Multidimensional screening with substitutable attributes and costly verification*

Click to access latest version.

Orestis Vravosinos†

March 5, 2025

Abstract

A principal must decide whether to accept or reject an agent. The principal can verify at a cost the value of a composite measure of the agent's training and talent. The measure does not reveal training and talent separately. The agent can present evidence of training but not of talent. Although favorable, evidence can make the principal ascribe the value of the composite measure to training, thereby negatively affecting his assessment of the agent's talent. Thus, verification may distort the agent's incentives to present evidence. Indeed, when the composite measure is less sensitive to talent than talent is valuable to the principal, a conflict arises between the two evaluation methods: (i) verification and (ii) asking for evidence. The optimal mechanism leads to three types of errors, all favoring high- over low-training agents: (i) It rejects some worthy low-training agents, while (ii) accepting some unworthy high-training ones without verification and (iii) also accepting some unworthy medium-training ones after verification.

Keywords: evidence game, signal jamming, manipulation, test, scoring, underdisclosure, multidimensional screening, costly verification
JEL classification codes: D82, D83

*I am grateful to Erik Madsen for his continued guidance and terrifyingly insightful comments. The paper has also immensely benefited from discussions with Debraj Ray. I also thank Agata Farina, Alessandro Bonatti, Alexander Bloedel, Alexander Jakobsen, Annie Liang, Ayça Kaya, Basil Williams, Benjamin Bernard, Bobby Pakzad-Hurson, Curtis Taylor, Dana Foarta, David Pearce, Dilip Abreu, Doron Ravid, Frank Yang, Gonzalo Cisternas, Guillaume Fréchette, Guillaume Haeringer, Idione Meneghel, Inga Deimen, Johannes Hörner, Joseph Root, Laura Doval, Leeat Yariv, Luciano Pomatto, Mariann Ollár, Maren Vairo, Marina Halac, Moritz Meyer-ter-Vehn, Peter Caradonna, Sevgi Yuksel, Yeon-Koo Che, and participants in the 35th Stony Brook International Conference on Game Theory and seminars at NYU for helpful comments and discussions.

†New York University; e-mail: orestis.vravosinos@nyu.edu.

1 Introduction

In many settings, a candidate's suitability for a position depends on multiple valuable qualities, such as education, training, knowledge, intelligence, or adaptability. The candidate has hard evidence on some qualities (e.g., education) but not others (e.g., intelligence). I refer to the qualities that the candidate has evidence on as *training* and the ones that she does not have evidence on as *talent*. While the evaluator cannot ask for evidence of talent, he can however try to verify talent at a cost. Nevertheless, in many real-world environments, the evaluator cannot verify talent in isolation; instead, he can verify (the value of) a *composite measure* of training and talent without being able to disentangle the individual contribution of each of the two to the composite measure. This makes evidence of training critical for the evaluator to extract information about the candidate's talent through verification of the composite measure. For instance, standardized college admission tests pick up a combination of talent and training, which makes information about an applicant's training crucial when the admissions committee tries to extract information about the applicant's talent from the test score.

Ideally, the evaluator would seamlessly combine evidence and verification, using evidence to learn about training and verification to gauge talent, conditional on what he has learned about training through evidence. In practice, however, this may not be straightforward. Although presenting all her evidence to convince the evaluator of her training is, in principle, in the candidate's best interest, verification can distort her incentives to present evidence. Specifically, evidence of training may lead the evaluator to attribute the composite measure to training, thereby negatively affecting his assessment of the candidate's talent. Therefore, to manipulate how the evaluator interprets the composite measure, the candidate may strategically withhold evidence of training.

This can create a conflict between the two evaluation tools: (i) verification and (ii) asking for evidence of training. Under what circumstances does the conflict arise? When it does, how does the evaluator use evidence and verification to optimally evaluate the candidate while taking the conflict into account? These are the questions that this paper aims to answer.

The tension between verification and asking for evidence is ubiquitous. A college applicant may downplay her parental support or how much effort she has exerted to portray her academic performance and standardized test scores as results of her brilliance rather than effort or supportive background, thereby aiming to get admitted by a college that values talent and potential. For example, she can hide her background or how intensively she has studied in the past by (i) overstating the struggles that she has gone through, (ii) not mentioning tutoring or extracurricular activities, (iii) withholding information on her parents' education or professions, or even (iv) hiding her race.[1] A job candidate

[1]Indeed, there is evidence that applicants may not only hide their race but even misrepresent it.

might downplay her background and prior effort to make the employer attribute her achievements and pre-employment test results, such as aptitude or skill tests, to talent and hire her. An employee may understate how long she took to complete a task to make the employer attribute her productivity to ability (i.e., the rate at which her work hours translate into value to the firm) and promote her. This strategy can pay off if promotion decisions rely mostly on the employer's beliefs concerning the employee's ability—because the importance of ability increases (relative to the importance of working long hours) if the employee is promoted. An academic on the job market may strategically withhold certain results, saving them to answer audience questions later to appear exceptionally adept at thinking on her feet.

This way of thinking is so fundamental that children also seem to follow it. Students often eagerly proclaim they have not studied hard for an exam—not only when they have performed poorly but also when they have performed exceptionally well. By stressing their low effort or even understating it, they may be trying to have their score attributed to their presumptive brilliance. The desire to project "effortless perfection" has been documented among university students, who often deliberately hide how hard they study (Travers et al., 2015; Casale et al., 2016).

Despite how fundamental this way of thinking is, to the best of my knowledge, no prior work has studied the following problem: evaluating people when—to affect how a composite measure of their various virtues is interpreted—they can strategically withhold evidence that both (i) is, in principle, favorable to them and (ii) contains useful information for the evaluator. I study the problem in the following principal-agent setting. In the baseline setting, the agent has a bidimensional type.[2] The first dimension is her *training* (e.g., a college or job applicant's socioeconomic background, effort, and training, an employee's effort, a researcher's knowledge) and the second is her *talent* (e.g., a college or job applicant's innate ability, an employee's efficiency or managerial skills, a researcher's ability to think fast).[3] The agent can present hard evidence to prove any part of her training but cannot prove her training is not even higher than what the evidence she has presented suggests. She cannot unilaterally prove anything about her talent.

The value of the agent to the principal (i) is non-decreasing in both training and talent and (ii) can be positive or negative. The principal ultimately wants to make a

In a 2021 survey, 34% of white Americans admitted to lying about being a racial minority on their college application (see https://www.intelligent.com/34-of-white-college-students-lied-about-their-race-to-improve-chances-of-admission-financial-aid-benefits). 48% of people who lied claimed to be Native American, and 3/4 of those who lied were accepted by the colleges that they lied to.

[2]Section 3.5 generalizes the results to types of any finite dimension.

[3]Although plausibly endogenous in some cases (e.g., when a college admissions committee decides whether to admit an applicant who can choose to withhold evidence of effort), I solve the problem for exogenous evidence and then extend the model to allow for endogenous evidence production. Section 5.3 shows that the structure of the optimal mechanism remains qualitatively the same even if evidence is endogenous (i.e., produced by the agent before her interaction with the principal), as long as the principal cannot influence evidence production by committing to a mechanism before the agent produces evidence.

binary choice: accept the agent (and receive the value of the agent as payoff) or reject her (and receive payoff 0). He does so by committing to a mechanism that asks the agent to (i) present evidence of training and (ii) make a cheap talk statement about her talent. Conditional on the evidence presented and the cheap talk statement made, the mechanism then (i) either verifies the value of a composite measure of the agent's training and talent and then accepts or rejects her conditional on that value or (ii) makes the acceptance or rejection decision without verification. The latter option is relevant because verification is (possibly) costly to perform. The composite measure is an increasing scalar function of the agent's training and talent. The agent wants to get accepted independently of her type.

If the composite measure measures exactly what the principal values in an agent (so that the principal's preference is to accept the agent if and only if the composite measure is high enough), then the value of verification is apparent. But what happens if the principal values talent (relative to training) to a different degree than the composite measure reflects talent (relative to training)? In other words, what if the principal's marginal rate of substitution between talent and training differs from the marginal rate of substitution between talent and training in the composite measure (i.e., holding fixed the value of the composite measure)?

If the composite measure is more sensitive to talent than talent is valuable to the principal, verification does not create incentives for the agents to withhold evidence. Then, the principal can ask for evidence and at the same time verify the value of the composite measure without having to worry about the agent withholding evidence. The main result concerns the optimal screening mechanism in the opposite case: when the composite measure is *less* sensitive to talent than talent is valuable to the principal. In that case, the optimal evaluation scheme never combines evidence and verification in the evaluation of a certain agent. Rather, it asks for evidence of training only to accept some high-training agents *without* verification. The optimal mechanism favors high- over low-training agents: (i) It accepts some high-training agents—including unworthy ones (i.e., who give the principal a negative payoff when accepted)—without verifying their composite measure but rather only by asking them for a certain level of evidence of training; and (ii) among agents who do not meet that threshold level of evidence, (iia) it accepts (after verification) some unworthy agents with high training but low talent while (iib) rejecting some worthy agents with high talent but low training.

Remarkably, this is the structure of the optimal mechanism in the extreme case where the principal *only* values talent (i.e., his payoff for accepting the agent is increasing in talent and constant in training).[4] The principal still optimally favors high-training agents even though training is worthless to him. He does so for two reasons: (i) to save on

[4]When the principal's payoff for accepting the agent depends only on talent, the composite measure is automatically less sensitive to talent than talent is valuable to the principal.

verification costs by accepting high-training agents without verifying their composite measure and (ii) because agents can withhold evidence of training to imitate more talented agents.

There is an important interaction between the two forces. The first force pushes towards high-training agents being favored—even in the extreme case where training is worthless—only to the extent that the second force is also present. To see this, assume that the principal verifies every agent's composite measure. Because agents can withhold evidence of training, the principal will optimally accept some unworthy high-training agents, who can withhold evidence of training to manipulate the interpretation of the composite measure. Now allow the principal to also choose which agents' composite measure to verify. Then, if verification is costly, he may optimally choose to not verify some high-training agents' composite measure, accepting them without verification. While this means that the principal will accept even more unworthy high-training agents that he would if he chose to verify everyone's composite measure, this increase in false acceptances can be more than counterbalanced by the decrease in verification costs.

The results capture a stark contrast in the difficulty of hiring different types of employees. When training (that can be proven through hard evidence) is most valuable, the hiring process is easy. On the other hand, when talent—which is assessed through a composite measure that is also sensitive to training—is most valuable, the hiring process is flawed, favoring candidates with high training at the expense of equally or more valuable candidates with great talent but limited training.

The results have implications for hiring, promotions, and college admissions. In the context of promotions, training can be understood as the employee's effort, and talent can be understood as her efficiency (i.e., the rate at which effort translates into productivity or value to the firm) or managerial skills.[5] The employer can verify the employee's productivity. Then, the payoff to the principal from accepting (i.e., promoting) the employee is the difference between her productivity in the new position (if promoted) and her productivity in her current position. The payoff is, as assumed, non-decreasing in effort and efficiency if both effort and efficiency have a (weakly) higher marginal productivity in the higher position. This is indeed the case if the higher position comes with increased responsibilities that allow the employee's effort and talent to have a larger impact. Talent being (relative to effort) more important in the higher position than in the current one is also a natural assumption. Then, the composite measure (i.e., current productivity) is less sensitive to talent than talent is valuable to the employer, which means some hard-working employees are (optimally) promoted—either with or without their productivity verified—although their promotion destroys firm value. At the same time, some talented but not hard-working employees are not promoted to managerial

[5]The employee has chosen effort in a previous stage (see section 5.3 for a discussion of endogenous evidence production) and can show or hide how much effort she has exerted.

positions, although their promotion would generate value for the firm.

Consider, now, hiring by a prestigious employer. Evidence is the candidate's CV quality (e.g., high school quality, undergraduate institution quality and GPA, awards, distinctions, reference letters), and talent is her ability and drive not captured by the evidence. Verification amounts to letting a less prestigious employer hire the candidate—with the option to poach the candidate later at a cost after observing her performance with that employer. In the optimal mechanism, Ivy-Leaguers are immediately hired by prestigious employers, whereas worthy candidates with less impressive credentials have to go through less prestigious employers to prove their worth before they land a prestigious position. If the candidates' performance in the less prestigious position is less sensitive to talent than talent is valuable in the more prestigious position, worthy candidates with low credentials are at a disadvantage not only in the first stage of hiring by the prestigious employer but also in the poaching stage.

Lastly, the results have implications for affirmative action in college admissions (i.e., trying to control for applicants' unequal backgrounds). Affirmative action is not very effective if both of the following conditions are satisfied: (i) College applicants can to a large extent hide their privilege, education, and preparation and (ii) standardized test scores reflect talent (e.g., relative to socioeconomic background and prior education, training, and test preparation) less than colleges value talent. If both conditions hold, the optimal admissions policy requires roughly the same test score from every applicant for admission—regardless of background. However, if any of the two conditions fails, affirmative action is effective, and we should expect its reversal to significantly reduce diversity in college admissions.

The hiring and college admissions applications combined illustrate how inequalities can be perpetuated. When standardized tests are under-sensitive to talent, college applicants from privileged backgrounds with superior access to high-quality education and extensive preparation have an advantage over equally or more worthy candidates from disadvantaged backgrounds. Upon graduation, those from prestigious institutions have an advantage in the labor market over more worthy candidates from less prestigious institutions.

After a discussion of related literature, section 2 presents the model. Section 3 characterizes incentive-compatible mechanisms and then solves the principal's problem. Section 4 discusses applications. Section 5 presents extensions of the model. Section 6 concludes. Proofs are gathered in Appendix A.

Related literature. This paper contributes to the multidimensional screening literature (see, e.g., Armstrong, 1996; Rochet and Choné, 1998; Rochet and Stole, 2003). Although duality approaches have proven useful in verifying a mechanism's optimality (Rochet and Choné, 1998; Carroll, 2017; Daskalakis et al., 2017; Cai et al., 2019), full characterizations of multidimensional screening problems remain challenging. Partial characterizations have, for example, been obtained (i) for the case where the principal can use costly instruments in screening (Yang, 2022) or (ii) that derive sufficient conditions for menus with specific characteristics to be optimal for a multiproduct monopolist (Haghpanah and Hartline, 2021; Yang, 2023). I advance this literature by proposing a novel and insightful multidimensional screening problem and deriving a full characterization under general assumptions. My analysis does not rely on ironing procedures (see, e.g., Mussa and Rosen, 1978; Myerson, 1981; Rochet and Choné, 1998) or the duality approach. Instead, I show that the principal's problem can be reduced to maximizing a linear (and thus convex) and continuous functional over a (convex and compact) space of monotone functions. Bauer's maximum principle then implies an extreme point solves the problem.[6] The proof proceeds using properties of extreme points of spaces of monotone functions. In that sense, my paper is also related to recent papers that characterize extreme points of spaces of monotone functions (see, e.g., Kleiner et al., 2021; Yang and Zentefis, 2024; Yang and Yang, 2025).

This paper also fits into the literature on models with costly verification. A main difference between my model and existing models with costly verification is that in existing work, verification amounts to either the revelation of the agent's one-dimensional type (see, e.g., Townsend, 1979; Gale and Hellwig, 1985; Dunne and Loewenstein, 1995; Ben-Porath et al., 2014; Bizzotto et al., 2020; Erlanson and Kleiner, 2020; Halac and Yared, 2020; Li, 2020; Kattwinkel and Knoepfle, 2023) or the revelation of one dimension of the agent's multi-dimensional type (see, e.g., Glazer and Rubinstein, 2004; Carroll and Egorov, 2019; Li, 2021). Therefore, the interpretation of the verification result is not influenced by the agent's initial disclosure as in my model, where the substitutability between the different dimensions is key.

Nevertheless, the composite measure that verification reveals is not entirely new to the literature. It is reminiscent of the signal jamming problem in career concern models (see, e.g., Holmström, 1999). Still, in these models the main force is the agent's incentives to exert effort in order to influence the principal's learning (though costless observation of the agent's productivity) of the agent's talent. Here, I focus on information transmission and testing.[7] I show that if the principal can ask for hard evidence of effort, the signal jamming problem is mitigated if productivity is sensitive enough to talent—compared to the principal's preferences for accepting (e.g., promoting) the agent. However, when productivity is *not* sensitive enough to talent, the signal jamming problem persists even if the principal can ask for evidence of effort. Agents have incentives to withhold evidence, which they should be paid information rents to reveal.

The paper has links to a few other strands of the literature, particularly persuasion

[6] Manelli and Vincent (2007) also use Bauer's maximum principle to study a multidimensional screening problem.

[7] Other differences from career concerns models is that the principal has commitment power and chooses whether to test the agent at a cost.

games (Viscusi, 1978; Grossman, 1981; Milgrom, 1981), evidence games (see, e.g., Shin, 1994; Dziuda, 2011; Hart et al., 2017), and models with signal manipulation (Frankel and Kartik, 2019, 2022; Perez-Richet and Skreta, 2022; Ball, 2025) or costly lying (e.g., Kartik, 2009; Sobel, 2020).

2 A model of multidimensional screening with substitutable attributes and costly verification

There are an agent (she) and a principal (he). The agent is privately informed of her bidimensional type (e,t), which has a full-support density $f : [0,1]^2 \to \mathbb{R}_{++}$. e is the agent's *evidence*. An agent of type (e,t) can present any level of evidence $r \in [0,e]$. By presenting evidence r she proves that her e is at least r. However, she cannot prove that she is not withholding evidence. t is the agent's *talent*, which she cannot unilaterally prove anything about.[8] The principal can test the agent by paying a cost $c \geq 0$.

The testing technology. Testing the agent amounts to observing a deterministic signal $\sigma(e,t) \in [0,1]$ of the agent's type (e,t). $\sigma : [0,1]^2 \to [0,1]$ is increasing and continuous in e and t. The assumption of a deterministic signal is not uncommon. In fact, it is more general than the assumption that the test reveals one of the dimensions of the agent's type, which is for example made in Glazer and Rubinstein (2004), Carroll and Egorov (2019), and Kattwinkel and Knoepfle (2023).[9]

Payoffs. Ultimately, the principal must decide whether to accept or reject the agent. He receives (gross of testing costs) Bernoulli payoff $u(e,t)$ from accepting an agent of type (e,t), where $u : [0,1]^2 \to \mathbb{R}$ is non-decreasing and continuous in e and t. If he rejects the agent, he receives payoff normalized to 0. An isocurve of the principal's (gross) payoff is given by $I_u(\overline{u}) := \{(e,t) \in [0,1]^2 : u(e,t) = \overline{u}\}$.[10] The agent's Bernoulli payoff is equal to 1 if accepted and 0 if rejected.

Canonical examples. In a linear specification, $u(e,t) := \gamma_u e + (1 - \gamma_u)t - \underline{q}$, where $\gamma_u \in [0,1]$ measures how much the principal values e versus t, and $\underline{q} \in (0,1)$ measures the threshold quality that the agent needs to have to be of (positive) value to the principal.

[8] It is straightforward to see that the model also captures the case where evidence also measures a combination of two qualities. Let type (e,t) be able to present any level of evidence $r \in [0,\varepsilon(e,t)]$, where $\varepsilon(e,t)$ is increasing in e and t. Then, we can redefine the agent's type to be $(\widetilde{e},t)$, where $\widetilde{e} := \varepsilon(e,t)$.

[9] That is, if we allow σ to be constant in e or t (but not both). If σ is constant in e (and, thus, reveals t exactly), the optimal mechanism is the same as under pro-t biased testing (see section 3). If σ is constant in t (and, thus, reveals e exactly), the test is useless, and there exists an optimal mechanism that only asks for evidence.

[10] $I_u(\overline{u})$ is assumed to be a curve for any $\overline{u}$. This is the case if, for example, $u(e,t)$ is increasing in e or t.

Similarly, $\sigma(e,t) \coloneqq \gamma_s e + (1-\gamma_s)t$, where $\gamma_s \in (0,1)$ measures how sensitive the test is to e relative to t. In a Cobb-Douglas specification, $u(e,t) \coloneqq e^{\gamma_u} t^{1-\gamma_u} - \underline{q}$ and $\sigma(e,t) \coloneqq e^{\gamma_s} t^{1-\gamma_s}$ with $\gamma_u \in [0,1]$ and $\gamma_s, \underline{q} \in (0,1)$. No parametric assumptions are imposed on u or σ. For simplicity in depiction, all figures use the linear specification.

The principal's problem. To decide whether to accept the agent, the principal commits to a direct mechanism $M \equiv \langle T,P \rangle$ that specifies: (i) the probability $T(e,t) \in [0,1]$ with which the principal will test the agent if she presents evidence e and sends cheap talk message t and (ii) the probability $P(e,t,s)$, which should be non-decreasing in $s \in [0,1]$, with which the principal will accept the agent after the agent has presented evidence e and sent cheap talk message t and the test has returned result $s \in [0,1]$.[11] If no test is performed, $s = \emptyset$ and the agent is accepted with probability $P(e,t,\emptyset)$. Notice that (e,t) refers to the message sent by the agent. When necessary to avoid confusion, we will denote by (e',t') the agent's message to distinguish it from the agent's type, which in those cases will be denoted by (e,t). Overall, the principal designs a mechanism $M \equiv \langle T,P \rangle$, where $T : [0,1]^2 \to [0,1]$ and $P : [0,1]^2 \times ([0,1] \cup \{\emptyset\}) \to [0,1]$ with $P(e,t,s)$ non-decreasing in $s \in [0,1]$, and (breaking the agent's indifferences in his favor) an agent response rule $\phi : [0,1]^2 \to [0,1]^2$ to maximize

$$\int_0^1 \int_0^1 \left\{ \underbrace{\begin{bmatrix} T(\phi(e,t))P(\phi(e,t),\sigma(e,t)) \\ +[1-T(\phi(e,t))]P(\phi(e,t),\emptyset) \end{bmatrix}}_{\substack{\text{total probability that } (e,t) \text{ is accepted,} \\ \text{either (i) after getting tested or} \\ \text{(ii) without a getting tested}}} u(e,t) - c\underbrace{T(\phi(e,t))}_{\substack{\text{probability} \\ \text{that}(e,t) \\ \text{is tested}}} \right\} f(e,t)dtde$$

subject to the agent's incentive compatibility (IC) constraint

$$\phi(e,t) \in \underset{(\widehat{e},\widehat{t}) \leq (e,1)}{\arg\max} \underbrace{\left\{ T(\widehat{e},\widehat{t})P(\widehat{e},\widehat{t},\sigma(e,t)) + (1-T(\widehat{e},\widehat{t}))P(\widehat{e},\widehat{t},\emptyset) \right\}}_{\text{total probability that } (e,t) \text{ is accepted if she reports } (\widehat{e},\widehat{t})}.$$

3 Optimal multidimensional screening with substitutable attributes and costly verification

This section characterizes incentive-compatible (IC) mechanisms and then solves the principal's problem.

[11]The condition that $P(e,t,s)$ be non-decreasing in $s \in [0,1]$ can be understood as an incentive-compatibility condition in a model where $\sigma(e,t)$ is the (maximum) score that agent type (e,t) can achieve, and the agent can intentionally underperform.

3.1 Simplifying the class of mechanisms

Before characterizing IC mechanisms, we show that we can without loss restrict the class of mechanisms that we need to consider.

Truthful mechanisms are without loss. The first simplification is that the principal can without loss of optimality restrict attention to truthful mechanisms (i.e., mechanisms that induce truth-telling). To see why, notice that the correspondence $(e,t) \mapsto \{(e',t') \in [0,1]^2 : e' \le e\}$, which maps each agent type (e,t) to the messages she can send, satisfies the Nested Range Condition of Green and Laffont (1986), who show that under this condition, the set of implementable social choice functions coincides with the set of truthfully implementable social choice functions.[12] Therefore, we define IC mechanisms as follows.

