit candidates are worse-off in* Γ_4 *than in* $\widetilde{\Gamma}_1$ *if and only if the manager is harsh.*

Proof. The result is a direct consequence of the Envelope Theorem. ■

Dominating Information Generation in the Extended Model. The main qualitative features of the game of information generation, namely Γ_4, are preserved in the game Γ_3, in which both fit and unfit candidates exert effort, provided that unfit candidates' effort is sufficiently costly. For any game Γ_3, defined by a triplet (F, C_q, C_u), we define a set of games indexed by $\lambda \in [0, 1]$, and denoted by $\Gamma_3(\lambda)$, where the only difference between Γ_3 and $\Gamma_3(\lambda)$ is that, if the cost function of unfit candidates in the former is C_u, in the later it is given by $\lambda^{-1} C_u$, for all $\lambda \in (0, 1]$, whereas $\Gamma_3(0)$ corresponds to Γ_4 with distribution F and cost function C_q for fit candidates. Let $\underline{\underline{\kappa}}(\lambda) := \inf_{s \in (0,1)} F(s, \theta_q^*(s))/F(s, \theta_u^*(s; \lambda))$ and $\overline{\overline{\kappa}}(\lambda) := \sup_{s \in (0,1)} (1 - F(s, \theta_q^*(s)))/(1 - F(s, \theta_u^*(s; \lambda)))$, where $\theta_u^*(\cdot; \lambda)$ is the best response of unfit candidates with cost function $\lambda^{-1} C_u$ for all $\lambda \in (0, 1]$, and $\theta_u^*(\cdot; 0) = \underline{\theta}$.

Proposition 8 *Assume F.2-F.6, C.1-C.4, and* $f > 0$*. Then, there exists* $\overline{\lambda} > 0$ *such that for all* $\lambda \in (0, \overline{\lambda})$*, (i) there exists* $\widetilde{\widetilde{\kappa}}(\lambda) \in (\underline{\underline{\kappa}}(\lambda), \overline{\overline{\kappa}}(\lambda))$ *such that (27) holds in the game* $\Gamma_3(\lambda)$*, mutatis mutandis, replacing* $\widetilde{\widetilde{\kappa}}$*,* $\underline{\underline{\kappa}}$*, and* $\overline{\overline{\kappa}}$*, with* $\widetilde{\widetilde{\kappa}}(\lambda)$*,* $\underline{\underline{\kappa}}(\lambda)$*, and* $\overline{\overline{\kappa}}(\lambda)$*, respectively; and (ii) if F satisfies FMP, then the manager is uniformly conservative.*

8.4 Other Patterns

The following example illustrates how other patterns for soft and harsh standards may emerge in the game Γ_3, contrasting the soft-harsh and harsh-soft cut-off characterizations that arise in the pure information undermining and pure information generation games (Propositions 1 and 5, respectively).

[28] The proof reveals that the equilibrium of Γ_4 is indeed unique for all $\kappa < \widetilde{\kappa}$ and all $\kappa > \widetilde{\widetilde{\kappa}}$.

Example 6 *Consider the cumulative distribution functions indexed by* $\theta \in \Theta = [0, 1]$*,*

$$F(z, \theta) = F(z, 0) - \theta \left(1 - \frac{\theta}{2}\right)(F(z, 0) - F(z, 1)),$$

for all $z \in [0, 1]$*, where* $F(\cdot, 0)$ *and* $F(\cdot, 1)$ *are distribution functions on* $[0, 1]$ *with first, second and third-order derivatives that are continuous real functions over* $[0, 1]$*, and such that* $d(f(\cdot, 1)/f(\cdot, 0))/ds > 0$ *and* $f(1, 0) > 0$*.*

Unfit candidates' cost function is given by $C_u(\theta) = \frac{1}{\lambda}\frac{1}{2}\theta^2$ *for all* $\theta \in [0, 1]$*, with* $\lambda \in (0, 1]$*. Fit candidates are endowed with* $\underline{\theta}_q \in (0, 1)$ *and their cost function is given by* $C_q(\theta; \underline{\theta}_q) = \frac{1}{2}(\theta - \underline{\theta}_q)^2$ *for all* $\theta \in [\underline{\theta}_q, 1]$*.*

Claim 1 *In Example 6, (i) the manager is uniformly confirmative for high enough* $\underline{\theta}_q$*, (ii) the manager is uniformly conservative for small enough* λ*, and (iii) for some distributions* F *and intermediate constellations of* λ *and* $\underline{\theta}_q$*, none of the cut-off patterns in Proposition 1 and 5 is observed.*

In particular, in the appendix we provide an example with multiple cut-offs where the manager turns from soft to harsh or vice-versa as κ increases.

9 Discussion

The quality of a manager's information determines the efficiency of the decisions he makes. Unfit candidates can prevent the manager from learning their abilities by exerting costly effort. More undermining of the manager's information environment by unfit candidates leads to less accurate decision making. Standards accomplish the dual job of maximizing efficiency (i.e., minimizing wrong decisions given the evidence) and simultaneously deterring effort by unfit candidates. Therefore, this paper highlights a trade-off: the manager gives up ex-post efficiency for better information ex-ante. The benefits from inducing a better quality of information distort the evidentiary standard below (above) its ex-post efficient level when the unfit candidates' efforts are strategic complements (substitutes) of the standard. In turn, strategic complementarity (substitutability) develops in the region of the domain of the induced signal distribution that exhibits strict submodularity (supermodularity). As a consequence, the manager always sets confirmative standards when his beliefs about the candidate's qualification are relatively extreme and often conservative standards when his beliefs are moderate.

Our result of mixed standard distortions, contingent on prior beliefs about candidates, may seem somewhat controversial. The commitment to a standard may not need to be explicit and it may be built on reputation. However, fairness considerations may arise: our results suggest that fit candidates with exogenously determined priors may be subject to harsher standards than those that can be justified based on ex-post efficiency arguments. Our model does not consider fairness; it instead focuses on ex-ante optimality, in anticipation of the incentives of standards over the hidden activities undertaken by unfit candidates to undermine hard evidence. Future work that addresses fairness issues is left for future research.

There is a myriad of real world problems in which agents devote a large amount of resources to influence the choice made by a committee. In his motivation example for conservatism in decisions by committees, Li (2001) observes that "most of the evidence concerning effectiveness of a new drug is provided by its producer, not by the panelists." More broadly, information undermining by interested parties is ubiquitous. In the context of decisions made by committees, our main results suggest that the use of conservative decision rules does not always benefit a committee of multiple agents. The committee must balance the incentives that the standard provides to both sides in generating more reliable evidence. Although conservatism is beneficial in inducing its members to gather more precise evidence, it may fail in deterring unfit candidates from mimicking fit candidates. When the latter effect dominates the former, conservatism reduces the ex-ante welfare of the committee. Our analysis suggests that confirmative standards might be optimal in those scenarios. An interesting extension to this paper could study how standards are set in presence of both free riding by committee members and information undermining by interested parties.

Our model does not address competition among candidates for a given position. In practice, managers choose from a pool of candidates using a recruitment process of several rounds. Competition can be introduced in our setting by running up a tournament among the candidates. The purpose of a standard is to help in selecting further the pool of candidates. The candidate who leads to the highest posterior belief given the evidence among the selected candidates wins the tournament. In this sequential game, the same standard distortion is applied to the entire pool of candidates. Hence, the optimal distortion is no longer contingent on the individual prior belief about each candidate's qualification. The competition between the selected candidates for the offered position tends to distort the candidates' levels of effort upward. Hence, dissuading information undermining is a priori more difficult (in the sense of a greater compromise by the decision maker) in this environment, making inferences weaker. But if there are several rounds associated with the decision process, then several pieces of hard evidence will be available for each candidate. The provision of more than one piece of hard evidence may cause the decision maker set softer standards. In turn, this softness leads to less information undermining by the pool of candidates. In future work we plan to study the interaction between these two opposite effects on information undermining. Competition could alternatively be introduced by modelling search efforts by the manager. In this case, the relative weights given to wrongful hiring and wrongful rejection errors could be endogenized.

Finally, Skaperdas and Vaidya (2012) model competition within a court setting. The plaintiff and the defendant compete to gather costly evidence so as to influence the verdict of the court. In contrast to our work, the incentives of the parties attempting to influence the decision maker (the judge) are assumed to be outside his control. Exploring the theoretical implications of relaxing this assumption in a court setting would also be a natural next step.

Appendix: Omitted Results and Proofs

This appendix contains the non-trivial proofs omitted from the text.

Proofs of Section 2

Proof of Remark 1. By Assumptions F.1 and F.3, for all $z \in (0,1)$ and $\theta < \overline{\theta}$,

$$0 > h(z,\theta) = \int_0^z \frac{\partial h(z',\theta)}{\partial z} dz',$$

where the equality follows from the fact that $h(0,\theta) = 0$ for all $\theta \in \Theta$. Thus there is $z' \in (0,z)$ such that $\frac{\partial h(z',\theta)}{\partial z} < 0$; and similarly, there is $z'' \in (z,1)$ such that $\frac{\partial h(z'',\theta)}{\partial z} > 0$. Therefore, $m(z',\theta) < 0$ and $m(z'',\theta) > 0$, and by continuity of m, there is $z''' \in (z',z'')$ such that $m(z''',\theta) = 0$. Indeed, z''' is the unique root because Assumption F.4 implies that $\frac{\partial^2 f(z''',\theta)}{\partial\theta\partial z} > 0$, yielding $\frac{\partial m(z''',\theta)}{\partial z} > 0$ and hence, $m(\cdot,\theta)$ is strictly increasing on an open interval containing z'''. Furthermore, Assumption F.2 and Proposition 5 in Milgrom (1981) imply that $m(\cdot,\theta)$ is increasing for all $\theta < \overline{\theta}$. Thus, $\widetilde{s} : \left(\underline{\theta}, \overline{\theta}\right) \to (0,1)$ maps θ to the unique root of $m(\cdot,\theta)$, for all $\theta < \overline{\theta}$. □

Proof of Remark 2. The Implicit Function Theorem, Young's Theorem, (5), and (6) yield

$$\frac{\partial s^*(\theta,\kappa)}{\partial\theta} = \frac{g(s^*_{\theta,\kappa},\theta)}{f(s^*_{\theta,\kappa},\theta)} \left(\frac{\partial g(s^*_{\theta,\kappa},\theta)}{\partial s}\right)^{-1} \frac{\partial f(s^*_{\theta,\kappa},\theta)}{\partial\theta}$$

for all $(\theta,\kappa) \in \Theta \times (g(0,\underline{\theta}), g(1,\underline{\theta}))$. The result is a direct consequence of this equality. □

Proof of Lemma 1. By the Implicit Function Theorem, $\widetilde{s}$ is continuous over Θ°, with

$$\frac{d\widetilde{s}(\theta)}{d\theta} = -\frac{\partial m(\widetilde{s}(\theta),\theta)}{\partial\theta} \left(\frac{\partial m(\widetilde{s}(\theta),\theta)}{\partial s}\right)^{-1}$$

for all $\theta \in \Theta^\circ$. In particular, $\frac{\partial m(\widetilde{s}(\theta),\theta)}{\partial s} > 0$ and hence $\frac{d\widetilde{s}(\theta)}{d\theta}$ is finite for all $\theta \in \Theta^\circ$, by Assumptions F.1 and F.4.

Let $\hat{s}$ be the global maximizer θ^*. From Assumption C.1, $\hat{s} \in (0,1)$, and $\theta^*(\hat{s}) > \underline{\theta}$. Further, $\frac{d\theta^*(\hat{s})}{ds} = 0$ and hence, $\widetilde{s}(\theta^*(\hat{s})) = \hat{s}$. Notice that the uniqueness of this maximizer follows from the fact that any other hypothetical maximizer $\hat{s}' \neq \hat{s}$ would also need to satisfy $\widetilde{s}(\theta^*(\hat{s})) = \hat{s}'$, contradicting that $\widetilde{s}$ is a function.

