Multidimensional screening of strategic candidates*

Click to access latest version.

Orestis Vravosinos†

June 16, 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, the optimal mechanism never asks for evidence an agent whose composite measure it verifies. In the optimal mechanism, errors favoring high- over low-training agents arise because (i) verification creates incentives for the agent to withhold evidence of training and (ii) the principal saves on verification costs by accepting an agent with high-training without verifying the composite measure. The two forces are complements in inducing errors favoring high- over low-training agents.

Keywords: multidimensional screening, persuasion game, evidence game, costly verification, verifiable disclosure, signal-jamming, costly lying, signal manipulation
JEL classification codes: C72, D82, D83, D86, I23, I24, J41, M12, M51

*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, Christoph Carnehl, Curtis Taylor, Dana Foarta, David Pearce, Dhruva Bhaskar, 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, Susanne Goldlücke, Yeon-Koo Che, and audiences at the 35th Stony Brook International Conference on Game Theory, Mississippi State University, New York University, University of Chicago, and University of Georgia.

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

"My plan was to leave one copy [of textbooks] at home and one at school. This was less about the inconvenience of carrying books back and forth than it was about appearing as if I didn't need to study at home. [...] I went home each day conspicuously empty-handed. At night, holed up in my bedroom with my duplicate textbooks, I solved and re-solved every quadratic equation, I memorized Latin declensions and reviewed names, dates, history of all those Greek wars and battles and gods and goddesses. The next day, I'd arrive at school fortified with all that I had learned but no indication that I had studied."

—Bill Gates, *Source Code: My Beginnings* (2025)

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 through some costly procedure. 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.[1] Specifically, evidence of training may lead the evaluator to attribute the composite measure to training, thereby negatively affecting his assessment

[1]Absent a strong negative dependence between training and talent, evidence of training is indeed good news about the candidate's suitability. But even if there is a strong negative relationship between the two, absent verification, any incentive-compatible evaluation scheme that utilizes evidence should reward the candidate for presenting it. Therefore, presenting evidence is always weakly in the candidate's best interest.

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 incentives to present 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 about her parents' education or professions, or even (iv) hiding her race.[2] A job candidate 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

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

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.[3] 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).[4] The agent can present hard evidence to prove any part of her training. However, she 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 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 when verification is 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?

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

[3]The results are generalized to types of any finite dimension in the Online Appendix.

[4]Although training is 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 training and then extend the model to endogenize it. Section 5.3 shows that the structure of the optimal mechanism remains qualitatively the same even if training is endogenous (i.e., produced by the agent before her interaction with the principal), as long as the principal cannot influence training by committing to a mechanism before the agent receives training.

certain agent. Rather, it asks for evidence of training only to accept a high-training agent *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 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).[5] The principal still optimally favors high-training agents even though training is worthless to him. He does so because of two forces: (i) to save on verification costs by accepting high-training agents without verifying their composite measure and (ii) due to the strategic incentives of agents to withhold evidence of training when the principal verifies the value of a composite measure that is under-sensitive to talent.

There is an important interaction between these two forces. Each of the two forces *individually* causes the principal to optimally favor high- over low-training agents. Namely, when the composite measure is overly sensitive to talent—in which case the second force is absent—the cost of verification induces the principal to accept some unworthy high-training agents, including unworthy ones, without verification. Similarly, when the composite measure is under-sensitive to talent, the optimal mechanism makes mistakes in favor of high-training and against low-training agents even when verification is free. When the two forces are *combined* (i.e., the composite measure is under-sensitive to talent and verification is costly), the second force reduces the effectiveness of verification. This causes the principal to accept even more agents without verification to save on verification costs, thereby exacerbating the errors the principal makes by accepting agents without verification. The two forces are complements in inducing errors in favor of high- and against low-training agents.

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 too 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 as

[5]In that case, the composite measure is automatically less sensitive to talent than talent is valuable to the principal.

her efficiency or managerial skills. The employer can verify the employee's productivity in the current position. Then, the payoff to the principal from promoting the employee is the difference between her productivity in the new position and her productivity in the current position. It is natural to assume that efficiency and managerial skills are more important in the higher than in the current position. Then, the composite measure (i.e., current productivity) is less sensitive to talent than talent is valuable to the employer. The optimal promotion scheme then promotes some hard-working employees—either with or without monitoring their productivity—although their promotion destroys firm value. At the same time, some talented but less hard-working employees are not promoted, although their promotion would benefit the firm.

Consider, now, hiring by a prestigious employer. Evidence is the candidate's CV quality, 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 go through less prestigious employers to prove their worth before landing 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, the prestigious employer makes errors in the poaching stage. This means that 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., screening for talent, controlling 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 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). The solution to the multiproduct monopolist's problem is famously elusive and complex. The optimal mechanism may involve lotteries to allocate goods (Manelli and Vincent, 2006), possibly uncountably infinitely many of them (Daskalakis et al., 2017). Even in the case of two goods with additive and independent values, the optimal mechanism is unknown except for some special cases (Manelli and Vincent, 2006).

I advance this literature by proposing a novel and insightful multidimensional screening problem with a remarkably simple solution. The optimal mechanism is completely described by (i) the level of training that the agent needs to prove through evidence to be accepted without verification and (ii) the composite measure that is required for acceptance in case the agent cannot present the required evidence of training to be accepted without verification.[6] 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.[7] 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

[6] In the more general case of n dimensions of talent and m dimensions of training (see the Online Appendix), the optimal mechanism is again remarkably simple: It consists of (i) an m-dimensional upper set that describes what evidence of training the agent needs to present to be accepted without verification and (ii) the composite measure that is required for acceptance without verification.

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

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 multidimensional type (see, e.g., Glazer and Rubinstein, 2004; Carroll and Egorov, 2019; Li, 2021).[8] 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 verification.[9] 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 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 $f:[0,1]^2 \to \mathbb{R}_{++}$. No other assumption is imposed on f; any form of stochastic dependence between e and t is allowed. 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

[8] Assuming that the principal can choose a single dimension of the agent's type to verify (at no cost), Carroll and Egorov (2019) derive conditions under which the principal can fully learn the agent's type (i) by allowing the agent to have a say over which dimension is verified, which reveals information about the agent's type and (ii) by allowing for randomization over which dimension is verified. In my setting, there is a unique composite measure that the principal can verify at a cost, and there are no gains from randomization.

[9] Additional differences from career concerns models are the following: The principal has commitment power, and verification is costly.

prove anything about.[10]

Verification. By paying a cost $c \geq 0$, the principal can verify the value of a composite measure of the agent's evidence and talent. $\sigma(e,t) \in [0,1]$ is the *composite measure* of the agent's type (e,t). $\sigma : [0,1]^2 \to [0,1]$ is increasing and continuous in e and t. $I_\sigma(s) \coloneqq \{(e,t) \in [0,1]^2 : \sigma(e,t) = s\}$ denotes an iso-composite-measure curve.

Payoffs. Ultimately, the principal must decide whether to accept or reject the agent. He receives (gross of verification 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. $I_u(\overline{u}) \coloneqq \{(e,t) \in [0,1]^2 : u(e,t) = \overline{u}\}$ denotes an indifference curve of the principal.[11] The agent's Bernoulli payoff is equal to 1 if accepted and 0 if rejected.

Parametric examples. In a linear specification, $u(e,t) \coloneqq \gamma_u e + (1-\gamma_u)t - \underline{q}$, 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. Similarly, $\sigma(e,t) \coloneqq \gamma_s e + (1-\gamma_s)t$, where $\gamma_s \in (0,1)$ measures how sensitive the composite measure 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 verify the composite measure if the agent 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 composite measure is $s \in [0,1]$.[12] If the composite measure is not verified, $s = \emptyset$ 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 $P : [0,1]^2 \times ([0,1] \cup \{\emptyset\}) \to [0,1]$

[10] It is straightforward to see that the model also captures the case where evidence 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 $\widetilde{e} \coloneqq \varepsilon(e,t)$ and t measures the part of talent not captured by evidence $\widetilde{e}$.