Definition 1. A mechanism $M \equiv \langle T,P \rangle$ is IC if for every $(e,t) \in [0,1]^2$

$$(e,t) \in \underset{(e',t') \in [0,e] \times [0,1]}{\arg\max} \left\{ T(e',t')P(e',t',\sigma(e,t)) + (1 - T(e',t'))P(e',t',\emptyset) \right\}.$$

Pass-or-fail tests are without loss. Next, we can constrain attention to mechanisms with threshold acceptance policies conditional on testing; that is, mechanisms such that

$$P(e,t,s) = \begin{cases} 0 & \text{if } s < \sigma(e,t) \\ P_{at}(e,t) & \text{if } s \ge \sigma(e,t) \end{cases} \tag{1}$$

for any (e,t) and some $P_{at} : [0,1]^2 \to [0,1]$, where *at* is a mnemonic for the probability of accepting the agent *after testing* (given that the threshold composite measure is met). If type (e,t) reports her type truthfully and is tested, she is then accepted with probability $P_{at}(e,t)$. Notice that the threshold is set exactly equal to the composite measure that a truthfully-reporting agent can achieve. To see why constraining attention to such mechanisms is without loss of optimality, observe that among all mechanisms that (conditional on testing) accept type (e,t) with probability $P_{at}(e,t)$, the one that satisfies equation (1) minimizes incentives of other types to imitate (e,t).[13]

We can further restrict attention to mechanisms that accept the agent with certainty if she meets the composite measure threshold (i.e., $P_{at}(e,t) = 1$ for every (e,t)). To see why,

[12]Essentially, the principal implements a social choice function $g : [0,1]^2 \to [0,1]^2 \times [0,1]^{[0,1]}$, where $g_1(e,t)$ the probability of testing, $g_2(e,t)$ the probability of acceptance conditional on not testing, and $g_3(e,t,\cdot)$ a self-map on $[0,1]$ that (conditional on testing) maps the test result s to the probability $g_3(e,t,s)$ of acceptance.

[13]Namely, accepting the agent with even higher probability for performing above $\sigma(e,t)$ will result in the same probability of accepting type (e,t) after testing her and only give additional incentives to other agents to imitate (e,t). Similarly, there is no reason to accept the agent for composite measures lower than $\sigma(e,t)$. Particularly, this argument holds when we compare all mechanisms that test (e,t) with the same probability, and, thus have the same testing costs.

denote the total probability with which agent (e,t) is accepted if she truthfully reports her type by $\Pi(e,t) \coloneqq (1 - T(e,t))P(e,t,\emptyset) + T(e,t)P_{at}(e,t)$, and define outcome-equivalent mechanisms as follows.

Definition 2. A mechanism $M' \equiv \langle T',P' \rangle$ is outcome-equivalent to a mechanism $M \equiv \langle T,P \rangle$ if for every (e,t), $\Pi(e,t) = \Pi'(e,t)$, where $\Pi(e,t) \equiv (1-T(e,t))P(e,t,\emptyset)+T(e,t)P_{at}(e,t)$ and $\Pi'(e,t) \equiv (1 - T'(e,t))P'(e,t,\emptyset) + T'(e,t)P'_{at}(e,t)$.

Lemma 1 shows that when testing is costly, an agent who is tested and passes the test is accepted with probability 1 in any optimal mechanism. When testing is free, it is still without loss to constrain attention to mechanisms that accept the agent with probability 1 when she passes the test.

Lemma 1. Given any IC mechanism M, there exists an IC mechanism $M' \equiv \langle T',P' \rangle$ with $P'_{at}(e,t) = 1$ for every (e,t) that is outcome-equivalent to M. Also, for $c > 0$, in any optimal mechanism $M \equiv \langle T,P \rangle$, $P_{at}(e,t) = 1$ for any (e,t) such that $T(e,t) > 0$.[14]

The intuition behind this result is as follows. The only reason to test an agent before accepting her—rather than accept her without a test—is to prevent others from imitating her. The total probability with which each agent is accepted is the sum of (i) the probability $(1 - T(e,t))P(e,t,\emptyset)$ of being accepted without getting tested and (ii) the probability $T(e,t)P_{at}(e,t)$ of being accepted after getting tested (and passing the test). Simply put, if the principal pays the cost to test an agent, he may as well assign as large a part as possible of the total probability of accepting her to the case where he accepts her after a test.

Specifically, if agent (e,t) is not accepted with certainty after passing the test (i.e., $P_{at}(e,t) < 1$ and $T(e,t) > 0$), we can (i) increase the probability $P_{at}(e,t)$ with which she is accepted conditional on getting tested (and passing the test), (ii) decrease the probability $T(e,t)$ with which she is tested, and (iii) decrease (if positive) the probability $P(e,t,\emptyset)$ of accepting her conditional on not testing her, keeping fixed both (a) the probability $(1 - T(e,t))P(e,t,\emptyset)$ of accepting her without testing and (b) the probability $T(e,t)P_{at}(e,t)$ of accepting her after testing. By doing so, we (i) keep fixed the total probability $\Pi(e,t)$ of accepting (e,t), (ii) do not change the incentives of other types to imitate (e,t), since any agent imitating (e,t) will be accepted with probability $(1 - T(e,t))P(e,t,\emptyset)$ (if she only has at least as much evidence as (e,t) but cannot test as high as her) or $\Pi(e,t)$ (if she can also test as high as (e,t)), and (iii) reduce the probability of testing (e,t), thereby limiting testing costs.

[14]Strictly put, $P_{at}(e,t)$ can be lower than 1 for a zero-measure set of (e,t) with $T(e,t) > 0$. For (e,t) with $T(e,t) = 0$, the value of $P_{at}(e,t)$ does not matter, so we can again set $P_{at}(e,t) = 1$ without loss.

3.2 Incentive-compatible mechanisms

Given what we have seen, we constrain attention to truthful mechanisms with pass-or-fail tests. Let $\tau(e,s)$ be implicitly given by $\sigma(e,\tau(e,s)) = s$. $\tau(e,s)$ gives the level of talent that an agent with evidence e should have to achieve composite measure (exactly) s. $\tau(e,s)$ is well-defined for (e,s) such that $s \in [0,1]$ and $e \in [\underline{e}(s),\overline{e}(s)]$, where $\underline{e}(s) \coloneqq \min\{e \in [0,1] : \sigma(e,1) \geq s\}$ and $\overline{e}(s) \coloneqq \max\{e \in [0,1] : \sigma(e,0) \leq s\}$.[15] Proposition 1 characterizes IC mechanisms.

Proposition 1. A mechanism $M \equiv \langle T,P\rangle$ is IC if and only if

(i) $\Pi(e,t)$ is non-decreasing in t for every $e \in [0,1]$,

(ii) $\Pi(e, \tau(e,s))$ is non-decreasing in e over $e \in [\underline{e}(s),\overline{e}(s)]$ for every $s \in [0,1]$, and

(iii) $(1 - T(e,t))P(e,t,\emptyset) \leq \Pi(e,0)$ for every $(e,t) \in [0,1]^2$,

where $\Pi(e,t) \equiv (1 - T(e,t))P(e,t,\emptyset) + T(e,t)$ is the probability with which agent (e,t) is accepted if she truthfully reports her type.

Figure 1 schematically summarizes IC conditions (i) and (ii) of Proposition 1.

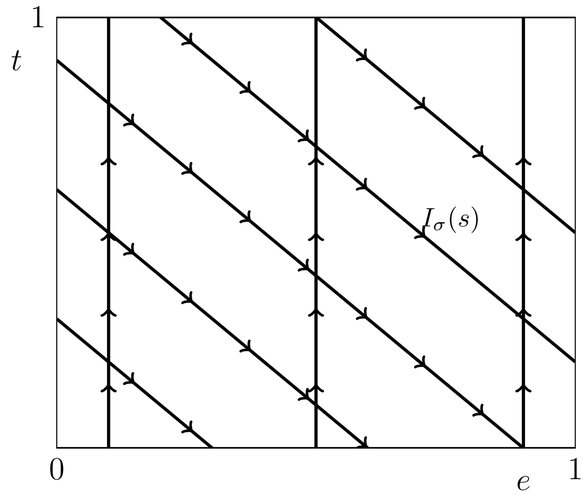
Condition (i) is necessary and sufficient to ensure that an agent (e,t) does not want to reveal her evidence but under-report her talent to imitate agent (e,t') with $t' < t$, pass (e,t')'s test (since the composite measure is increasing in talent), and get accepted with probability $\Pi(e,t')$.

Condition (iii) is necessary and sufficient to ensure that no untalented agent $(e,0)$ has incentives to over-report her talent, imitating an agent (e,t)—whose composite measure she *cannot* achieve—and possibly getting accepted in case she is not tested. Put differently, among agents with the same level of evidence e, in order to accept talented agents more frequently (than the untalented agent $(e,0)$), the principal needs to test them with high enough probability to prevent agent $(e,0)$ from imitating them. Conditions (i) and (iii) combined also imply that $\Pi(e,t) \geq \Pi(e,0) \geq (1 - T(e,t'))P(e,t',\emptyset)$ for every e,t,t', so no agent has incentives to present all her evidence but overstate her talent.

Last, condition (ii) is necessary and sufficient to ensure that agents do not want to withhold some of their evidence in order to overstate their talent, thereby imitating agents whose composite measure they *can* achieve. Namely, an agent (e,t) does not want to imitate an agent (e',t') with less evidence $e' < e$, more talent $t' > t$, and equal composite measure $\sigma(e',t') = \sigma(e,t)$ to get accepted with probability $\Pi(e',t')$ instead of $\Pi(e,t)$. Notice that for any possible level of evidence $e' < e$ that agent (e,t) may reveal, if it is not

[15] $\underline{e}(s)$ (resp. $\overline{e}(s)$) is the minimum (resp. maximum) level of evidence that an agent can have while achieving composite measure (exactly) s. That is, agents with evidence lower than $\underline{e}(s)$ score less than s even if they have talent $t = 1$. Analogously, agents with evidence higher than $\overline{e}(s)$ score more than s even if they have talent $t = 0$.

Figure 1: Directions of (weak) increase in $\Pi(e,t)$ in IC mechanisms



Note: the arrowed lines show the directions in which $\Pi(e,t)$ is non-decreasing in IC mechanisms.

profitable for (e,t) to overstate her talent so much that she will fail the test if tested, then because of condition (i), she will want to overstate her talent as much as possible (making sure that she will be able to pass the test), up to the point where $\sigma(e',t') = \sigma(e,t)$.

We have so far seen that conditions (i), (ii), and (iii) are necessary and sufficient for the agent not to have incentives to deviate in any of the following three ways: (a) present all her evidence but under-report her talent, (b) present all her evidence but overstate her talent, or (c) withhold some of her evidence and overstate her talent, imitating agents whose composite measure she *can* achieve. To see why they are necessary and sufficient for IC, it remains to observe that these conditions also rule out the fourth type of deviations by the agent: hiding evidence and overstating talent to imitate agents whose composite measure she *cannot* achieve. To see this, notice that conditions (i), (ii), and (iii) combined imply that $\Pi(e,t) \geq \Pi(e,0) \geq \Pi(e',0) \geq (1 - T(e',t'))P(e',t',\emptyset)$ for any $e' < e$, where the second inequality follows from conditions (i) and (ii) combined, ensuring that (e,t) does not want to withhold evidence and overstate her talent so much (to a point where $\sigma(e',t') > \sigma(e,t)$) that she fails the test.

Condition (iii) of Proposition 1 is satisfied with equality. Lemma 2 shows that when testing is costly and some talented agents are (optimally) accepted with higher probability than untalented ones with the same level of evidence, the optimal mechanism satisfies condition (iii) of Proposition 1 with equality. Under free testing or when it is not optimal to accept talented agents with higher probability, it is still without loss to constrain attention to mechanisms that satisfy condition (iii) of Proposition 1 with equality.

Lemma 2. Given any IC mechanism $M \equiv \langle T,P \rangle$, there exists an IC mechanism $M' \equiv$

$\langle T',P'\rangle$ with $(1-T'(e,t))P'(e,t,\emptyset)=\Pi'(e,0)$ for every (e,t) that is outcome-equivalent to M and has at most as high testing costs as M. For $c>0$, if also $\Pi(e,t)>\Pi(e,0)$ for a positive measure of agent types, then M' has lower testing costs than M.

Here is the intuition behind this result. Take any IC mechanism $M\equiv\langle T,P\rangle$. When $\Pi(e,0)>(1-T(e,t))P(e,t,\emptyset)$, it means that untalented agent $(e,0)$ strictly prefers to not overstate her talent. This strict preference is due to overtesting of talented agents. Namely, $T(e,t)$ can be reduced and $P(e,t,\emptyset)$ can be increased keeping $\Pi(e,t)\equiv(1-T(e,t))P(e,t,\emptyset)+T(e,t)$ fixed while maintaining $(1-T(e,t))P(e,t,\emptyset)\leq\Pi(e,0)$, so that condition (iii) of Proposition 1 is still satisfied.[16] Conditions (i) and (ii) of Proposition 1 are also still satisfied since Π does not change. Then, talented agents are tested with lower but high enough probability to prevent untalented agents from imitating them.

From now on, we constrain attentions to mechanisms with $(1-T(e,t))P(e,t,\emptyset)=\Pi(e,0)$, or equivalently, $\Pi(e,t)=\Pi(e,0)+T(e,t)$, for every (e,t). The total probability of accepting the agent has two components: (i) a base probability $\Pi(e,0)$ of accepting the agent for her evidence without a test and (ii) an additional probability $T(e,t)$ of accepting the agent for her talent, which through testing allows her to differentiate herself from less talented agents with the same level of evidence.

3.3 Optimal screening under free testing

We are now ready to characterize the optimal mechanisms under free testing (i.e., $c=0$). The principal's objective function is $\int_0^1\int_0^1\Pi(e,t)u(e,t)f(e,t)dtde$, which can be written as

$$\int_0^1\int_{\underline{e}(s)}^{\overline{e}(s)}\Pi(e,\tau(e,s))u(e,\tau(e,s))f(e,\tau(e,s))deds, \tag{2}$$

where instead of integrating over e and t, we integrate over composite measure s and e. The principal's problem amounts to choosing $\Pi(e,\tau(e,s))$, seen as a function of (e,s), non-decreasing in s (condition (i) of Proposition 1) and e (condition (ii) of Proposition 1) to maximize (2), which is linear (and, thus, convex) in Π.[17] Bauer's maximum principle then implies that there exists an extreme Π (i.e., an extreme point of the space of non-decreasing functions from $\{(e,s)\in[0,1]^2: e\in[\underline{e}(s),\overline{e}(s)]\}$ to $[0,1]$) that maximizes (2). It is an extension of Lemma 2.7 in Börgers (2015) that an extreme Π maps each (e,s) to either 0 or 1.

[16] Notice that by decreasing $T(e,t)$ and increasing $P(e,t,\emptyset)$ while keeping $\Pi(e,t)$ fixed, we increase $(1-T(e,t))P(e,t,\emptyset)$. This is possible to do while maintaining $(1-T(e,t))P(e,t,\emptyset)\leq\Pi(e,0)$ because $\Pi(e,0)>(1-T(e,t))P(e,t,\emptyset)$ to start with.

[17] Condition (iii) of Proposition 1 is immaterial, since testing is free. As implied by Lemma 2, any Π that satisfies conditions (i) and (ii) of Proposition 1 can be implemented with T and P such that $(1-T(e,t))P(e,t,\emptyset)=\Pi(e,0)$ for every (e,t). However, there are many other ways to implement any Π that satisfies conditions (i) and (ii). For example, setting $P(e,t,\emptyset)=0$ and $T(e,t)=\Pi(e,t)$ for every (e,t) (i.e., nobody is ever accepted without a test) automatically satisfies condition (iii) of Proposition 1.

Lemma 3. Let $c = 0$. There exists an optimal deterministic mechanism (i.e., an optimal mechanism where $\Pi(e,t) \in \{0,1\}$ for all (e,t)).

3.3.1 Testing technology biased in favor of talent

We are now ready to derive the optimal mechanism. Consider first the case where the testing technology is biased in favor of talent in the sense that the test is more sensitive to talent than talent is valuable to the principal.[18]

Definition 3. σ is pro-t biased if for every composite measure $s \in [0,1]$ there exists e_s such that for every (e,t), if $e > e_s$ (resp. $e < e_s$) and $\sigma(e,t) = s$, then $u(e,t) > c$ (resp. $u(e,t) < c$).

This is a single-crossing condition. It says that iso-test-score curves cross the principal's indifference curve $I_u(c)$ "from below" (see Figure 3(a)). Here is the intuition behind the definition. Because the test is overly sensitive to talent, it is too generous towards those with high talent and low evidence and too strict towards those with low talent and high evidence. Therefore, among all agents with the same composite measure (if tested), the principal would like to accept those with high but not those with low evidence.

Clearly, if the principal's payoff from accepting the agent is increasing along iso-test-score curves, σ is pro-t biased. This is the case if the principal's MRS of talent for evidence is higher (in absolute value) than the test's MRS of talent for evidence.

Claim 1. If $u(e,\tau(e,s))$ is increasing in e over $e \in [\underline{e}(s),\overline{e}(s)]$ for every $s \in [0,1]$, then σ is pro-t biased (for any c). The condition is satisfied if $\frac{\partial u(e,t)/\partial e}{\partial u(e,t)/\partial t} > \frac{\partial \sigma(e,t)/\partial e}{\partial \sigma(e,t)/\partial t}$ for every (e,t).

First, notice that if the principal only values evidence, σ is automatically pro-t biased. Then, he can trivially achieve the first best without the need to test—much like in the case where talent was absent from the model. Namely, accepting every agent with sufficient evidence to be of positive value to the principal is IC, because it does not create incentives for agents to withhold and/or overstate their talent as only evidence is rewarded.

Allowing for the principal to also value talent, Proposition 2 shows that when testing is (i) free and (ii) pro-t biased, the principal can still achieve the full information benchmark.

Proposition 2. Let $c = 0$, and assume that σ is pro-t biased. Then, $\Pi(e,t) = \mathbf{I}\left(u(e,t) \geq 0\right)$ is IC, so the principal achieves the full information first-best.

If the principal also values talent but less strongly than the composite measure depends on talent, agents do not have incentives to withhold evidence in order to overstate their talent when the principal accepts every agent of positive value. To achieve the first-best, the principal needs to both ask for evidence and conduct the test. Namely, because there

[18] We define pro-t biased testing for any testing cost c. The optimal mechanism under costly testing is studied in section 3.4.

exist e,t,t' with $t' > t$ such that $u(e,t') > 0 > u(e,t)$, he needs to use the test to accept (e,t') but not (e,t). Figure 3(a) presents the optimal mechanism under pro-t biased testing.

Lemma 2 restricts attention to the following way of implementing the first-best Π: setting $T(e,t) = \mathbf{I}(u(e,t) \geq 0 \wedge u(e,0) < 0)$ and $P(e,t,\emptyset) = \mathbf{I}(u(e,0) \geq 0)$. That is, agents who are not valuable to the principal truthfully report their type and are rejected without getting tested. Agents who are valuable but cannot prove so by presenting evidence e such that $u(e,0) > 0$ (which would prove that even if they have $t = 0$, they are valuable) are tested and then accepted. Finally, agents who can prove that they are valuable by presenting evidence e such that $u(e,0) \geq 0$ do so and are accepted without a test. Clearly, since testing is free, $T(e,t) = \mathbf{I}(u(e,t) \geq 0)$ and $P(e,t,\emptyset) = 0$ for every (e,t) is, for example, also optimal, as is testing every agent and accepting only the valuable ones.

3.3.2 Testing technology biased in favor of evidence

Consider now the case where the testing technology is biased in favor of evidence in the sense that the test is more sensitive to evidence than evidence is valuable to the principal—or equivalently, the test is less sensitive to talent than talent is valuable to the principal.[19]

Definition 4. σ is pro-e biased if for every composite measure $s \in [0,1]$ there exists e_s such that for every (e,t), if $e < e_s$ (resp. $e > e_s$) and $\sigma(e,t) = s$, then $u(e,t) > c$ (resp. $u(e,t) < c$).

This is again a single-crossing condition. It says that iso-test-score curves cross the principal's indifference curve $I_u(c)$ "from above" (see Figure 3(b)). Here is the intuition behind the definition. Because the test is overly sensitive to evidence, it is too generous towards those with high evidence and low talent and too strict towards those with low evidence and high talent. Therefore, among all agents with the same composite measure (if tested), the principal would like to accept those with low but not those with high evidence.

Clearly, if the principal's payoff from accepting the agent is decreasing along iso-test-score curves, σ is pro-e biased. This is the case if the principal's MRS of talent for evidence is lower (in absolute value) than the test's MRS of talent for evidence.

Claim 2. If $u(e,\tau(e,s))$ is decreasing in e over $e \in [\underline{e}(s),\overline{e}(s)]$ for every $s \in [0,1]$, then σ is pro-e biased (for any c). The condition is satisfied if $\frac{\partial u(e,t)/\partial e}{\partial u(e,t)/\partial t} < \frac{\partial \sigma(e,t)/\partial e}{\partial \sigma(e,t)/\partial t}$ for every (e,t).

The first-best is no longer achievable.[20] Indeed, Figure 2 shows that accepting (almost) every agent with $u(e,t) > 0$ and rejecting (almost) every agent with $u(e,t) < 0$ is not IC,

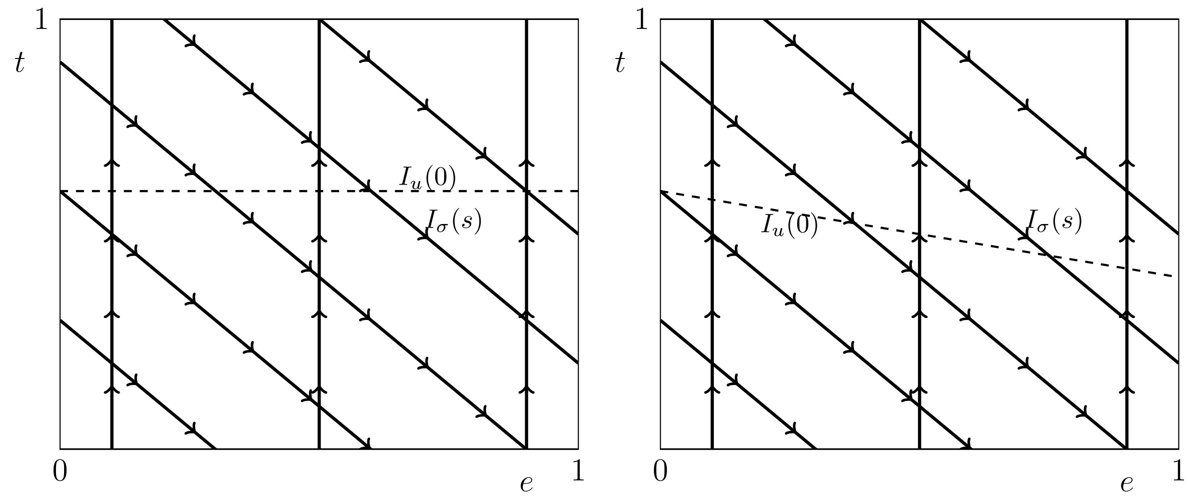
[19] We define pro-e biased testing for any testing cost c. The optimal mechanism under costly testing is studied in section 3.4.