Suppose there exists $s' \neq \hat{s}$ such that $\widetilde{s}(\theta^*(s')) = s'$. If s' is not a local extreme of θ^*, then s' is a tangency point between θ^* and the inverse of $\widetilde{s}$.[29] But this would imply $0 = d\theta^*(s')/ds = (d\widetilde{s}(\theta^*(s'))/d\theta)^{-1}$, which leads to a contradiction because $\frac{d\widetilde{s}(\theta)}{d\theta}$ is finite for all $\theta \in \Theta^\circ$. Furthermore, s' cannot be a local minimum of θ^* as this would imply that for some $\theta > \theta^*(s')$, there are $s'' < s' < s'''$ with $\theta^*(s'') = \theta^*(s''') = \theta$ and such that $d\theta^*(s'')/ds < 0$ and $d\theta^*(s''')/ds > 0$, implying $\widetilde{s}(\theta) < s'' < s''' < \widetilde{s}(\theta)$, a contradiction. Therefore s' can only be a local maximum of θ^*. But this would imply that there is a local minimum of θ^* in the interval $(\min\{s', \hat{s}\}, \max\{s', \hat{s}\})$, contradicting that θ^* does not have local minima in $(0, 1)$. We conclude that θ^* intersects $\widetilde{s}$ only once at $\hat{s}$ and hence, the thesis of the lemma follows immediately. □

Nash equilibrium in the static game. The analysis of the Nash equilibrium of the static game Γ_1 reveals the robustness of the typical comparative statics from the non-strategic setting (when unfit candidates cannot alter their signal distribution) to strategic interaction in the static game. The magnitude of these effects, however, is curtailed by strategic behavior.

Proposition 9 *Assume F.1-F.3 and C.1. The game Γ_1 has a unique Nash equilibrium in pure strategies. The equilibrium standard is weakly increasing in κ. For all equilibrium $(s_{NE}, \theta_{NE}) \in D^\circ$, the magnitude of this effect is smaller than in a non-strategic problem where fit and unfit candidates have fixed signal distributions $F(\cdot, \overline{\theta})$ and $F(\cdot, \theta_{NE})$, respectively.*

Proof. **Existence.** Both $U(\cdot, \cdot)$ and $V(\cdot, \cdot)$ are continuous functions. The MRLP guarantees that $-V(\cdot, \theta)$ is quasiconcave. Furthermore, since $h(s, \cdot)$ is increasing and $C''(\theta) > 0$ for all $\theta \in \Theta^\circ$, we have that $-U(s, \cdot)$ is quasiconcave. Therefore, Γ_1 has at least one Nash equilibrium in pure strategies (see, e.g., Theorem 2.2 in Reny (2005)).

Uniqueness. Uniqueness follows from (12) and the fact that $g(\cdot, \theta^*(\cdot))$ is strictly increasing.

Comparative statics. Since $g(\cdot, \theta^*(\cdot))$ is strictly increasing, the manager's equilibrium standard is weakly increasing in κ.

Magnitudes. From (12), in an equilibrium $(s_{NE}, \theta_{NE}) \in D^\circ$, the rate of change of the equilibrium standard with respect to κ is $\left(\frac{dg(s_{NE}, \theta^*(s_{NE}))}{ds}\right)^{-1}$. In contrast, in the non-strategic problem, the rate of change of the optimal standard with respect to κ is $\left(\frac{\partial g(s_{NE}, \theta^*(s_{NE}))}{\partial s}\right)^{-1}$. From (11), the former rate of change is smaller. ■

We briefly discuss why the equilibrium standard is less sensitive to changes in κ in the strategic setting than in purely statistic problems. The effect of the change in s (induced by a change in κ) on the likelihood ratio function is reinforced by the change in unfit candidates' effort, regardless of whether the equilibrium is located in a submodular or supermodular region of F. In other words, the strategic substitutability (complementarity)

[29]The inverse of $\widetilde{s}$ then could be defined over an open interval containing s' because the tangency occuring under the working hypothesis would imply that $d\widetilde{s}(\theta^*(s'))/d\theta \neq 0$.

of the unfit candidate's effort with respect to the manager's standard buffers the effect of κ on the manager's standard but it does not reverse its direction.

Proofs of Section 3

Remark 9 *Assume F.1, F.3, and C.1. The game Γ_2 has at least one subgame perfect equilibrium in pure strategies.*

Proof. Since $U(s,\cdot)$ is continuous and strictly convex on Θ for each $s \in [0,1]$, we have a unique effort $\theta^*(s) \in \Theta$ such that $U(s,\theta^*(s)) < U(s,\theta)$ for all $\theta \in \Theta$. The Implicit Function Theorem and (8) yield the continuity of θ^*. It follows that $V(\cdot,\theta^*(\cdot))$ is continuous over $[0,1]$. Thus, the Weirestrass Theorem yields the result. ■

Proof of Lemma 2. Since s^*_{SP} is an interior minimizer of $V(\cdot,\theta^*(\cdot))$, from (15), it satisfies

$$g(s^*_{SP},\theta^*(s^*_{SP})) = \kappa\left(1+\frac{h(s^*_{SP},\theta^*(s^*_{SP}))}{f(s^*_{SP},\theta^*(s^*_{SP}))}\frac{d\theta^*(s^*_{SP})}{ds}\right). \tag{29}$$

F.1-F.3 imply that $h(s,\theta)$ is strictly negative for all $(s,\theta) \in D^\circ$. Thus the sign of the second term on the right-hand side of (29) is the same as the sign of $-\frac{d\theta^*(s^*_{SP})}{ds}$, which in turn, from condition (9), is the same as the sign of $\frac{\partial f(s^*_{SP},\theta^*(s^*_{SP}))}{\partial\theta}$. Thus, $g(s^*_{SP},\theta^*(s^*_{SP})) < \kappa$ if and only if F is strictly submodular at $(s^*_{SP},\theta^*(s^*_{SP}))$. Furthermore, since $g(\cdot,\theta^*(s^*_{SP}))$ is strictly increasing, we have that either $g(s,\theta^*(s^*_{SP})) = \kappa$ for some $s > s^*_{SP}$, or the minimizer of $V(\cdot,\theta^*(s^*_{SP}))$ is $s = 1$. Thus, $s^*_{SP} < s^*(\theta^*(s^*_{SP}),\kappa)$ if and only if F is strictly submodular at $(s^*_{SP},\theta^*(s^*_{SP}))$.

The argument for the case in which F is strictly supermodular at $(s^*_{SP},\theta^*(s^*_{SP}))$ is analogous. □

The next remark, stating that the endpoints of the pseudo likelihood ratio function are the same as those of the likelihood function in the classical statistical problem, will be useful to establish that the problems that are test-worthy in the classical statistical problem are the same as those that are test-worthy under pure information undermining.

Remark 10 *Assume F.1-F.3 and C.1. Then,*

$$\lim_{s\to 0} v(s) = \underline{\kappa} \quad \textit{and} \quad \lim_{s\to 1} v(s) = \overline{\kappa}.$$

Proof. Since $F(s,\theta) > 0$ for all $s > 0$, there exists $\delta > 0$ such that, for all $s \in (0,\delta)$,

$$v(s) > g(s,\theta^*(s)) \geq \frac{F(s,\overline{\theta})}{F(s,\theta^*(s))}, \tag{30}$$

where the weak inequality follows from the fact that MLRP implies reverse hazard order stochastic dominance. On the other hand, by L'Hôpital:

$$\lim_{s\to 0} \frac{F(s,\overline{\theta})}{F(s,\theta^*(s))} = \lim_{s\to 0} v(s). \tag{31}$$

Statements (30) and (31) imply $\lim_{s\to 0} v(s) = g(0,\underline{\theta})$.

The argument proving $\lim_{s\to 1} v(s) = g(1,\underline{\theta})$ is analogous, using $\frac{1-F(s,\overline{\theta})}{1-F(s,\theta^*(s))}$ instead of $\frac{F(s,\overline{\theta})}{F(s,\theta^*(s))}$ and the hazard order instead of the reverse hazard order and hence, it is omitted. ■

Proofs of Section 4

Proof of Lemma 3. We prove that $\mathcal{S}^*$ is weakly increasing using an indirect argument. Consider $\kappa' > \kappa$, $s \in \mathcal{S}^*(\kappa)$, and $s' \in \mathcal{S}^*(\kappa')$. Then,

$$\begin{aligned} F(s,\overline{\theta}) - \kappa F(s,\theta^*(s)) &\leq F(s',\overline{\theta}) - \kappa F(s',\theta^*(s')) \\ F(s',\overline{\theta}) - \kappa' F(s',\theta^*(s')) &\leq F(s,\overline{\theta}) - \kappa' F(s,\theta^*(s)). \end{aligned}$$

Adding these inequalities yields

$$(\kappa' - \kappa)\left(F(s,\theta^*(s)) - F(s',\theta^*(s'))\right) \leq 0,$$

which implies $F(s,\theta^*(s)) \leq F(s',\theta^*(s'))$. Now suppose $s' < s$; since densities are strictly positive, we have $F(s',\overline{\theta}) < F(s,\overline{\theta})$. Thus,

$$F(s',\overline{\theta}) - \kappa F(s',\theta^*(s')) < F(s,\overline{\theta}) - \kappa F(s,\theta^*(s)),$$

contradicting that $s \in \mathcal{S}^*(\kappa)$.

Now we show that $\mathcal{S}^*(\underline{\kappa}) = \{0\}$. Notice that

$$V(0,\underline{\theta};\underline{\kappa}) < V(s,\underline{\theta};\underline{\kappa}) \leq V(s,\theta^*(s);\underline{\kappa}),$$

for all $s > 0$, where the strict inequality follows from the fact that $s = 0$ is the unique best response of the manager to the effort $\theta = \underline{\theta}$ (by the strict MLRP). The weak inequality follows from the MLRP.

Since $\mathcal{S}^*$ is weakly increasing, we conclude $\mathcal{S}^*(\kappa) = \{0\}$ for all $\kappa \in (0,\underline{\kappa}]$. An analogous argument proves that $\mathcal{S}^*(\kappa) = \{1\}$ for all $\kappa \in [\overline{\kappa},\infty)$.

Now we prove that $\mathcal{S}^*(\kappa) \subset (0,1)$ for all $\kappa \in (\underline{\kappa},\overline{\kappa})$: pick any $\kappa \in (\underline{\kappa},\overline{\kappa})$; since $\lim_{s\to 0} \frac{dV(s,\theta^*(s))}{ds} =$

$\lim_{s\to 0} f(0,\overline{\theta}) - \kappa \lim_{s\to 0} f(0,\underline{\theta}) < 0$,[30] we can find an $s \in (0,1)$ such that $\frac{dV(s',\theta^*(s'))}{ds} < 0$ for all $s' \in (0,s)$ (see (15)). Thus $V(s,\theta^*(s)) < V(0,\underline{\theta})$, and hence $0 \notin \mathcal{S}^*(\kappa)$. An analogous argument shows that $1 \notin \mathcal{S}^*(\kappa)$ for arbitrary $\kappa \in (\underline{\kappa},\overline{\kappa})$.

Finally, we show that $\mathcal{S}^*$ is strictly increasing over $(\underline{\kappa},\overline{\kappa})$. Notice that for all $\kappa \in (\underline{\kappa},\overline{\kappa})$, $s \in \mathcal{S}^*(\kappa)$ only if the right hand side of (15) is equal to 0. Thus, s can only be an element of $\mathcal{S}^*(\kappa)$ for only one $\kappa \in (\underline{\kappa},\overline{\kappa})$. □

Proof of Proposition 1. For any game Γ_2, let

$$\widetilde{\kappa} := \sup\{\kappa \in (0,\infty) : \sup \mathcal{S}^*(\kappa) < \hat{s}\},$$

where $\hat{s}$ is the modularity-switch point.

Part 1. We first prove that $\widetilde{\kappa} < \overline{\kappa}$:

Case 1. $\overline{\kappa} = \infty$: For any κ,

$$\min_{s\in[0,\hat{s}]} \{V(1,\underline{\theta};\kappa) - V(s,\theta^*(s);\kappa)\} = \min_{s\in[0,\hat{s}]} \left\{1 - F(s,\overline{\theta}) - \kappa(1 - F(s,\theta^*(s)))\right\}.$$

This expression is negative for a large enough κ°, thus $s \notin \mathcal{S}^*(\kappa^\circ)$ for all $s \in [0,\hat{s}]$. Since $\mathcal{S}^*$ is weakly increasing (Lemma 3), $\mathcal{S}^*(\kappa') \cap [0,\hat{s}] = \emptyset$ for all $\kappa' > \kappa^\circ$. Thus, $\widetilde{\kappa} \le \kappa^\circ < \infty$.