[11] $I_u(\overline{u})$ 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.

[12] 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) composite measure that agent type (e,t) can achieve, and the agent can intentionally manipulate her composite measure downwards.

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

$$\int_0^1 \int_0^1 \left\{ \left[\begin{array}{c} \overbrace{T(\phi(e,t))P(\phi(e,t),\sigma(e,t))}^{\substack{\text{probability that } (e,t) \text{ is} \\ \text{accepted after verification}}} \\ + \underbrace{[1-T(\phi(e,t))]P(\phi(e,t),\emptyset)}_{\substack{\text{probability that } (e,t) \text{ is} \\ \text{accepted without verification}}} \end{array} \right] u(e,t) - c \underbrace{T(\phi(e,t))}_{\substack{\text{probability of verification of} \\ (e,t)\text{'s composite measure}}} \right\} f(e,t) dt de$$

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

$$\phi(e,t) \in \underset{(e',t') \leq (e,1)}{\arg\max} \underbrace{\{T(e',t')P(e',t',\sigma(e,t)) + (1-T(e',t'))P(e',t',\emptyset)\}}_{\text{total probability that } (e,t) \text{ is accepted if she reports } (e',t')}.$$

3 Optimal multidimensional screening with substitutable attributes and costly verification

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

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 $(e,t) \mapsto \{(e',t') \in [0,1]^2 : e' \leq e\}$, 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.[13] Therefore, we define IC mechanisms as follows.

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

$$(e,t) \in \underset{(e',t') \leq (e,1)}{\arg\max} \{T(e',t')P(e',t',\sigma(e,t)) + (1-T(e',t'))P(e',t',\emptyset)\}.$$

[13]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 verification, $g_2(e,t)$ the probability of acceptance conditional on no verification, and $g_3(e,t,\cdot)$ a self-map on $[0,1]$ that (conditional on verification) maps the composite measure s to the probability $g_3(e,t,s)$ of acceptance.

Mechanisms that are deterministic after verification are without loss. Next, we can constrain attention to mechanisms with threshold acceptance policies after verification; that is, mechanisms such that

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

for any (e,t) and some $P_{at} : [0,1]^2 \to [0,1]$, where *at* is a mnemonic for the probability of acceptance *after verification* (provided that the threshold composite measure $\sigma(e,t)$ is met). If type (e,t) reports her type truthfully, then if the composite measure is verified, she is accepted with probability $P_{at}(e,t)$. Notice that the threshold is set exactly equal to the composite measure that a truthfully-reporting agent will achieve. To see why constraining attention to such mechanisms is without loss of optimality, observe that among all mechanisms that (conditional on verification) 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).[14]

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, 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 verification is costly, in any optimal mechanism, an agent whose composite measure is verified is accepted with probability 1 if she passes the appropriate threshold. When verification is free, it is still without loss to constrain attention to mechanisms that accept the agent with probability 1 if she passes the appropriate composite measure threshold.

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 $T(e,t) > 0$.[15]

[14] 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) in case of verification 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 with the same verification policy T and thus equal verification costs.

[15] 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.

The intuition behind this result is as follows. The only reason to accept an agent after verification—rather than accept her without verification—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 acceptance without verification and (ii) the probability $T(e,t)P_{at}(e,t)$ of acceptance after verification (provided that the composite measure passes the threshold.). But then, simply put, if the principal pays for verification, he may as well set $P_{at}(e,t) = 1$ to assign as large a part as possible of the total probability of acceptance to the case of acceptance after verification.

In more detail, if agent (e,t) is not accepted with certainty after meeting the threshold (i.e., $P_{at}(e,t) < 1$ and $T(e,t) > 0$), we can (i) increase the probability $P_{at}(e,t)$ of acceptance in case of verification, (ii) decrease the probability $T(e,t)$ of verification, and (iii) decrease (if positive) the probability $P(e,t,\emptyset)$ of acceptance in case of no verification, keeping fixed both (a) the probability $(1 - T(e,t))P(e,t,\emptyset)$ of acceptance without verification and (b) the probability $T(e,t)P_{at}(e,t)$ of acceptance after verification. By doing so, we (i) keep fixed the total probability $\Pi(e,t)$ of accepting (e,t), (ii) reduce the probability $T(e,t)$ of verification of (e,t)'s composite measure, thereby limiting verification costs, and (iii) do not change the incentives of other agents 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 has a lower composite measure than (e,t)) or $\Pi(e,t)$ (if her composite measure is at least as high as (e,t)'s), and neither $(1 - T(e,t))P(e,t,\emptyset)$ nor $\Pi(e,t)$ has changed.

3.2 Incentive-compatible mechanisms

Given what we have seen, we constrain attention to truthful mechanisms that are deterministic after verification. 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 $\overline{e}(s) \coloneqq \max\{e \in [0,1] : \sigma(e,0) \leq s\}$.[16] 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 $(e,t) \in [0,1]^2$,

[16] $\underline{e}(s)$ (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 $t = 0$.

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.

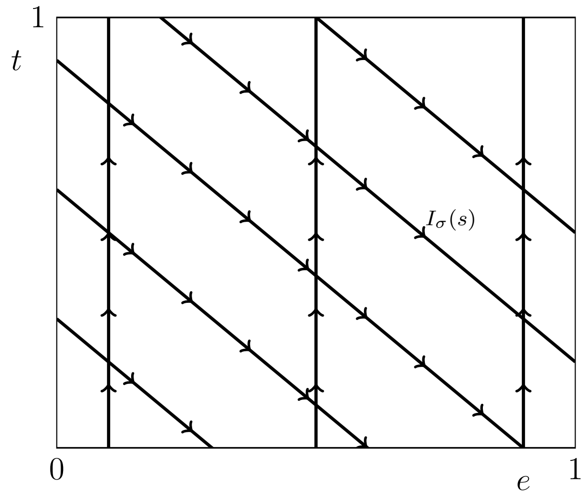
Condition (i) is necessary and sufficient to ensure that an agent (e,t) does not want to reveal her evidence but under-report her talent to imitate agent (e,t') with $t' < t$, meet (e,t')'s composite measure threshold in case of verification (since the composite measure is increasing in talent), and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

Condition (iii) is necessary and sufficient to ensure that an untalented agent $(e,0)$ does not want to over-report her talent, imitating an agent (e,t)—whose composite measure is higher—and possibly getting accepted without verification. 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 verify the talented agents' composite measure 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 (e,t) wants to present all her evidence but overstate her talent to be $t' > t$ and get accepted with probability $(1 - T(e,t'))P(e,t',\emptyset)$ instead of $\Pi(e,t)$.

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 present, if it is not profitable for (e,t) to overstate her talent so much that she will fail to meet the composite measure threshold in case of verification, then because of condition (i), she will want to overstate her talent as much as possible, making sure that she will be able to meet the composite measure threshold), 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 misreport her type 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 kind of misreport by the agent: withholding 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 every e,e',t,t' with $e' < e$, 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 cannot meet the composite measure threshold in case of verification.

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.

Condition (iii) of Proposition 1 is satisfied with equality. Lemma 2 shows that when verification 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 verification 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 verification 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 verification costs than M.