[20] σ being pro-e biased is not necessary for this conclusion. The conclusion still applies as long as the conditions in definition 4 is satisfied for a positive measure of $s \in [0,1]$.

Figure 2: *Not* achieving the first-best: testing technology biased in favor of evidence

(a) Only talent valuable to the principal

(b) Test less sensitive to talent than talent valuable to the principal



Note: the arrowed lines represent the directions of (weak) increase in $\Pi(e,t)$ in any IC mechanism. The dashed lines represent the principal's indifference curve $I_u(0)$.

as it creates incentives for agents with $u(e,t) < 0$ to withhold evidence and imitate more talented and valuable agents.

But what *can* actually be achieved when the test is less sensitive to talent than talent is valuable to the principal? Proposition 3 describes the optimal mechanism when testing is (i) free and (ii) pro-e biased. Proposition 3 shows that in the optimal mechanism, agent (e,t) is accepted if and only if $\sigma(e,t) \geq s^*$.

Proposition 3. Let $c = 0$, and assume that σ is pro-e biased. Then, there exists an optimal mechanism with $\Pi(e,t) = \mathbf{I}(\sigma(e,t) \geq s^*)$.

Lemma 2 restricts attention to the following way of implementing this Π: setting $T(e,t) = \mathbf{I}(\sigma(e,t) \geq s^* \wedge e \leq \overline{e}(s^*))$ and $P(e,t,\emptyset) = \mathbf{I}(e > \overline{e}(s^*))$. That is, agents who cannot achieve composite measure at least s^* truthfully report their type and are rejected without getting tested. Agents who can achieve that composite measure and cannot prove this by presenting evidence $e > \overline{e}(s^*)$ (which would prove that even if they have $t = 0$, they can achieve composite measure s^*) are tested and then accepted. Finally, agents who can prove that they can meet the composite measure threshold by presenting evidence $e \geq \overline{e}(s^*)$ do so and are accepted without a test. Clearly, since testing is free, $T(e,t) = \mathbf{I}(\sigma(e,t) \geq s^*)$ and $P(e,t,\emptyset) = 0$ for every (e,t) is, for example, also optimal, as is testing every agent and accepting only those that pass the composite measure threshold s^*.

Finding the optimal mechanism is remarkably simple. It amounts to maximizing a continuous function of one variable over a closed interval. The principal needs to find

$s^* \in \arg\max_{\widetilde{s}\in[0,1]} v(\widetilde{s})$, where

$$v(\widetilde{s}) := \int_{\widetilde{s}}^{1} \int_{\underline{e}(s)}^{\overline{e}(s)} \widetilde{u}(e,s)\widetilde{f}(e,s)deds$$

and $\widetilde{u}(e,s) := u(e,\tau(e,s))$ and $\widetilde{f}(e,s) := f(e,\tau(e,s))$.[21] When $s^* \in (0,1)$, it solves $\int_{\underline{e}(s^*)}^{\overline{e}(s^*)} u(e,\tau(e,s^*))f(e,\tau(e,s^*))de = 0$. The principal effectively chooses a threshold composite measure s^* and accepts every agent who can achieve this score. In choosing this threshold, he balances the Type I (i.e., rejecting agents who lie above $I_u(0)$) and Type II (i.e., accepting agents who lie below $I_u(0)$) errors. This trade-off can be seen in Figure 3(b).

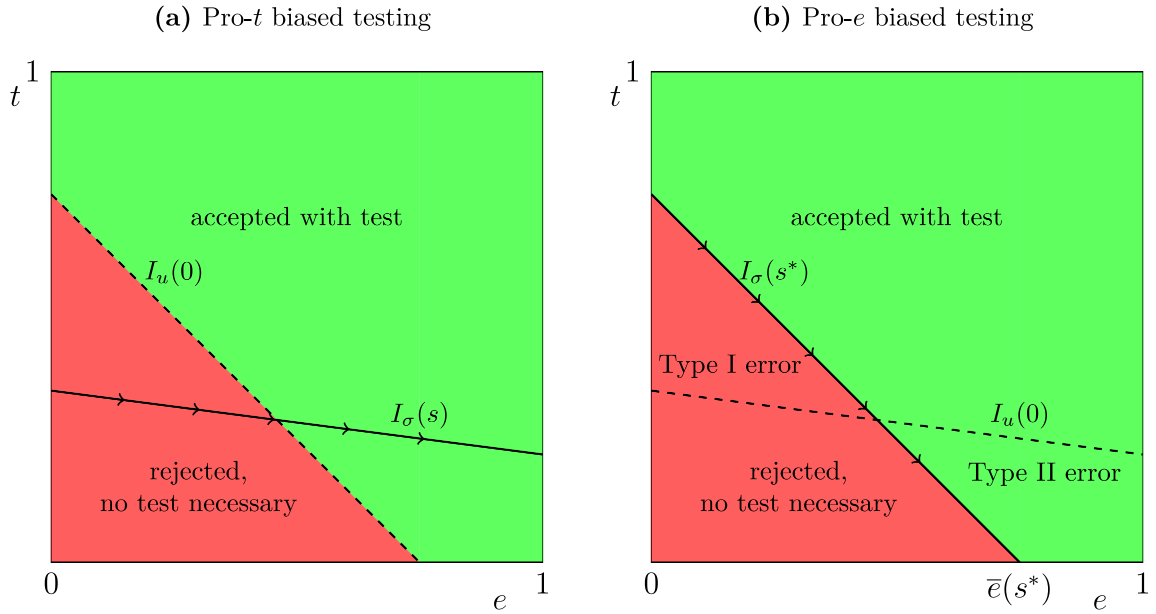
Here is a sketch of the proof of Proposition 3. Because σ is pro-e biased, for any two types of zero value to the principal $(e,t),(e',t') \in I_u(0)$ with $e' > e$, $\sigma(e',t') \geq \sigma(e,t)$. But then, if $\sigma(e',t') \geq \sigma(e,t)$ and $e' > e$, IC requires $\Pi(e',t') \geq \Pi(e,t)$. In other words, $\Pi(e,t)$ has to be non-decreasing as e increases along the $I_u(0)$ curve. Therefore, in any deterministic IC mechanism, there exists a threshold type on the $I_u(0)$ curve such that agents on the $I_u(0)$ curve with more (resp. less) evidence than the threshold type are accepted (resp. rejected). Next, observe that IC requires that $\Pi(e,t)$ be non-decreasing along iso-test-score curves (condition (ii) of Proposition 1). Thus, having fixed $\Pi(e,t)$ along the $I_u(0)$ curve, keeping $\Pi(e,t)$ constant along iso-test-score curves maximizes the principal's payoff. That is because, on the part of an iso-test-score curve that lies below (resp. above) $I_u(0)$, the principal wants to make $\Pi(e,t)$ as low (resp. high) as possible but is constrained to set $\Pi(e,t)$ at least (resp. most) equal to its value on the curve $I_u(0)$ for that specific composite measure level. Condition (i) of Proposition 1 is automatically satisfied.

Discussion. When seen against the results under pro-t biased testing (see Proposition 2), Proposition 3 reveals a stark contrast in the difficulty of hiring different types of employees. When skills and knowledge that can be proven through hard evidence are most valuable, the hiring process is easy. On the other hand, when talent is most valuable and assessed through tests and interviews that are overly sensitive to the candidate's training and preparation, the hiring process is flawed. It favors unworthy candidates with advanced training at the expense of those with limited training who are, however, more valuable to the firm.

The revealed difference in the difficulty of hiring talented versus well-trained employees can be partly the reason behind the fact that firm survival rates increase with firm age and size (Evans, 1987; Dunne and Hughes, 1994; Farinas and Moreno, 2000; Agarwal and Gort, 2002; Bartelsman et al., 2005). To the extent that start-ups often face new

[21]The principal's problem reduces to this because all mechanisms with $\Pi(e,t) = \mathbf{I}(\sigma(e,t) \geq s^*)$ and appropriate T are IC.

Figure 3: The optimal mechanism under free testing



Note: the dashed line represents the principal's indifference curve $I_u(0)$: the principal is indifferent between accepting and rejecting agents on that curve. The arrowed line represents an iso-test-score curve (at an arbitrary level s in the left panel and at the optimal level s^* in the right panel). The green (resp. red) area denotes the set of agents who are accepted (resp. rejected) in the optimal mechanism. The Type I error corresponds to the part of the red area that lies above the dashed line. The Type II error corresponds to the part of the green area that lies below the dashed line.

challenges without the established procedures or clearly defined roles of older and larger firms, the success of a start-up will depend crucially on the ability of its employees to adapt and learn new tasks fast (i.e., $u(e,t)$ is very sensitive to t). On the other hand, the continued success of an established firm—where each employee's tasks are more clearly and narrowly defined—will depend (relatively) more on employee training, knowledge, and expertise (i.e., $u(e,t)$ is relatively more sensitive to e). Thus, hiring should be harder in start-ups than in established firms.

3.4 Optimal screening under costly testing

When testing is costly, the principal needs compare the benefit of testing to its cost. The benefit of testing increased accuracy: it allows the principal to accept talented agents with higher probability that untalented ones. The principal's objective function is $\int_0^1\int_0^1 [\Pi(e,t)u(e,t) - cT(e,t)]\, f(e,t)dtde$. By Lemma 2, condition (iii) of Proposition 1 is satisfied with equality by the optimal mechanism, so in the objective function we can substitute $T(e,t) = \Pi(e,t) - \Pi(e,0)$ to write the objective function only in terms of Π as follows:

$$\int_0^1 \int_{\underline{e}(s)}^{\overline{e}(s)} \left[\Pi(e,\tau(e,s))(u(e,\tau(e,s)) - c) + c\Pi(e,0)\right] f(e,\tau(e,s))deds, \tag{3}$$

which is again linear in Π, so by Bauer's maximum principle, there exists an extreme Π—among all Π that are non-decreasing in s and e—that solves the principal's problem.

Lemma 4. There exists an optimal deterministic mechanism.

3.4.1 Testing technology biased in favor of talent

Proposition 4 characterizes the optimal mechanism under pro-t biased testing, generalizing Proposition 2 by allowing for possibly costly testing (i.e., $c \geq 0$).

Proposition 4. If σ is pro-t biased, then there exists an optimal mechanism with $\Pi(e,t) = \mathbf{I}(u(e,t) \geq c \text{ or } e \geq e^*)$ and $T(e,t) = \mathbf{I}(u(e,t) \geq c \text{ and } e < e^*)$ for some $e^* \in [0,1]$.

The principal's problem amounts to choosing a threshold level of evidence $e^* \in \arg\max_{\widetilde{e}\in[0,1]} v(\widetilde{e})$,[22] where

$$v(\widetilde{e}) := \int_0^1 \int_0^{\widetilde{e}} (u(e,t) - c)\mathbf{I}(u(e,t) \geq c)f(e,t)dedt + \int_0^1 \int_{\widetilde{e}}^1 u(e,t)f(e,t)dedt.$$

Every agent with evidence $e \geq e^*$ evidence is accepted without a test, while agents with evidence $e < e^*$ are tested and accepted if their value $u(e,t)$ to the principal is higher than the cost c of testing. The remaining agents are rejected without getting tested. Figure 4(a) presents the structure of the optimal mechanism.

When $e^* \in (0,1)$, the first-order condition is

$$v'(e^*) = \int_0^1 (u(e^*,t) - c)\mathbf{I}(u(e^*,t) \geq c)f(e^*,t)dt - \int_0^1 u(e^*,t)f(e^*,t)dt = 0,$$

or equivalently

$$\begin{aligned} v'(e^*) = & \overbrace{- \int_0^1 u(e^*,t)\mathbf{I}(u(e^*,t) \leq 0)f(e^*,t)dt}^{>0\text{: gain from rejection of unworthy agents (ii)}} \overbrace{- \int_0^1 u(e^*,t)\mathbf{I}(0 < u(e^*,t) < c)f(e^*,t)dt}^{>0\text{: loss from rejection of worthy agents (iii)}} \\ & - \underbrace{c\int_0^1 \mathbf{I}(u(e^*,t) \geq c)f(e^*,t)dt}_{>0\text{: loss from increase in testing costs (i)}} = 0. \end{aligned}$$

An increase in the threshold e^* would lead to: (i) increased testing costs by making additional agents who lie above $I_u(c)$ get tested before being accepted (who were accepted

[22]The principal's problem reduces to this because all mechanisms with $\Pi(e,t) = \mathbf{I}(u(e,t) \geq c \text{ or } e \geq e^*)$ and $T(e,t) = \mathbf{I}(u(e,t) \geq c \text{ and } e < e^*)$ for some $e^* \in [0,1]$ are IC.

without a test before the increase in e^*), (ii) the rejection without a test of additional agents who lie below $I_u(0)$ (who were accepted without a test before the increase in e^*), but also (iii) the rejection without a test of additional agents who lie below $I_u(c)$ but above $I_u(0)$ (who were accepted without a test before the increase in e^*). Channels (i) and (iii) negatively affect the principal's payoff, while channel (ii) tends to increase his payoff. In choosing the optimal threshold e^*, the principal trades off testing costs (i.e., effect (i)) with the increase in accuracy (i.e., the net effect of (ii) and (iii)).

Comparative statics. We now briefly discuss some comparative statics. For simplicity, assume that $e^* \in (0,1)$ is unique with the second-order condition of the principal's problem satisfied strictly and that some agents are optimally tested.[23] First, an increase in c causes the (combined) magnitude of channels (i) and (iii) to increase without affecting the magnitude of channel (ii).[24] Thus, e^* is decreasing in c; the more costly testing is, the more high-evidence agents are accepted without a test. Particularly, $v'(e)$ is decreasing in c with $\partial v'(e)/\partial c = -\int_0^1 \mathbf{I}(u(e^*,t) \geq c) f(e^*,t)dt < 0$, and by the Implicit Function Theorem $de^*/dc = -\partial v'(e)/\partial c|_{e=e^*}/v''(e^*) < 0$. Second, the principal's optimal payoff is decreasing in c. Third, since the principal's objective function is independent of the testing technology σ, the optimal mechanism and payoff are the same under any two pro-t biased testing technologies with the same testing cost c.

3.4.2 Testing technology biased in favor of evidence

Proposition 5 characterizes the optimal mechanism under pro-e biased testing, generalizing Proposition 3 by allowing for possibly costly testing (i.e., $c \geq 0$).

Proposition 5. If σ is pro-e biased, then there exists an optimal mechanism with $\Pi(e,t) = \mathbf{I}(\sigma(e,t) \geq s^* \text{ or } e \geq e^*)$ and $T(e,t) = \mathbf{I}(\sigma(e,t) \geq s^* \text{ and } e < e^*)$ for some $(e^*,s^*) \in [0,1]^2$.

Finding an optimal mechanism is again remarkably simple. The principal's problem amounts to choosing threshold composite measure and evidence levels $(e^*,s^*) \in \arg\max_{(\widetilde{e},\widetilde{s})\in[0,1]^2} v(\widetilde{e},\widetilde{s})$, where

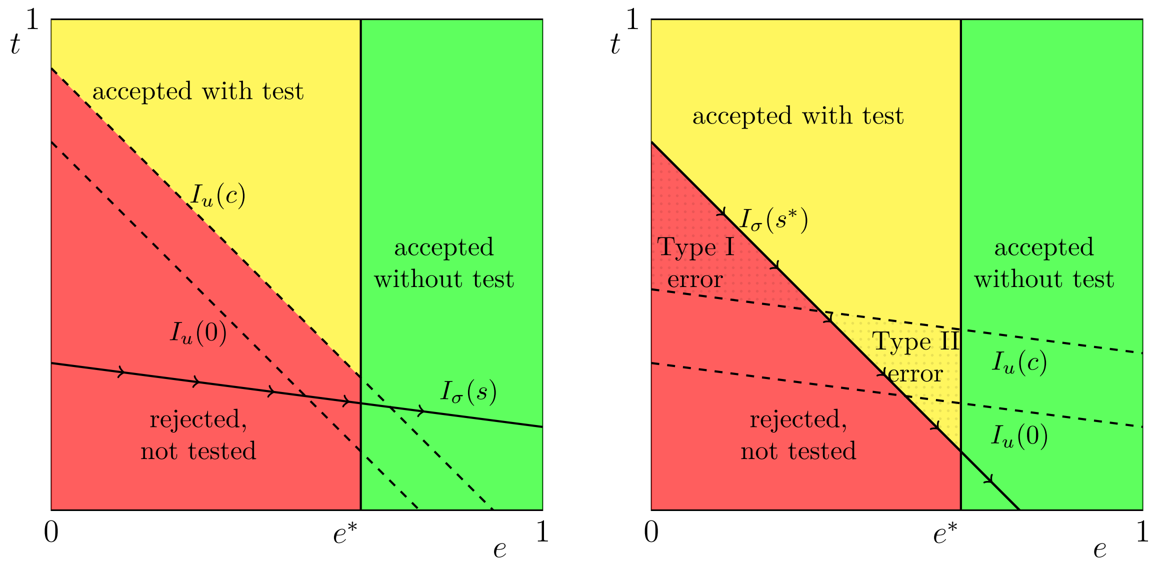
$$v(\widetilde{e},\widetilde{s}) := \int_{\widetilde{s}}^1 \int_{\underline{e}(s)}^{\max\{\min\{\overline{e}(s),\widetilde{e}\},\underline{e}(s)\}} \left(\widetilde{u}(e,s) - c\right) \widetilde{f}(e,s)deds + \int_0^1 \int_{\max\{\underline{e}(s),\widetilde{e}\}}^{\max\{\overline{e}(s),\widetilde{e}\}} \widetilde{u}(e,s)\widetilde{f}(e,s)deds,$$

[23] Namely, $e^* > 0$ and $u(e,t) > c$ for a positive measure of agents with $e < e^*$. This rules out the case $u(e,t) = e - \underline{q}$, where the principal only cares about evidence, in which case he does not test.

[24] In more detail, the partial derivative of $-c\int_0^1 \mathbf{I}(u(e^*,t) \geq c)f(e^*,t)dt$ with respect to c is $-\int_0^1 \mathbf{I}(u(e^*,t) \geq c)f(e^*,t)dt + cf(e^*,t')$ where t' is such that $u(e^*,t') = c$. The partial derivative of $\int_0^1 u(e^*,t)\mathbf{I}(0 < u(e^*,t) < c)f(e^*,t)dt$ with respect to c is $u(e^*,t')f(e^*,t') = cf(e^*,t') > 0$, which cancels out with the corresponding term in the derivative of $-c\int_0^1 \mathbf{I}(u(e^*,t) \geq c)f(e^*,t)dt$.

Figure 4: The optimal mechanism under costly testing

(a) Pro-t biased testing (b) Pro-e biased testing



Note: the dashed line represents the principal's indifference curve $I_u(c)$: the principal is indifferent between (i) accepting after testing and (ii) rejecting without testing agents on that curve. The arrowed line represents an iso-test-score curve (at an arbitrary level s in the left panel and at the optimal level s^* in the right panel). The green area denotes the set of agents who are accepted without getting tested in the optimal mechanism. The red one denotes the set of agents who are rejected without getting tested in the optimal mechanism. The yellow area denotes the set of agents who are accepted after getting tested in the optimal mechanism.

and $\tilde{u}(e,s) \equiv u(e,\tau(e,s))$ and $\tilde{f}(e,s) \equiv f(e,\tau(e,s))$.[25] Every agent with evidence $e \geq e^*$ is accepted without a test, while agents with evidence $e < e^*$ are tested and accepted if their composite measure is at least $\sigma(e,t) \geq s^*$. The remaining agents are rejected without getting tested. Figure 4(b) presents the structure of the optimal mechanism under pro-e biased testing.

When $\underline{e}(s^*) < e^*$ (i.e., some agents are accepted after being tested) and $e^*,s^* \in (0,1)$,[26] the first-order conditions are

$$v_1(e^*,s^*) = \overbrace{-\int_0^{s^*} \tilde{u}(e^*,s)\mathbf{I}(\tilde{u}(e^*,s) \leq 0)\tilde{f}(e^*,s)ds}^{>0\text{: gain from rejection of unworthy agents (ii)}} \overbrace{-\int_0^{s^*} \tilde{u}(e^*,s)\mathbf{I}(\tilde{u}(e^*,s) > 0)\tilde{f}(e^*,s)ds}^{>0\text{: loss from rejection of worthy agents (iii)}} - \underbrace{c\int_{s^*}^1 \tilde{f}(e^*,s)ds}_{\substack{>0\text{: loss from increase}\\ \text{in testing costs (i)}}} = 0$$

[25]The principal's problem reduces to this because all mechanisms with $\Pi(e,t) = \mathbf{I}(\sigma(e,t) \geq s^* \text{ or } e \geq e^*)$ and $T(e,t) = \mathbf{I}(\sigma(e,t) \geq s^* \text{ and } e < e^*)$ for some $(e^*,s^*) \in [0,1]^2$ are IC.

[26]Notice that $e^* \leq \overline{e}(s^*)$ (for if $e^* > \overline{e}(s^*)$ and $c > 0$, reducing e^* would increase $v(e^*,s^*)$). For $c = 0$, $e^* = 1$ without loss.

$$v_2(e^*,s^*) = \underbrace{-\int_{\underline{e}(s^*)}^{e^*} \min\{\widetilde{u}(e,s^*)-c,0\}\widetilde{f}(e,s^*)de}_{>0:\text{ gain from decrease in Type II error}} \underbrace{-\int_{\underline{e}(s^*)}^{e^*} \max\{\widetilde{u}(e,s^*)-c,0\}\widetilde{f}(e,s^*)de}_{>0:\text{ loss from increase in Type I error}} = 0,$$

given that optimality requires $e^* \leq \overline{e}(s^*)$.[27] The principal chooses s^* considering the trade-off between Type I and Type II errors—conditional on the fact that only agents with evidence $e \geq e^*$ are accepted without a test; for agents with evidence $e < e^*$, the principal chooses between accepting after testing and rejecting without testing. The Type I error is due to the fact that the principal rejects without testing some agents whom he would prefer to accept after testing. The Type II error is due to the fact that the principal tests and accepts some agents whom he would prefer to reject without testing. An increase in the threshold e^* would lead to: (i) increased testing costs by making additional agents who lie above $I_\sigma(s^*)$ get tested before being accepted (who were accepted without a test before the increase in e^*), (ii) the rejection without a test of additional agents who lie below $I_u(0)$ (who were accepted without a test before the increase in e^*), but also possibly (iii) the rejection without a test of additional agents who lie below $I_\sigma(s^*)$ but above $I_u(0)$ (who were accepted without a test before the increase in e^*).[28] Channels (i) and (iii) negatively affect the principal's payoff, while channel (ii) tends to increase his payoff. In choosing the optimal threshold e^*, the principal trades off testing costs (i.e., effect (i)) with the increase in accuracy (i.e., the net effect of (ii) and (iii), which is a net gain).