Case 2. $\overline{\kappa} < \infty$: Recall that $\mathcal{S}^*(\overline{\kappa}) = \{1\}$. In particular, $V(1,\underline{\theta};\overline{\kappa}) < \min_{s\in[0,\hat{s}]} V(s,\theta^*(s);\overline{\kappa})$. Notice that $V(s,\theta;\cdot)$ is continuous over $(0,\infty)$, for all $(s,\theta) \in [0,1]\times\Theta$, and, by the Maximum Theorem, so it is $\min_{s\in[0,\hat{s}]} V(s,\theta^*(s);\cdot)$. Thus, for small enough $\delta > 0$, we have $V(1,\underline{\theta};\overline{\kappa}-\delta) < \min_{s\in[0,\hat{s}]} V(s,\theta^*(s);\overline{\kappa}-\delta)$. Therefore, $\sup \mathcal{S}^*(\overline{\kappa}-\delta) > \hat{s}$ and since $\mathcal{S}^*$ is weakly increasing (Lemma 3), we conclude that $\widetilde{\kappa} \le \overline{\kappa}-\delta < \overline{\kappa}$.

Part 2. Now we show that the manager is ex-post efficient for all $\kappa \in [\overline{\kappa},\infty)$: From Lemma 3, $\mathcal{S}^*(\kappa) = \{1\}$ for all $\kappa \ge \overline{\kappa}$. Recall that $\theta^*(1) = \underline{\theta}$. From (5), $s^*(\underline{\theta},\kappa) = 1$ for all $\kappa \ge \overline{\kappa}$. Thus, the manager is ex-post efficient in equilibrium for all $\kappa \ge \overline{\kappa}$.

Part 3. Now we show that the manager is harsh for all $\kappa \in (\widetilde{\kappa},\overline{\kappa})$: consider any $\kappa \in (\widetilde{\kappa},\overline{\kappa})$ and $s \in \mathcal{S}^*(\kappa)$; since $\mathcal{S}^*$ is strictly increasing over $(\underline{\kappa},\overline{\kappa})$ (see Lemma 3), we have that $s \in (\hat{s},1)$. Lemmas 1 and 2 imply that the manager is harsh if $s \in (\hat{s},1)$. And if $s = 1$ the manager is harsh because $s^*(\underline{\theta},\kappa) < 1$ for $\kappa < \overline{\kappa}$.

Noting that $\widetilde{\kappa} = \inf\{\kappa \in (0,\infty) : \inf \mathcal{S}^*(\kappa) > \hat{s}\}$, an argument analogous to that of Part 1 shows that $\widetilde{\kappa} > \underline{\kappa}$. Similarly, arguments analogous to those in Parts 2 and 3 yield that the manager is ex-post efficient for all $\kappa \in (0,\underline{\kappa}]$ and soft for all $\kappa \in (\underline{\kappa},\widetilde{\kappa})$, respectively. □

[30] Remark 10 implies that either $\lim_{s\to 0} f(s,\overline{\theta}) = 0$ or $\lim_{s\to 0} h(s,\theta^*(s)) \frac{\partial f(s,\theta^*(s))}{d\theta} \frac{1}{f(s,\theta^*(s))} \left(C''(\theta^*(s)) + \frac{\partial h(s,\theta^*(s))}{\partial s}\right)^{-1} = 0$; either way yields that $\lim_{s\to 0} \frac{dV(s,\theta^*(s))}{ds} < 0$.

Proofs of Section 5

Proof of Lemma 4. If F does not have a neutral signal, then $\frac{d\tilde{s}(\theta)}{d\theta} \neq 0$ at some $\theta \in \Theta^\circ$. Consider $\theta' \in \Theta^\circ$ such that $\kappa' := g(\tilde{s}(\theta'), \theta') \neq 1$.[31] We construct the level curve of $V(\cdot, \cdot; \kappa')$ attaining $V(\tilde{s}(\theta'), \theta'; \kappa')$, denoted by LC, implicitly defining a function $s \mapsto \theta_{LC}(s)$ by

$$F(s, \overline{\theta}) - \kappa' F(s, \theta_{LC}(s)) = F(\tilde{s}(\theta'), \overline{\theta}) - \kappa' F(\tilde{s}(\theta'), \theta'),$$

with domain $D_{LC} \subseteq [0,1]$ and slope

$$\frac{d\theta_{LC}(s)}{ds} = \frac{f(s, \overline{\theta}) - \kappa' f(s, \theta_{LC}(s))}{\kappa' h(s, \theta_{LC}(s))}. \tag{32}$$

for all $s \in D_{LC}$. We observe the following:

(i) $V(\tilde{s}(\theta'), \theta'; \kappa') < V(1, \underline{\theta}; \kappa')$, because the following statements are equivalent

$$\begin{aligned} F(\tilde{s}(\theta'), \overline{\theta}) - \kappa' F(\tilde{s}(\theta'), \theta') &< 1 - \kappa' \\ \kappa' &< \frac{1 - F(\tilde{s}(\theta'), \overline{\theta})}{1 - F(\tilde{s}(\theta'), \theta')} \\ g(\tilde{s}(\theta'), \theta') &< \frac{1 - F(\tilde{s}(\theta'), \overline{\theta})}{1 - F(\tilde{s}(\theta'), \theta')}, \end{aligned}$$

and the last statement holds because $F(\cdot, \overline{\theta})$ strictly stochastically dominates $F(\cdot, \theta')$ according to the hazard rate order. An analogous argument, but using instead the fact that $F(\cdot, \overline{\theta})$ strictly stochastically dominates $F(\cdot, \theta')$ according to the reverse hazard rate order, reveals that $V(\tilde{s}(\theta'), \theta'; \kappa') < V(0, \underline{\theta}; \kappa')$. Thus, $(0, \theta)$ and $(1, \theta)$ are not in the level curve defined above for all $\theta \in \Theta$, and hence, $D_{LC} \subset (0, 1)$.

(ii) Since

$$V(\tilde{s}(\theta'), \underline{\theta}; \kappa') < V(\tilde{s}(\theta'), \theta'; \kappa') < V(0, \underline{\theta}; \kappa'),$$

there exists $s_1 \in (0, \tilde{s}(\theta'))$ such that $(s_1, \underline{\theta}) \in LC$. It follows that $\frac{d\theta_{LC}(s)}{ds} \geq 0$ for all $s \in (s_1, \tilde{s}(\theta'))$: if not, then θ_{LC} would be strictly decreasing over some interval (s_2, s_3), with $s_1 < s_2 < s_3 < \tilde{s}(\theta')$, which would imply that for all $s \in (s_2, s_3)$ there would be $s' \in (s_1, s_2)$ and $s'' \in (s_3, \tilde{s}(\theta'))$ such that $(s, \theta_{LC}(s)), (s', \theta_{LC}(s)), (s'', \theta_{LC}(s)) \in LC$, contradicting the quasiconvexity of $V(\cdot, \theta_{LC}(s); \kappa')$.[32] Furthermore, from (32) and the strict MLRP, we cannot have $\frac{d\theta_{LC}(\cdot)}{ds} = 0$ over any interval (s_4, s_5) with $s_1 < s_4 < s_5 < \tilde{s}(\theta')$. Thus, $\theta_{LC}(\cdot)$ is strictly increasing over $(s_1, \tilde{s}(\theta'))$.

[31] Note that $\frac{dg(\tilde{s}(\theta), \theta)}{d\theta} > (<) 0$ if and only if $\frac{d\tilde{s}(\theta)}{d\theta} > (<) 0$.

[32] The quasiconvexity of $V(\cdot, \theta_{LC}(s); \kappa')$ implies that this function attains $V(\tilde{s}(\theta'), \theta'; \kappa')$ at at most two different standards.

(iii) Analogously, we can show that there exists $s_6 \in (\tilde{s}(\theta'), 1)$ such that $(s_6, \underline{\theta}) \in LC$ and $\theta_{LC}(\cdot)$ is strictly decreasing over $(\tilde{s}(\theta'), s_6)$ for some $s_6 \in (\tilde{s}(\theta'), 1)$. The quasiconvexity and continuity of $V(\cdot, \underline{\theta}; \kappa')$ imply that $V(\cdot, \underline{\theta}; \kappa') > V(\tilde{s}(\theta'), \theta'; \kappa')$ over $[0, s_1) \cup (s_6, 1]$ and hence, $D_{LC} = [s_1, s_6]$.

From (i)-(iii), we can construct a cost function C_F satisfying Assumption C.1 and such that $\theta^*(s) \geq \theta_{LC}(s)$ for all $s \in D_{LC}$ with equality only at $s = \tilde{s}(\theta')$, and hence $\mathcal{S}^*(\kappa') = \{\tilde{s}(\theta')\}$.[33] Thus, in a game Γ_2 with distribution F and cost function C_F, we have $\hat{s} = \tilde{s}(\theta')$ and $\widetilde{\kappa} = \kappa' \neq 1$; and hence, by Corollary 1, the manager is conservative at $\kappa \in (\min\{1, \kappa'\}, \max\{1, \kappa'\})$, contradicting uniform confirmativism. □

Claim 2 *Assume F.1-F.4 and C.1. If F has the FMP, then*

(i) F has a neutral signal $s^\star$,

(ii) $\theta^*(s) = \theta^*(s_f(s))$ for all $s \in (0,1)$ and all cost function satisfying Assumption C.1,

(iii) $F(s,\theta) - F(s_f(s), \theta) = F(s, \theta') - F(s_f(s), \theta')$ for all $\theta, \theta' \in \Theta^\circ$ and for all $s \in (0,1)$, and

(iv) we have

$$V(s, \theta^*(s); 1) = V(s_f(s), \theta^*(s_f(s)); 1)$$

for all $s \in [0,1]$ and all cost function satisfying Assumption C.1.

Proof. (i) Consider a distribution F that satisfies FMP, and the working hypothesis: $\widetilde{s}(\theta) \neq \widetilde{s}(\theta')$ for some $\theta, \theta' \in \Theta^\circ$. Since $\widetilde{s}(\theta)$ is the only fixed point of $s_f(\cdot, \theta)$, there exists $s' \neq \widetilde{s}(\theta')$ such that

$$h(\widetilde{s}(\theta'), \theta) = h(s', \theta).$$

But then FMP implies that $h(\widetilde{s}(\theta'), \theta') = h(s', \theta')$, contradicting that $\widetilde{s}(\theta')$ is the only fixed point of $s_f(\cdot, \theta')$. Thus, $\widetilde{s}$ is constant and hence F has a neutral signal.

(ii) The second statement is immediate from (8).

(iii) The statement follows from

$$F(s,\theta) - F(s_f(s), \theta) = \int_{\underline{\theta}}^{\theta} h(s,t)dt + F(s, \underline{\theta}) - \int_{\underline{\theta}}^{\theta} h(s_f(s), t)dt - F(s_f(s), \underline{\theta}) = F(s, \underline{\theta}) - F(s_f(s), \underline{\theta})$$

for all $s \in [0,1]$ and $\theta \in \Theta$.

(iv) Follows directly from (ii) and (iii), setting $\theta = \overline{\theta}$ and $\theta' = \theta^*(s)$. ■

[33] Defining such a cost function requires $C_F'(\theta_{LC}(s)) \leq -h(s, \theta_{LC}(s))$ for all $s \in D_{LC}$ with equality only at $s = \tilde{s}(\theta')$. For instance, $C_F(\theta) = p_0(\theta - \underline{\theta})^{p_1} + p_2(\theta - \underline{\theta})^2$, for all $\theta \in \Theta$, with an appropriate choice of $p_0, p_2 > 0$ and $p_1 > 2$, achieves this.