Here is the intuition behind this result. Take any IC mechanism $M \equiv \langle T,P \rangle$. When $\Pi(e,0) > (1-T(e,t))P(e,t,\emptyset)$ for some $t > 0$ and e, it means that untalented agent $(e,0)$ strictly prefers to not overstate her talent to be t. This strict preference is due to over-verification of the talented agent (e,t)'s composite measure. We can decrease $T(e,t)$ and increase $P(e,t,\emptyset)$ 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.[17] Conditions (i) and (ii) of Proposition 1 are also still satisfied since Π has not changed. Then, (e,t)'s composite measure is verified with lower but still high enough

[17]Notice that because M is IC, condition (i) of Proposition 1 implies that $\Pi(e,t) \equiv (1-T(e,t))P(e,t,\emptyset) + T(e,t) \geq \Pi(e,0)$, which combined with $\Pi(e,0) > (1-T(e,t))P(e,t,\emptyset)$ implies that $T(e,t) > 0$ to start with, so we can decrease $T(e,t)$. Also, if $P(e,t,\emptyset) = 1$ to start with, then we keep $P(e,t,\emptyset)$ fixed as we decrease $T(e,t)$. Notice also that by decreasing $T(e,t)$ and increasing (or keeping fixed, if equal to 1) $P(e,t,\emptyset)$ while keeping $\Pi(e,t)$ fixed, we increase $(1-T(e,t))P(e,t,\emptyset)$. This is feasible 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.

probability to prevent $(e,0)$ from imitating (e,t).

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). In an IC mechanism without excessive verification, the total probability of acceptance has two components: (i) a base probability $\Pi(e,0)$ of accepting the agent for her evidence without verification and (ii) an additional probability $T(e,t)$ of accepting the agent for her talent, which through verification allows her to differentiate herself from less talented agents with the same level of evidence.

3.3 Optimal screening under free verification

We are now ready to characterize the optimal mechanisms under free verification (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 e and the composite measure s. 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.[18] 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 a standard result that an extreme Π maps each (e,s) to either 0 or 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 Composite measure biased in favor of talent

We are now ready to derive the optimal mechanism. Consider first the case where the composite measure is biased in favor of talent in the sense that it is more sensitive to talent than talent is valuable to the principal. In the linear and Cobb-Douglas specifications (see section 2), σ is pro-t biased if and only if $\gamma_\sigma < \gamma_u$, in which case, the composite measure assigns lower weight to e and higher weight to t compared to the principal's preferences. A pro-t biased σ can be defined more generally as follows.[19]

[18] Condition (iii) of Proposition 1 is immaterial, since verification 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 verification) automatically satisfies condition (iii) of Proposition 1.

[19] We define a pro-t biased composite measure for any verification cost c. The optimal mechanism under costly verification is studied in section 3.4.

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-composite-measure 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 composite measure 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, the principal's payoff from accepting the agent is higher (resp. lower) than the verification cost for agents with high (resp. low) evidence.

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

Allowing for the principal to also value talent, Proposition 2 shows that when verification is (i) free and (ii) the composite measure 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.

Assume the principal tries to achieve the first-best without worrying about the possibility that agents may withhold evidence to manipulate the interpretation of the composite measure. Proposition 2 says that if the principal values talent but less strongly than the composite measure depends on talent, agents do not have incentives to withhold evidence, and so the principal was correct to assume that all agents will present their evidence. To achieve the first-best, the principal needs to both ask for evidence and verify the composite measure: Because there exist e,t,t' with $t' > t$ such that $u(e,t') > 0 > u(e,t)$, he needs to use verification to accept (e,t') but not (e,t). Figure 3(a) presents the optimal mechanism when verification is free and the composite measure is pro-t biased.[20]

[20]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

3.3.2 Composite measure biased in favor of evidence

Consider now the case where the composite measure is biased in favor of evidence in the sense that the composite measure is more sensitive to evidence than evidence is valuable to the principal—or equivalently, the composite measure is less sensitive to talent than talent is valuable to the principal. In the linear and Cobb-Douglas specifications (see section 2), σ is pro-t biased if and only if $\gamma_\sigma > \gamma_u$. A pro-e biased σ can be defined more generally as follows.[21]

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-composite-measure curves cross the principal's indifference curve $I_u(c)$ "from above" (see Figure 3(b)). Claim 2 is analogous to Claim 1.

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

But what *can* actually be achieved when the composite measure is less sensitive to talent than talent is valuable to the principal? Proposition 3 describes the optimal mechanism when verification is free and the composite measure is pro-e biased. In the optimal mechanism, agent (e,t) is accepted if and only if $\sigma(e,t) \geq s^*$.[23]

---

principal truthfully report their type and are rejected without verification. Agents who are valuable but cannot prove it by presenting evidence e such that $u(e,0) > 0$ (which would prove that even if they have $t = 0$, they are valuable) are accepted after verification. Finally, agents who can prove they are valuable by presenting evidence e such that $u(e,0) \geq 0$ do so and are accepted without verification. Clearly, since verification 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 always verifying and accepting only the valuable agents.

[21] We define pro-e biased composite measures for any verification cost c. The optimal mechanism under costly verification is studied in section 3.4.

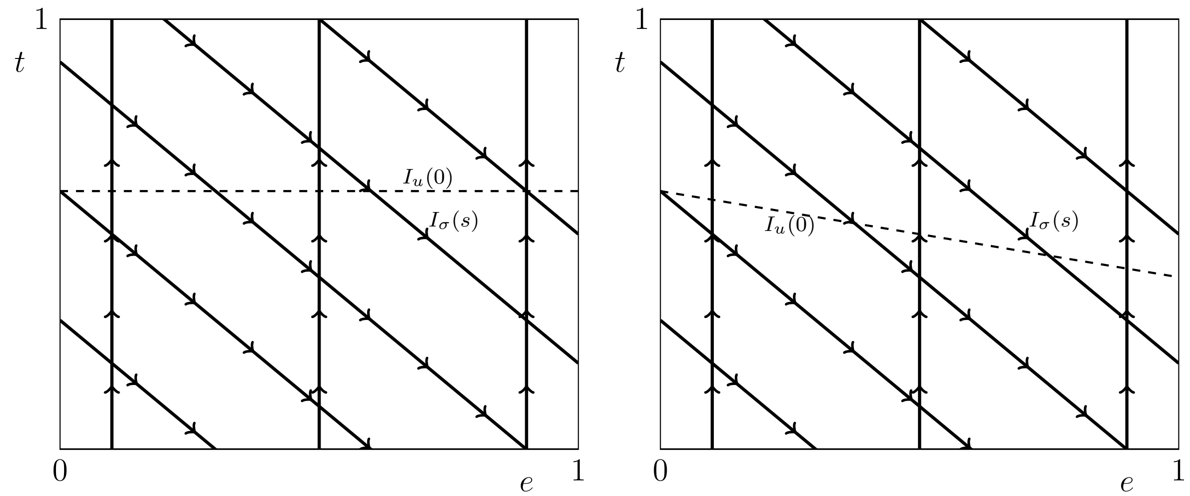
[22] σ being pro-e biased is not necessary for this conclusion. The conclusion still applies as long as the condition in definition 4 is satisfied for a positive measure of $s \in [0,1]$.

[23] 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 verification. 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 accepted after verification. 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 verification. Clearly, since verification 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 always verifying the composite measure and accepting only the agents that pass the composite measure threshold s^*.

Figure 2: *Not* achieving the first-best: composite measure 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 $I_u(0)$.

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

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 $s^* \in \arg\max_{s_{min} \in [0,1]} \int_{s_{min}}^{1} \int_{\underline{e}(s)}^{\overline{e}(s)} \widetilde{u}(e,s)\widetilde{f}(e,s)deds$, where $\widetilde{u}(e,s) := u(e,\tau(e,s))$ and $\widetilde{f}(e,s) := f(e,\tau(e,s))$.[24] 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, notice that IC requires that $\Pi(e,t)$ be non-decreasing along iso-composite-measure 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-composite-measure curves maximizes the principal's payoff. That is because, on the part of an iso-composite-measure

[24]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.

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 a pro-t biased composite measure (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 a composite measure (e.g., interview or test performance) that is 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.