Comparative statics. We now briefly discuss some comparative statics. For simplicity, assume that $s^*,e^* \in (0,1)$ are unique with the second-order condition of the principal's problem satisfied strictly and that some agents are optimally tested. Denote by $J(e^*,s^*)$ the Jacobian matrix of the first derivatives evaluated at (e^*,s^*), which is by assumption negative definite. Particularly, $v_{11}(e^*,s^*),v_{22}(e^*,s^*) < 0$ and $\det(J(e^*,s^*)) > 0$. Also, $v_{12}(e^*,s^*) = v_{21}(e^*,s^*) = -\left(\widetilde{u}(e^*,s^*) - c\right)\widetilde{f}(e^*,s^*) > 0$. First, the total derivatives of $e^*$and s^* with respect to c are:

$$\frac{de^*}{dc} \propto \overbrace{-v_{1c}(e^*,s^*)v_{22}(e^*,s^*)}^{\substack{<0:\text{ direct effect of } c \text{ on}\\ e^* \text{ due to increase in}\\ \text{marginal testing costs}}} + \overbrace{v_{2c}(e^*,s^*)v_{12}(e^*,s^*)}^{\substack{>0:\text{ indirect effect of } c\\ \text{on } e^* \text{ through direct}\\ \text{effect of } c \text{ on } s^*}},$$
$$\frac{ds^*}{dc} \propto \underbrace{-v_{2c}(e^*,s^*)v_{11}(e^*,s^*)}_{\substack{>0:\text{ direct effect of } c \text{ on}\\ s^* \text{ due to increase in}\\ \text{marginal testing costs}}} + \underbrace{v_{1c}(e^*,s^*)v_{21}(e^*,s^*)}_{\substack{<0:\text{ indirect effect of } c\\ \text{on } s^* \text{ through direct}\\ \text{effect of } c \text{ on } e^*}},$$

where $v_{1c}(e^*,s^*) = -\int_{s^*}^{1} \widetilde{f}(e^*,s)ds < 0$ and $v_{2c}(e^*,s^*) = \int_{\underline{e}(s^*)}^{e^*} \widetilde{f}(e,s^*)de > 0$ are the partial derivatives of v_1 and v_2 with respect to c. An increase in the (marginal) cost c of testing

[27]If $e^* > \overline{e}(s^*)$, decreasing e^* to make it equal to $\overline{e}(s^*)$ would decrease testing costs without changing the set of agents who are accepted, thereby increasing the principal's payoff.

[28]Channel (iii) is not necessarily present.

tends to directly cause (i) e^* to decrease by magnifying the testing cost savings associated with a decrease in e^* and (ii) s^* to increase by magnifying the testing cost savings associated with an increase in e^*.[29] However, an increase in s^* tends to cause e^* to increase by reducing the marginal increase in testing costs associated with an increase in e^*. Conversely, an increase in e^* tends to cause s^* to increase by increasing the marginal (with respect to s^*) Type II error. Therefore, although an increase in c tends to directly cause e^* to fall and s^* to rise, the interaction between e^* and s^* works in the opposite direction making the net effect ambiguous. Still, we know that if s^* decreases in response to an increase in c, e^* should also decrease and—the contrapositive—if e^* increases in response to an increase in c, s^* should also increase. Second, the principal's optimal payoff is decreasing in c. Third, the optimal payoff is higher under less pro-e biased testing technologies. Namely, take any two pro-e biased testing technologies σ' and σ. If all iso-test-score curves of σ cross the iso-test-score curves of σ' from above (i.e., σ is less pro-e biased than σ'), the principal's optimal payoff is higher under σ' than under σ.[30] Fourth, the principal's payoff should tend to increase with the correlation between evidence and talent. A strong (positive) correlation between e and t means that there are not many agents with high (resp. low) talent and low (resp. high) evidence, which implies that both Type I and Type II errors are small. As e and t become perfectly (positively) correlated, the principal achieves the first-best just by asking for evidence—regardless of his preferences and the testing technology.

Implementation of the optimal mechanism. We have so far restricted (without loss) attention to truth-telling mechanisms. However, the optimal mechanism under pro-e biased testing can be implemented in the following simple way. The principal gives the agent two paths to getting accepted: (i) provide evidence e^* and you will be accepted without a test or (ii) take a test without providing any evidence, score at least s^*, and you will be accepted. The first option is not always provided (e.g., when testing is free, the first option is not necessary in the optimal mechanism). Asking for evidence is useful to the principal (as long as $e^* < 1$). Last, notice that a similarly simple implementation of the optimal mechanism under pro-t biased testing is not possible. In that case, the principal needs to ask for evidence also from agents who are tested.[31]

[29]Put differently, an increase in c can be seen to increase the marginal (with respect to s^*) Type II error and decrease the marginal Type II error, thereby tending to make s^* increase to equalize the magnitudes to the two errors.

[30]Comparative statics of s^* and e^* with respect to σ would have little value, since the optimal composite measure thresholds (which are determined simultaneously with the optinal evidence thresholds) under different testing technologies are not comparable, as they can only be interpreted with respect to their corresponding testing technologies.

[31]These observations on the implementation of optimal mechanisms also imply that under free testing (i.e., $c = 0$), if the principal (optimally) asks for evidence—which he does not need to do under pro-e biased testing, then he most likely values evidence (i.e., $u(e,t)$ is increasing in e).

3.5 $(m+n)$-dimensional screening with substitutable attributes and costly verification

We now generalize the results allowing for multiple dimensions of evidence and talent. Let the agent's type be $(e_1,e_2,\ldots,e_m,t_1,t_2,\ldots,t_n)$ with full-support density $f:[0,1]^{m+n}\to\mathbb{R}_{++}$. $(e_1,e_2,\ldots,e_m)$ are different dimensions of evidence and $(t_1,t_2,\ldots,t_m)$ are different dimensions of talent. The agent can present any combination of evidence $\boldsymbol{r}\in[\boldsymbol{0},\boldsymbol{e}]$. The testing technology $\sigma:[0,1]^{m+n}\to[0,1]$ is continuous and increasing. $u(\boldsymbol{e},\boldsymbol{t})$ is continuous and non-decreasing. It follows by the same arguments as before that truthful mechanisms with pass-or-fail tests are still without loss.

Lemma 5 makes the following additional observation: among agents with the same evidence and composite measure, IC mechanisms cannot screen for different dimensions of talent. That $\Pi(\boldsymbol{e},\boldsymbol{t})=\Pi(\boldsymbol{e},\boldsymbol{t}')$ for every $\boldsymbol{e},\boldsymbol{t},\boldsymbol{t}'$ with $\sigma(\boldsymbol{e},\boldsymbol{t})=\sigma(\boldsymbol{e},\boldsymbol{t}')$ is necessary to ensure that no agent has incentives to present all her evidence but misreport her talent to imitate an agent with the same composite measure.

Lemma 5. If a mechanism $M\equiv\langle T,P\rangle$ is IC, then $\Pi(\boldsymbol{e},\boldsymbol{t})=\Pi(\boldsymbol{e},\boldsymbol{t}')$ for every $\boldsymbol{e},\boldsymbol{t},\boldsymbol{t}'$ such that $\sigma(\boldsymbol{e},\boldsymbol{t})=\sigma(\boldsymbol{e},\boldsymbol{t}')$.

Therefore, we restrict attention to mechanisms with $\Pi(\boldsymbol{e},\boldsymbol{t})=\Pi(\boldsymbol{e},\boldsymbol{t}')$ for every $\boldsymbol{e},\boldsymbol{t},\boldsymbol{t}'$ such that $\sigma(\boldsymbol{e},\boldsymbol{t})=\sigma(\boldsymbol{e},\boldsymbol{t}')$. Lemma 6 shows that we can further restrict attention to mechanisms that treat agents with the same evidence and composite measure exactly the same way with respect to testing and acceptance probabilities conditional on test results.

Lemma 6. Given any IC mechanism M, there exists an IC mechanism $M'\equiv\langle T',P'\rangle$ with $T'(\boldsymbol{e},\boldsymbol{t})=T'(\boldsymbol{e},\boldsymbol{t}')$ and $P'(\boldsymbol{e},\boldsymbol{t},\emptyset)=P'(\boldsymbol{e},\boldsymbol{t}',\emptyset)$ for every $\boldsymbol{e},\boldsymbol{t},\boldsymbol{t}'$ such that $\sigma(\boldsymbol{e},\boldsymbol{t})=\sigma(\boldsymbol{e},\boldsymbol{t}')$ that is outcome-equivalent to M. Also, for $c>0$, in any optimal mechanism $M\equiv\langle T,P\rangle$, $T(\boldsymbol{e},\boldsymbol{t})=T(\boldsymbol{e},\boldsymbol{t}')$ for almost every $\boldsymbol{e},\boldsymbol{t},\boldsymbol{t}'$ such that $\sigma(\boldsymbol{e},\boldsymbol{t})=\sigma(\boldsymbol{e},\boldsymbol{t}')$.

The intuition behind this result goes as follows. The only reason to test an agent before accepting her—rather than accept her without a test—is to prevent others from imitating her. Take any agent $(\boldsymbol{e},\boldsymbol{t})$ who contemplates which of the agents in the set $X(\boldsymbol{e}',s)\coloneqq\{(\boldsymbol{e}',\boldsymbol{t}'):\sigma(\boldsymbol{e}',\boldsymbol{t}')=s\}$, where $\boldsymbol{e}'\leq\boldsymbol{e}$, to imitate. By Lemma 5, Π is the same for every agent in $X(\boldsymbol{e},s)$, so if $\sigma(\boldsymbol{e},\boldsymbol{t})\geq s$, then agent $(\boldsymbol{e},\boldsymbol{t})$'s payoff from imitating an agent in $X(\boldsymbol{e}',s)$ does not depend on which particular agent she chooses to imitate. If, on the other hand, $\sigma(\boldsymbol{e},\boldsymbol{t})<s$, agent $(\boldsymbol{e},\boldsymbol{t})$'s payoff from imitating an agent $(\boldsymbol{e}',\boldsymbol{t}')\in X(\boldsymbol{e}',s)$ is increasing (resp. decreasing) in $P(\boldsymbol{e}',\boldsymbol{t}',\emptyset)$ (resp. $T(\boldsymbol{e}',\boldsymbol{t}')$). Thus, the principal can decrease $T(\boldsymbol{e}',\boldsymbol{t}')$ and increase $P(\boldsymbol{e}',\boldsymbol{t}',\emptyset)$ for every agent $(\boldsymbol{e}',\boldsymbol{t}')\in X(\boldsymbol{e}',s)$ with $T(\boldsymbol{e}',\boldsymbol{t}')>\inf_{(\boldsymbol{e}'',\boldsymbol{t}'')\in X(\boldsymbol{e}',s)}T(\boldsymbol{e}'',\boldsymbol{t}'')$ (and thus, $P(\boldsymbol{e}'',\boldsymbol{t}',\emptyset)<\sup_{(\boldsymbol{e}'',\boldsymbol{t}'')\in X(\boldsymbol{e}',s)}P(\boldsymbol{e}',\boldsymbol{t}'',\emptyset)$) keeping Π fixed. There is no point in testing (or accepting without testing) two agents with the same evidence and composite measure with different probabilities, as doing so does not reduce incentives of others to imitate them and makes the principal incur higher than necessary testing costs.

Thus, we can restrict attention to mechanisms with $\Pi(\boldsymbol{e},\boldsymbol{t}) = \Pi(\boldsymbol{e},\boldsymbol{t}')$, $T(\boldsymbol{e},\boldsymbol{t}) = T(\boldsymbol{e},\boldsymbol{t}')$, $P'(\boldsymbol{e},\boldsymbol{t},\emptyset) = P'(\boldsymbol{e},\boldsymbol{t}',\emptyset)$, and $P(\boldsymbol{e},\boldsymbol{t},s) = P(\boldsymbol{e},\boldsymbol{t}',s)$ for every $\boldsymbol{e},\boldsymbol{t},\boldsymbol{t}',s$ such that $\sigma(\boldsymbol{e},\boldsymbol{t}) = \sigma(\boldsymbol{e},\boldsymbol{t}')$.[32] In other words, the principal can constrain attention to mechanisms that ask agents only for evidence and claim about their composite measure (rather than a whole profile of talent dimensions). The principal designs a mechanism $M \equiv \langle T,P \rangle$, where $T : [0,1]^{m+1} \to [0,1]$ and $P : [0,1]^{m+1} \times ([0,1] \cup \{\emptyset\}) \to [0,1]$. Proposition 6 generalizes the IC characterization of Proposition 1 to the case of $(m+n)$-dimensional screening.

Proposition 6. A mechanism $M \equiv \langle T,P \rangle$ is IC if and only if

(i) $\Pi(\boldsymbol{e},s)$ is non-decreasing in s over $s \in [\sigma(\boldsymbol{e},\boldsymbol{0}),\sigma(\boldsymbol{e},\boldsymbol{1})]$ for every $\boldsymbol{e} \in [0,1]^m$,

(ii) $\Pi(\boldsymbol{e},s)$ is non-decreasing in $\boldsymbol{e}$ over $\boldsymbol{e} \in \{\boldsymbol{e} \in [0,1]^m : \sigma(\boldsymbol{e},\boldsymbol{0}) \leq s \leq \sigma(\boldsymbol{e},\boldsymbol{1})\}$ for every $s \in [0,1]$, and

(iii) $(1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset) \leq \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0}))$ for every $(\boldsymbol{e},s) \in [0,1]^{m+1}$,

where $\Pi(\boldsymbol{e},s) \coloneqq (1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset) + T(\boldsymbol{e},s)$ is the probability with which agent an agent is accepted if she truthfully reports her evidence $\boldsymbol{e}$ and composite measure s.

The intuition behind the conditions is analogous to the one behind the conditions of Proposition 1. Notice that there are no IC condition on the comparison between the values of T, P, or Π for agent types $(\boldsymbol{e},s)$ and $(\boldsymbol{e}',s')$ such that $\boldsymbol{e}$ and $\boldsymbol{e}'$ are incomparable (i.e., $\boldsymbol{e} \ngeq \boldsymbol{e}'$ and $\boldsymbol{e} \nleq \boldsymbol{e}'$). That is, because neither agent type has the evidence to imitate the other.

Lemma 7 generalizes Lemma 2 to show that when testing is costly and some talented agents are (optimally) accepted with higher probability than untalented ones with the same evidence, then the optimal mechanism satisfies condition (iii) of Proposition 6 with equality. Under free testing or when it is not optimal to accept talented agents with higher probability, it is still without loss to constrain attention to mechanisms that satisfy condition (iii) of Proposition 6 with equality.

Lemma 7. Given any IC mechanism $M \equiv \langle T,P \rangle$, there exists an IC mechanism $M' \equiv \langle T',P' \rangle$ with $(1 - T'(\boldsymbol{e},s))P'(\boldsymbol{e},s,\emptyset) = \Pi'(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0}))$ for every $(\boldsymbol{e},s)$, $s \in [\sigma(\boldsymbol{e},\boldsymbol{0}),\sigma(\boldsymbol{e},\boldsymbol{1})]$ that is outcome-equivalent to M and has at most as high testing costs as M. For $c > 0$, if also $\Pi(\boldsymbol{e},s) > \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0}))$ for a positive measure of $(\boldsymbol{e},s)$'s, then M' has lower testing costs than M.

[32] That $P(\boldsymbol{e},\boldsymbol{t},s) = P(\boldsymbol{e},\boldsymbol{t}',s)$ for every $s \in [0,1]$ when $\sigma(\boldsymbol{e},\boldsymbol{t}) = \sigma(\boldsymbol{e},\boldsymbol{t}')$ follows already from restricting attention to pass-or-fail tests.

Define $\widetilde{f}(\boldsymbol{e},s) := \int_{\boldsymbol{t}\in[0,1]^n} \mathbf{I}(\sigma(\boldsymbol{e},\boldsymbol{t}) = s)f(\boldsymbol{e},\boldsymbol{t})d\boldsymbol{t}$, the probability density of agents with evidence $\boldsymbol{e}$ and composite measure s, and $\widetilde{u}(\boldsymbol{e},s) := \mathbb{E}_{\mathbf{t}}[u(\boldsymbol{e},\boldsymbol{t})|\sigma(\boldsymbol{e},\boldsymbol{t}) = s] = \int_{\mathbf{t}\in[0,1]^n} u(\mathbf{e},\mathbf{t})\mathbf{I}(\sigma(\mathbf{e},\mathbf{t}) = s)f(\mathbf{e},\mathbf{t})d\mathbf{t}/\widetilde{f}(\boldsymbol{e},s)$, the principal's expected payoff from accepting all agents with evidence $\boldsymbol{e}$ and composite measure s.[33] Assume that $\widetilde{u}(\boldsymbol{e},s)$ is non-decreasing in s. The principal's objective function is $\int_0^1 \cdots \int_0^1 \int_{\sigma(\boldsymbol{e},\mathbf{0})}^{\sigma(\boldsymbol{e},\mathbf{1})} [\Pi(\boldsymbol{e},s)\widetilde{u}(\boldsymbol{e},s) - cT(\boldsymbol{e},s)]\, \widetilde{f}(\boldsymbol{e},s)ds\, de_1 \cdots de_m$. By Lemma 7, condition (iii) of Proposition 6 is satisfied with equality by the optimal mechanism, so in the objective function we can substitute $T(\boldsymbol{e},s) = \Pi(\boldsymbol{e},s) - \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))$. Then, the objective function reads

$$\int_0^1 \cdots \int_0^1 \int_{\sigma(\boldsymbol{e},\mathbf{0})}^{\sigma(\boldsymbol{e},\mathbf{1})} [\Pi(\boldsymbol{e},s)(\widetilde{u}(\boldsymbol{e},s) - c) + c\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))]\, \widetilde{f}(\boldsymbol{e},s)dsde_1 \cdots de_m, \tag{4}$$

which is linear in Π, so by Bauer's maximum principle, there exists an extreme Π—among all Π that are non-decreasing in s and $\boldsymbol{e}$—that solves the principal's problem.

Lemma 8. There exists an optimal deterministic mechanism.

3.5.1 Testing technology biased in favor of talent

The definition of pro-t biased testing generalizes to the case of $(m + n)$-dimensional screening as follows.

Definition 5. σ is pro-t biased if for every $\boldsymbol{e},\boldsymbol{e}' \in [0,1]^m$ and every composite measure $s \in [\max\{\sigma(\boldsymbol{e},\mathbf{0}),\sigma(\boldsymbol{e}',\mathbf{0})\}, \min\{\sigma(\boldsymbol{e},\mathbf{1}),\sigma(\boldsymbol{e}',\mathbf{1})\}]$, if $\widetilde{u}(\boldsymbol{e},s) \geq c \geq \widetilde{u}(\boldsymbol{e}',s)$ with at least one inequality holding strictly, then $\boldsymbol{e}' \not\geq \boldsymbol{e}$.

Generalizing Proposition 4, Proposition 7 derives the optimal mechanism under pro-t biased testing.

Proposition 7. If σ is pro-t biased, then there exists an optimal mechanism with $\Pi(\boldsymbol{e},s) = \mathbf{I}(\widetilde{u}(\boldsymbol{e},s) > c \text{ or } \boldsymbol{e} \in E^*)$ and $T(\boldsymbol{e},s) = \mathbf{I}(\widetilde{u}(\boldsymbol{e},s) > c \text{ and } \boldsymbol{e} \notin E^*)$ for some upper set E^* of $[0,1]^m$ (i.e., $E^* \subseteq [0,1]^m$ such that for any $\boldsymbol{e} \in E^*$ and $\boldsymbol{e}' \in [0,1]^m$, if $\boldsymbol{e}' \geq \boldsymbol{e}$, then $\boldsymbol{e}' \in E^*$).

Clearly, if $c = 0$, $E^* = \emptyset$ without loss. If $c > 0$, $E^* \supseteq \{\boldsymbol{e} \in [0,1]^m : \widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) > c\}$. This has to be true, because among the agents who are accepted, an agent should be tested only if this will prevent others from imitating her. Any agent who has enough evidence to imitate an agent $(\boldsymbol{e},\mathbf{0})$ with $\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) > c$ and get accepted also has composite measure at least as high as $(\boldsymbol{e},\mathbf{0})$ can. Therefore, $(\boldsymbol{e},\mathbf{0})$ with $\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) > c$ should not be tested.

[33] For $(\boldsymbol{e},s)$ such that $s = \sigma(\boldsymbol{e},\mathbf{0})$, $\widetilde{u}(\boldsymbol{e},s) \equiv u(\boldsymbol{e},\mathbf{0})$). For simplicity, the following Regularity Assumption is made: $\widetilde{u}(\boldsymbol{e},s)$ is increasing in s.

3.5.2 Testing technology biased in favor of evidence

The definition of pro-e biased testing generalizes to the case of $(m+n)$-dimensional screening as follows.

Definition 6. σ is pro-e biased if for every $\boldsymbol{e}$,$\boldsymbol{e}' \in [0{,}1]^m$ and every composite measure $s \in [\max\{\sigma(\boldsymbol{e},\boldsymbol{0}),\sigma(\boldsymbol{e}',\boldsymbol{0})\}, \min\{\sigma(\boldsymbol{e},\boldsymbol{1}),\sigma(\boldsymbol{e}',\boldsymbol{1})\}]$, if $\tilde{u}(\boldsymbol{e},s) \geq c \geq \tilde{u}(\boldsymbol{e}',s)$ with at least one inequality holding strictly, then $\boldsymbol{e}' \geq \boldsymbol{e}$.

Generalizing Proposition 5, Proposition 8 derives the optimal mechanism under pro-e biased testing.

Proposition 8. If σ is pro-e biased, then there exists an optimal mechanism with $\Pi(\boldsymbol{e},s) = \mathbf{I}(s \geq s^* \text{ or } \boldsymbol{e} \in E^*)$ and $T(\boldsymbol{e},s) = \mathbf{I}(s \geq s^* \text{ and } \boldsymbol{e} \notin E^*)$ for some $s^* \in [0{,}1]$ and some upper set E^* of $[0{,}1]^m$.

If $c > 0$, then $E^* \supseteq \{\boldsymbol{e} \in [0{,}1]^m : \tilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) > s^*\}$, which again has to be true, because among the agents who are accepted, an agent should be tested only if this will prevent others from imitating her.

4 Applications

In this section, I use the model to discuss hiring by prestigious employers, promotion decisions, college admissions, and academic job market hiring.

4.1 Hiring by prestigious employers

A job candidate's evidence e is her CV quality (e.g., high school quality, undergraduate institution quality and GPA, awards, distinctions, reference letters). t is her ability and drive not captured by e. A prestigious employer wants to decide whether to hire the candidate. Testing amounts to letting some other employer hire the candidate. Testing is costly because if the employer wants to then hire the candidate, he will have to poach her at a cost.

In the optimal mechanism, Ivy-Leaguers with high credentials get immediately hired by prestigious employers thanks to their evidence. On the other hand, talented candidates with education from lower-ranked institutions and lower grades have to go through less prestigious employers to prove their worth before they land a prestigious position. Also, if the candidates' performance in the less prestigious position is less sensitive to talent than talent is valuable in the more prestigious position—a natural assumption, then worthy candidates with low credentials are at a disadvantage also in the poaching stage.

4.2 Promotions

An employee of efficiency t has exerted effort e. $\sigma(e,t)$ is the employee's productivity, increasing in e and in t. The employee can provide evidence or not on e by, for example, working at the office or from home. Testing (by the employer/manager) amounts to verifying the employee's productivity $\sigma(e,t)$. The value to the principal of the agent who is not promoted (i.e., continues to work in her current position) is $\sigma(e,t)$. His value of the agent if promoted is $\widetilde{u}(e,t)$. Then, his problem is equivalent to the one in section 2 with $u(e,t) \coloneqq \widetilde{u}(e,t) - \sigma(e,t)$, as long as the difference $\widetilde{u}(e,t) - \sigma(e,t)$ is non-decreasing in both e and t.[34] This condition on the difference can be interpreted to say that both effort and talent have a (weakly) higher marginal return in the higher position, which comes with increased responsibilities that allow the employee's talent and effort to have a larger impact.