Proof of Lemma 5. Suppose F satisfies the FMP. From Claim 2, $V(s, \theta^*(s); 1) = V(s_f(s), \theta^*(s_f(s)); 1)$ for all $s \in [0, 1]$ and all cost function. That is, each fixed-matched pair of standards yields the same expected loss to the manager at $\kappa = 1$ for every cost function satisfying Assumption C.1. Thus, $\tilde{\kappa} = 1$ and hence the manager is uniformly confirmative by Corollary 2. $\square$

Proof of Proposition 2. *Necessity.* Consider F with a neutral signal $s^\star$. Then $g(s^\star, \theta) = 1$ for all $\theta \in \Theta$ and $v(s^\star) = 1$ for every cost function. Suppose that there exists a pair (s', θ') with $s' \in (0, s^\star)$ and $\theta \in \Theta^\circ$ such that $F(s', \overline{\theta}) - F(s', \theta') < F(s_f(s', \theta'), \overline{\theta}) - F(s_f(s', \theta'), \theta')$ and

$$f(s', \overline{\theta}) > f(s', \theta') - h(s', \theta') \frac{\partial f(s', \theta')}{\partial \theta} \left(\frac{\partial h(s', \theta')}{\partial \theta} \right)^{-1}. \tag{33}$$

We show that we can find a cost function satisfying Assumption C.1 such that $\mathcal{S}^*(1) = \{s'\}$. Similarly to the argument in the proof of Lemma 4, we construct the level curve of $V(\cdot, \cdot; 1)$, attaining $V(s', \theta'; 1)$. This level curve defines a function $s \mapsto \theta_{LC_2}(s)$ over a set $D_{LC_2} \subset [0, 1]$ mapping s to the effort $\theta_{LC_2}(s)$ that yields the same expected loss to the manager as (s', θ'). We observe the following:

(i) since $V(0, \underline{\theta}; 1) = V(1, \underline{\theta}; 1) = 0 > V(s', \theta'; 1)$, we have that $(0, \theta)$ and $(1, \theta)$ are not in the level curve for all $\theta \in \Theta$ and therefore $D_{LC_2} \subset (0, 1)$.

(ii) the slope of $\theta_{LC_2}(\cdot)$ is given by

$$\frac{d\theta_{LC_2}(s)}{ds} = \frac{f(s, \overline{\theta}) - f(s, \theta_{LC_2}(s))}{h(s, \theta_{LC_2}(s))}$$

and hence, $d\theta_{LC_2}(\cdot)/ds > (<)0$ at $(s, \theta_{LC_2}(s))$ for all $s \in D_{LC_2}$ such that $s < (>)s^\star$; and

In order to construct a cost function C_{F_2} satisfying Assumption C.1 such that $\mathcal{S}^*(1) = \{s''\}$, C_{F_2} has to yield that $\theta^*(s) \geq \theta_{LC_2}(s)$ for all $s \in D_{LC_2}$ with equality only at $s = s'$. That is, C_{F_2} must satisfy $C'_{F_2}(\theta_{LC_2}(s)) \leq -h(s, \theta_{LC_2}(s))$ for all $s \in D_{LC_2}$ with equality only at $s = s'$. The level curve $\theta_{LC_2}(s)$ and $\theta^*(s)$ have to be tangent at (s', θ''), i.e., $\frac{d\theta_{LC_2}(s')}{ds} = \frac{d\theta^*(s')}{ds}$:

$$\frac{f(s', \overline{\theta}) - f(s', \theta')}{h(s', \theta')} = -\frac{\partial f(s', \theta')}{\partial \theta} \left(C''_{F_2}(\theta') + \frac{\partial h(s', \theta')}{\partial \theta} \right)^{-1},$$

which can be satisfied for a strictly positive $C''_{F_2}(\theta')$ if and only if (33) holds. Thus, with C_{F_2} we have that $\mathcal{S}^*(1) = \{s'\}$, and therefore, since $\mathcal{S}^*$ is strictly increasing over $(\underline{\kappa}, \overline{\kappa})$, $\widetilde{\kappa} > 1$. By Corollary 2, the manager is not uniformly confirmative.

An analogous argument shows that the manager cannot be uniformly confirmative if (20) is not satisfied for some $(s', \theta') \in (s^\star, 1) \times \Theta^\circ$.

Sufficiency. If (19) implies (20), then (18) holds for all critical point of $V(\cdot, \theta^*(\cdot); 1)$ for all cost function satisfying Assumption C.1, and the argument in the proof of Lemma 5 implies that the manager is uniformly confirmative for all such cost function. □

Claim 3 *Assume F.2-F.3. Any family of distributions F such that $\frac{\partial f(z,\theta)}{\partial \theta} = \alpha(z)\beta(\theta)$ for all $z \in (0,1)$ and $\theta \in \Theta^\circ$, for some real functions $\alpha : (0,1) \to \mathbb{R}$ and $\beta : \Theta^\circ \to \mathbb{R}$, satisfies the FMP.*

Proof. Let $a(s) := \int_0^s \alpha(z)dz$ for all $s \in [0,1]$. From the hypothesis,

$$h(s,\theta) = \int_0^s \frac{\partial h(z,\theta)}{\partial z} dz = \beta(\theta) \int_0^s \alpha(z)dz = \beta(\theta)a(s),$$

for all $s \in (0,1)$ and $\theta \in \Theta^\circ$. Since Assumptions F.2-F.3 imply $h(s,\theta) < 0$ for all $s \in (0,1)$ and $\theta \in \Theta^\circ$, $\beta(\theta) \neq 0$ for all $\theta \in \Theta^\circ$. Thus, $\beta(\theta)a(s) = \beta(\theta)a(s')$ for some $\theta \in \Theta^\circ$ implies $\beta(\theta')a(s) = \beta(\theta')a(s')$ for all $\theta' \in \Theta^\circ$ and $s, s' \in [0,1]$. ■

Proofs of Section 6

Proof of Remark 5. Let θ^*_λ be the best response of the candidate if the effort cost function is $\lambda^{-1}C$ and

$$D_\lambda(s) := f(s, \theta^*_\lambda(s)) - \frac{h(s, \theta^*_\lambda(s)) \frac{\partial f(s, \theta^*_\lambda(s))}{\partial \theta}}{\lambda^{-1}C''(\theta^*_\lambda(s)) + \frac{\partial h(s, \theta^*_\lambda(s))}{\partial \theta}}$$

for all $s \in [0,1]$. Since $f > 0$, there exists $\overline{\lambda}_1 > 0$ such that for all $\lambda \in (0, \overline{\lambda}_1)$, we have that $\min_{s\in[0,1]} \{D_\lambda(s)\} > 0$. Thus for all $\lambda \in (0, \overline{\lambda}_1)$, $dv(\cdot)/ds > 0$ is equivalent to

$$\min_{s\in[0,1]} \left\{ \frac{1}{f(s,\overline{\theta})} \frac{\partial f(s,\overline{\theta})}{\partial s} - \frac{1}{D_\lambda(s)} \frac{dD_\lambda(s)}{ds} \right\} > 0.$$

Indeed,

$$\lim_{\lambda\to 0} \min_{s\in[0,1]} \left\{ \frac{1}{f(s,\overline{\theta})} \frac{\partial f(s,\overline{\theta})}{\partial s} - \frac{1}{D_\lambda(s)} \frac{dD_\lambda(s)}{ds} \right\} = \min_{s\in[0,1]} \left\{ \frac{1}{f(s,\overline{\theta})} \frac{\partial f(s,\overline{\theta})}{\partial s} - \frac{1}{f(s,\underline{\theta})} \frac{\partial f(s,\underline{\theta})}{\partial s} \right\} > 0,$$

where the inequality is guaranteed by (1). Thus, there exists $\overline{\lambda} > 0$ such that for all $\lambda \in (0, \overline{\lambda})$, we have $dv(\cdot)/ds > 0$. □

Proof of Remark 7. <u>Case 1</u>. Suppose $dF(\cdot, \theta^*(\cdot))/ds > 0$ over $(0,1)$. For all $s \in (\underline{s}, \overline{s})$, if s is a critical point of $V(\cdot, \theta^*(\cdot))$, then s is a local maximum and hence $s \notin \mathcal{S}^*(\kappa)$ for any $\kappa \in (0,\infty)$. Let

$$\kappa^* := \sup\{\kappa \in (0,\infty) : \sup \mathcal{S}^*(\kappa) \leq \underline{s}\} \quad \text{and} \quad \kappa_* := \inf\{\kappa \in (0,\infty) : \inf \mathcal{S}^*(\kappa) \geq \overline{s}\}.$$

We observe that $\kappa^* = \kappa_*$: if $\kappa^* < \kappa_*$, then for all $\kappa \in (\kappa^*, \kappa_*)$ we have that $\mathcal{S}^*(\kappa) \cap (\underline{s}, \overline{s}) \neq \emptyset$, contradicting that $s \in (\underline{s}, \overline{s})$ implies that $s \notin \mathcal{S}^*(\kappa)$ for any $\kappa \in (0, \infty)$. On the other hand, if $\kappa^* > \kappa_*$, then for any $\kappa \in (\kappa_*, \kappa^*)$, $\sup \mathcal{S}^*(\kappa) \leq \underline{s}$ and $\inf \mathcal{S}^*(\kappa) \geq \overline{s}$, a contradiction.

Thus, for all $\kappa' < \kappa^*$ we have $\sup \mathcal{S}^*(\kappa') \leq \underline{s}$, and for all $\kappa'' > \kappa^*$ we have $\inf \mathcal{S}^*(\kappa'') \geq \overline{s}$. Hence the thesis holds for $\kappa = \kappa^*$ and all $\delta \in (0, \overline{s} - \underline{s})$.

Case 2. Suppose $dF(s, \theta^*(s))/ds \leq 0$ for some $s \in (0,1)$. Then, for all $\kappa \in (0, \infty)$, we have $dV(\cdot, \theta^*(\cdot); \kappa)/ds > 0$ over $(s - \varepsilon, s + \varepsilon)$ for some $\varepsilon \in (0, \min\{s, 1-s\})$. Thus, $s \notin \mathcal{S}^*(\kappa)$ for any $\kappa \in (0, \infty)$. Analogously to the argument in Case 1, we can define

$$\kappa'^* := \sup\{\kappa \in (0, \infty) : \sup \mathcal{S}^*(\kappa) \leq s - \varepsilon\} \quad \text{and} \quad \kappa'_* := \inf\{\kappa \in (0, \infty) : \inf \mathcal{S}^*(\kappa) \geq s + \varepsilon\}.$$

The rest of the argument is analogous to Case 1, leading to the statement that thesis holds for $\kappa = \kappa'^*$ and all $\delta \in (0, 2\varepsilon)$. $\square$

Proofs of Section 7

Proof of Proposition 3. First we prove part (i). From Remark 10, $\lim_{s \to 0} v(s) = \underline{\kappa}$. For all $\kappa \in (\underline{\kappa}, \widetilde{\kappa})$, the elements of $\mathcal{S}^*(\kappa)$ are critical points of $V(\cdot, \theta^*(\cdot))$ that are smaller than $\hat{s}$. For all equilibrium (s^*_{SP}, θ^*) such that $s^*_{SP} \in (0, \min_{\theta \in [\underline{\theta}, \theta^*(\hat{s})]} \widetilde{s}(\theta))$, we have $v(s^*_{SP}) > g(s^*_{SP}, \theta^*(s^*_{SP})) > g(s^*_{SP}, \underline{\theta})$. Hence for all $\kappa \in (\underline{\kappa}, \underline{\overline{\kappa}}]$ and equilibrium (s^*_{SP}, θ^*) with $s^*_{SP} \in \mathcal{S}^*(\kappa) \cap (0, \min_{\theta \in [\underline{\theta}, \theta^*(\hat{s})]} \widetilde{s}(\theta))$, we have that the root of $g(\cdot, \underline{\theta}) = \kappa$ is strictly greater than s^*_{SP}. Therefore fit candidates are better off at (s^*_{SP}, θ^*) than in the classical statistical problem, and by the envelope theorem, so are unfit candidates.

Now we prove part (ii). From Remark 10, $\lim_{s \to 1} v(s) = \overline{\kappa}$. Assumption F.5, Assumption C.2 and straightforward computations yield that

$$\lim_{s \to 1} \frac{dv(s)}{ds} > \lim_{s \to 1} \frac{dg(s, \theta^*(s))}{ds} > 0, \tag{34}$$

where the second inequality follows from (11). Hence v is increasing as s goes to 1. It follows that there exists $\underline{\overline{\kappa}} < \overline{\kappa}$ such that $\mathcal{S}^*$ is singled value for all $\kappa > \underline{\overline{\kappa}}$, and we denote its unique element by $s^*_{SP}(\kappa)$.