3.4 Optimal screening under costly verification

When verification is costly, the principal needs to compare the benefit of verification to its cost. The benefit of verification is 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:

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

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 e—that solves the principal's problem.

Lemma 4. There exists an optimal deterministic mechanism.

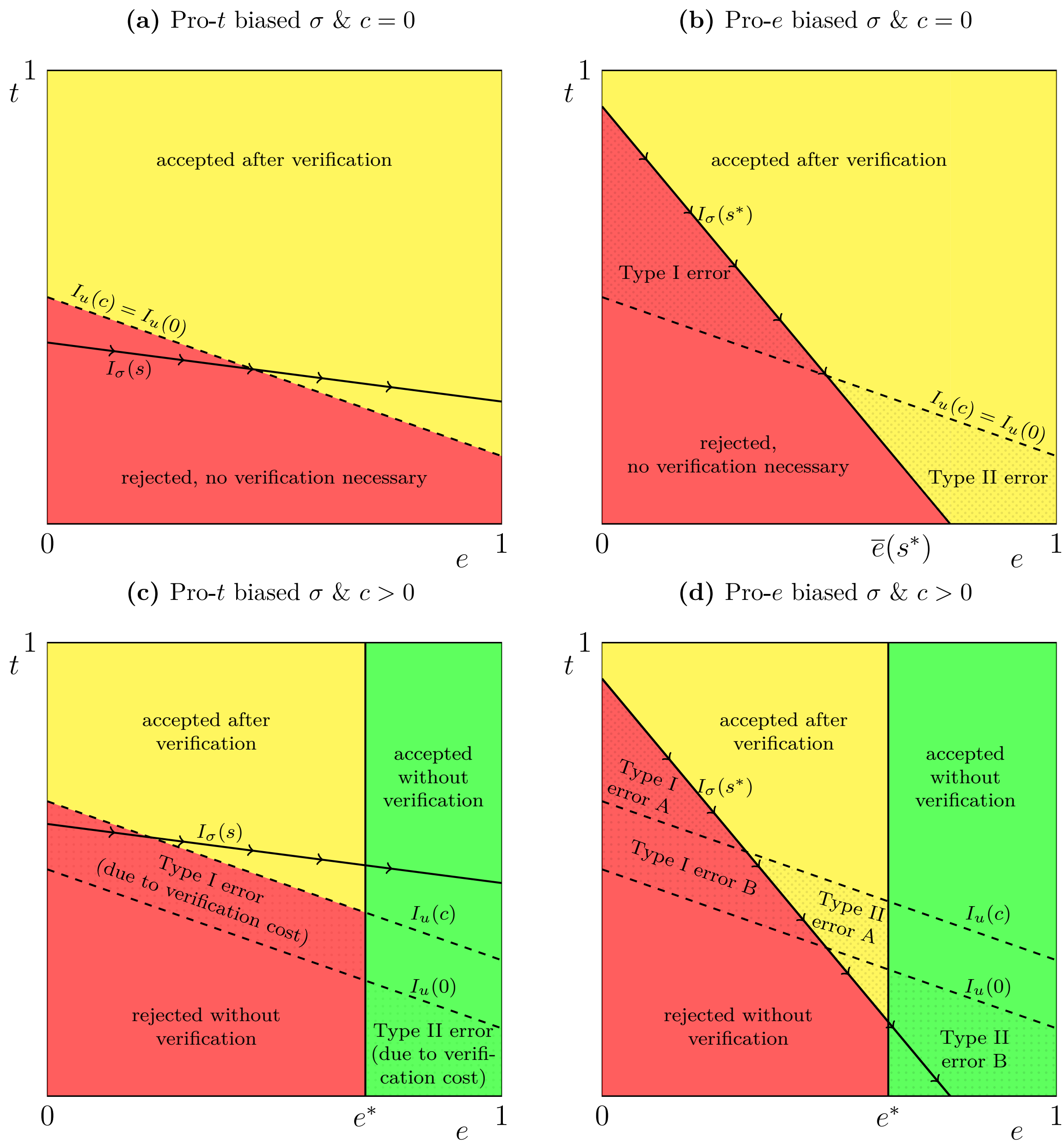
3.4.1 Composite measure biased in favor of talent

Proposition 4 characterizes the optimal mechanism under a pro-t biased composite measure, generalizing Proposition 2 by allowing for possibly costly verification (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 $e^* \in$

Figure 3: Optimal bidimensional screening with substitutable attributes and possibly costly verification



Note: the dashed line $I_u(c)$ represents the principal's indifference curve at utility level c: the principal is indifferent between (i) accepting after verification and (ii) rejecting without verification agents on that curve. The dashed line $I_u(0)$ represents the principal's indifference curve at utility level 0. The arrowed line represents an iso-composite-measure curve, at an arbitrary level s in panels (a) and (c), and at the optimal level s^* in panel (b) and (d). The green area denotes the set of agents who are accepted without verification. The yellow area denotes the set of agents who are accepted after verification. The red area denotes the set of agents who are rejected without verification. Although s^* is used in both panels (b) and (d), s^* in panel (b) can be different from s^* in panel (d). In panel (d), $I_u(0)$ can intersect the vertical line at e^* above or below the point where $I_\sigma(s^*)$ intersects the vertical line at e^*.

$\arg\max_{e_{min}\in[0,1]} v^{\text{pro-}t}(e_{min})$,[25] where

$$v^{\text{pro-}t}(e_{min}) \coloneqq \underbrace{\int_0^1\int_0^{e_{min}} (u(e,t)-c)\mathbf{I}(u(e,t)\geq c)f(e,t)dedt}_{\substack{\text{payoff from agents accepted after verification}\\ \text{net of verification costs}}} + \underbrace{\int_0^1\int_{e_{min}}^1 u(e,t)f(e,t)dedt}_{\substack{\text{payoff from agents}\\ \text{accepted without verification}}}.$$

Every agent with evidence $e \geq e^*$ evidence is accepted without verification, while agents with evidence $e < e^*$ are accepted after verification if their value $u(e,t)$ to the principal is higher than the cost c of verification. The remaining agents are rejected without verification. Figure 3(c) presents the structure of the optimal mechanism.

Denote by $v_e^{\text{pro-}t}(e_{min})$ the derivative of $v^{\text{pro-}t}(e_{min})$ with respect to e_{min}. When $e^* \in (0,1)$, the first-order condition is

$$\begin{aligned} v_e^{\text{pro-}t}(e^*) &= \int_0^1 (u(e^*,t)-c)\mathbf{I}(u(e^*,t)\geq c)f(e^*,t)dt - \int_0^1 u(e^*,t)f(e^*,t)dt \\ &= -\int_0^1 \min\{u(e^*,t),c\}f(e^*,t)dt = 0, \end{aligned}$$

or equivalently

$$\begin{aligned} v_e^{\text{pro-}t}(e^*) = \overbrace{-\int_0^1 u(e^*,t)\mathbf{I}(u(e^*,t)\leq 0)f(e^*,t)dt}^{>0\text{: gain from decrease in Type II error (ii)}} - \overbrace{\int_0^1 u(e^*,t)\mathbf{I}(0<u(e^*,t)<c)f(e^*,t)dt}^{>0\text{: loss from increase in Type I error (iii)}} \\ - \underbrace{c\int_0^1 \mathbf{I}(u(e^*,t)\geq c)f(e^*,t)dt}_{>0\text{: loss from increase in verification costs (i)}} = 0. \end{aligned}$$

An increase in the threshold e^* would lead to: (i) increased verification costs by making additional agents who lie above $I_u(c)$ get accepted after verification (who were accepted without verification before the increase in e^*), (ii) a decrease in the Type II error, but also (iii) an increase in the Type I error. 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 verification costs (i.e., effect (i)) with 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 verification is used for a positive measure of agents.[26] First, an increase in c causes the (combined) magnitude of channels (i) and (iii) to increase without

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

[26] Namely, $u(e,t) > c$ 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 use verification.

affecting the magnitude of channel (ii).[27] Thus, e^* is decreasing in c; the more costly verification is, the more high-evidence agents are accepted without verification. Particularly, $v_e^{\text{pro-}t}(e)$ is decreasing in c with $\partial v_e^{\text{pro-}t}(e)/\partial c = -\int_0^1 \mathbf{I}(u(e^*,t) \geq c) f(e^*,t)dt < 0$, and by the Implicit Function Theorem $de^*/dc = -\partial v_e^{\text{pro-}t}(e)/\partial c|_{e=e^*}/v_{ee}^{\text{pro-}t}(e^*) < 0$. Second, the principal's optimal payoff is decreasing in c. Third, since the principal's objective function is independent of σ, the optimal mechanism and payoff are the same under any two pro-t biased composite measures with the same cost c.