Under differentiability and given Claims 1 and 2, the test is pro-t (resp. pro-e) biased if for every (e,t), $\partial u(e,t)/\partial e/(\partial u(e,t)/\partial t)$ is higher (resp. lower) than $\partial\sigma(e,t)/\partial e/(\partial\sigma(e,t)/\partial t)$, or equivalently,

$$\frac{\partial\widetilde{u}(e,t)/\partial e}{\partial\widetilde{u}(e,t)/\partial t} \overset{\text{(resp. } <)}{>} \frac{\partial\sigma(e,t)/\partial e}{\partial\sigma(e,t)/\partial t},$$

that is, if the marginal rate of substitution of effort for talent is higher (resp. lower) in the production function of the new position than of the current one.

4.3 College admissions and standardized testing

A college applicant's evidence e is her high school quality, grades, private tutoring received, awards, and extracurricular activities. t is her "natural" ability or drive that is not captured by e. The college wants to decide whether to admit the applicant or not. Testing amounts to requiring the applicant to take the standardized test.[35]

In the optimal mechanism, if the standardized test is not sensitive enough to talent, students can withhold evidence, which makes admission decisions imperfect at the expense of students with low evidence (e.g., those with limited access to quality education, tutoring, extracurricular activities, and opportunities to participate in competitions). Particularly, if colleges want diversity and only value talent (trying to control for the applicants' unequal backgrounds), the above problem is necessarily present under standardized testing to the extent that applicants can pretend to be from a more modest background than they

[34] $u(e,t)$ could also be defined as $u(e,t) \coloneqq \widetilde{u}(e,t) - \sigma(e,t) - \underline{q}$, where $\underline{q}$ is the threshold productivity differential for the promotion to be beneficial to the firm (e.g., $\underline{q}$ could be the productivity differential of another employee who could be promoted instead).

[35] In this setting, the college does not condition the requirement to take a test on the candidate's report. However, when the college requires a test score, the optimal mechanism takes the same form as in the case of $c = 0$.

actually are. Students from a privileged background have an advantage over equally good—or even somewhat better—students from a more modest background.

Even if universities do not only value talent in applicants (but instead value the total ability of the candidate, part of which is due to nurture), students from advantaged backgrounds are still favored over equally able students from disadvantaged backgrounds when the test is not sensitive enough to talent (compared to the total ability that colleges care about).

4.4 Academic job market talks

An academic job market candidate's research topic is comprised of a "mass" $b > 1$ of (uncountably infinitely many) problems.[36] $e \in [0,1]$ is the candidate's knowledge, the mass of problems which she has found answers to. t is her ability to think on her feet. More concretely, it is the probability with which she finds an answer on the spot to a problem that she has not already solved. After the candidate presents answers to a mass $e' \in [0,e]$ of problems and makes a claim about t, the hiring committee may test her. Testing amounts to posing to the candidate countably infinitely many problems randomly sampled from the mass of problems that the candidate has not already disclosed answers to.[37] Thus, if she presents answers to mass $e' \in [0,e]$ of problems and is tested, she will answer proportion $p(e,t,r) \coloneqq [e - e' + (b - e)t]/(b - e')$ of the problems posed to her. This is the sum of (i) the proportion $(e - e')/(b - e')$ of problems sampled from the set of problems that the candidate already has answers to (but has not disclosed them) and (ii) the proportion $(b - e)/(b - e')$ of problems sampled from the set of problems that the candidate does not already have answers to multiplied by the proportion t to which the candidate will find answers on the spot. $u(e,t)$ is the hiring committee's surplus from hiring the candidate (compared to the committee's outside option). Observing e' and $p(e,t,e')$ is equivalent to observing e' and $\sigma(e,t) \coloneqq e + (b - e)t$, so the committee's problem is equivalent to the problem that we have studied.

Different testing technologies can be interpreted as different values of b. An increase in the mass b of the universe of problems makes it less likely that the candidate will be asked a question that she already has an answer to (but has not presented), thereby making the test more sensitive to talent and increasing the principal's payoff.

[36]The analysis can apply to presentations more generally (e.g., by a start-up founder to a venture capital firm).

[37]This can be understood as there being a set of problems with cardinality equal to the cardinality of $\mathbb{R}$. There is no interdependence among the problems (e.g., the agent having the answer to a problem x carries no information with regard to whether she also has the answer to a problem y). Also, the agent is equally likely to have or find an answer to any of the problems. Thus, there is no need to identify problems with an index.

5 Extensions and robustness

This section first discusses optimal screening under alternative evidence structures. Then, it studies two extensions of the model: (i) one where the principal has to pay a cost *before* the agent reports her type in order to design the test, which he will then (after the agent reports her type) choose whether to administer at an additional cost and (ii) the case where evidence is not exogenous but rather endogenously produced by the agent before she interacts with the principal.

5.1 Optimal screening under alternative evidence structures

I study optimal screening under three alternative scenarios: (i) The agent cannot withhold evidence, (ii) the agent can also present evidence of talent, or (iii) the agent cannot present evidence (on either dimension of her type).[38]

5.1.1 Optimal screening when the agent cannot withhold evidence

Assume that e is observed by the principal. Then, given that the test is at least somewhat sensitive to t, testing reveals t and, thus, the agent's type completely. The principal's problem is decoupled: he can solve it for each e separately.[39] It is easy to see that the optimal mechanism is described by Proposition 9.

Proposition 9. Assume that the agent cannot withhold evidence. In the optimal mechanism, for every level of evidence $e \in [0,1]$, if

(i) $u_{\text{accept}}(e) > \max\{u_{\text{test}}(e), 0\}$, then every agent with evidence e is accepted without a test,

(ii) $0 > \max\{u_{\text{accept}}(e), u_{\text{test}}(e)\}$, then every agent with evidence e is rejected without getting tested,

(iii) $u_{\text{test}}(e) \geq \max\{u_{\text{accept}}(e), 0\}$, then an agent with evidence e is tested and accepted if $u(e,t) \geq c$; otherwise, she is rejected without getting tested,

where $u_{\text{accept}}(e) \coloneqq \int_0^1 u(e,t)f(t)dt$ and $u_{\text{test}}(e) \coloneqq \int_0^1 (u(e,t) - c)\mathbf{I}(u(e,t) \geq c)f(t)dt$.

This implies that under pro-e (resp. pro-t biased testing), if two agents (e_1,t_1) and (e_2,t_2), $e_2 > e_1$, both need to be tested (based on their level of evidence) to get accepted,

[38] The case where the agent can present evidence on t but not on e is a relabeling of the main model.

[39] If the agent's type is (e_p,e,t) distributed over $[0,1]^3$, where e_p is the publicly observed part of evidence and e is the part that can be hidden, the principal can solve the problem for each e_p separately. The optimal mechanism is a collection mechanisms like the one described in section 3: one mechanism for each value of e_p. In the case of pro-e biased testing, for example, the optimal evidence $e^*(e_p)$ and composite measure cutoffs $s^*(e_p)$ depend on the observable part of evidence e_p.

then the composite measure threshold that (e_1,t_1) needs to meet is lower (resp. higher) than the composite measure threshold that (e_2,t_2) needs to meet.[40] For example, if the principal only values talent (i.e., $u(e,t) = t - \underline{q}$ for some $\underline{q} \in (0,1)$), agents with less evidence do not need to test as high (as those with more evidence) to get accepted. This is in stark contrast with the optimal mechanism where agents can withhold evidence, in which case every agent faces the same composite measure cutoff.

This analysis implies the following for college admissions. If (i) college applicants can to a large extent hide their privilege and (ii) standardized tests reflect talent less than colleges value talent, then every applicant will have to achieve roughly the same test score to get admitted, and affirmative action (i.e., trying to control for unequal backgrounds, measured by e) will not be very effective in admitting a diverse class of talented students. If any of the two condition fails, affirmative action is effective. Particularly, if (i) college applicants *cannot* withhold evidence of privilege and (ii) standardized tests reflect talent less than colleges value talent, then applicants from disadvantaged backgrounds will face lower test score cutoffs, and affirmative action is effective. If standardized tests are sensitive enough to talent (compared to college preferences), then testing does not create incentives for applicants to hide evidence of privilege (even if they can do so), and affirmative action is effective regardless of whether college applicants can hide evidence of privilege or not. Thus, if condition (i) or (ii) fails, a reversal of affirmative action would have significant effects on diversity in college admissions.

5.1.2 Optimal screening when the agent can also present evidence of talent

Consider the case where—apart from e—t also has an evidence structure. That is, agent (e,t) can report any $(e',t') \leq (e,t)$ but not $e' > e$ or $t' > t$. Then, the principal can achieve the full information first-best, inducing—without testing—every agent to present all her evidence on both e and t. The conclusion is the same if t is observed (at no cost) by the principal and e is evidence.

The comparison between this and the main model emphasizes (i) the difference in peoples' incentives to present evidence that is in principle (i.e., absent testing) favorable to them and (ii) how these incentives shape the principal's problem of evaluating them. The existence of an agent characteristic that is valuable to the principal but which (i) the agent cannot provide evidence on and (ii) the principal can only imperfectly test using a test that is overly (compared to the principal's preferences) sensitive to other (valuable) agent characteristics creates incentives for the agent to understate those other characteristics that she can actually provide favorable evidence on. This problem vanishes (i) if the agent can provide evidence on every characteristic (or if those that she cannot provide evidence on are observed by the principal) or (ii) if the test is sensitive only to

[40] To see this, notice that under pro-e (resp. pro-t biased testing), if $u(e_1,t_1) = u(e_2,t_2) = c$ and $e_2 > e_1$, then $\sigma(e_1,t_1) < \sigma(e_1,t_1)$ (resp. $\sigma(e_1,t_1) > \sigma(e_1,t_1)$).

the characteristic that the agent cannot provide evidence on.

These results are consistent with the finding that hiding one's effort is particularly prevalent among younger individuals. Effortless perfection (i.e., the need to seem perfect without apparent effort) and hiding one's effort have been documented among university students (Travers et al., 2015; Casale et al., 2016).The psychology literature has emphasized personality traits that may lie behind this finding. Namely, hiding effort has been identified as a unique expression of perfectionistic self-presentation (Flett et al., 2016). My model hints towards an alternative (or complementary) interpretation of this finding. If as an individual progresses in her career, her talent is revealed during all the evaluation stages that she goes through, then individuals that are further along in their career paths should have reduced incentives to hide their hard work.

5.1.3 Optimal screening when the agent cannot present evidence

Consider the case where the agent can present evidence on neither e nor t. That is, agent (e,t) can report any $(e',t') \in [0,1]^2$. We can still restrict attention to truthful mechanisms with pass-or-fail tests. Proposition 10 characterizes IC mechanisms.

Proposition 10. Assume that the agent cannot present evidence. A mechanism $M \equiv \langle T,P \rangle$ is IC if and only if

(i) $\Pi(e,t)$ is non-decreasing in t for every e,

(ii) $\Pi(e,\tau(e,s))$ is constant in e over $e \in [\underline{e}(s),\overline{e}(s)]$ for every $s \in [0,1]$, and

(iii) $(1-T(e,t))P(e,t,\emptyset) \leq \Pi(0,0)$ for every (e,t),

where $\Pi(e,t) \equiv (1-T(e,t))P(e,t,\emptyset) + T(e,t)$.

Condition (i) is identical to the one in Proposition 10, where e is evidence. Condition (iii) is stronger (when combined with the other two conditions) than the corresponding condition (iii) of Proposition 10. It ensures that the *least* talented agent with the *least* evidence does not have incentives to over-report her talent and/or evidence to imitate an agent (e,t) whose composite measure is *higher*.[41] Put differently, in order to accept some agents with higher probability (than agent (0,0)), the principal needs to test those agents with high enough probability to prevent agent (0,0) from imitating them to get accepted in case she is not tested. The condition is stricter than the one in Proposition 10 because now agents can also imitate agents with higher e to get accepted in the case that they are not tested. Thus, that agents cannot present evidence on e enhances the need to test.

Last, condition (ii) makes sure that agents do not want to under- or over-report their e to imitate agents whose composite measure is *not higher* than their own. Namely, an agent

[41] Combined with conditions (i) and (ii), this also means that no other agent has incentives to imitate an agent whose composite measure is higher.

(e,t) does not want to imitate an agent (e',t') with evidence $e' > e$ (resp. $e' < e$), talent $t' < t$ (resp. $t' > t$), and equal composite measure $\sigma(e',t') = \sigma(e,t)$ in order to get accepted with probability $\Pi(e',t')$ instead of $\Pi(e,t)$. Notice that for any possible level of evidence e' that agent (e,t) may reveal, because of condition (i), she will want to report her talent to be as high as possible (making sure that she will be able to pass the test), up to the point where $\sigma(e',t') = \sigma(e,t)$. The condition is stricter than the one in Proposition 10 because now agents can not only understate but also overstate e. This nullifies the advantage that agents with high e have (relative to agents with the same composite measure but lower e) when they can present evidence.

The probability of getting accepted without a test is the same for everyone. Lemma 9 shows that when testing is costly and some agents are (optimally) accepted with higher probability than other ones, the optimal mechanism satisfies condition (iii) of Proposition 1 with equality. Under free testing or when it is not optimal to accept some agents with higher probability, it is still without loss to constrain attention to mechanisms that satisfy condition (iii) of Proposition 10 with equality.

Lemma 9. Assume that the agent cannot present evidence. Given any IC mechanism $M \equiv \langle T,P \rangle$, there exists an IC mechanism $M' \equiv \langle T',P' \rangle$ with $(1 - T'(e,t))P'(e,t,\emptyset) = \Pi'(0,0)$ for every (e,t) that is outcome-equivalent to M and has at most as high testing costs as M. For $c > 0$, if also $\Pi(e,t) > \Pi(0,0)$ for a positive measure of agent types, then M' has lower testing costs than M.

By Lemma 9 $\Pi(e,t) = \Pi(0,0) + T(e,t)$. Thus, the principal's objective function, $\int_0^1 \int_0^1 [\Pi(e,t)u(e,t) - cT(e,t)]\, f(e,t)dtde$, can be written as

$$\int_0^1 \int_{\underline{e}(s)}^{\overline{e}(s)} [\Pi(e,\tau(e,s))(u(e,\tau(e,s)) - c) + c\Pi(0,0)]\, f(e,\tau(e,s))deds, \tag{5}$$

which is linear in Π, so by Bauer's maximum principle, there exists an extreme Π (among $\Pi(e,\tau(e,s))$ that are constant in e and non-decreasing in s) that solves the principal's problem. Proposition 11 describes that extreme optimal mechanism.

Proposition 11. Assume that the agent cannot present evidence. There exists an optimal mechanism with $\Pi(e,t) = \mathbf{I}(\sigma(e,t) \geq s^*)$ and $T(e,t) = \Pi(e,t) - \Pi(0,0)$ for some $s^* \in (0,1)$. That is, either

(i) $s^* = 0$, and every agent is accepted without a test or

(ii) $s^* > 0$, and each agent (e,t) is (a) accepted after getting tested if $\sigma(e,t) \geq s^*$ or (b) rejected without getting tested if $\sigma(e,t) < s^*$.

The inability of agents to present evidence on one of their attributes limits the set of IC mechanisms, thereby decreasing—in most cases—the principal's optimal payoff.

Assume for simplicity that the optimal mechanism is unique. When the testing technology is pro-e biased, if some—but not all–agents are optimally accepted without a test when e is actually evidence (i.e., the optimal evidence threshold for acceptance without a test lies strictly between 0 and 1), then the principal's payoff is lower if evidence is not available to the agents. When the testing technology is pro-t biased, if not all agents are optimally accepted without a test when e is actually evidence, then the principal's payoff is lower if evidence is not available to the agents. Particularly, the principal now has to choose s^* trading-off Type I and Type II errors even when σ is pro-t biased. Pro-t biased tests are not inherently better than pro-e biased tests when the agents cannot present evidence on e. Regardless of whether it is pro-t or -e biased, the more closely the test aligns with the principal's preferences, the higher the principal's optimal payoff is.

The comparison between the baseline model and the case where the agent cannot present evidence implies the following about the "signal jamming" problem that arises in career concern models (see, e.g., Holmström, 1999). Under—as in career concerns models—free monitoring of the employee's productivity,[42] if the employer can ask for hard evidence of effort, then the signal jamming problem is mitigated if productivity is sensitive enough to talent—compared to the employer's preferences for accepting (e.g., promoting) the employee. However, when productivity is *not* sensitive enough to talent, the signal jamming problem persists even if the employer can ask for evidence of effort. Agents have incentives to withhold evidence, which they should be paid information rents to reveal.

5.2 Costly test design

Treating the testing technology σ as exogenous is reasonable in several applications. For example, in hiring by prestigious employers (section 4.1), the employee's production function in the less prestigious position is not chosen by the prestigious employer. In promotion decisions (section 4.2), the employee's production function in the current position depends on her current job description and responsibilities, which should mostly reflect the firm's regular operating needs rather than support the employer's promotion decisions.

However, in other cases (e.g., hiring decisions where testing amounts to actual tests and interviews), the principal may be able to choose the testing technology. How does his problem change in that case? Let there be a cost $C(\sigma)$ that the principal needs to pay before the interaction with the agent, so that she can use testing technology σ during the interaction with the agent. Indeed, it is reasonable that the principal needs to design a test (if she designs a test at all) *before* the interaction with the agent due to time constraints and the complexity of designing a test. During the interaction with the

[42]The argument still holds as long as monitoring is not too costly. If it is, then no monitoring occurs, so there can be no signal jamming either.

agent, the principal can only choose whether to administer the test at cost c. Then, the principal's problem can be solved in two steps: (i) finding the optimal mechanism for each possible testing technology $\sigma \in \Sigma$, and then (ii) choosing the optimal testing technology $\sigma^* \in \Sigma$ form the set Σ of conceivable testing technologies. The solution to the first step is the one we have already described.[43]

If tests that are more sensitive to talent are more expensive to devise, our results imply that as long as the test is under-sensitive (compared to the principal's preferences) to talent, there are gains from increasing its sensitivity to it, which the principal will have to compare to the cost of making the test more sensitive to talent. The principal will want to make the tests at most as sensitive to talent as his preferences are, since tests that are overly sensitive to talent are as effective as those that are exactly aligned with the principal's preferences.[44]

However, when agents cannot present evidence on any of their attributes, the principal can always gain from finely calibrating the test's sensitivity (to the agent's attributes) to align it with his preferences. Regardless of whether it is pro-t or -e biased, the more closely the test aligns with the principal's preferences, the higher the principal's payoff is.

5.3 Endogenous evidence production

If the agent produces evidence before the interaction with the principal, in some applications, the principal may be able to affect the agent's evidence production by committing to a mechanism *before* the agent produces evidence. Indeed, in promotion decisions (section 4.2), the employer may use the prospect of promotion to incentivize the employee to exert effort.[45] Treating evidence as exogenous is more in line with other applications. For instance, in hiring decisions (section 4.1), a single employer has little labor market power to affect the candidate's (effort to obtain) credentials. Similarly, in college admissions (section 4.3), a single college cannot affect how hard high school students study.

Our characterization of the optimal mechanism then still applies—even if evidence is endogenous, as long as the principal cannot influence evidence production by committing ex ante to a mechanism. Let the agent's talent t follow a distribution with density g and support [0,1]. Taking as given the principal's mechanism, summarized by evidence and composite measure thresholds (e^*,s^*), the agent exerts costly effort $x \in \mathbb{R}_+$ to produce

[43]That is, assuming that Σ contains only pro-t and pro-e biased testing technologies (and possibly a testing technology that exactly coincides with the principal's preferences). Also, it is easy to see that there are no gains from designing multiple tests to the extent that all tests have the same administration cost c.

[44]Still, if there is uncertainty over a test's properties, in a robust approach, the principal will not need to worry about making the test overly sensitive to talent.

[45]Still, if promotions are not the main motive for the employee to exert effort (e.g., a bonus could be the main motive), effort can still be taken as approximately exogenous. For example, if the employee can obtain a higher position by changing employers, the prospect of promotion in the current company may not significantly affect her effort.

evidence.[46] Exerting effort x has cost $C_t(x)$, non-decreasing in x. Evidence is distributed, conditional on x, according to density function $h_x(e)$ with support [0,1]. Denote by $x^*(t)$ the equilibrium level of effort by type t. An equilibrium is a fixed point (x^*,e^*,s^*) where $x^* : [0,1] \to \mathbb{R}_+$ is a best-response to (e^*,s^*) and (e^*,s^*) is a best-response to x^* (i.e., the thresholds (e^*,s^*) that solve the principal's problem when the agent's type has density $f(e,t) = g(t)h_{x^*(t)}(e)$). (x^*,e^*,s^*) can be interpreted as a symmetric equilibrium where each of multiple "effort-taking" principals chooses thresholds (e^*,s^*).

While a detailed analysis of endogenous evidence production is beyond the scope of this paper, the following observation shows the importance of the fact that the optimal mechanism has been characterized under minimal assumptions on the agent's type distribution (i.e., that it admits a full-support density). In equilibrium, by exerting effort x, agent t will earn expected payoff $\int_{\min\{e^*,\varepsilon(t,s^*)\}}^1 h_x(e)de - C_t(x)$, where $\varepsilon(t,s)$ is implicitly given by $\sigma(\varepsilon(t,s),t) = s$. If, for example, $c = 0$, $e^* = 1$ and so $x^*(t) = 0$ for every $t \geq \tau(0,s^*)$ or $t \leq \tau(1,s^*)$. That is, agents so talented that they are accepted even without evidence and agents so untalented that they are rejected even if they present evidence $e = 1$ do not exert effort. More generally, the agent's incentives to exert effort—and, thus, effort itself—will often be non-monotone in t.[47] Thus, evidence and talent may be stochastically dependent in complicated ways.

6 Conclusion

This paper has proposed a model of multidimensional screening, where an agent (she) with two attributes—evidence and talent—presents evidence (i.e., verifiably discloses possibly part of her evidence) and is (possibly) tested at a cost by the principal (he), who then decides whether to accept or reject the agent. The agent cannot unilaterally prove anything about her talent. The test delivers a signal (i.e., the composite measure)—increasing in both evidence and talent—of the agent's type and the principal (weakly) values both evidence and talent in an agent. If the principal is going to test the agent, then the agent may have incentives to withhold evidence—although the principal values evidence—to influence how the principal interprets the test result. Particularly, she may

[46]Notice that the optimal mechanism can always be summarized by these two thresholds. Under pro-t biased testing, there is only an evidence threshold.

[47]For example, let $x \in [0,1]$ with $C_t(x) := \xi(t)x^2/2$, where $\xi(t) > 0$ is decreasing in t, $u(e,t) := \gamma_u e + (1-\gamma_u)t - \underline{q}$, $\sigma(e,t) := \gamma_s e + (1-\gamma_s)t$, where $1 > \gamma_s > \gamma_u$, and $H_x(e) := \int_0^e h_x(y)dy = -2xe(1-e) + e(2-e)$. Then,

$$x^*(t) = \begin{cases} 0 & \text{if } t \leq \dfrac{s^* - \gamma_s}{1-\gamma_s} \\ \dfrac{2[s^* - (1-\gamma_s)t][\gamma_s - s^* + (1-\gamma_s)t]}{\gamma_s^2 \xi(t)} & \text{if } t \in \left(\dfrac{s^* - \gamma_s}{1-\gamma_s}, \dfrac{s^*}{1-\gamma_s}\right) \\ 0 & \text{if } t \geq \dfrac{s^*}{1-\gamma_s}. \end{cases}$$

want to withhold evidence to make the principal attribute the test result to talent, thereby overestimating her talent.