Let $\mathcal{U}(\kappa)$ and $\underline{\mathcal{U}}(\kappa)$ be the expected loss to unfit candidates in the equilibrium of Γ_2 and in the outcome of the classical statistical problem, respectively, for all $\kappa \in [\underline{\kappa}, \underline{\overline{\kappa}})$, i.e.,

$$\mathcal{U}(\kappa) := C(\theta^*(s^*_{SP}(\kappa))) + F(s^*_{SP}(\kappa), \theta^*(s^*_{SP}(\kappa))) \quad \text{and} \quad \underline{\mathcal{U}}(\kappa) := F(s^*(\underline{\theta}, \kappa), \underline{\theta}). \tag{35}$$

The derivatives of $\mathcal{U}$ and $\underline{\mathcal{U}}$ are

$$\frac{d\mathcal{U}(\kappa)}{d\kappa} = f(s^*_{SP}(\kappa), \theta^*(s^*_{SP}(\kappa)))\frac{ds^*_{SP}(\kappa)}{d\kappa} = f(s^*_{SP}(\kappa), \theta^*(s^*_{SP}(\kappa)))\left(\frac{dv}{ds}(s^*_{SP}(\kappa))\right)^{-1} \tag{36}$$

and

$$\frac{d\underline{\mathcal{U}}(\kappa)}{d\kappa} = f(s^*(\underline{\theta}, \kappa), \underline{\theta})\frac{ds^*(\underline{\theta}, \kappa)}{d\kappa} = f(s^*(\underline{\theta}, \kappa), \underline{\theta})\left(\frac{\partial g(s^*(\underline{\theta}, \kappa), \underline{\theta})}{\partial s}\right)^{-1}, \tag{37}$$

respectively, for all $\kappa \in [\underline{\kappa}, \overline{\kappa})$.

A sufficient condition for the thesis is $\lim_{\kappa\to\overline{\kappa}} \frac{d\mathcal{U}(\kappa)}{d\kappa} < \lim_{\kappa\to\overline{\kappa}} \frac{d\underline{\mathcal{U}}(\kappa)}{d\kappa}$, which, under the assumption $f(1, \underline{\theta}) > 0$, is equivalent to

$$\lim_{\kappa\to\overline{\kappa}} \frac{dv(s^*_{SP}(\kappa))}{ds} > \lim_{\kappa\to\overline{\kappa}} \frac{\partial g(s^*(\underline{\theta}, \kappa), \underline{\theta})}{\partial s}. \tag{38}$$

Since $\lim_{\kappa\to\overline{\kappa}} s^*_{SP}(\kappa) = \lim_{\kappa\to\overline{\kappa}} s^*(\underline{\theta}, \kappa) = 1$, (38) holds if and only if

$$\lim_{s\to1} \frac{dv(s)}{ds} > \lim_{s\to1} \frac{\partial g(s, \underline{\theta})}{\partial s}. \tag{39}$$

In order to see that this is the case, notice that

$$\lim_{s\to1} \frac{dg(s, \theta^*(s))}{ds} > \lim_{s\to1} \frac{\partial g(s, \underline{\theta})}{\partial s}, \tag{40}$$

which follows from (11),

$$\lim_{s\to1} \frac{dg(s, \theta^*(s))}{ds} = \frac{\partial g(1, \underline{\theta})}{\partial s} + \frac{g(1, \underline{\theta})}{f(1, \underline{\theta})}\left(\frac{\partial f(1, \underline{\theta})}{\partial \theta}\right)^2\left(C''(\underline{\theta}) + \frac{\partial h(1, \underline{\theta})}{\partial \theta}\right)^{-1} > \frac{\partial g(1, \underline{\theta})}{\partial s}.$$

The inequality (39) is implied by (34) and (40). □

Proofs of Section 8.1

Proof of Lemma 8. Observe that

$$\begin{aligned}
\frac{d\mathcal{V}(s)}{ds} &= f(s, \theta^*_q(s)) + h(s, \theta^*_q(s))\frac{d\theta^*_q(s)}{ds} - \kappa\left(f(s, \theta^*_u(s)) + h(s, \theta^*_u(s))\frac{d\theta^*_u(s)}{ds}\right) \\
&= f(s, \theta^*_u(s))\left(g(s, \theta^*_q(s), \theta^*_q(s)) - \kappa\right) + h(s, \theta^*_q(s))\frac{d\theta^*_q(s)}{ds} - \kappa h(s, \theta^*_u(s))\frac{d\theta^*_u(s)}{ds} \\
&= f(s, \theta^*_u(s))\left(g(s, \theta^*_q(s), \theta^*_q(s)) - \kappa\right) + h(s, \theta^*_u(s))\frac{d\theta^*_u(s)}{ds}\left(r(s) - \kappa\right)
\end{aligned}$$

for all $s \in (0, 1) \setminus \{\hat{s}\}$. Recall that the manager is soft (harsh) in an equilibrium with standard $s^*_{SP} \in (0, 1)$ if $g(s^*_{SP}, \theta^*_q(s^*_{SP}), \theta^*_q(s^*_{SP})) < (>)\kappa$. Thus, parts (i) and (ii) follow immediately from the second and third equalities, respectively. □

Proofs of Section 8.2

Proof of Proposition 4. In the sequel, when convenient, we note explicitly the dependence of D_q, w, r, and θ_q^* on the initial advantage of fit candidates, and write $D_q(s;\underline{\theta}_q)$, $w(s;\underline{\theta}_q)$, $r(s;\underline{\theta}_q)$ and $\theta_q^*(s;\underline{\theta}_q)$ instead of $D_q(s)$, $w(s)$, $r(s)$, and $\theta_q^*(s)$, respectively, for all $s \in [0,1]$ and $\underline{\theta}_q \in \Theta$.

The proof hinges on the following lemmata:

Lemma 10 *Assume F.2, F.3, F.5, F.6, C.1, C.3, and $f > 0$. Then, there exists $\underline{\underline{\theta}}_q \in \Theta^\circ$ such that for all $\underline{\theta}_q \in \left(\underline{\underline{\theta}}_q, \overline{\theta}\right)$, $D_q(s;\underline{\theta}_q) > 0$ for all $s \in (0,1)$.*

Proof. The Maximum Theorem, Assumption F.5 and Assumption C.3 guarantee that $\min_{s\in[0,1]} D_q(s)$ varies continuously with $\underline{\theta}_q$. Further, $h(s,\overline{\theta}) = 0$ for all $s \in [0,1]$ and, from the hypothesis, $\min_{s\in[0,1]} f(s,\overline{\theta}) > 0$. Thus, there exists $\underline{\underline{\theta}}_q \in \Theta^\circ$ such that $\min_{s\in[0,1]} D_q(s) > 0$ for all $\underline{\theta}_q \in \left(\underline{\underline{\theta}}_q, \overline{\theta}\right)$. ■

Lemma 11 *Assume F.2-F.6, C.1-C.4, and $f > 0$. Then, there exists $\underline{\underline{\theta}}_q \in \Theta^\circ$ such that for all $\underline{\theta}_q \in \left(\underline{\underline{\theta}}_q, \overline{\theta}\right)$, there exists $\hat{s}(\underline{\theta}_q) \in (0,1)$ such that if $(s_{SP}^*, \theta_q^*, \theta_u^*)$ is a subgame perfect equilibrium of Γ_3, then the manager is soft if $s_{SP}^* \in (0, \hat{s}(\underline{\theta}_q))$ and harsh if $s_{SP}^* \in (\hat{s}(\underline{\theta}_q), 1)$.*

Proof. First,

$$\lim_{(s,\underline{\theta}_q)\to(0,\overline{\theta})} \frac{dw(s;\underline{\theta}_q)}{ds} > \lim_{(s,\underline{\theta}_q)\to(0,\overline{\theta})} \frac{dg(s,\theta_q^*(s;\underline{\theta}_q),\theta_u^*(s))}{ds} \tag{41}$$

and

$$\lim_{(s,\underline{\theta}_q)\to(1,\overline{\theta})} \frac{dw(s;\underline{\theta}_q)}{ds} > \lim_{(s,\underline{\theta}_q)\to(1,\overline{\theta})} \frac{dg(s,\theta_q^*(s;\underline{\theta}_q),\theta_u^*(s))}{ds}. \tag{42}$$

Thus, there exists $0 < \delta_1 < \delta_2 < 1$ and $\underline{\underline{\theta}}_{q_1} \in \Theta^\circ$ such that,

$$w(s;\underline{\theta}_q) \begin{cases} > g(s,\theta_q^*(s;\underline{\theta}_q),\theta_u^*(s)) & \text{if } s \in (0,\delta_1) \\ < g(s,\theta_q^*(s;\underline{\theta}_q),\theta_u^*(s)) & \text{if } s \in (\delta_2,1) \end{cases}$$

for all $\underline{\theta}_q \in \left(\underline{\underline{\theta}}_{q_1}, \overline{\theta}\right)$.

Second,

$$\lim_{(s,\underline{\theta}_q)\to(\hat{s}_u,\overline{\theta})} \frac{dw(s;\underline{\theta}_q)}{ds} < \lim_{(s,\underline{\theta}_q)\to(\hat{s}_u,\overline{\theta})} \frac{dg(s,\theta_q^*(s;\underline{\theta}_q),\theta_u^*(s))}{ds}. \tag{43}$$

Thus, there exist $\underline{\underline{\theta}}_{q_2} \in \Theta^\circ$, $\delta_3 \in (\delta_1, \hat{s}_u)$ and $\delta_4 \in (\hat{s}_u, \delta_2)$ such that, for all $\underline{\theta}_q \in \left(\underline{\underline{\theta}}_{q_2}, \overline{\theta}\right)$, there exists

$\hat{s}(\underline{\theta}_q) \in (\delta_3, \delta_4)$ satisfying

$$w(s;\underline{\theta}_q)\begin{cases} > g(s,\theta_q^*(s;\underline{\theta}_q),\theta_u^*(s)) & \text{if } s \in (\delta_3, \hat{s}(\underline{\theta}_q)) \\ = g(s,\theta_q^*(s;\underline{\theta}_q),\theta_u^*(s)) & \text{if } s = \hat{s}(\underline{\theta}_q) \\ < g(s,\theta_q^*(s;\underline{\theta}_q),\theta_u^*(s)) & \text{if } s \in (\hat{s}(\underline{\theta}_q), \delta_4). \end{cases}$$

Third, in the game Γ_2 defined by F and $C = C_u$, for all $s \in [\delta_1, \delta_3]$, either $v(s) > g(s, \theta^*(s))$ or $D_u(s) \leq 0$, and $v(s) < g(s, \theta^*(s))$ for all $s \in [\delta_4, \delta_2]$. Thus, if $s \in [\delta_1, \delta_3] \cup [\delta_4, \delta_2]$ and $D_u(s) > 0$, then

$$w(s;\underline{\theta}_q)\begin{cases} > g(s,\theta_q^*(s;\underline{\theta}_q),\theta_u^*(s)) & \text{if } s \in [\delta_1, \delta_3] \\ < g(s,\theta_q^*(s;\underline{\theta}_q),\theta_u^*(s)) & \text{if } s \in [\delta_4, \delta_2] \end{cases}$$

for all $\underline{\theta}_q \in \left(\underline{\theta}_{q_3}, \overline{\theta}\right)$ for a large enough $\underline{\theta}_{q_3} \in \Theta^\circ$.