3.4.2 Composite measure biased in favor of evidence

Proposition 5 characterizes the optimal mechanism under a pro-e biased composite measure, generalizing Proposition 3 by allowing for possibly costly verification (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 $(e^*,s^*) \in [0,1]^2$.

The principal's problem amounts to choosing threshold evidence and composite measure levels $(e^*,s^*) \in \arg\max_{(e_{min},s_{min})\in[0,1]^2} v^{\text{pro-}e}(e_{min},s_{min})$, where

$$v^{\text{pro-}e}(e_{min},s_{min}) := \overbrace{\int_{s_{min}}^1 \int_{\underline{e}(s)}^{\max\{\min\{\overline{e}(s),e_{min}\},\underline{e}(s)\}} (\widetilde{u}(e,s) - c)\, \widetilde{f}(e,s)deds}^{\text{payoff from agents accepted after verification net of verification costs}} \\ + \underbrace{\int_0^1 \int_{\max\{\underline{e}(s),e_{min}\}}^{\max\{\overline{e}(s),e_{min}\}} \widetilde{u}(e,s)\widetilde{f}(e,s)deds}_{\text{payoff from agents accepted without verification}},$$

and $\widetilde{u}(e,s) \equiv u(e,\tau(e,s))$ and $\widetilde{f}(e,s) \equiv f(e,\tau(e,s))$.[28] Every agent with evidence $e \geq e^*$ is accepted without verification, while agents with evidence $e < e^*$ are accepted after verification if their composite measure is at least $\sigma(e,t) \geq s^*$. The remaining agents are rejected without verification.

Figure 3(d) presents the structure of the optimal mechanism when the composite measure is pro-e biased. The principal makes four types of errors. Type I error A is due to the fact that the principal rejects without verification some agents whom he would prefer to accept after verification. Type I error B is due to the fact that the principal rejects without verification some agents whom he would prefer to accept without verification. Type II error A is due to the fact that the principal accepts after verification some agents

[27]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 $-c\int_0^1 \mathbf{I}(u(e^*,t) \geq c)f(e^*,t)dt$.

[28]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.

whom he would prefer to reject without verification. Last, Type II error B is due to the fact that the principal accepts without verification some agents whom he would prefer to reject without verification.

When $\underline{e}(s^*) < e^*$ (i.e., some agents are accepted after verification) and $e^*,s^* \in (0,1)$,[29] the first-order conditions are

$$v_e^{\text{pro-e}}(e^*,s^*) = \overbrace{\underbrace{-\int_0^{s^*} \widetilde{u}(e^*,s)\mathbf{I}(\widetilde{u}(e^*,s) \leq 0)\widetilde{f}(e^*,s)ds}_{>0\text{: gain from rejection of unworthy agents (ii)}} \underbrace{- \int_0^{s^*} \widetilde{u}(e^*,s)\mathbf{I}(\widetilde{u}(e^*,s) > 0)\widetilde{f}(e^*,s)ds}_{>0\text{: loss from rejection of worthy agents (iii)}}}^{>0\text{: net gain from decrease in Type II error (A+B)}}$$
$$- \underbrace{c\int_{s^*}^{1} \widetilde{f}(e^*,s)ds}_{\substack{>0\text{: loss from increase}\\ \text{in verification costs (i)}}} = 0$$

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

The principal chooses s^*, determining which agents are accepted after verification and which are rejected without verification. In choosing s^*, he considers the trade-off between Type I error A and Type II error A.[30]

An increase in the threshold e^* would lead to: (i) increased verification costs by making additional agents who lie above $I_\sigma(s^*)$ are accepted after verification (who were accepted without verification before the increase in e^*), (ii) the rejection without verification of additional agents who lie below $I_u(0)$ (who were accepted without verification before the increase in e^*), but also possibly (iii) the rejection without verification of additional agents who lie below $I_\sigma(s^*)$ but above $I_u(0)$ (who were accepted without verification before the increase in e^*).[31] 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 verification costs (i.e., effect (i)) with accuracy (i.e., the net effect of (ii) and (iii)).

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 verification is used for a positive measure of agents. Denote by $J(e^*,s^*)$ the Jacobian matrix of the first derivatives evaluated at (e^*,s^*),

[29]Notice that $e^* \leq \overline{e}(s^*)$ (for if $e^* > \overline{e}(s^*)$ and $c > 0$, reducing e^* would increase $v(e^*,s^*)$).

[30]I define the two errors taking as *given* that for $e < e^*$, the only way to accept an agent is after verification. The optimal mechanism also makes errors due to the verification cost. The cost contributes to the rejection of agents who would be worth accepting if verification was free. As s^* increases, part of Type II error A turns into Type I error B, which benefits the principal, who prefers to reject without verification (rather than accept after verification) agents who lie below $I_u(c)$.

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

which is by assumption negative definite. Particularly, $v_{ee}^{\text{pro-}e}(e^*,s^*),v_{ss}^{\text{pro-}e}(e^*,s^*) < 0$ and $\det(J(e^*,s^*)) > 0$. Also, $v_{es}^{\text{pro-}e}(e^*,s^*) = v_{se}^{\text{pro-}e}(e^*,s^*) = -\left(\widetilde{u}(e^*,s^*) - c\right)\widetilde{f}(e^*,s^*) > 0$.

First, the total derivatives of $e^*$and s^* with respect to c are:

$$\frac{de^*}{dc} \propto \overbrace{-v_{ec}^{\text{pro-}e}(e^*,s^*)v_{ss}^{\text{pro-}e}(e^*,s^*)}^{\substack{\text{<0: direct effect of } c \text{ on} \\ e^* \text{ due to increase in} \\ \text{marginal verification costs}}} + \overbrace{v_{sc}^{\text{pro-}e}(e^*,s^*)v_{es}^{\text{pro-}e}(e^*,s^*)}^{\substack{\text{>0: indirect effect of } c \\ \text{on } e^* \text{ through direct} \\ \text{effect of } c \text{ on } s^*}},$$

$$\frac{ds^*}{dc} \propto \underbrace{-v_{sc}^{\text{pro-}e}(e^*,s^*)v_{ee}^{\text{pro-}e}(e^*,s^*)}_{\substack{\text{>0: direct effect of } c \text{ on} \\ s^* \text{ due to increase in} \\ \text{marginal verification costs}}} + \underbrace{v_{ec}^{\text{pro-}e}(e^*,s^*)v_{se}^{\text{pro-}e}(e^*,s^*)}_{\substack{\text{<0: indirect effect of } c \\ \text{on } s^* \text{ through direct} \\ \text{effect of } c \text{ on } e^*}},$$

where $v_{ec}^{\text{pro-}e}(e^*,s^*) = -\int_{s^*}^{1}\widetilde{f}(e^*,s)ds < 0$ and $v_{sc}^{\text{pro-}e}(e^*,s^*) = \int_{\underline{e}(s^*)}^{e^*}\widetilde{f}(e,s^*)de > 0$ are the partial derivatives of $v_e^{\text{pro-}e}$ and $v_s^{\text{pro-}e}$ with respect to c. An increase in the cost c of verification tends to directly cause (i) e^* to decrease by magnifying the verification cost savings associated with a decrease in e^* and (ii) s^* to increase by magnifying the verification cost savings associated with an increase in e^*.[32] However, an increase in s^* tends to cause e^* to increase by (i) reducing the marginal increase in verification costs associated with an increase in e^* and (ii) increasing the marginal net decrease in the Type II errors A and B 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 A. 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.