This problem arises when the test (score) is less sensitive to talent than talent is valuable to the principal. In that case, the optimal mechanism features three types of inefficiencies, all of which favor high-evidence agents over low-evidence ones: (i) It accepts some unworthy agents without testing them but rather only by asking them to present a certain level of evidence, and among agents who cannot meet that evidence threshold, (ii) it accepts after testing some unworthy agents with medium-evidence and low talent, and (iii) it rejects some worthy agents with high talent but low evidence. Remarkably, this is the structure of the optimal mechanism even when the principal *only* values talent. The principal still optimally rewards evidence even though it is worthless to him.

The results indicate how less worthy individuals with high credentials or effort to show are favored—by an optimal and objective evaluation mechanism—over more worthy ones, who have however lower credentials (or effort to show). Ivy-Leaguers are immediately hired by prestigious employers, while those from more modest backgrounds have to go through less prestigious employers to prove their worth before landing a prestigious position. Even controlling for the fact that they need to first take a less prestigious position, they may still be at a disadvantage when trying to transition to a more prestigious one. Hard-working employees with mediocre managerial skills are promoted to managerial positions over less hard-working ones who would, however, make better managers.

Last, in college admissions, high school students from privileged backgrounds have an advantage over equally good or even better students from modest backgrounds—even if colleges value diversity and try to control for the applicants' unequal backgrounds but their evaluation mechanisms (e.g., standardized tests) are sensitive to the applicant's prior training. Affirmative action (i.e., trying to control for college applicants' unequal backgrounds) has limited effectiveness if two conditions are satisfied: (i) Applicants have considerable room to hide evidence of privilege, and (ii) standardized test scores reflect talent less than colleges value talent. If any of the two condition fails, then affirmative action is effective, and we should expect its reversal to have significant effects on diversity in college admissions.

References

Agarwal, R. and Gort, M. (2002). Firm and product life cycles and firm survival. *American Economic Review*, 92(2):184–190.

Armstrong, M. (1996). Multiproduct nonlinear pricing. *Econometrica*, 64(1):51–75.

Ball, I. (2025). Scoring strategic agents. *American Economic Journal: Microeconomics*, 17(1):97–129.

Bartelsman, E., Scarpetta, S., and Schivardi, F. (2005). Comparative analysis of firm demographics and survival: evidence from micro-level sources in OECD countries. *Industrial and Corporate Change*, 14(3):365–391.

Ben-Porath, E., Dekel, E., and Lipman, B. L. (2014). Optimal allocation with costly verification. *American Economic Review*, 104(12):3779–3813.

Bizzotto, J., Rüdiger, J., and Vigier, A. (2020). Testing, disclosure and approval. *Journal of Economic Theory*, 187:105002.

Börgers, T. (2015). *An introduction to the theory of mechanism design*. Oxford University Press, USA.

Cai, Y., Devanur, N. R., and Weinberg, S. M. (2019). A duality-based unified approach to bayesian mechanism design. *SIAM Journal on Computing*, 50(3):STOC16–160.

Carroll, G. (2017). Robustness and separation in multidimensional screening. *Econometrica*, 85(2):453–488.

Carroll, G. and Egorov, G. (2019). Strategic communication with minimal verification. *Econometrica*, 87(6):1867–1892.

Casale, S., Fioravanti, G., Rugai, L., Flett, G. L., and Hewitt, P. L. (2016). The interpersonal expression of perfectionism among grandiose and vulnerable narcissists: Perfectionistic self-presentation, effortless perfection, and the ability to seem perfect. *Personality and Individual Differences*, 99:320–324.

Daskalakis, C., Deckelbaum, A., and Tzamos, C. (2017). Strong duality for a multiple-good monopolist. *Econometrica*, 85(3):735–767.

Dessein, W., Frankel, A., and Kartik, N. (2024a). Test-optional admissions. *arXiv preprint arXiv:2304.07551*.

Dessein, W., Frankel, A., and Kartik, N. (2024b). The test-optional puzzle.

Dunne, P. and Hughes, A. (1994). Age, size, growth and survival: Uk companies in the 1980s. *The Journal of Industrial Economics*, 42(2):115–140.

Dunne, S. A. and Loewenstein, M. A. (1995). Costly verification of cost performance and the competition for incentive contracts. *The RAND Journal of Economics*, pages 690–703.

Dziuda, W. (2011). Strategic argumentation. *Journal of Economic Theory*, 146(4):1362–1397.

Erlanson, A. and Kleiner, A. (2020). Costly verification in collective decisions. *Theoretical Economics*, 15(3):923–954.

Evans, D. S. (1987). The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries. *The Journal of Industrial Economics*, 35(4):567–581.

Farinas, J. C. and Moreno, L. (2000). Firms' growth, size and age: A nonparametric approach. *Review of Industrial organization*, 17:249–265.

Flett, G. L., Nepon, T., Hewitt, P. L., Molnar, D. S., and Zhao, W. (2016). Projecting perfection by hiding effort: supplementing the perfectionistic self-presentation scale with a brief self-presentation measure. *Self and Identity*, 15(3):245–261.

Frankel, A. and Kartik, N. (2019). Muddled information. *Journal of Political Economy*, 127(4):1739–1776.

Frankel, A. and Kartik, N. (2022). Improving information from manipulable data. *Journal of the European Economic Association*, 20(1):79–115.

Gale, D. and Hellwig, M. (1985). Incentive-compatible debt contracts: The one-period problem. *The Review of Economic Studies*, 52(4):647–663.

Glazer, J. and Rubinstein, A. (2004). On optimal rules of persuasion. *Econometrica*, 72(6):1715–1736.

Green, J. R. and Laffont, J.-J. (1986). Partially verifiable information and mechanism design. *The Review of Economic Studies*, 53(3):447–456.

Grossman, S. J. (1981). The Informational Role of Warranties and Private Disclosure about Product Quality. *The Journal of Law and Economics*, 24(3):461–483.

Haghpanah, N. and Hartline, J. (2021). When is pure bundling optimal? *The Review of Economic Studies*, 88(3):1127–1156.

Halac, M. and Yared, P. (2020). Commitment versus flexibility with costly verification. *Journal of Political Economy*, 128(12):4523–4573.

Hart, S., Kremer, I., and Perry, M. (2017). Evidence games: Truth and commitment. *American Economic Review*, 107(3):690–713.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The review of Economic studies*, 66(1):169–182.

Kartik, N. (2009). Strategic communication with lying costs. *The Review of Economic Studies*, 76(4):1359–1395.

Kattwinkel, D. and Knoepfle, J. (2023). Costless information and costly verification: A case for transparency. *Journal of Political Economy*, 131(2):504–548.

Kleiner, A., Moldovanu, B., and Strack, P. (2021). Extreme points and majorization: Economic applications. *Econometrica*, 89(4):1557–1593.

Li, Y. (2020). Mechanism design with costly verification and limited punishments. *Journal of Economic Theory*, 186:105000.

Li, Y. (2021). Mechanism design with financially constrained agents and costly verification. *Theoretical Economics*, 16(3):1139–1194.

Manelli, A. M. and Vincent, D. R. (2007). Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *Journal of Economic theory*, 137(1):153–185.

Milgrom, P. (2008). What the seller won't tell you: Persuasion and disclosure in markets. *Journal of Economic Perspectives*, 22(2):115–131.

Milgrom, P. R. (1981). Good News and Bad News: Representation Theorems and Applications. *The Bell Journal of Economics*, 12(2):380–391.

Mussa, M. and Rosen, S. (1978). Monopoly and product quality. *Journal of Economic theory*, 18(2):301–317.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of operations research*, 6(1):58–73.

Perez-Richet, E. and Skreta, V. (2022). Test design under falsification. *Econometrica*, 90(3):1109–1142.

Rochet, J.-C. and Choné, P. (1998). Ironing, sweeping, and multidimensional screening. *Econometrica*, 66(4):783–826.

Rochet, J.-C. and Stole, L. A. (2003). The economics of multidimensional screening. *Econometric Society Monographs*, 35:150–197.

Shin, H. S. (1994). News Management and the Value of Firms. *The RAND Journal of Economics*, 25(1):58.

Sobel, J. (2020). Lying and deception in games. *Journal of Political Economy*, 128(3):907–947.

Townsend, R. M. (1979). Optimal contracts and competitive markets with costly state verification. *Journal of Economic theory*, 21(2):265–293.

Travers, L. V., Randall, E. T., Bryant, F. B., Conley, C. S., and Bohnert, A. M. (2015). The cost of perfection with apparent ease: Theoretical foundations and development of the effortless perfectionism scale. *Psychological Assessment*, 27(4):1147.

Viscusi, W. K. (1978). A Note on "Lemons" Markets with Quality Certification. *The Bell Journal of Economics*, 9(1):277–279.

Yang, F. (2022). Costly multidimensional screening. *arXiv preprint arXiv:2109.00487*.

Yang, F. (2023). Nested bundling.

Yang, F. and Yang, K. H. (2025). Multidimensional monotonicity and economic applications. *arXiv preprint arXiv:2502.18876*.

Yang, K. H. and Zentefis, A. K. (2024). Monotone function intervals: Theory and applications. *American Economic Review*, 114(8):2239–2270.

A Proofs

Proof of Lemma 1 Take an IC mechanism $M \equiv \langle T,P \rangle$. Construct the mechanism $M' \equiv \langle T',P' \rangle$ with (i) $P'_{at}(e,t) = 1$, (ii) $T'(e,t) = T(e,t)P_{at}(e,t) \leq T(e,t)$, and (iii) $P'(e,t,\emptyset) = (1 - T(e,t))P(e,t,\emptyset)/(1 - T'(e,t))$ for any (e,t).[48]

We have then that (a) $T'(e,t)P'_{at}(e,t) = T(e,t)P_{at}(e,t)$, (b) $(1 - T'(e,t))P'(e,t,\emptyset) = (1 - T(e,t))P(e,t,\emptyset)$ and (c) $\Pi'(e,t) = \Pi(e,t)$ for any (e,t). (a)-(c) combined imply that the problem of every agent type under M' is the same as it was under M, so M' is also IC. (c) means that M' is outcome-equivalent to M.

Last, to see why the second part is true, notice that for $c > 0$, M' saves on testing costs compared to M if there exists (a positive measure of types) (e,t) with $T(e,t) > 0$ and $P_{at}(e,t) < 1$. **Q.E.D.**

Proof of Proposition 1 Denote the total probability with which type (e,t) is accepted if she reports $(\widehat{e},\widehat{t})$ (with $\widehat{e} \leq e$) by

$$\widetilde{P}(\widehat{e},\widehat{t};e,t) := (1 - T(\widehat{e},\widehat{t}))P(\widehat{e},\widehat{t},\emptyset) + T(\widehat{e},\widehat{t})\mathbf{I}\left(\sigma(e,t) \geq \sigma(\widehat{e},\widehat{t})\right).$$

Also, define condition (iii') (a strengthening of condition (iii)) to say that $(1 - T(e,t))P(e,t,\emptyset) \leq \Pi(e',0)$ for every e,t,e' with $e \leq e'$.

Step 1: I first show that condition (i) is necessary for IC by showing the contrapositive. Assume that for some e,t_1,t_2 with $t_2 > t_1$, $\Pi(e,t_2) < \Pi(e,t_1)$. Then, IC of type (e,t_2) is violated, since $\widetilde{P}(e,t_1;e,t_2) = \Pi(e,t_1) > \Pi(e,t_2)$, that is, (e,t_2) can imitate (e,t_1) to (reach

[48] In $P'(e,t,\emptyset)$, if $T'(e,t) = 1$, cancel $(1 - T(e,t))$ in the numerator and $(1 - T'(e,t))$ in the denominator.

(e,t_1)'s composite measure threshold and) get accepted with higher probability that she would if she truthfully reported her type.

Step 2: I now show that condition (iii') is necessary for IC by showing the contrapositive.[49] Assume that for some e,e',t with $e' \geq e$, $(1-T(e,t))P(e,t,\emptyset) > \Pi(e',0)$. Then, IC of type $(e',0)$ is violated, since $\widetilde{P}(e,t;e',0) \geq (1-T(e,t))P(e,t,\emptyset) > \Pi(e',0)$, that is, $(e',0)$ can imitate (e,t) to get accepted with higher probability that she would if she truthfully reported her type (even if her composite measure is lower than (e,t)'s).

Step 3: I now show that provided that (i) and (iii') are satisfied, $\Pi(r,\tau(r,\sigma(e,t)))$ being non-decreasing in r over $r \in [\underline{e}(\sigma(e,t)),e]$ for every (e,t) is necessary and sufficient for IC.

IC of type (e,t) is satisfied if and only if

$$\max_{(\widehat{e,t})\leq(e,1)} \left[(1-T(\hat{e},\hat{t}))P(\hat{e},\hat{t};\emptyset) + T(\hat{e},\hat{t})\mathbf{I}\left(\sigma(e,t) \geq \sigma(\hat{e},\hat{t})\right)\right] = \Pi(e,t). \tag{6}$$

Assume that conditions (i) and (iii') are satisfied. Then, $\Pi(e,t) \geq \Pi(e,0) \geq (1 - T(\hat{e},\hat{t}))P(\hat{e},\hat{t},\emptyset)$ for any $(\hat{e},\hat{t})$ with $\hat{e} \leq e$. Therefore, (6) is equivalent to

$$\max_{(\widehat{e,t})\in\{(x,y)\in[0,1]^2:x\leq e \text{ and } \sigma(e,t)\geq\sigma(x,y)\}} \left[(1-T(\hat{e},\hat{t}))P(\hat{e},\hat{t};\emptyset) + T(\hat{e},\hat{t})\right] = \Pi(e,t). \tag{7}$$

Given that $\Pi(e,t)$ is non-decreasing in t (condition (i)), (7) can equivalently be written as

$$\max_{r\in[\underline{e}(\sigma(e,t)),e]} \{[1-T(r,\tau(r,\sigma(e,t)))]P(r,\tau(r,\sigma(e,t)),\emptyset) + T(r,\tau(r,\sigma(e,t)))\} = \Pi(e,t)$$

or equivalently,

$$e \in \operatorname*{arg\,max}_{r\in[\underline{e}(\sigma(e,t)),e]} \Pi(r,\tau(r,\sigma(e,t))). \tag{8}$$

Thus, IC is satisfied for every type if and only if for every (e,t), (8) is satisfied. This is true if and only if $\Pi(r,\tau(r,\sigma(e,t)))$ is non-decreasing in r for $r \in [\underline{e}(\sigma(e,t)),e]$ for every (e,t).

That the latter is sufficient for (8) to hold for every (e,t) is immediate. I show necessity by showing the contrapositive. Assume that for some (e,t), $\Pi(r,\tau(r,\sigma(e,t)))$ is *not* non-decreasing in r for $r \in [\underline{e}(\sigma(e,t)),e]$. That is, for some (e,t) there exist r_1,r_2 with $\underline{e}(\sigma(e,t)) \leq r_1 < r_2 \leq e$ such that $\Pi(r_2,\tau(r_2,\sigma(e,t))) < \Pi(r_1,\tau(r_1,\sigma(e,t)))$. Then,

$$r_2 \notin \operatorname*{arg\,max}_{x\in[\underline{e}(\sigma(e,t)),r_2]} \Pi(x,\tau(x,\sigma(e,t))).$$

Namely, IC of type $(r_2,\tau(r_2,\sigma(e,t)))$ is violated, as she prefers to imitate type $(r_1,\tau(r_1,\sigma(e,t)))$.

Step 4: It is easy to see that $\Pi(r,\tau(r,\sigma(e,t)))$ being non-decreasing in r over $r \in$

[49] That $P(e,0,\emptyset) = \Pi(e,0)$ follows from $T(e,0) = 0$.

$[\underline{e}(\sigma(e,t)),e]$ for every (e,t) is equivalent to condition (ii).

Step 5: Finally, notice that provided that conditions (i) and (ii) hold, conditions (iii) and (iii') are equivalent. That (iii') implies (iii) is immediate. We will show that the opposite direction also holds. Assume that conditions (i), (ii), and (iii) hold. Then, for any e,e',t with $e' \geq e$

$$\Pi(e',0) \geq \Pi(e,\tau(e,\sigma(e',0))) \geq \Pi(e,0) \geq (1-T(e,t))P(e,t,\emptyset),$$

where the first inequality follows from condition (ii),[50] the second from condition (i), and the third from condition (iii). **Q.E.D.**

Proof of Lemma 2 Take any IC mechanism $M \equiv \langle T,P\rangle$. Condition (iii) of Proposition 1 says that $\Pi(e,0) \geq (1-T(e,t))P(e,t,\emptyset)$ for any (e,t). Then, construct the mechanism $M' := \langle T',P'\rangle$ with[51]

$$\begin{aligned}
T'(e,t) &:= \Pi(e,t) - \Pi(e,0) = (1-T(e,t))P(e,t,\emptyset) + T(e,t) - \Pi(e,0) \\
&\leq \Pi(e,0) + T(e,t) - \Pi(e,0) = T(e,t), \quad \text{and} \\
P'(e,t,\emptyset) &:= \frac{\Pi(e,0)}{1-\Pi(e,t)+\Pi(e,0)} \geq \frac{(1-T(e,t))P(e,t,\emptyset)}{1-\Pi(e,t)+(1-T(e,t))P(e,t,\emptyset)} = P(e,t,\emptyset)
\end{aligned}$$

for every (e,t), where the inequalities follow from $\Pi(e,0) \geq (1-T(e,t))P(e,t,\emptyset)$.

By construction we have that $\Pi'(e,t) = \Pi(e,t)$ for every (e,t), so M' satisfies conditions (i) and (ii) of Proposition 1. By construction, we also have that for every (e,t)

$$\Pi'(e,0) = \Pi(e,0) = (1-T'(e,t))P'(e,t,\emptyset),$$

so M' also satisfies condition (iii) of Proposition 1. Therefore, M' is IC.

Last, to see why the second part is true, notice that for $c > 0$, M' saves on testing costs compared to M if there exists (a positive measure of) (e,t) with $P(e,t,\emptyset)(1-T(e,t)) < \Pi(e,0)$, since $T'(e,t) < T(e,t)$ for such (e,t). **Q.E.D.**

Proof of Lemmata 3 and 4 I prove the more general Lemma 4. It is useful to look at the principal's choice as a function $\Pi(e,\tau(e,s))$ of (e,s). The objective function (3) is linear (and, thus, convex) in Π. By the Dominated Convergence Theorem, it is also continuous in Π in the product topology. Denote by $\mathcal{P}$ the space of non-decreasing functions from $\{(e,s) \in [0,1]^2 : e \in [\underline{e}(s),\overline{e}(s)]\}$ to $[0,1]$. $\mathcal{P}$ is convex. It is also compact in the product topology, since (i) by Tychonoff's Theorem, the space $\overline{\mathcal{P}} := [0,1]^{\{(e,s)\in[0,1]^2:e\in[\underline{e}(s),\overline{e}(s)]\}}$ is compact as the product of compact spaces, and (ii) $\mathcal{P}$ is a closed subset of $\overline{\mathcal{P}}$. By Bauer's

[50]The first inequality assumes that $e \geq \underline{e}(\sigma(e',0))$. If this is not the case, using conditions (i) and (ii) iteratively, we can still show that $\Pi(e',0) \geq \Pi(e,0)$.

[51]For (e,t) such that $\Pi(e,t) = 1$ and $\Pi(e,0) = 0$, set $P'(e,t,\emptyset) = 0$.

maximum principle, it follows that there exists a maximizing function $(e,s) \to \Pi(e,\tau(e,s))$ that is an extreme point of $\mathcal{P}$. Last, a function $(e,s) \to \Pi(e,\tau(e,s))$ is an extreme point of $\mathcal{P}$ if and only if $\Pi(e,\tau(e,s)) \in \{0,1\}$ for all (e,s) is its domain.[52] The proof of this part is analogous to the one of Lemma 2.7 in Börgers (2015).

If direction: consider any non-decreasing (in e and s) Π with $\Pi(e,\tau(e,s)) \in \{0,1\}$ for all (e,s), and take any function $g : \{(e,s) \in [0,1]^2 : e \in [\underline{e}(s),\overline{e}(s)]\} \to \mathbb{R}$ such that $g(e^*,s^*) \neq 0$ for some (e^*,s^*). If $g(e^*,s^*) > 0$ and $\Pi(e^*,\tau(e^*,s^*)) = 0$, then $\Pi(e^*,\tau(e^*,s^*)) - g(e^*,s^*) < 0$. If $g(e^*,s^*) > 0$ and $\Pi(e^*,\tau(e^*,s^*)) = 1$, then $\Pi(e^*,\tau(e^*,s^*)) + g(e^*,s^*) > 1$. Similarly, if $g(e^*,s^*) < 0$ and $\Pi(e^*,\tau(e^*,s^*)) = 0$, then $\Pi(e^*,\tau(e^*,s^*)) + g(e^*,s^*) < 0$. If $g(e^*,s^*) < 0$ and $\Pi(e^*,\tau(e^*,s^*)) = 1$, then $\Pi(e^*,\tau(e^*,s^*)) - g(e^*,s^*) > 1$. Thus, Π is an extreme point.

Only if direction: now consider any non-decreasing (in e and s) Π with $\Pi(e^*,\tau(e^*,s^*)) \notin \{0,1\}$ for some (e^*,s^*). Construct function g as follows. $g(e,s) \coloneqq \Pi(e,\tau(e,s))$ for every (e,s) such that $\Pi(e,\tau(e,s)) \leq 1/2$ and $g(e,s) \coloneqq 1 - \Pi(e,\tau(e,s))$ for every (e,s) such that $\Pi(e,\tau(e,s)) > 1/2$. $g(e^*,s^*) \in (0,1)$, so $g \neq 0$. Consider the function $(e,s) \mapsto \Pi(e,\tau(e,s)) + g(e,s)$. Take any e_1,e_2,s with $e_2 \geq e_1$ and observe that if $\Pi(e_2,\tau(e_2,s)) > 1/2$, then

$$\Pi(e_2,\tau(e_2,s)) + g(e_2,s) = 1 \geq \Pi(e_1,\tau(e_1,s)) + g(e_1,s)$$

since by construction $\Pi(e,\tau(e,s)) + g(e,s) \leq 1$ for every (e,s), while if $\Pi(e_2,\tau(e_2,s)) \leq 1/2$, then (since Π is non-decreasing) $\Pi(e_1,\tau(e_1,s)) \leq \Pi(e_2,\tau(e_2,s)) \leq 1/2$, and so

$$\Pi(e_2,\tau(e_2,s)) + g(e_2,s) = 2\Pi(e_2,\tau(e_2,s)) \geq 2\Pi(e_1,\tau(e_1,s)) = \Pi(e_1,\tau(e_1,s)) + g(e_1,s).$$

Similarly, it can be seen that $\Pi(e,\tau(e,s_2)) + g(e,s_2) \geq \Pi(e,\tau(e,s_1)) + g(e,s_1)$ for any s_1,s_2,e with $s_2 \geq s_1$. Also $\Pi(e,\tau(e,s)) + g(e,s) \in [0,1]$ for every (e,s). Therefore, the function $(e,s) \mapsto \Pi(e,\tau(e,s)) + g(e,s)$ lies in the set of non-decreasing functions from $\{(e,s) \in [0,1]^2 : e \in [\underline{e}(s),\overline{e}(s)]\}$ to $[0,1]$. Similarly, it can be seen that the function $(e,s) \mapsto \Pi(e,\tau(e,s)) - g(e,s)$ lies in the set of non-decreasing functions from $\{(e,s) \in [0,1]^2 : e \in [\underline{e}(s),\overline{e}(s)]\}$ to $[0,1]$. Thus, Π is not an extreme point. **Q.E.D.**

Proof of Claims 1 and 2 The first part is immediate. The total derivative of $u(e,\tau(e,s))$ with respect to e is equal to

$$\begin{aligned}\frac{du(e,\tau(e,s))}{de} &= \frac{\partial u(e,\tau(e,s))}{\partial e} + \frac{\partial \tau(e,s)}{\partial e}\left.\frac{\partial u(e,t)}{\partial t}\right|_{t=\tau(e,s)} \\ &= \frac{\partial u(e,t)}{\partial e} + \frac{\partial \sigma(e,t)/\partial e}{\partial \sigma(e,t)/\partial t}\left.\frac{\partial u(e,t)}{\partial t}\right|_{t=\tau(e,s)},\end{aligned}$$

[52] More precisely, this should hold for almost all (e,s) is its domain (see Börgers, 2015).

and the second part follows. **Q.E.D.**

Proof of Proposition 2 We need to show that $\Pi(e,t) = \mathbf{I}\left(u(e,t) > 0\right)$ satisfies conditions (i) and (ii) of Proposition 1.