It follows that for all $\underline{\theta}_q > \underline{\theta}_q := \max\left\{\underline{\theta}_{q_1}, \underline{\theta}_{q_2}, \underline{\theta}_{q_3}\right\}$, $\hat{s}(\underline{\theta}_q)$ is the only root of $w(\cdot;\underline{\theta}_q) - g(\cdot, \theta_q^*(\cdot;\underline{\theta}_q), \theta_u^*(\cdot))$ over $(0,1)$, and for any equilibrium $(s_{SP}^*, \theta_q^*, \theta_u^*)$, we have $w(s_{SP}^*;\underline{\theta}_q) > g(s_{SP}^*, \theta_q^*(s_{SP}^*;\underline{\theta}_q), \theta_u^*(s_{SP}^*))$ if $s_{SP}^* < \hat{s}(\underline{\theta}_q)$ and $w(s_{SP}^*;\underline{\theta}_q) < g(s_{SP}^*, \theta_q^*(s_{SP}^*;\underline{\theta}_q), \theta_u^*(s_{SP}^*))$ if $s_{SP}^* > \hat{s}(\underline{\theta}_q)$. The thesis of the lemma follows immediately. ■

Lemma 12 *Assume F.2, F.3, F.5, F.6, C.1-C.4, and $f > 0$. Then, there exists $\underline{\theta}_q \in \Theta^\circ$ such that for all $\underline{\theta}_q \in \left(\underline{\theta}_q, \overline{\theta}\right)$, Lemma 3 holds in Γ_3, mutatis mutandis, replacing $\underline{\kappa}$ and $\overline{\kappa}$ with $\underline{\kappa}(\underline{\theta}_q)$ and $\overline{\kappa}(\underline{\theta}_q)$, respectively.*

Proof. From Lemma 10, $F(\cdot, \theta_q^*(\cdot;\underline{\theta}_q))$ is strictly increasing for high enough $\underline{\theta}_q$. Then, the argument showing that $\mathcal{S}^*$ is increasing in the proof of Lemma 3 holds here as well.

We prove that $\mathcal{S}^*(\underline{\kappa}(\underline{\theta}_q)) = \{0\}$ (the proof of the statement $\mathcal{S}^*(\overline{\kappa}(\underline{\theta}_q)) = \{1\}$ is analogous). First notice that

$$\lim_{(s,\underline{\theta}_q)\to(0,\overline{\theta})} \frac{dw(s;\underline{\theta}_q)}{ds} = \lim_{s\to 0} \frac{dv(s)}{ds} > \lim_{s\to 0} \frac{dg(s,\theta_u^*(s))}{ds} > 0. \tag{44}$$

Thus, there exists $\underline{\theta}_{q_1} \in \Theta^\circ$ and $\delta > 0$ such that for all $\underline{\theta}_q \in \left(\underline{\theta}_{q_1}, \overline{\theta}\right)$, $\mathcal{V}(0;\underline{\kappa}(\underline{\theta}_q),\underline{\theta}_q) < \mathcal{V}(s;\underline{\kappa}(\underline{\theta}_q),\underline{\theta}_q)$ for all $s \in (0,\delta)$, where $\mathcal{V}(\cdot;\kappa,\underline{\theta}_q)$ is the expected loss to the manager $\mathcal{V}(s)$ for prior κ when the initial advantage of fit candidates is $\underline{\theta}_q$.

Second, notice that from Lemma 3 and Assumption F.5, we know that

$$\lim_{\underline{\theta}_q\to\overline{\theta}} \mathcal{V}(0;\underline{\kappa}(\underline{\theta}_q),\underline{\theta}_q) < \lim_{\underline{\theta}_q\to\overline{\theta}} \mathcal{V}(s;\underline{\kappa}(\underline{\theta}_q),\underline{\theta}_q)$$

for all $s > 0$. Let $\widetilde{\mathcal{V}}(\cdot,\underline{\theta}_q) := \mathcal{V}(\cdot;\underline{\kappa}(\underline{\theta}_q),\underline{\theta}_q)$ and notice that $\widetilde{\mathcal{V}}(0,\cdot)$ and $\min_{s\in[\delta,1]} \widetilde{\mathcal{V}}(s,\cdot)$ are continuous functions. Thus, there exists $\underline{\theta}_{q_2}$ such that $\widetilde{\mathcal{V}}(0,\underline{\theta}_q) < \min_{s\in[\delta,1]} \widetilde{\mathcal{V}}(s,\underline{\theta}_q)$ for all $\underline{\theta}_q > \underline{\theta}_{q_2}$. Thus, $\widetilde{\mathcal{V}}(0,\underline{\theta}_q) < \widetilde{\mathcal{V}}(s,\underline{\theta}_q)$ for all

$s > 0$ and $\underline{\theta}_q > \max\left\{\underline{\theta}_{q_1}, \underline{\theta}_{q_2}\right\}$; that is, $\mathcal{S}^*(\underline{\kappa}(\underline{\theta}_q)) = \{0\}$ for all $\underline{\theta}_q > \max\left\{\underline{\theta}_{q_1}, \underline{\theta}_{q_2}\right\}$.

For $\kappa > \underline{\kappa}(\underline{\theta}_q)$, we have $d\mathcal{V}(0)/ds < 0$ and hence $0 \notin \mathcal{S}^*(\kappa)$.[34] Analogously, for $\kappa < \overline{\kappa}(\underline{\theta}_q)$, we have $d\mathcal{V}(1)/ds > 0$ and hence $1 \notin \mathcal{S}^*(\kappa)$.

Finally, an argument analogous to that in the corresponding part of the proof of Lemma 3 reveals that $\mathcal{S}^*$ is strictly increasing over $[\underline{\kappa}(\underline{\theta}_q), \overline{\kappa}(\underline{\theta}_q)]$. ■

Proof of Proposition 4. The proof of part (i) is analogous to the proof of Proposition 1, with $\hat{s}(\underline{\theta}_q)$ defined in Lemma 11, $\underline{\kappa}(\underline{\theta}_q)$, and $\overline{\kappa}(\underline{\theta}_q)$ playing the role of $\hat{s}$, $\underline{\kappa}$, and $\overline{\kappa}$, respectively.[35] In particular, Lemma 11 plays the role of Lemmas 1 and 2, and Lemma 12 plays the role of Lemma 3.

Now we prove part (ii). If F satisfies FMP, then, by L'Hôpital's rule

$$r(s;\underline{\theta}_q) = \frac{\left(\frac{dh(s,\underline{\theta}_q)}{ds} - \frac{dh(s,\underline{\theta}_q)}{ds}\frac{df(s,\underline{\theta}_q)}{d\theta}\left(C_q''(\underline{\theta}_q) + \frac{dh(s,\underline{\theta}_q)}{d\theta}\right)^{-1}\right)\frac{df(s,\underline{\theta}_q)}{d\theta}\left(C_q''(\underline{\theta}_q) + \frac{dh(s,\underline{\theta}_q)}{d\theta}\right)^{-1}}{\left(\frac{dh(s,\underline{\theta})}{ds} - \frac{dh(s,\underline{\theta})}{ds}\frac{df(s,\underline{\theta})}{d\theta}\left(C_u''(\underline{\theta}) + \frac{dh(s,\underline{\theta})}{d\theta}\right)^{-1}\right)\frac{df(s,\underline{\theta})}{d\theta}\left(C_u''(\underline{\theta}) + \frac{dh(s,\underline{\theta})}{d\theta}\right)^{-1}}$$

for $s = 0, 1$; and

$$r(s^\star;\underline{\theta}_q) = \frac{h(s^\star, \theta_q^*(s^\star;\underline{\theta}_q))\frac{\partial^2 f(s^\star,\theta_q^*(s^\star;\underline{\theta}_q))}{\partial\theta\partial s}\left(C_q''(\theta_q^*(s^\star;\underline{\theta}_q)) + \frac{dh(s^\star,\theta_q^*(s^\star;\underline{\theta}_q))}{d\theta}\right)^{-1}}{h(s^\star, \theta_u^*(s^\star))\frac{\partial^2 f(s^\star,\theta_u^*(s^\star))}{\partial\theta\partial s}\left(C_q''(\theta_u^*(s^\star)) + \frac{dh(s^\star,\theta_u^*(s^\star))}{d\theta}\right)^{-1}}.$$

It follows that r is well defined over $[0,1]$; furthermore $\max_{[0,1]} r(s;\cdot)$ is continuous and converges to 0 as $\underline{\theta}_q \to \overline{\theta}$. Therefore, for large enough $\underline{\theta}_q$, $r(s) < \kappa$ for all $s \in [0,1]$ and $\kappa \geq \underline{\kappa}(\underline{\theta}_q)$. From part (ii) of Lemma 8, the manager is soft (harsh) in an equilibrium $(s_{SP}^*, \theta_q^*(s_{SP}^*), \theta_u^*(s_{SP}^*))$ if $s_{SP}^* < (>) s^\star$. Thus, there exists $\underline{\theta}_{q_3} < \overline{\theta}$ such that $\hat{s}(\underline{\theta}_q)$ defined in the proof of Lemma 11 is equal to $s^\star$ for all $\underline{\theta}_q \in (\underline{\theta}_{q_3}, \overline{\theta})$. Further, (i)-(iii) in Claim 2 still hold in the extended model, so the analogous to (iv) in that claim holds as well.[36] Hence, if $\kappa = 1$, then $\mathcal{V}(s) = \mathcal{V}(s_f(s))$ for all $s \in [0,1]$. Therefore

$$\widetilde{\kappa}(\underline{\theta}_q) := \sup\{\kappa \in (0,\infty) : \sup \mathcal{S}^*(\kappa) < \hat{s}(\underline{\theta}_q)\} = 1$$

by an argument analogous to that in the proof of Lemma 5.

Proofs of Section 8.3

Proof of Lemma 9. Since s_{SP}^* is an interior minimizer of $\mathcal{W}(\cdot, \theta_q^*(\cdot))$, it satisfies $\tilde{v}(s_{SP}^*) = \kappa$. Recall that $h(s,\theta)$ is strictly negative for all $s \in (0,1)$ and $\theta \in \Theta^\circ$, thus $\tilde{g}(s_{SP}^*, \theta_q^*(s_{SP}^*)) > \tilde{v}(s_{SP}^*) = \kappa$ if and only if F

[34] Assumption F.5 yields that $\lim_{s\to 0} w(s;\underline{\theta}_q) = \underline{\kappa}(\underline{\theta}_q)$ and $\lim_{s\to 1} v(s;\underline{\theta}_q) = \overline{\kappa}(\underline{\theta}_q)$; and thus $\mathcal{S}^*(\kappa) \subset (0,1)$ for all $\kappa \in (\underline{\kappa}(\underline{\theta}_q), \overline{\kappa}(\underline{\theta}_q))$.

[35] Case 1 in Part 1 in the the proof of Proposition 1 is not necessary here as the assumption $f > 0$ rules out the possibility that $\overline{\kappa}(\underline{\theta}_q) = \infty$ for all $\underline{\theta}_q \in \Theta^\circ$.

[36] In particular, part (ii) of Claim 2 holds for both θ_q^* and θ_u^*.

is strictly submodular at $(s^*_{SP}, \theta^*_q(s^*_{SP}))$. Since $\tilde{g}(\cdot, \theta^*_q(s^*_{SP}))$ is strictly increasing, either $\tilde{g}(s, \theta^*_q(s^*_{SP})) = \kappa$ for some $s < s^*_{SP}$, or the minimizer of $\mathcal{W}(\cdot, \theta^*_q(s^*_{SP}))$ is 0. Thus, $s^*_{SP} > s^*(\theta^*_q(s^*_{SP}), \kappa)$ if and only if F is strictly submodular at $(s^*_{SP}, \theta^*_q(s^*_{SP}))$.

The argument for the case in which F is strictly supermodular at $(s^*_{SP}, \theta^*_q(s^*_{SP}))$ is analogous. □

The following lemma is analogous to Lemma 3 in the benchmark model.

Lemma 13 *Assume F.1-F.3, F.6, and C.3.* $\mathcal{S}^*$ *is weakly increasing over* $(0, \infty)$ *and strictly increasing over* $(\underline{\underline{\kappa}}, \overline{\overline{\kappa}})$. *Further,*

$$\mathcal{S}^*(\kappa) \begin{cases} = \{0\} & \text{if } 0 < \kappa < \underline{\underline{\kappa}} \\ \subset (0,1) & \text{if } \underline{\underline{\kappa}} < \kappa < \overline{\overline{\kappa}} \\ = \{1\} & \text{if } \overline{\overline{\kappa}} < \kappa. \end{cases}$$

Proof. The proof that $\mathcal{S}^*$ is weakly increasing is indirect and analogous to the one in the proof of Lemma 3, so we omit it.

By definition,

$$\underline{\underline{\kappa}} \leq \frac{F(s, \theta^*_q(s))}{F(s, \underline{\theta})}$$
$$0 \leq F(s, \theta^*_q(s)) - \underline{\underline{\kappa}} F(s, \underline{\theta})$$

for all $s \in (0,1)$. Thus $0 \in \mathcal{S}^*(\underline{\underline{\kappa}})$ and since $\mathcal{S}^*$ is weakly increasing, $\mathcal{S}^*(\kappa) = \{0\}$ for all $\kappa < \underline{\underline{\kappa}}$.