Second, the principal's optimal payoff is decreasing in c. Third, the optimal payoff is higher under less pro-e biased composite measures. Namely, take any two pro-e biased composite measures σ' and σ. If all iso-composite-measure curves of σ cross the iso-composite-measure curves of σ' from above (i.e., σ is more pro-e biased than σ'), the principal's optimal payoff is higher under σ' than under σ.[33] Fourth, the principal's payoff is expected 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 sensitivity of the composite measure to e or t.

Implementation of the optimal mechanism. We have so far restricted (without loss) attention to truth-telling mechanisms. However, the optimal mechanism under a pro-e

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

[33] Comparative statics of s^* with respect to σ would have little value, since optimal composite measure thresholds under different composite measures are not comparable.

composite measure 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 verification or (ii) without providing any evidence, ask the principal to verify your composite measure, and if it is at least s^*, you will be accepted. The first option is not always provided (e.g., when verification is free). A similarly simple implementation of the optimal mechanism under a pro-t composite measure is not possible. In that case, the principal needs to ask for evidence also from agents whose composite measure he verifies.[34]

3.5 The effects of the verification cost and evidence-bias of the composite measure on the principal's decision errors

Looking at Figure 3, one can see how the verification cost and the pro-e bias of the composite measure induce the principal to make errors in favor of high- and against low-evidence agents. When we compare Figures 3(a) and 3(c) (see sections 3.3.1 and 3.4.1), we see that the verification cost *alone* gives rise to such errors. Similarly, when we compare Figures 3(a) and 3(b) (see sections 3.3.1 and 3.3.2), we see that the bias of the composite measure in favor of evidence also gives rise to such errors *even* when verification is free. But how do the two forces (i.e., verification cost and pro-e bias of the composite measure) interact in inducing errors in favor of high- and against low-evidence agents? Proposition 6 shows that the two forces are *complements*: The pro-e bias of the composite measure exacerbates the errors due to the verification cost by decreasing the threshold level of evidence required for acceptance without verification, as can be seen when we compare Figures 3(c) and 3(d).

Proposition 6. Take any pair of pro-t and pro-e biased composite measures. For any optimal evidence threshold $e^*_{\text{pro-}t} \in \arg\max_{e_{min}} v^{\text{pro-}t}(e_{min})$ under the pro-t biased measure and any optimal evidence and composite measure thresholds $(e^*_{\text{pro-}e}, s^*_{\text{pro-}e}) \in \arg\max_{(e_{min}, s_{min})} v^{\text{pro-}e}(e_{min}, s_{min})$ under the pro-e biased measure, (i) the evidence threshold is higher under the pro-t biased measure: $e^*_{\text{pro-}t} \geq e^*_{\text{pro-}e}$, or (ii) both evidence thresholds are optimal under both measures: $e^*_{\text{pro-}t} \in \arg\max_{e_{min}} v^{\text{pro-}e}(e_{min}, s^*_{\text{pro-}e})$ and $e^*_{\text{pro-}e} \in v^{\text{pro-}t}(e_{min})$.[35]

To see why, we can examine Figures 3(c) and 3(d). Both under a pro-t and a pro-e biased composite measure, an increase in the evidence threshold e^* causes (i) agents who lie below $I_\sigma(s^*)$ to move from the green area (i.e., from getting accepted without verification) to the red area (i.e., to getting rejected without verification) and (ii) agents

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

[35] If $e^*_{\text{pro-}e} \in (0,1)$, then the derivative $v_e^{\text{pro-}t}(e^*_{\text{pro-}e}) > 0$, and thus, if $v^{\text{pro-}t}(e_{min})$ is single-peaked in e_{min}, then $e^*_{\text{pro-}t} > e^*_{\text{pro-}e}$.

who lie above $I_u(c)$ to move from the green to the yellow area (i.e., to getting accepted after verification). The difference lies in the effect of an increase in e^* on agents who lie above $I_\sigma(s^*)$ and below $I_u(c)$. Under a pro-t biased composite measure, an increase in e^* causes those agents to move from the green area to the red area. On the other hand, under a pro-e biased composite measure, it causes them to move from the green area to the yellow area. Because those agents who lie above $I_\sigma(s^*)$ and below $I_u(c)$ deliver payoff less than c when accepted, it is better that they be moved to the red area (and deliver payoff zero) rather than to the yellow area (and deliver a net negative payoff, since the gross payoff from accepting them cannot cover the verification cost). Therefore, an increase in e^*, which increases verification, is more attractive to the principal under a pro-t rather than a pro-e biased composite measure.

Simply put, because verification does not distort the incentives of agents to present evidence under a pro-t composite measure but it *does* distort incentives under a pro-e biased composite measure, verification is more effective under a pro-t rather than a pro-e biased composite measure. Thus, fewer agents are optimally accepted without verification under a pro-t than under a pro-e biased composite measure.

4 Applications

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

4.1 Hiring for prestigious positions

A job candidate's evidence e is her CV quality. t is her ability and drive not captured by e. An employer wants to decide whether to hire the candidate for prestigious position. Verification works as follows: The employer has the option to (i) let another employer hire the candidate for some less prestigious position, (ii) observe her performance in that position, and (iii) decide whether to poach her at a cost higher than the cost of hiring her from the beginning.

In the optimal mechanism, candidates with high credentials are immediately hired for prestigious positions. On the other hand, talented candidates with low credentials have to go through less prestigious employers to prove their worth before landing a prestigious position (see Figure 3(c-d)). Also, if the candidates' performance in the less prestigious position is less sensitive to talent than talent is valuable in the more prestigious position, then worthy candidates with low credentials are at a disadvantage also in the poaching stage (see Figure 3(d)).

4.2 Promotions

An employee is characterized by efficiency t and hardworkingness e. $\sigma(e,t)$ is the employee's productivity in her current positions. The employee can provide or withhold evidence on e by, for example, choosing the hours she works at the office or from home. The employer can verify the employee's productivity $\sigma(e,t)$. If the agent continues to work in her current position, the employer's payoff is $\sigma(e,t)$. If the agent is promoted, the employer's payoff is $v(e,t)$. The employer's problem is equivalent to the one in section 2 with $u(e,t) := v(e,t) - \sigma(e,t)$, as long as the difference $v(e,t) - \sigma(e,t)$ is non-decreasing in both e and t.[36] This condition has a natural interpretation: Both effort and talent have a weakly higher marginal return in the higher position, which comes with increased responsibilities that allow the employee's efficiency and hardworkingness to have a larger impact.

Under differentiability and given Claims 1 and 2, the composite measure 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,

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

This condition also has a natural interpretation: In the production function of the higher position, the relative importance of efficiency and talent (relative to hardworkingness) is higher than in the current position.

4.3 College admissions and standardized testing

A college applicant's evidence e is her prior training, preparation, and parental support. t is her talent or drive that is not captured by e. The college wants to decide whether to admit the applicant or not. Verification amounts to requiring the applicant to submit her standardized test score.[37]

In the optimal mechanism, if the standardized test is not sensitive enough to talent, applicants have incentives to withhold evidence, which makes admission decisions imperfect at the expense of students with low prior training, preparation, and parental support (e.g., 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

[36] $u(e,t)$ could also be defined as $u(e,t) := v(e,t) - \sigma(e,t) - \underline{q}$, 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).