Condition (i): Since $\Pi(e,t) \in \{0,1\}$ for every (e,t), it suffices to show that for any (e,t), if $\Pi(e,t) = 1$, then $\Pi(e,t') = 1$ for every $t' \geq t$. Indeed, we have that for any (e,t)

$$\Pi(e,t) = 1 \implies u(e,t) > 0 \implies u(e,t') > 0 \text{ for every } t' \geq t,$$

where the second implication follows since $u(e,t)$ is non-decreasing in t.

Condition (ii): Similarly, it suffices to show that for any (r,s), if $\Pi(r,\tau(r,s)) = 1$, then $\Pi(r',\tau(r',s)) = 1$ for every $r' \in [r, \overline{e}(s)]$. Indeed, we have that for any (r,s), $\Pi(r,\tau(r,s)) = 1$ implies that $u(r,\tau(r,s)) > 0$, which in turn implies that $u(r',\tau(r',s)) > 0$ for every $r' \in [r, \overline{e}(s)]$.

To see why the last part follows, assume instead that $u(r',\tau(r',s)) \leq 0$ for some $r' \in [r, \overline{e}(s)]$. Particularly, it must be $r' > r$. Since σ is pro-t biased, there exists e_s such that if $e > e_s$ (resp. $e \leq e_s$) and $\sigma(e,t) = s$, then $u(e,t) > 0$ (resp. $u(e,t) \leq 0$). We have that $u(r',\tau(r',s)) \leq 0$, so σ being pro-t biased implies that $r' \leq e_s$. But $r' > r$, so $r < e_s$, and since $\sigma(r,\tau(r,s)) = s$, σ being pro-t biased implies that $u(r,\tau(r,s)) \leq 0$, a contradiction. **Q.E.D.**

Proof of Proposition 3 *Step 1:* In definition 4 of pro-e biased testing, for s such that $u(e,t) > c = 0$ (resp. $u(e,t) \leq 0$) for every $(e,t) \in I_\sigma(s)$, e_s is not uniquely defined. In that case, for s such that $u(e,t) > 0$ (resp. $u(e,t) \leq 0$) for every $(e,t) \in I_\sigma(s)$, set $e_s = \overline{e}(s)$ (resp. $e_s = \underline{e}(s)$). We will show that (under pro-e biased testing) e_s is non-decreasing in s. Take any $\underline{s},\overline{s} \in [0,1]$ with $\overline{s} > \underline{s}$, and define $S \coloneqq (e_{\overline{s}},e_{\underline{s}}) \cap [\underline{e}(\overline{s}), \overline{e}(\overline{s})] \cap [\underline{e}(\underline{s}), \overline{e}(\underline{s})]$.

Step 1, case 1: If $S = \emptyset$, then $e_{\underline{s}} \leq e_{\overline{s}}$. To see this, consider the following two subcases.

Step 1, case 1(a): if $\underline{e}(\overline{s}) \geq \overline{e}(\underline{s})$, then $e_{\underline{s}} \leq \overline{e}(\underline{s}) \leq \underline{e}(\overline{s}) \leq e_{\overline{s}}$, so $e_{\underline{s}} \leq e_{\overline{s}}$, a contradiction.

Step 1, case 1(b): if $\underline{e}(\overline{s}) < \overline{e}(\underline{s})$, then $S = (e_{\overline{s}},e_{\underline{s}}) \cap [\underline{e}(\overline{s}), \overline{e}(\underline{s})]$. Since $S = \emptyset$, either $\underline{e}(\overline{s}) \geq e_{\underline{s}}$ or $\overline{e}(\underline{s}) \leq e_{\overline{s}}$. If $\underline{e}(\overline{s}) \geq e_{\underline{s}}$, then $e_{\underline{s}} \leq \underline{e}(\overline{s}) \leq e_{\overline{s}}$, so $e_{\underline{s}} \leq e_{\overline{s}}$, a contradiction. Similarly, if $\overline{e}(\underline{s}) \leq e_{\overline{s}}$, then $e_{\underline{s}} \leq \overline{e}(\underline{s}) \leq e_{\overline{s}}$, so $e_{\underline{s}} \leq e_{\overline{s}}$, a contradiction.

Step 1, case 2: We now prove by contradiction that if $S \neq \emptyset$, then $e_{\underline{s}} \leq e_{\overline{s}}$. To this end, assume that $S \neq \emptyset$ and $e_{\underline{s}} > e_{\overline{s}}$. Given that $S \neq \emptyset$, we can take some $e^* \in S$. Since $e^* \in [\underline{e}(\underline{s}), \overline{e}(\underline{s})]$ and σ is continuous, there exists $t^* \in [0,1]$ such that $\sigma(e^*,t^*) = \underline{s}$. Since σ is pro-e biased and $e^* < e_{\underline{s}}$, it follows that $u(e^*,t^*) > 0$. Similarly, since σ is pro-e biased, $e^* > e_{\overline{s}}$, and $e^* \in [\underline{e}(\overline{s}), \overline{e}(\overline{s})]$, there exists $t^{**} \in [0,1]$ such that $\sigma(e^*,t^{**}) = \overline{s}$ and $u(e^*,t^{**}) \leq 0$. Also, because $\overline{s} > \underline{s}$ and $\sigma(e,t)$ is increasing in t, $t^{**} > t^*$. Overall, we have $t^{**} > t^*$ and $u(e^*,t^*) > 0 \geq u(e^*,t^{**})$, a contradiction to $u(e,t)$ being non-decreasing in t.

Step 2: Given e_s, define also t_s implicitly given by $\sigma(e_s,t_s) = s$. We have then that for every composite measure $s \in [0,1]$, (e_s,t_s) is the "threshold" agent who lies on the iso-test-score curve $I_\sigma(s)$. That is, any other agent (e,t) on that iso-test-score curve with $e < e_s$ (resp. $e > e_s$) gives—if accepted—a positive (resp. negative) payoff to the principal.

We divide the problem of finding an optimal IC mechanism in three parts. First, we fix an arbitrary "partial" IC mechanism $s \mapsto \Pi(e_s,t_s)$ for every $s \in [0,1]$. Then, we complete that partial IC mechanism (i.e., we assign a value to $\Pi(e,t)$ for every (e,t) for which $\Pi(e,t)$ has no been assigned a value in the first step), so that the complete mechanism is IC and optimal given the fixed partial mechanism. Finally, we find an optimal partial mechanism.

Step 3: Fix the value of $\Pi(e_s,t_s)$ for every $s \in [0,1]$ such that these values are part of some IC mechanism.[53] Given that e_s is non-decreasing in s, by Proposition 1, the values of $\Pi(e_s,t_s)$ are part of some IC mechanism if and only if $\Pi(e_s,t_s)$ is non-decreasing in s. Therefore, by Proposition 4, there exists an optimal mechanism with $\Pi(e_s,t_s) = \mathbf{I}(s \geq \underline{s})$ for some $\underline{s} \in [0,1]$.

Step 4: It follows then that for IC to be satisfied by the complete mechanism, it must be that (i) $\Pi(e,t) = 1$ for every (e,t) such that $e > e_s$ and $\sigma(e,t) = s$ for some $s \geq \underline{s}$ and (ii) $\Pi(e,t) = 0$ for every (e,t) such that $e < e_s$ and $\sigma(e,t) = s$ for some $s < \underline{s}$. Also, since (e_s,t_s) is the "threshold" agent, the principal wants to make $\Pi(e,t)$ as high (resp. low) as possible for every (e,t) such that $e < e_s$ (resp. $e > e_s$). Thus, given the IC constraint, it is optimal to set (i) $\Pi(e,t) = 1$ for every (e,t) such that $e < e_s$ and $\sigma(e,t) = s$ for some $s \geq \underline{s}$ and (ii) $\Pi(e,t) = 0$ for every (e,t) such that $e > e_s$ and $\sigma(e,t) = s$ for some $s < \underline{s}$. **Q.E.D.**

Proof of Proposition 4 By IC conditions (i) and (ii) of Proposition 1, any IC mechanism has $\Pi(e,0)$ non-decreasing in e. Thus, given Lemma 4, there exists an optimal mechanism with $\Pi(e,0) = \mathbf{I}(e \geq e^*)$ for some $e^* \in [0,1]$. The objective function (3) then becomes

$$\int_0^1 \int_{\min\{\underline{e}(s),e^*\}}^{\min\{\overline{e}(s),e^*\}} \left[\Pi(e,\tau(e,s))(u(e,\tau(e,s)) - c)\right] f(e,\tau(e,s))deds + \int_0^1 \int_{e^*}^1 u(e,t)f(e,t)dedt.$$

The mechanism affects the second term only through e^*. Given e^*, setting $\Pi(e,t) = \mathbf{I}\left(u(e,t) > c \text{ or } e \geq e^*\right)$ maximizes the first term and—given that σ is pro-t biased—makes the mechanism IC, since it satisfies conditions (i) and (ii) of Proposition 1. $T(e,t) = \mathbf{I}(u(e,t) \geq c \text{ and } e < e^*)$ is backed out from Lemma 4. **Q.E.D.**

Proof of Proposition 5 By IC conditions (i) and (ii) of Proposition 1, any IC mechanism has $\Pi(e,0)$ non-decreasing in e. Thus, given Lemma 4, there exists an optimal mechanism

[53] That is, fix the value of $\Pi(e_s,t_s)$ for every $s \in [0,1]$ to be such that there exists IC $\Pi : [0,1]^2 \to [0,1]$ that agrees with the values of $\Pi(e_s,t_s)$ for every $s \in [0,1]$.

with $\Pi(e,0) = \mathbf{I}(e \geq e^*)$ for some $e^* \in [0,1]$. The objective function (3) then becomes

$$\int_0^1 \int_{\min\{\underline{e}(s),e^*\}}^{\min\{\overline{e}(s),e^*\}} \left[\Pi(e,\tau(e,s))(u(e,\tau(e,s)) - c)\right] f(e,\tau(e,s)) deds \\ + \int_0^1 \int_{e^*}^1 u(e,t)f(e,t)dedt.$$

The mechanism affects the second term only through e^*. Given e^*, maximizing the first term is equivalent to the problem studied by Proposition 3 with the principal's payoff function given by $u(e,t) - c$. Thus, for $e < e^*$, $\Pi(e,t) = \mathbf{I}(\sigma(e,t) \geq s^*$ for some $s^* \in [0,1]$ maximizes the first term (under the IC conditions, when the problem is restricted to $(e,t) < (e^*,1)$). The complete mechanism then has $\Pi(e,t) = \mathbf{I}(\sigma(e,t) \geq s^*$ or $e \geq e^*)$, which satisfies conditions (i) and (ii) of Proposition 1. $T(e,t) = \mathbf{I}(\sigma(e,t) \geq s^*$ and $e < e^*)$ is backed out from Lemma 4. **Q.E.D.**

Proof of Lemma 5 Take any two agents $(\boldsymbol{e},\boldsymbol{t})$ and $(\boldsymbol{e},\boldsymbol{t}')$ with $\sigma(\boldsymbol{e},\boldsymbol{t}) = \sigma(\boldsymbol{e},\boldsymbol{t}')$. $(\boldsymbol{e},\boldsymbol{t})$'s IC requires $\Pi(\boldsymbol{e},\boldsymbol{t}) \geq \Pi(\boldsymbol{e},\boldsymbol{t}')$. $(\boldsymbol{e},\boldsymbol{t}')$'s IC requires $\Pi(\boldsymbol{e},\boldsymbol{t}') \geq \Pi(\boldsymbol{e},\boldsymbol{t})$. **Q.E.D.**

Proof of Lemma 6 Take any IC mechanism M. Construct the mechanism $M' \equiv \langle T',P' \rangle$ with[54]

$$T'(\boldsymbol{e},\boldsymbol{t}) \coloneqq \inf_{\boldsymbol{t}' \text{ s.t. } \sigma(\boldsymbol{e},\boldsymbol{t}')=\sigma(\boldsymbol{e},\boldsymbol{t})} T(\boldsymbol{e},\boldsymbol{t}) \leq T(\boldsymbol{e},\boldsymbol{t}), \quad \text{and}$$
$$P'(\boldsymbol{e},\boldsymbol{t},\emptyset) \coloneqq \frac{\Pi(\boldsymbol{e},\boldsymbol{t}) - T'(\boldsymbol{e},\boldsymbol{t})}{1 - T'(\boldsymbol{e},\boldsymbol{t})} \geq \frac{\Pi(\boldsymbol{e},\boldsymbol{t}) - T(\boldsymbol{e},\boldsymbol{t})}{1 - T(\boldsymbol{e},\boldsymbol{t})} = P(\boldsymbol{e},\boldsymbol{t},\emptyset)$$

for every $(\boldsymbol{e},\boldsymbol{t})$. Then, $\Pi'(\boldsymbol{e},\boldsymbol{t}) = (1 - T'(\boldsymbol{e},\boldsymbol{t}))P'(\boldsymbol{e},\boldsymbol{t},\emptyset) + T'(\boldsymbol{e},\boldsymbol{t}) = \Pi(\boldsymbol{e},\boldsymbol{t})$ for every $(\boldsymbol{e},\boldsymbol{t})$, where the second equality follows by construction of M'. Thus, M' is outcome-equivalent to M. Given that M is IC, outcome-equivalence implies that under M', no agent has incentives to imitate an agent with composite measure that is not higher than their own.

It remains to show that under mechanism M', no agent has incentives to imitate an agent with higher composite measure than her own. Take any agent $(\boldsymbol{e},\boldsymbol{t})$, evidence $\boldsymbol{e}' \leq \boldsymbol{e}$, and talent $\boldsymbol{t}'$. It holds that

$$\begin{aligned}
\Pi'(\boldsymbol{e},\boldsymbol{t}) = \Pi(\boldsymbol{e},\boldsymbol{t}) &\geq \sup_{\tilde{\boldsymbol{t}} \text{ s.t. } \sigma(\boldsymbol{e}',\tilde{\boldsymbol{t}})=\sigma(\boldsymbol{e}',\boldsymbol{t}')} \left\{(1 - T(\boldsymbol{e}',\tilde{\boldsymbol{t}}))P(\boldsymbol{e}',\tilde{\boldsymbol{t}},\emptyset)\right\} \\
&= \sup_{\tilde{\boldsymbol{t}} \text{ s.t. } \sigma(\boldsymbol{e}',\tilde{\boldsymbol{t}})=\sigma(\boldsymbol{e}',\boldsymbol{t}')} \left\{\Pi(\boldsymbol{e}',\tilde{\boldsymbol{t}}) - T(\boldsymbol{e}',\tilde{\boldsymbol{t}})\right\} \\
&= \Pi(\boldsymbol{e}',\boldsymbol{t}') + \sup_{\tilde{\boldsymbol{t}} \text{ s.t. } \sigma(\boldsymbol{e}',\tilde{\boldsymbol{t}})=\sigma(\boldsymbol{e}',\boldsymbol{t}')} \left\{-T(\boldsymbol{e}',\tilde{\boldsymbol{t}})\right\} \\
&= \Pi(\boldsymbol{e}',\boldsymbol{t}') - \inf_{\tilde{\boldsymbol{t}} \text{ s.t. } \sigma(\boldsymbol{e}',\tilde{\boldsymbol{t}})=\sigma(\boldsymbol{e}',\boldsymbol{t}')} T(\boldsymbol{e}',\tilde{\boldsymbol{t}})
\end{aligned}$$

[54] For $\boldsymbol{e}$ such that $\inf_{\boldsymbol{t}' \text{ s.t. } \sigma(\boldsymbol{e},\boldsymbol{t}')=\sigma(\boldsymbol{e},\boldsymbol{t})} T(\boldsymbol{e},\boldsymbol{t}) = 1$, set $P'(\boldsymbol{e},\boldsymbol{t},\emptyset) = P(\boldsymbol{e},\boldsymbol{t},\emptyset)$.

$$= \Pi'(\boldsymbol{e}',\boldsymbol{t}') - T'(\boldsymbol{e}',\boldsymbol{t}') = (1 - T'(\boldsymbol{e}',\boldsymbol{t}'))P'(\boldsymbol{e}',\boldsymbol{t}',\emptyset),$$

where (i) the first equality follows by construction of M', (ii) the inequality by IC of M, (iii) the second equality by definition of Π, (iv) the third equality by Lemma 5 and IC of M, which together imply that $\Pi(\boldsymbol{e}',\tilde{\boldsymbol{t}}) = \Pi(\boldsymbol{e}',\boldsymbol{t}')$ for every $\tilde{\boldsymbol{t}}$ such that $\sigma(\boldsymbol{e}',\tilde{\boldsymbol{t}}) = s$, (v) the fifth inequality by construction of M', and the final equality by definition of Π'. We have thus shown that for any agent $(\boldsymbol{e},\boldsymbol{t})$, $\Pi'(\boldsymbol{e},\boldsymbol{t}) \geq (1 - T'(\boldsymbol{e}',\boldsymbol{t}'))P'(\boldsymbol{e}',\boldsymbol{t}',\emptyset)$ for every $(\boldsymbol{e}',\boldsymbol{t}') \leq (\boldsymbol{e},\boldsymbol{1})$, so under mechanism M', no agent has incentives to imitate an agent with higher composite measure than her own.

For $c > 0$, M' also minimizes testing costs. **Q.E.D.**

Proof of Proposition 6 The proof proceeds like the proof of Proposition 1 and is, thus, omitted. **Q.E.D.**

Proof of Lemma 7 The proof proceeds like the proof of Lemma 2.

Take any IC mechanism $M \equiv \langle T,P \rangle$. Condition (iii) of Proposition 6 says that $\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) \geq (1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset)$ for any $(\boldsymbol{e},s)$. Then, construct the mechanism $M' := \langle T',P' \rangle$ with[55]

$$\begin{aligned}
T'(\boldsymbol{e},s) &:= \Pi(\boldsymbol{e},s) - \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) = (1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset) + T(\boldsymbol{e},s) - \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) \\
&\leq \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) + T(\boldsymbol{e},s) - \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) = T(\boldsymbol{e},s), \quad \text{and} \\
P'(\boldsymbol{e},s,\emptyset) &:= \frac{\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0}))}{1 - \Pi(\boldsymbol{e},s) + \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0}))} \geq \frac{(1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset)}{1 - \Pi(\boldsymbol{e},s) + (1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset)} = P(\boldsymbol{e},s,\emptyset)
\end{aligned}$$

for every $(\boldsymbol{e},s)$, where the inequalities follow from $\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) \geq (1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset)$.

By construction we have that $\Pi'(\boldsymbol{e},s) = \Pi(\boldsymbol{e},s)$ for every $(\boldsymbol{e},s)$, so M' satisfies conditions (i) and (ii) of Proposition 1. By construction, we also have that for every $(\boldsymbol{e},s)$

$$\Pi'(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) = \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) = (1 - T'(\boldsymbol{e},s))P'(\boldsymbol{e},s,\emptyset),$$

so M' also satisfies condition (iii) of Proposition 1. Therefore, M' is IC.

Last, to see why the second part is true, notice that for $c > 0$, M' saves on testing costs compared to M if there exists (a positive measure of) $(\boldsymbol{e},s)$ with $P(\boldsymbol{e},s,\emptyset)(1 - T(\boldsymbol{e},s)) < \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0}))$, since $T'(\boldsymbol{e},s) < T(\boldsymbol{e},s)$ for such $(\boldsymbol{e},s)$. **Q.E.D.**

Proof of Proposition 7 Let $M \equiv \langle T,P \rangle$ be an optimal deterministic mechanism with $\Pi(\boldsymbol{e},s) \equiv (1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset) + T(\boldsymbol{e},s)$. Define $E^* := \{\boldsymbol{e} \in [0,1]^m : \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) = 1\}$ (so $\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) = 0$ for every $\boldsymbol{e} \notin E^*$). Given that M is IC, conditions (i) and (ii) of Proposition 6 combined imply that E^* is an upper set of $[0,1]^m$. To see this, take any $\boldsymbol{e} \in E^*$

[55] For $(\boldsymbol{e},s)$ such that $\Pi(\boldsymbol{e},s) = 1$ and $\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) = 0$, set $P'(\boldsymbol{e},s,\emptyset) = 0$.

and any $\boldsymbol{e}' \in [0,1]^m$. If $\boldsymbol{e}' \geq \boldsymbol{e}$, then $\Pi(\boldsymbol{e}',\sigma(\boldsymbol{e}',\mathbf{0})) \geq \Pi(\boldsymbol{e},\sigma(\boldsymbol{e}',\mathbf{0})) \geq \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) = 1$, so $\Pi(\boldsymbol{e}',\sigma(\boldsymbol{e}',\mathbf{0})) = 1$ and thus $\boldsymbol{e}' \in E^*$. The first inequality follows from condition (ii) and $\boldsymbol{e}' \geq \boldsymbol{e}$. The second inequality follows from condition (i), $\boldsymbol{e}' \geq \boldsymbol{e}$, and σ being increasing.[56]

Also, condition (i) of Proposition 6 implies that $\Pi(\boldsymbol{e},s) = 1$ for every $\boldsymbol{e} \in E^*$ and every $s \in [0,1]$. Then, the principal's objective function (4) can be written as

$$\int_{e\in E^*}\int_{\sigma(\boldsymbol{e},\mathbf{0})}^{\sigma(\boldsymbol{e},\mathbf{1})} \widetilde{u}(\boldsymbol{e},s)\widetilde{f}(\boldsymbol{e},s)dsd\boldsymbol{e} + \int_{e\notin E^*}\int_{\sigma(\boldsymbol{e},\mathbf{0})}^{\sigma(\boldsymbol{e},\mathbf{1})} \left[\Pi(\boldsymbol{e},s)(\widetilde{u}(\boldsymbol{e},s)-c)\right]\widetilde{f}(\boldsymbol{e},s)dsd\boldsymbol{e}.$$

The first term depends on the mechanism M only through E^*. The second term depends on the mechanism M only through the values of Π for $\boldsymbol{e} \notin E^*$. Setting $\Pi(\boldsymbol{e},s) = \mathbf{I}(\widetilde{u}(\boldsymbol{e},s) > c)$ for every $\boldsymbol{e} \notin E^*$ maximizes the second term. It is also IC.

To show this, we first prove that $\Pi(\boldsymbol{e},s) = \mathbf{I}(\widetilde{u}(\boldsymbol{e},s) > c \text{ or } \boldsymbol{e} \in E^*)$ satisfies condition (i) of Proposition 6. Take any $\boldsymbol{e},s,s'$ with $s' > s$. It suffices to show that $\Pi(\boldsymbol{e},s') = 0$ implies $\Pi(\boldsymbol{e},s) = 0$. If $\Pi(\boldsymbol{e},s') = 0$, then $\widetilde{u}(\boldsymbol{e},s') \leq c$ and $\boldsymbol{e} \notin E^*$. Since $\widetilde{u}(\boldsymbol{e},s)$ is non-decreasing in s, $\widetilde{u}(\boldsymbol{e},s) \leq \widetilde{u}(\boldsymbol{e},s') \leq c$. Therefore, $\Pi(\boldsymbol{e},s) = 0$.