Also by definition, for all $\kappa > \underline{\underline{\kappa}}$, there exists $s \in (0,1)$ such that

$$\kappa > \frac{F(s, \theta^*_q(s))}{F(s, \underline{\theta})}$$
$$0 > F(s, \theta^*_q(s)) - \kappa F(s, \underline{\theta}),$$

and therefore $0 \notin \mathcal{S}^*(\kappa)$.

An analogous argument shows that $1 \in \mathcal{S}^*(\overline{\overline{\kappa}})$, $\mathcal{S}^*(\kappa) = \{1\}$ for all $\kappa > \overline{\overline{\kappa}}$ and that $1 \notin \mathcal{S}^*(\kappa)$ for all $\kappa < \overline{\overline{\kappa}}$.

Finally, the argument to prove that $\mathcal{S}^*$ is strictly increasing over $(\underline{\underline{\kappa}}, \overline{\overline{\kappa}})$ is analogous to the corresponding argument in the proof of Lemma 3. ■

Proof of Proposition 5. First, we establish that $\tilde{\tilde{\kappa}} \in (\underline{\underline{\kappa}}, \overline{\overline{\kappa}})$. Let $\mathcal{W}(\cdot, \cdot; \kappa) : [0,1] \times [\underline{\theta}_q, \overline{\theta}] \to \mathbb{R}$ be the expected loss function to the manager in Γ_4 for all $\kappa \in (0, \infty)$. Observe that

$$\mathcal{W}(\hat{s}_q, \theta^*_q(\hat{s}_q); \tilde{\tilde{\kappa}}) < \mathcal{W}(s, \theta^*_q(\hat{s}_q); \tilde{\tilde{\kappa}}) < \mathcal{W}(s, \theta^*_q(s); \tilde{\tilde{\kappa}}),$$

for all $s \in [0,1] \setminus \{\hat{s}_q\}$, where the first inequality follows from the fact that $\hat{s}_q$ is the unique minimizer of $\mathcal{W}(\cdot, \theta_q^*(\hat{s}_q); \tilde{\tilde{\kappa}})$ and the second follows from observing that $\mathcal{W}(s, \cdot; \tilde{\tilde{\kappa}})$ is decreasing for all $s \in (0,1)$ and that $\hat{s}_q$ is the unique maximizer of $\theta_q^*(\cdot)$. Thus, $\mathcal{S}^*(\tilde{\tilde{\kappa}}) = \{\hat{s}_q\}$, and since in the proof of Lemma 13 it is shown that $0 \in \mathcal{S}^*(\underline{\underline{\kappa}})$ and $1 \in \mathcal{S}^*(\overline{\overline{\kappa}})$, we conclude $\tilde{\tilde{\kappa}} \in (\underline{\underline{\kappa}}, \overline{\overline{\kappa}})$.

Now we prove that the manager is ex-post efficient for all $\kappa \in \big(0, \underline{\underline{\kappa}}\big)$. Observe that $\underline{\underline{\kappa}} \le \tilde{g}(0, \underline{\theta}_q)$:

$$\underline{\underline{\kappa}} \le \lim_{s \to 0} \frac{F(s, \theta_q^*(s))}{F(s, \underline{\theta})} = \lim_{s \to 0} \frac{f(s, \theta_q^*(s)) + h(s, \theta_q^*(s)) \frac{d\theta_q^*(s)}{ds}}{f(s, \underline{\theta})} \le \lim_{s \to 0} \tilde{g}(s, \theta_q^*(s)) = \tilde{g}(0, \underline{\theta}_q).$$

Since the ex-post optimal standard is weakly increasing in κ (see the analysis of $s^*(\theta, \cdot)$ in Section 2.2), $s = 0$ is the only ex-post optimal standard for $\kappa = \tilde{g}(0, \underline{\theta}_q)$, and $\underline{\underline{\kappa}} < \tilde{g}(0, \underline{\theta}_q)$, we have that the manager is ex-post efficient for all $\kappa < \underline{\underline{\kappa}}$.

The argument proving that the manager is harsh for all $\kappa \in (\underline{\underline{\kappa}}, \tilde{\tilde{\kappa}})$ is analogous to the argument showing that the manager is harsh for all $\kappa \in (\tilde{\kappa}, \overline{\kappa})$ in the proof of Proposition 1 in the benchmark model, but using Lemma 9, Lemma 7 for $i = q$, and Lemma 13, instead of Lemmas 1-3, and it is omitted.

The arguments showing that the manager is ex-post efficient for all $\kappa \in \big(\overline{\overline{\kappa}}, \infty\big)$ and soft for all $\kappa \in (\tilde{\tilde{\kappa}}, \overline{\overline{\kappa}})$ are analogous to the arguments showing that the manager is ex-post efficient for all $\kappa \in \big(0, \underline{\underline{\kappa}}\big)$ and harsh for all $\kappa \in (\underline{\underline{\kappa}}, \tilde{\tilde{\kappa}})$, respectively. Finally, the manager is ex-post efficient if $\kappa = \tilde{\kappa}$ because of Lemma 9. □

The following corollary states that in the pure information generation game, commitment to standards typically involves a mix of confirmativism and conservatism when the manager faces candidates with different priors of being fit.

Corollary 5 *Assume F.1-F.4, F.6, and C.3. For every game* Γ_4 *and* $\kappa \notin \{\underline{\underline{\kappa}}, \overline{\overline{\kappa}}\}$ *the manager is*

1. *conservative at* κ *if and only if* $\underline{\underline{\kappa}} < \kappa < \min\{1, \tilde{\tilde{\kappa}}\}$ *or* $\max\{1, \tilde{\tilde{\kappa}}\} < \kappa < \overline{\overline{\kappa}}$*, and*

2. *confirmative at* κ *if and only if* $\min\{1, \tilde{\tilde{\kappa}}\} < \kappa < \max\{1, \tilde{\tilde{\kappa}}\}$.

The following result is a direct consequence of Corollary 5. It provides a necessary and sufficient condition for uniform conservatism and rules out uniform confirmativism.

Corollary 6 *Assume F.1-F.4, F.6, and C.3. For any* Γ_4*, (i) the manager is uniformly conservative if and only if* $\tilde{\tilde{\kappa}} = 1$ *and (ii) the manager cannot be uniformly confirmative.*

Proof of Proposition 6. *Sufficiency.* If F has a neutral signal $s^\star$, then $\hat{s}_q = s^\star$ and $\tilde{\tilde{\kappa}} = \tilde{g}(\hat{s}_q, \theta_q^*(\hat{s}_q)) = 1$. Corollary 6 implies that the manager is uniformly conservative.

Necessity. If F does not have a neutral signal, then $\frac{d\tilde{s}(\theta)}{d\theta} \neq 0$ and hence $\frac{d\widetilde{g}(\tilde{s}(\theta),\theta)}{d\theta} \neq 0$ for some $\theta \in \Theta^\circ$. Thus, $\widetilde{g}(\tilde{s}(\theta'),\theta') \neq 1$ for some $\theta' \in \Theta^\circ$. Consider any $\underline{\theta}_q \in (\underline{\theta},\theta')$ and cost function $C_q(\theta) = \alpha\frac{1}{2}(\theta - \underline{\theta}_q)^2$, with $\alpha = \frac{-h(\widetilde{s}(\theta'),\theta')}{\theta'-\underline{\theta}}$. For this cost function, $C'_q(\theta') = -h(\widetilde{s}(\theta'),\theta')$, and hence $\hat{s}_q = \tilde{s}(\theta')$ and $\widetilde{\widetilde{\kappa}} = \widetilde{g}(\tilde{s}(\theta'),\theta') \neq 1$. Corollary 6 implies that the manager is not uniformly conservative. □

Proof of Proposition 7. We first prove part (i). If $\underline{\underline{\kappa}} < \widetilde{g}(0,\underline{\theta}_q)$, for all $\kappa \in (\underline{\underline{\kappa}},\widetilde{g}(0,\underline{\theta}_q))$, the optimal standard in the information-generation classical statistical problem is $s = 0$. On the other hand, $\mathcal{S}^*(\kappa) \subset (0,1)$ for all $\kappa \in (\underline{\underline{\kappa}},\widetilde{g}(0,\underline{\theta}_q))$ (see Lemma 13). It follows that unfit candidates are worse off in Γ_4. Fit candidates are also worse-off in Γ_4 as $F(s,\theta) > 0$ for all $s > 0$ and $\theta \in \Theta$.

Now consider the alternative case, $\underline{\underline{\kappa}} = \widetilde{g}(0,\underline{\theta}_q)$. If $0 < s \in \mathcal{S}^*(\underline{\underline{\kappa}})$, then there exists $\delta > 0$ such that for all $\kappa \in (\underline{\underline{\kappa}},\underline{\underline{\kappa}}+\delta)$, $\inf \mathcal{S}^*(\kappa) > s$ whereas the optimal standards in the information-generation classical problem, by continuity of $s^*_\kappa := \arg\min_{s\in[0,1]} \mathcal{W}(s,\underline{\theta}_q;\kappa)$ as a function of κ, is less than s, and also $F(s',\theta^*_q(s')) > F(s^*_\kappa,\underline{\theta}_q)$ for all $s' \in \mathcal{S}^*(\kappa)$ for all $\kappa \in (\underline{\underline{\kappa}},\underline{\underline{\kappa}}+\delta)$. Finally, if $\mathcal{S}^*(\underline{\underline{\kappa}}) = \{0\}$ then, the argument is analogous to the one in the proof of part (ii) of Proposition 3 (and it is omitted).

Now we provide the proof of part (ii). An argument analogous to the one provided in the proof of Proposition 5 reveals that $\overline{\overline{\kappa}} \geq \widetilde{g}(1,\underline{\theta}_q)$. If $\overline{\overline{\kappa}} > \widetilde{g}(1,\underline{\theta}_q)$, then we have that $\mathcal{S}^*(\kappa) \subset (0,1)$ for all $\kappa \in (\widetilde{g}(1,\underline{\theta}_q),\overline{\overline{\kappa}})$, whereas in the information-generation classical statistical problem, the optimal standard is 1. It follows that both fit and unfit candidates are better-off in Γ_4 than in the information-generation classical statistical problem.

Now we consider the complementary case, $\overline{\overline{\kappa}} = \widetilde{g}(1,\underline{\theta}_q)$. If $1 > s \in \mathcal{S}^*(\overline{\overline{\kappa}})$ then, there exists $\delta > 0$ such that for all $\kappa \in (\overline{\overline{\kappa}}-\delta,\overline{\overline{\kappa}})$, $\sup \mathcal{S}^*(\kappa) < s$ whereas the optimal standards in the information-generation classical problem, by continuity of $s^*_\kappa := \arg\min_{s\in[0,1]} \mathcal{W}(s,\underline{\theta}_q;\kappa)$ as a function of κ, is greater than s, and also $F(s',\theta^*_q(s')) < F(s^*_\kappa,\underline{\theta}_q)$ for all $s' \in \mathcal{S}^*(\kappa)$ for all $\kappa \in (\underline{\underline{\kappa}},\underline{\underline{\kappa}}+\delta)$. Finally, if $\mathcal{S}^*(\overline{\overline{\kappa}}) = \{1\}$ then, the argument is analogous to the one in the proof of part (i) of Proposition 3 (and it is omitted). □

Proof of Proposition 8. In the sequel, when convenient, we note explicitly the dependence of D_u, r and w on the unfit candidates' cost parameter λ, and write $D_u(s;\lambda)$, $r(s;\lambda)$, and $w(s;\lambda)$ instead of $D_u(s)$, $r(s)$, and $w(s)$, respectively, for all $s \in [0,1]$ and $\lambda \in [0,1]$. The proof hinges on the following lemmata:

Lemma 14 *Assume F.2, F.3, F.5, C.1, C.3, and $f > 0$. Then, there exists $\overline{\lambda}_1 > 0$ such that for all $\lambda \in [0,\overline{\lambda}_1)$, $D_u(s;\lambda) > 0$ for all $s \in (0,1)$.*