[37] 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 $c = 0$.

necessarily present under standardized testing to the extent that applicants can withhold evidence of prior training, preparation, and parental support. Students from advantaged backgrounds have an advantage over equally good—or even better—students from more modest backgrounds.

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.[38] $e \in [0,1]$ 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 verify the proportion of questions she can answer. Verification amounts to posing to the candidate countably infinitely many problems randomly sampled from the mass of problems that the candidate has not initially disclosed answers to.[39] Thus, if she presents answers to mass $e' \in [0,e]$ of problems, she will answer proportion $p(e,t,e') \coloneqq [e - e' + (b - e)t]/(b - e')$ 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. 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.

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 composite measure, which he can then choose to verify 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.

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

[39] The agent is equally likely to find an answer to any of the problems, so there is no need to identify problems with an index.

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).

5.1.1 Optimal screening when the agent cannot withhold evidence

In this section, it is assumed that e is observed by the principal.[40] Then, given that the composite measure is at least somewhat sensitive to t, verification reveals t. The principal's problem is decoupled: He can solve it for each e separately. It is easy to see that for each e, the principal needs to choose between (i) accepting without verification every agent with evidence e and (ii) accepting an agent with evidence e after verification if $u(e,t) \geq c$ and rejecting an agent with evidence e without verification if $u(e,t) < c$. When $\mathbb{E}_t\left[\min\{u(e,t), c\}|\, e\right] > 0$, option (i) delivers a higher payoff to the principal, while when $\mathbb{E}_t\left[\min\{u(e,t), c\}|\, e\right] < 0$, option (ii) is better. Proposition 7 describes the optimal mechanism.

Proposition 7. In the optimal mechanism, for each level of evidence $e \in [0,1]$, if

(i) $\mathbb{E}_t\left[\min\{u(e,t), c\}|\, e\right] > 0$, then $\Pi(e,t) = 1$ and $T(e,t) = 0$ for every $t \in [0,1]$ (i.e., every agent with evidence e is accepted without verification), and

(ii) $\mathbb{E}_t\left[\min\{u(e,t), c\}|\, e\right] \leq 0$, then $\Pi(e,t) = T(e,t) = \mathbf{I}(u(e,t) \geq c)$ for every $t \in [0,1]$ (i.e., an agent with evidence e is accepted after verification if $u(e,t) \geq c$; otherwise, she is rejected without verification),

where $\mathbb{E}_t\left[\min\{u(e,t), c\}|\, e\right]$ is the expectation of $\min\{u(e,t), c\}$ conditional on e.

Notice that because the agent cannot withhold evidence, the optimal mechanism does not depend on σ. Part (ii) of Proposition 7 implies that under a pro-e (resp. pro-t biased composite measure), if two agents (e_1,t_1) and (e_2,t_2), $e_2 > e_1$, both need to have their composite measures verified (based on their level of evidence) to get accepted, 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.[41] 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.

[40]We can also allow for only part of the evidence to be observed by the principal. If the agent's type is (e_p,e,t) distributed over $[0,1]^3$, where e_p is the publicly observed part and e is the part that can be hidden, the optimal mechanism is a collection mechanisms like the one described in section 3: one mechanism for each value of e_p.

[41]To see this, notice that under a pro-e (resp. pro-t) biased composite measure, 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. $\sigma(e_1,t_1) > \sigma(e_1,t_1)$).

Combined with the analysis of the baseline model, the results for the case where e is public information have implications for affirmative action in college admissions (i.e., screening for talent, trying to control for applicants' unequal backgrounds, measured by e). The baseline model has shown that 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 will not be very effective. The results of this section show that if any of the two condition fails, affirmative action is effective. Particularly, if college applicants *cannot* withhold evidence of an advantaged background and 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 verification 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 withhold evidence of an advantaged background. Thus, if condition (i) or (ii) fails, constraints on affirmative action would have significant effects on diversity in college admissions.

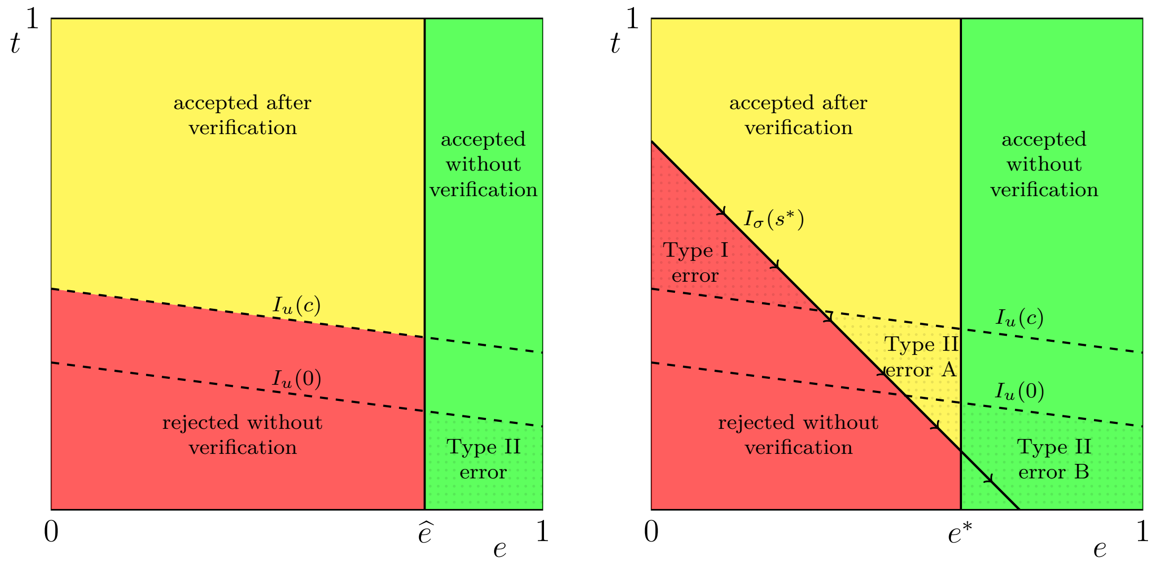
Comparison with baseline model. Assume now that the expectation of $\min\{u(e,t), c\}$ conditional on e crosses zero from below at most once. This means that if for a certain level of evidence e, the principal prefers to accept every agent with evidence e without verification rather than pay verification costs to reject some unworthy agents with evidence e, then for any higher level of evidence $e' > e$, the principal still prefers to accept every agent with evidence e' without verification. Clearly, a sufficient condition is that the expectation of $\min\{u(e,t), c\}$ conditional on e be increasing in e, which is satisfied as long as t does not stochastically depend on e "too negatively."[42] Then, Corollary 7.1 shows that the optimal mechanism coincides with the optimal mechanism of the baseline model—where the agent *can* withhold evidence—when the composite measure is pro-t biased. This means that as long as the composite measure is sensitive enough to talent, the ability of the agent to withhold evidence does not constrain the principal's ability to screen. Under a pro-t biased composite measure, the ability of the agent to withhold evidence constrains the principal's ability to screen only if the principal would want to switch from case (i) for some evidence level e to case (ii) for some evidence level $e' > e$ in Proposition 7, which would make those agents with evidence e' who are rejected to only present evidence e.

Corollary 7.1. Assume that there exists $\widehat{e}$ such that $\operatorname{sgn}\{\mathbb{E}_t[\min\{u(e,t), c\}|\, e]\} = \operatorname{sgn}\{e - \widehat{e}\}$. In the optimal mechanism, $\Pi(e,t) = \mathbf{I}(u(e,t) \geq c \text{ or } e \geq \widehat{e})$ and $T(e,t) = \mathbf{I}(u(e,t) \geq c \text{ and } e < \widehat{e})$, which coincides with the optimal mechanism of the baseline

[42] For example, it is sufficient that for any $e' > e$, the conditional distribution of t conditional on e' first-order stochastically dominates the conditional distribution of t conditional on e.