It remains to show that $\Pi(\boldsymbol{e},s) = \mathbf{I}(\widetilde{u}(\boldsymbol{e},s) > c \text{ or } \boldsymbol{e} \in E^*)$ satisfies condition (ii) of Proposition 6. Take any $\boldsymbol{e},\boldsymbol{e}',s$ with $\boldsymbol{e}' \geq \boldsymbol{e}$. We need to show that $\Pi(\boldsymbol{e}',s) = 0$ implies $\Pi(\boldsymbol{e},s) = 0$. If $\Pi(\boldsymbol{e}',s) = 0$, then $\widetilde{u}(\boldsymbol{e}',s) \leq c$ and $\boldsymbol{e}' \notin E^*$. It follows then that $\boldsymbol{e} \notin E^*$, since E^* is an upper set of $[0,1]^m$, $\boldsymbol{e}' \geq \boldsymbol{e}$, and $\boldsymbol{e}' \notin E^*$. It remains to show that $\widetilde{u}(\boldsymbol{e},s) \leq c$. We will show this by contradiction. Assume that $\widetilde{u}(\boldsymbol{e},s) > c$. Then, we have that $\widetilde{u}(\boldsymbol{e},s) > c \geq \widetilde{u}(\boldsymbol{e}',s)$, which, given that σ is pro-t biased, implies that $\boldsymbol{e}' \not\geq \boldsymbol{e}$, a contradiction. **Q.E.D.**

Proof of Proposition 8 Let $M \equiv \langle T,P \rangle$ be an optimal deterministic mechanism with $\Pi(\boldsymbol{e},s) \equiv (1-T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset) + T(\boldsymbol{e},s)$. Define $E^* := \{\boldsymbol{e} \in [0,1]^m : \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) = 1\}$ (so $\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) = 0$ for every $\boldsymbol{e} \notin E^*$). Given that M is IC, conditions (i) and (ii) of Proposition 6 combined imply that E^* is an upper set of $[0,1]^m$.

Also, condition (i) of Proposition 6 implies that $\Pi(\boldsymbol{e},s) = 1$ for every $\boldsymbol{e} \in E^*$ and every $s \in [0,1]$. Then, the principal's objective function (4) can be written as

$$\int_{e\in E^*}\int_{\sigma(\boldsymbol{e},\mathbf{0})}^{\sigma(\boldsymbol{e},\mathbf{1})} \widetilde{u}(\boldsymbol{e},s)\widetilde{f}(\boldsymbol{e},s)dsd\boldsymbol{e} + \int_{e\notin E^*}\int_{\sigma(\boldsymbol{e},\mathbf{0})}^{\sigma(\boldsymbol{e},\mathbf{1})} \left[\Pi(\boldsymbol{e},s)(\widetilde{u}(\boldsymbol{e},s)-c)\right]\widetilde{f}(\boldsymbol{e},s)dsd\boldsymbol{e}.$$

The first term depends on the mechanism M only through E^*. The second term depends on the mechanism M only through the values of Π for $\boldsymbol{e} \notin E^*$.

Take any $(\boldsymbol{e},s),(\boldsymbol{e}',s') \in I_{\widetilde{u}}(c) \setminus E^*$ with $s \neq s'$. That $(\boldsymbol{e},s),(\boldsymbol{e}',s') \in I_{\widetilde{u}}(c)$ means that $\widetilde{u}(\boldsymbol{e},s) = \widetilde{u}(\boldsymbol{e}',s') = c$. First, we show that if $\boldsymbol{e}' \not\geq \boldsymbol{e}$, then $s' < s$. Let $\boldsymbol{e}' \not\geq \boldsymbol{e}$:

Case 1: if $s' \in [\sigma(\boldsymbol{e},\mathbf{0}),\sigma(\boldsymbol{e},\mathbf{1})]$, then σ being pro-e biased implies that $\widetilde{u}(\boldsymbol{e},s') \leq c$. To

[56] If $\sigma(\boldsymbol{e}',\mathbf{0}) > \sigma(\boldsymbol{e},\mathbf{1})$, then $\Pi(\boldsymbol{e},\sigma(\boldsymbol{e}',\mathbf{0}))$ is not well-defined (since there is no agent with evidence $\boldsymbol{e}$ and composite measure $\sigma(\boldsymbol{e}',\mathbf{0})$) but the inequalities still follow if we use conditions (i) and (ii) iteratively.

see this, notice that if instead $\widetilde{u}(\boldsymbol{e},s') > c$, then we would have $\widetilde{u}(\boldsymbol{e},s') > c = \widetilde{u}(\boldsymbol{e}',s')$, so pro-$e$ biased testing would imply that $\boldsymbol{e}' \geq \boldsymbol{e}$, a contradiction. We then have that $\widetilde{u}(\boldsymbol{e},s') \leq c = \widetilde{u}(\boldsymbol{e},s)$. Particularly, $\widetilde{u}(\boldsymbol{e},s') < c = \widetilde{u}(\boldsymbol{e},s)$, because $\widetilde{u}(\boldsymbol{e},s') = \widetilde{u}(\boldsymbol{e},s) = \widetilde{u}(\boldsymbol{e}',s') = c$ is not possible by the Regularity Assumption. Given that $\widetilde{u}(\boldsymbol{e},s)$ is non-decreasing in s, $\widetilde{u}(\boldsymbol{e},s') < c = \widetilde{u}(\boldsymbol{e},s)$ implies that $s' < s$.

Case 2: if $s' < \sigma(\boldsymbol{e},\mathbf{0})$, then since $s \in [\sigma(\boldsymbol{e},\mathbf{0}),\sigma(\boldsymbol{e},\mathbf{1})]$, it follows that $s' < s$.

Case 3a: if $s' > \sigma(\boldsymbol{e},\mathbf{1})$ and $\sigma(\boldsymbol{e},\mathbf{1}) \in [\sigma(\boldsymbol{e}',\mathbf{0}),\sigma(\boldsymbol{e}',\mathbf{1})]$, then because $\sigma(\boldsymbol{e},\mathbf{1}) \geq s$ and $\widetilde{u}(\boldsymbol{e},s)$ is non-decreasing in s, it follows that $\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{1})) \geq \widetilde{u}(\boldsymbol{e},s) = c$. Thus, we have $\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{1})) \geq c = \widetilde{u}(\boldsymbol{e}',s') \geq \widetilde{u}(\boldsymbol{e}',\sigma(\boldsymbol{e},\mathbf{1}))$ with at least one inequality holding strictly (for otherwise $\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{1})) = \widetilde{u}(\boldsymbol{e}',\sigma(\boldsymbol{e},\mathbf{1})) = \widetilde{u}(\boldsymbol{e}',s') = c$ with $s' \neq \sigma(\boldsymbol{e},\mathbf{1})$ and $\boldsymbol{e} \neq \boldsymbol{e}'$, which is not possible by the Regularity Assumption), so pro-e biased testing implies that $\boldsymbol{e}' \geq \boldsymbol{e}$, a contradiction. Therefore, Case 3a is impossible.

Case 3b: if $s' > \sigma(\boldsymbol{e},\mathbf{1})$ and $\sigma(\boldsymbol{e},\mathbf{1}) < \sigma(\boldsymbol{e}',\mathbf{0})$, then by continuity and monotonicity of σ and because $\sigma(\boldsymbol{e},\mathbf{1}) \in [\sigma(\mathbf{0},\mathbf{0}),\sigma(\boldsymbol{e}',\mathbf{0}))$ there exists $\boldsymbol{e}'' \leq \boldsymbol{e}'$ such that $\sigma(\boldsymbol{e}'',\mathbf{0}) = \sigma(\boldsymbol{e},\mathbf{1})$. We have then that

$$\begin{aligned}
\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{1})) \geq \widetilde{u}(\boldsymbol{e},s) = c = \widetilde{u}(\boldsymbol{e}',s') &\geq \min_{\boldsymbol{t}:\sigma(\boldsymbol{e}',\boldsymbol{t})=s'} u(\boldsymbol{e}',\boldsymbol{t}) \\
&= u(\boldsymbol{e}', \operatorname*{arg\,min}_{\boldsymbol{t}:\sigma(\boldsymbol{e}',\boldsymbol{t})=s'} u(\boldsymbol{e}',\boldsymbol{t})) \geq u(\boldsymbol{e}'', \operatorname*{arg\,min}_{\boldsymbol{t}:\sigma(\boldsymbol{e}',\boldsymbol{t})=s'} u(\boldsymbol{e}',\boldsymbol{t})) \geq u(\boldsymbol{e}'',\mathbf{0}) \\
&= \mathbb{E}_{\mathbf{t}}\left[u(\boldsymbol{e}'',\boldsymbol{t})|\sigma(\boldsymbol{e}'',\boldsymbol{t}) = \sigma(\boldsymbol{e}'',\mathbf{0})\right] \equiv \widetilde{u}(\boldsymbol{e}'',\sigma(\boldsymbol{e}'',\mathbf{0})) = \widetilde{u}(\boldsymbol{e}'',\sigma(\boldsymbol{e},\mathbf{1}))
\end{aligned}$$

with at least one inequality holding strictly. The first line follows because $\sigma(\boldsymbol{e},\mathbf{1}) \geq s$, $\widetilde{u}(\boldsymbol{e},s)$ is non-decreasing in s, $(\boldsymbol{e},s),(\boldsymbol{e}',s') \in I_{\widetilde{u}}(c)$, and $\widetilde{u}(\boldsymbol{e}',s') \equiv \mathbb{E}_{\mathbf{t}}[u(\boldsymbol{e}',\boldsymbol{t})|\sigma(\boldsymbol{e}',\boldsymbol{t}) = s'] \geq \min_{\boldsymbol{t}:\sigma(\boldsymbol{e}',\boldsymbol{t})=s'} u(\boldsymbol{e}',\boldsymbol{t})$. The second line follows because $\boldsymbol{e}' \geq \boldsymbol{e}''$, $\operatorname*{arg\,min}_{\boldsymbol{t}:\sigma(\boldsymbol{e}',\boldsymbol{t})=s'} u(\boldsymbol{e}',\boldsymbol{t}) \geq \mathbf{0}$, and u is non-decreasing. The third line follows because, given that σ is increasing, the only value of $\boldsymbol{t}$ that makes $\sigma(\boldsymbol{e}'',\boldsymbol{t}) = \sigma(\boldsymbol{e}'',\mathbf{0})$ is $\boldsymbol{t} = \mathbf{0}$; also, $\sigma(\boldsymbol{e}'',\mathbf{0}) = \sigma(\boldsymbol{e},\mathbf{1})$.

Case 3c: if $s' > \sigma(\boldsymbol{e},\mathbf{1})$ and $\sigma(\boldsymbol{e},\mathbf{1}) > \sigma(\boldsymbol{e}',\mathbf{1})$, then we arrive at a contradiction since $s' > \sigma(\boldsymbol{e}',\mathbf{1})$ is not possible. Thus, Case 3c is impossible.

We have thus shown that for any $(\boldsymbol{e},s),(\boldsymbol{e}',s') \in I_{\widetilde{u}}(c) \setminus E^*$ with $s \neq s'$, if $\boldsymbol{e}' \ngeq \boldsymbol{e}$, then $s' < s$. This is equivalent to its contrapositive: for any $(\boldsymbol{e},s),(\boldsymbol{e}',s') \in I_{\widetilde{u}}(c) \setminus E^*$, if $s' > s$, then $\boldsymbol{e}' \geq \boldsymbol{e}$. Therefore, by conditions (i) and (ii) of Proposition 6, there exists $s^* \in [0,1]$ such that for any $(\boldsymbol{e},s) \in I_{\widetilde{u}}(c) \setminus E^*$, $\Pi(\boldsymbol{e},s) = \mathbf{I}(s \geq s^*)$.

It remains to find the values for $(\boldsymbol{e},s) \notin I_{\widetilde{u}}(c) \cup E^*$. Take any $(\boldsymbol{e},s)$ in $I_{\widetilde{u}}(c) \cup E^*$.

Case 1: If $\widetilde{u}(\boldsymbol{e}^*,s) = c$ for some $\boldsymbol{e}^*$ such that $s \in [\sigma(\boldsymbol{e}^*,\mathbf{0}),\sigma(\boldsymbol{e}^*,\mathbf{1})]$, then

Case 1a: if $\widetilde{u}(\boldsymbol{e},s) < c$ and $s \geq s^*$, then $\widetilde{u}(\boldsymbol{e}^*,s) = c > \widetilde{u}(\boldsymbol{e},s)$, so because σ is pro-e biased, $\boldsymbol{e} \geq \boldsymbol{e}^*$, and thus IC condition (ii) of Proposition 6 requires that $\Pi(\boldsymbol{e},s) \geq \Pi(\boldsymbol{e}^*,s) = 1$, which implies $\Pi(\boldsymbol{e},s) = 1$.

Case 1b: If $\widetilde{u}(\boldsymbol{e}',s) > c$ and $s < s^*$, then $\widetilde{u}(\boldsymbol{e},s) > c = \widetilde{u}(\boldsymbol{e}^*,s)$, so because σ is pro-e biased, $\boldsymbol{e}^* \geq \boldsymbol{e}$, and thus IC condition (ii) of Proposition 6 requires that $\Pi(\boldsymbol{e},s) \leq$

$\Pi(\boldsymbol{e}^*,s) = 0$, which implies $\Pi(\boldsymbol{e},s) = 0$.

Case 1c: If $\tilde{u}(\boldsymbol{e},s) < c$ and $s < s^*$, then set $\Pi(\boldsymbol{e},s) = 0$, which is what the principal would ideally want to do with $(\boldsymbol{e},s)$ if he was not constrained by IC.

Case 1d: If $\tilde{u}(\boldsymbol{e},s) > c$ and $s \geq s^*$, then set $\Pi(\boldsymbol{e},s) = 1$, which is what the principal would ideally want to do with $(\boldsymbol{e},s)$ if he was not constrained by IC.

Case 2: If $\tilde{u}(\mathbf{e}',s) < c$ for every $\mathbf{e}'$ such that $s \in [\sigma(\mathbf{e}',\mathbf{0}),\sigma(\mathbf{e}',\mathbf{1})]$, then it is easy to see that $s < s^*$. Set $\Pi(\boldsymbol{e},s) = 0$, which is what the principal would ideally want to do with $(\boldsymbol{e},s)$ if he was not constrained by IC.

Case 3: If $\tilde{u}(\mathbf{e}',s) > c$ for every $\mathbf{e}'$ such that $s \in [\sigma(\mathbf{e}',\mathbf{0}),\sigma(\mathbf{e}',\mathbf{1})]$, then it is easy to see that $s \geq s^*$. Set $\Pi(\boldsymbol{e},s) = 1$, which is what the principal would ideally want to do with $(\boldsymbol{e},s)$ if he was not constrained by IC.

Putting all the above cases together, we get that for $(\boldsymbol{e},s) \notin I_{\tilde{u}}(c) \cup E^*$, $\Pi(\boldsymbol{e},s) = \mathbf{I}(s \geq s^*)$. Combining this with the fact that for any $(\boldsymbol{e},s) \in I_{\tilde{u}}(c) \setminus E^*$, $\Pi(\boldsymbol{e},s) = \mathbf{I}(s \geq s^*)$ and given the definition of E^*, we get that for any $(\boldsymbol{e},s)$ such that $s \in [\sigma(\boldsymbol{e},\mathbf{0}),\sigma(\boldsymbol{e},\mathbf{1})]$, $\Pi(\boldsymbol{e},s) = \mathbf{I}(s \geq s^* \text{ or } \boldsymbol{e} \in E^*)$. To conclude the proof, notice that Π satisfies conditions (i) and (ii) of Proposition 6, and is thus IC. Therefore, by solving a relaxed problem when ignoring the IC constraints in cases 1c, 1d, 2, and 3, we have also solved the original problem. **Q.E.D.**

Proof of Proposition 9 Trivial, and, thus, omitted.

Proof of Proposition 10 Denote the total probability with which type (e,t) is accepted if she reports $(\hat{e},\hat{t})$ (with $\hat{e} \leq e$) by

$$\widetilde{P}(\hat{e},\hat{t};e,t) := (1 - T(\hat{e},\hat{t}))P(\hat{e},\hat{t},\emptyset) + T(\hat{e},\hat{t})\mathbf{I}\left(\sigma(e,t) \geq \sigma(\hat{e},\hat{t})\right).$$

Also, define condition (iii') (a strengthening of condition (iii)) to say that $(1 - T(e,t))P(e,t,\emptyset) \leq \Pi(e',0)$ for every e,t,e'.

Step 1: I first show that condition (i) is necessary for IC by showing the contrapositive. Assume that for some e,t_1,t_2 with $t_2 > t_1$, $\Pi(e,t_2) < \Pi(e,t_1)$. Then, IC of type (e,t_2) is violated, since $\widetilde{P}(e,t_1;e,t_2) = \Pi(e,t_1) > \Pi(e,t_2)$.

Step 2: I now show that condition (iii') is necessary for IC by showing the contrapositive. Assume that for some e,e',t, $(1 - T(e,t))P(e,t,\emptyset) > \Pi(e',0)$. Then, IC of type $(e',0)$ is violated, since $\widetilde{P}(e,t;e',0) \geq (1 - T(e,t))P(e,t,\emptyset) > \Pi(e',0)$.

Step 3: I now show that provided that (i) and (iii') are satisfied, $\Pi(r,\tau(r,\sigma(e,t)))$ being constant in r over $r \in [\underline{e}(\sigma(e,t)),e]$ for every (e,t) is necessary and sufficient for IC.

IC of type (e,t) is satisfied if and only if

$$\max_{(\hat{e},\hat{t}) \leq (1,1)} \left[(1 - T(\hat{e},\hat{t}))P(\hat{e},\hat{t};\emptyset) + T(\hat{e},\hat{t})\mathbf{I}\left(\sigma(e,t) \geq \sigma(\hat{e},\hat{t})\right)\right] = \Pi(e,t). \tag{9}$$

Assume that conditions (i) and (iii') are satisfied. Then, $\Pi(e,t) \geq \Pi(e,0) \geq (1 - T(\widehat{e},\widehat{t}))P(\widehat{e},\widehat{t},\emptyset)$ for any $(\widehat{e},\widehat{t})$. Therefore, (9) is equivalent to

$$\max_{(\widehat{e},\widehat{t})\in\{(x,y)\in[0,1]^2:\sigma(e,t)\geq\sigma(x,y)\}} \left[(1 - T(\widehat{e},\widehat{t}))P(\widehat{e},\widehat{t};\emptyset) + T(\widehat{e},\widehat{t})\right] = \Pi(e,t). \tag{10}$$

Given that $\Pi(e,t)$ is non-decreasing in t (condition (i)), (10) can equivalently be written as

$$\max_{r\in[\underline{e}(\sigma(e,t)),1]} \{[1 - T(r,\tau(r,\sigma(e,t)))]P(r,\tau(r,\sigma(e,t)),\emptyset) + T(r,\tau(r,\sigma(e,t)))\} = \Pi(e,t)$$

or equivalently,

$$e \in \operatorname*{arg\,max}_{r\in[\underline{e}(\sigma(e,t)),\overline{e}(\sigma(e,t))]} \Pi(r,\tau(r,\sigma(e,t))). \tag{11}$$

Thus, IC is satisfied for every type if and only if for every (e,t), (11) is satisfied. This is true if and only if $\Pi(r,\tau(r,\sigma(e,t)))$ is constant in r for $r \in [\underline{e}(\sigma(e,t)),\overline{e}(\sigma(e,t))]$ for every (e,t).

That the latter is sufficient for (11) to hold for every (e,t) is immediate. I show necessity by showing the contrapositive. Assume that for some (e,t), $\Pi(r,\tau(r,\sigma(e,t)))$ is *not* constant in r for $r \in [\underline{e}(\sigma(e,t)),1]$. That is, for some (e,t) there exist r_1, r_2 with $\underline{e}(\sigma(e,t)) \leq r_1 < r_2 \leq \overline{e}(\sigma(e,t))$ such that $\Pi(r_2,\tau(r_2,\sigma(e,t))) \neq \Pi(r_1,\tau(r_1,\sigma(e,t)))$. If $\Pi(r_2,\tau(r_2,\sigma(e,t))) < \Pi(r_1,\tau(r_1,\sigma(e,t)))$, IC of type $(r_2,\tau(r_2,\sigma(e,t)))$ is violated, as she prefers to imitate type $(r_1,\tau(r_1,\sigma(e,t)))$. If, instead, $\Pi(r_2,\tau(r_2,\sigma(e,t))) > \Pi(r_1,\tau(r_1,\sigma(e,t)))$, IC of type $(r_1,\tau(r_1,\sigma(e,t)))$ is violated, as she prefers to imitate type $(r_2,\tau(r_2,\sigma(e,t)))$.

Step 4: It is easy to see that $\Pi(r,\tau(r,\sigma(e,t)))$ being constant in r over $r \in [\underline{e}(\sigma(e,t)),\overline{e}(\sigma(e,t))]$ for every (e,t) is equivalent to condition (ii).

Step 5: Finally, notice that provided that conditions (i) and (ii) hold, conditions (iii) and (iii') are equivalent. **Q.E.D.**

Proof of Lemma 9 Take any IC mechanism $M \equiv \langle T,P\rangle$. Condition (iii) of Proposition 1 says that $\Pi(0,0) \geq (1 - T(e,t))P(e,t,\emptyset)$ for any (e,t). Then, construct the mechanism $M' \coloneqq \langle T',P'\rangle$ with[57]

$$\begin{aligned} T'(e,t) &\coloneqq \Pi(e,t) - \Pi(0,0) = (1 - T(e,t))P(e,t,\emptyset) + T(e,t) - \Pi(0,0) \\ &\leq \Pi(0,0) + T(e,t) - \Pi(0,0) = T(e,t), \quad \text{and} \\ P'(e,t,\emptyset) &\coloneqq \frac{\Pi(0,0)}{1 - \Pi(e,t) + \Pi(0,0)} \geq \frac{(1 - T(e,t))P(e,t,\emptyset)}{1 - \Pi(e,t) + (1 - T(e,t))P(e,t,\emptyset)} = P(e,t,\emptyset) \end{aligned}$$

for every (e,t), where the inequalities follow from $\Pi(0,0) \geq (1 - T(e,t))P(e,t,\emptyset)$.

By construction we have that $\Pi'(e,t) = \Pi(e,t)$ for every (e,t), so M' satisfies conditions

[57] If $\Pi(0,0) = 0$, then for (e,t) such that $\Pi(e,t) = 1$, set $P'(e,t,\emptyset) = 0$.

(i) and (ii) of Proposition 10. By construction, we also have that for every (e,t)

$$\Pi'(0,0) = \Pi(0,0) = (1 - T'(e,t))P'(e,t,\emptyset),$$

so M' also satisfies condition (iii) of Proposition 10. Therefore, M' is IC.

Last, to see why the second part is true, notice that for $c > 0$, M' saves on testing costs compared to M if there exists (a positive measure of) (e,t) with $P(e,t,\emptyset)(1 - T(e,t)) < \Pi(0,0)$, since $T'(e,t) < T(e,t)$ for such (e,t). **Q.E.D.**