Proof. The Maximun Theorem and Assumption F.5 guarantee that $\min_{s\in[0,1]} D_u(s;\lambda)$ varies continuously with λ. Further, $\min_{s\in[0,1]} D_u(s;\lambda) = \min_{s\in[0,1]} f(s,\underline{\theta}) > 0$ for $\lambda = 0$. Thus, there exists $\overline{\lambda}_1 > 0$ such that $\min_{s\in[0,1]} D_u(s;\lambda) > 0$ for all $\lambda \in [0,\overline{\lambda}_1)$. ■

Lemma 15 *Assume F.2-F.6, C.1-C.4, and $f > 0$. Then, there exists $\overline{\lambda} > 0$ such that for all $\lambda \in \left[0, \overline{\lambda}\right)$, there exists $\hat{s}(\lambda) \in (0,1)$ such that if $(s^*_{SP}, \theta^*_q, \theta^*_u)$ is a subgame perfect equilibrium of Γ_3, then the manager is harsh if $s^*_{SP} \in (0, \hat{s}(\lambda))$ and soft if $s^*_{SP} \in (\hat{s}(\lambda), 1)$.*

Proof. The proof is analogous to the proof of Lemma 11, using the fact that $w(\cdot;\lambda)$ converges to $\widetilde{v}$ (instead of v) and $g(\cdot, \theta^*_q(\cdot), \theta^*_u(\cdot))$ converges to $\widetilde{g}(\cdot, \theta^*_q(\cdot))$ (instead of $g(\cdot, \theta^*_u(\cdot))$), as $\lambda \to 0$ (instead of $\underline{\theta}_q \to \overline{\theta}$), where this convergence is established to be either pointwise or uniform for some intervals, according to what is needed in different parts of the argument. ■

Lemma 16 *Assume F.2, F.3, F.5, F.6, C.1-C.4, and $f > 0$. Then, there exists $\overline{\lambda} > 0$ such that for all $\lambda \in \left[0, \overline{\lambda}\right)$, Lemma 13 holds in $\Gamma_3(\lambda)$, mutatis mutandis, replacing $\underline{\underline{\kappa}}$ and $\overline{\overline{\kappa}}$ with $\underline{\underline{\kappa}}(\lambda)$ and $\overline{\overline{\kappa}}(\lambda)$, respectively.*

Proof. From Lemma 14, $F(\cdot, \theta^*_u(\cdot;\lambda))$ is strictly increasing for all $\lambda < \overline{\lambda}_1$. Then, an argument analogous to the one in the proof of Lemma 13 applies with $\underline{\underline{\kappa}}(\lambda)$ and $\overline{\overline{\kappa}}(\lambda)$ playing the role of $\underline{\kappa}(\underline{\theta}_q)$ and $\overline{\kappa}(\underline{\theta}_q)$, respectively. ■

Proof of Proposition 8. Let $\widetilde{\widetilde{\kappa}}(\lambda) := g(\hat{s}(\lambda), \theta^*_q(\hat{s}(\lambda)), \theta^*_u \hat{s}(\lambda))$ and notice that, since $\lim_{\lambda \to 0} \hat{s}(\lambda) = \hat{s}_q$, we have $\lim_{\lambda \to 0} \widetilde{\widetilde{\kappa}}(\lambda) = \widetilde{\widetilde{\kappa}}$. Similarly, $\lim_{\lambda \to 0} \underline{\underline{\kappa}}(\lambda) = \underline{\underline{\kappa}}$ and $\lim_{\lambda \to 0} \overline{\overline{\kappa}}(\lambda) = \overline{\overline{\kappa}}$. Thus, there exists $\overline{\lambda} > 0$ such that $\underline{\underline{\kappa}}(\lambda) < \widetilde{\widetilde{\kappa}}(\lambda) < \overline{\overline{\kappa}}(\lambda)$ for all $\lambda \in (0, \overline{\lambda})$.

For part (i), the above lemmatta allow us to provide an argument analogous to the one in the proof of Proposition 5, with Lemma 15 playing the role of Lemmas 9 and 7, and Lemma 16 playing the role of Lemma 13. The analogy is transparent so details are omitted.

The proof of part (ii) is analogous to the proof of part (ii) of Proposition 4, but now showing that $\min_{s \in [0,1]} r(s;\lambda)$ goes to ∞ as $\lambda \to 0$, which in turns implies that there exists $\overline{\lambda} > 0$ such that $r(s) > \kappa$ for all $s \in [0,1]$ and $\kappa \leq \overline{\overline{\kappa}}(\lambda)$. The rest of the argument is analogous and its details are omitted.

Proofs of Section 8.4

Proof of Claim 1. First notice that $C'_u(\theta) > C'_q(\theta; \underline{\theta}_q)$ for all $\theta \in [\underline{\theta}_q, 1]$ and $\underline{\theta}_q \in \Theta^\circ$, because $1/\lambda > 1 - \underline{\theta}_q$ for all $\lambda \in (0,1]$ and $\underline{\theta}_q \in (0,1)$.

Statements (i) and (ii) are consequences of Proposition 1 and 5, respectively, and the fact that F satisfies FMP.

We now provide functions $F(\cdot, 0)$ and $F(\cdot, 1)$, and a constellation of λ and $\underline{\theta}_q$ yielding statement (iii). Let $\Delta F(s) := F(s,0) - F(s,1)$ for all $s \in [0,1]$ and $s^\star$ be the neutral signal of F. The best response functions of unfit and fit candidates are respectively given by

$$\theta^*_u(s) = \frac{\Delta F(s)}{\Delta F(s) + 1/\lambda} \quad \text{and} \quad \theta^*_q(s) = \frac{\Delta F(s) + \underline{\theta}_q}{\Delta F(s) + 1}.$$

Furthermore,

$$r(s) = \lambda^2(1 - \underline{\theta}_q)^2 \left(\frac{\Delta F(s) + 1/\lambda}{\Delta F(s) + 1}\right)^3, \tag{45}$$

for all $s \in [0, 1]$.

Consider the distributions $F(s, 0) = s$ and $F(s, 1) = s^2$, for all $s \in [0, 1]$. The distribution F has a neutral signal $s^\star = 0.5$. Consider the parameters $\lambda = 0.7$ and $\underline{\theta}_q = 0.1$. For this specification, $s^*_{SP} = 0$ for all $\kappa \leq 0.90$, $s^*_{SP} \in (0, 1)$ for $\kappa \in (0.90, 1.10)$, and $s^*_{SP} = 1$ for $\kappa > 1.10$.

On the other hand, $r(s) > w(s)$ for $s \in (0, 0.36) \cup (0.90, 1)$ and $r(s) < w(s)$ for $s \in (0.36, 0.90)$; in particular, $r(0.36) = w(0.36) = 0.97$ and $r(0.90) = w(0.90) = 1.08$. Thus, the manager is harsh for $\kappa \in (0.90, 0.97)$, soft for $\kappa \in (0.97, 1)$, harsh for $\kappa \in (1, 1.08)$, and soft for $\kappa \in (1.08, 1.10)$. Otherwise, the manager is ex-post efficient. □

References

Alonso, R. and N. Matouschek (2008): "Optimal delegation," *The Review of Economic Studies*, 75, 259–293.

Butcher, J. N. (1994): "Psychological assessment of airline pilot applicants with the MMPI-2," *Journal of Personality Assessment*, 62, 31–44.

——— (2002): "Assessing pilots with the wrong stuff: A call for research on emotional health factors in commercial aviators," *International Journal of Selection and Assessment*, 10, 168–184.

Chassang, S., P. I. Miquel, and E. Snowberg (2012): "Selective trials: A principal-agent approach to randomized controlled experiments," *American Economic Review*, 102, 1279–1309.

Crawford, V. P. and J. Sobel (1982): "Strategic information transmission," *Econometrica*, 1431–1451.

DeGroot, M. H. (2005): *Optimal statistical decisions*, vol. 82, John Wiley & Sons.

Demougin, D. and C. Fluet (2008): "Rules of proof, courts, and incentives," *The RAND Journal of Economics*, 39, 20–40.

Di Tillio, A., M. Ottaviani, and P. N. Sorensen (2016): "Strategic Sample Selection," .

Di Tillio, A., M. Ottaviani, and P. N. Sørensen (2017): "Persuasion bias in science: can economics help?" *The Economic Journal*, 127.

Ederer, F., R. Holden, and M. Meyer (2017): "Gaming and strategic opacity in incentive provision," *Games and Economic Behavior*, forthcoming.

FANELLI, D. (2009): "How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data," *PloS one*, 4, e5738.

FISCHER, P. E. AND R. E. VERRECCHIA (2000): "Reporting bias," *The Accounting Review*, 75, 229–245.

FISHMAN, M. J. AND K. M. HAGERTY (1990): "The optimal amount of discretion to allow in disclosure," *The Quarterly Journal of Economics*, 427–444.

FRANKEL, A. AND N. KARTIK (2017): "Muddled information," *Journal of Political Economy*, forthcoming.

GANUZA, J., F. GOMEZ, AND J. PENALVA (2012): "Minimizing errors, maximizing incentives: Optimal court decisions and the quality of evidence," Tech. rep.

GENTZKOW, M. AND E. KAMENICA (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615.

GERLACH, H. (2013): "Self-reporting, investigation, and evidentiary standards," *Journal of Law and Economics*, 56, 1061–1090.

GILBERT, R. J. AND P. KLEMPERER (2000): "An equilibrium theory of rationing," *The RAND Journal of Economics*, 1–21.

GLAZER, J. AND A. RUBINSTEIN (2004): "On optimal rules of persuasion," *Econometrica*, 72, 1715–1736.

HART, O. AND J. MOORE (2007): "Incomplete contracts and ownership: Some new thoughts," *The American Economic Review*, 182–186.

HEDLUND, J. (2017): "Bayesian persuasion by a privately informed sender," *Journal of Economic Theory*, 167, 229–268.

KAPLOW, L. (2011): "On the optimal burden of proof," *Journal of Political Economy*, 119, 1104–1140.

KARLIN, S. AND H. RUBIN (1956): "The theory of decision procedures for distributions with monotone likelihood ratio," *The Annals of Mathematical Statistics*, 272–299.

KOLOTILIN, A. (2015): "Experimental design to persuade," *Games and Economic Behavior*, 90, 215–226.

KOLOTILIN, A., H. LI, AND W. LI (2013): "Optimal limited authority for principal," *Journal of Economic Theory*, 148, 2344–2382.

LAUX, V. (2008): "On the value of influence activities for capital budgeting," *Journal of Economic Behavior & Organization*, 65, 625–635.

Li, H. (2001): "A theory of conservatism," *Journal of Political Economy*, 109, 617–636.

Matouschek, N. (2004): "Ex post inefficiencies in a property rights theory of the firm," *Journal of Law, Economics, and Organization*, 20, 125–147.

Milgrom, P. R. (1981): "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 380–391.

Neyman, J. and E. S. Pearson (1933): "IX. On the problem of the most efficient tests of statistical hypotheses," *Phil. Trans. R. Soc. Lond. A*, 231, 289–337.

Perez, E. and D. Prady (2012): "Complicating to Persuade?" Tech. rep., HAL.

Prendergast, C. (1993): "A theory of "yes men"," *American Economic Review*, 757–770.

Rayo, L. and I. Segal (2010): "Optimal information disclosure," *Journal of Political Economy*, 118, 949–987.

Reny, P. J. (2005): "Non-cooperative games: Equilibrium existence," *The New Palgrave Dictionary of Economics.*

Rubinfeld, D. L. and D. E. Sappington (1987): "Efficient awards and standards of proof in judicial proceedings," *The RAND Journal of Economics*, 308–315.

Sanchirico, C. W. (2004): "Evidence Tampering," *Duke Law Journal*, 1215–1336.

——— (2010): "Detection avoidance and enforcement theory: survey and assessment," *U of Penn, Inst for Law & Econ Research Paper Series.*

Skaperdas, S. and S. Vaidya (2012): "Persuasion as a contest," *Economic Theory*, 51, 465–486.

Spence, M. (1973): "Job market signaling," *The Quarterly Journal of Economics*, 355–374.

Stephenson, M. C. (2008): "Evidentiary Standards and Information Acquisition in Public Law," *American Law and Economics Review*, 10, 351–387.

Tetenov, A. (2016): "An economic theory of statistical testing," cemmap working paper, Centre for Microdata Methods and Practice CWP50/16, London.