Figure 4: Optimal bidimensional screening with substitutable attributes, costly verification, and private or public evidence

(a) Pro-t biased σ or publicly observed e (b) Pro-e biased σ and privately observed e



model—where the agent *can* withhold evidence—when the composite measure is pro-t biased (see Proposition 4).

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

Consider the case where the agent can also present evidence on t. That is, agent (e,t) can report any $(e',t') \leq (e,t)$ but not $e' > e$ or $t' > t$. Then, the principal can clearly achieve the full information first-best without verification, inducing 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.

A comparison between this and the main model offers the following lesson. When there is a valuable quality that the agent cannot provide evidence on and the principal can only imperfectly verify through a composite measure that is overly (compared to the principal's preferences) sensitive to another valuable quality, the principal is constrained in his evaluation of the agent by the agent's incentives to withhold evidence on that other valuable quality. This problem vanishes if the agent can provide evidence on every quality (or if those that she cannot provide evidence on are observed by the principal) or if the composite measure is sensitive enough to the quality that the agent cannot provide evidence on.

These results are consistent with the finding that hiding one's effort is prevalent among younger individuals. University students have been found to have a desire to project "effortless perfection" by deliberately hiding how hard they study (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 analysis suggests that this result may be specific to younger people. If as a person progresses in her career, her talent is revealed through all the evaluation stages that she goes through, then people that are further in their career paths should have lower incentives to hide their hard work than students and early-career professionals.

5.1.3 Optimal screening when the agent cannot present evidence

In this section, it is assumed that 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 threshold acceptance policies after verification. Proposition 8 characterizes IC mechanisms.

Proposition 8. 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 8, where e is evidence. Condition (iii) is stronger (when combined with the other two conditions) than the corresponding condition (iii) of Proposition 8. 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) with a higher composite measure. The condition is stricter than the one in Proposition 8 because now agents can also imitate types with higher e to potentially get accepted without verification. Thus, that agents cannot present evidence on e enhances the need for verification. Last, condition (ii) ensures that an agent (e,t) does not want to imitate an agent (e',t') with evidence equal composite measure $\sigma(e',t') = \sigma(e,t)$ to get accepted with probability $\Pi(e',t')$ instead of $\Pi(e,t)$. The condition is stricter than the one in Proposition 8 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 verification is the same for everyone. Lemma 5 shows that we can constrain attention to mechanisms that satisfy condition (iii) of Proposition 8 with equality.

Lemma 5. 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 verification 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 verification costs than M.

By Lemma 5, $\Pi(e,t) = \Pi(0,0) + T(e,t)$. 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

$$\int_0^1 \int_{\underline{e}(s)}^{\overline{e}(s)} [\Pi(e,\tau(e,s))(u(e,\tau(e,s)) - c) + c\Pi(0,0)]\, f(e,\tau(e,s))deds, \tag{4}$$

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 9 describes that extreme optimal mechanism.

Proposition 9. 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 verification or

(ii) $s^* > 0$, and each agent (e,t) is (a) accepted after verification if $\sigma(e,t) \geq s^*$ or (b) rejected without verification 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.[43] Also, 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 composite measures are not inherently better than pro-e biased ones when the agent cannot present evidence. Regardless of whether it is pro-t or -e biased, the more closely the composite measure aligns with the principal's preferences, the higher the principal's optimal payoff is.

A comparison between this and the main model implies the following about the "signal-jamming" problem that arises in career concern models (see, e.g., Holmström, 1999), where the employer monitors the employee's productivity. If the employer can ask for evidence of effort (which the employee can provide at little to no cost), 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 undersensitive enough to talent, the signal-jamming problem persists even if the employer can ask for evidence of effort. In that case, agents have incentives to withhold evidence, which they should be paid information rents to reveal.

[43] In more detail, assume for simplicity that the optimal mechanism is unique. When the composite measure is pro-e biased, if some—but not all–agent types are optimally accepted without verification when the agent can present evidence on e (i.e., $e^* \in (0,1)$ in Proposition 5), then the principal's payoff is lower when the agent cannot present evidence. When the composite measure is pro-t biased, if not all agent types are optimally accepted without verification when the agent can present evidence on e (i.e., $e^* > 0$ in Proposition 4), then the principal's payoff is lower when the agent cannot present evidence.

5.2 Costly composite measure design

Treating the composite measure function σ as exogenous is reasonable in several applications. For example, in hiring for prestigious positions (section 4.1), the employee's production function for the other employer is not chosen by the employer hiring for the prestigious position. 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 verification amounts to tests and interviews), the principal may be able to choose how agent types map into composite measures. 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 verify the value of σ during the interaction with the agent. Indeed, it is reasonable that the principal needs to design a composite measure (if she designs one at all) *before* the interaction with the agent due to time constraints and the complexity of designing a composite measure. Then, the principal's problem can be solved in two steps: (i) finding the optimal mechanism for each possible $\sigma \in \Sigma$, and then (ii) choosing the optimal $\sigma^* \in \Sigma$ form the set Σ of conceivable composite measure functions. The solution to the first step is the one we have already described.[44]

If composite measures that are more sensitive to talent are more expensive to design, the results of section 3 imply that as long as the composite measure 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 composite measure more sensitive to talent. The principal will want to make the composite measure at most as sensitive to talent as his preferences are, since composite measures that are overly sensitive to talent are as effective as those that are exactly aligned with the principal's preferences. However, as section 5.1.3 has shown, when agents cannot present evidence, the principal always gains from finely calibrating the composite measure's sensitivity to the agent's attributes to make it align with his preferences—regardless of whether the composite measure is pro-t or -e biased.

5.3 Endogenous evidence production

If the agent produces evidence before the interaction with the principal, in some settings, the principal may be able to affect the agent's evidence production by committing to a

[44] That is, assuming that Σ contains only pro-t and pro-e biased composite measures (and possibly a composite measure that exactly matches the principal's preferences). Also, if the composite measures in Σ are totally ordered (i.e., any pair of iso-composite-measure curves of any two composite measures in Σ cross at most once), there are no gains from designing multiple composite measures to the extent that all composite measures are equally costly to verify.

mechanism *before* the agent produces evidence. Indeed, in promotion decisions (section 4.2), the employer may sometimes have the power to commit to promotion rules, using the prospect of promotion to incentivize the employee to exert effort. Of course, whether the employer wants to do that will depend on the extent to which using the prospect of promotion to incentivize effort interferes with the primary objective of promotions: assigning employees to the positions where they are most valuable. 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 has little power to 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 evidence.[45] 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., (e^*,s^*) 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 emphasizes 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, agents so talented that they are accepted even they have $e = 0$ and agents so untalented that they are rejected even if they have $e = 1$ do not exert effort. More generally, effort may be non-monotone in t. Thus, evidence and talent may sometimes be stochastically dependent in complicated ways.

6 Conclusion

This paper has proposed a model of multidimensional screening, where an agent (she) with two attributes—training and talent—chooses how much hard evidence of training to present. The principal (he) then verifies (possibly) at a cost the value of a composite measure of training and talent, who then decides whether to accept or reject the agent.

[45] The optimal mechanism can always be summarized by these two thresholds. Under a pro-t biased composite measure, there is only an evidence threshold.

The agent cannot unilaterally prove anything about her talent. The composite measure is increasing in both training and talent, and the principal (weakly) values both training and talent in an agent. If the principal is going to verify the value of the composite measure, then the agent may have incentives to withhold evidence or training—although the principal values training—to influence how the principal interprets the composite measure. Particularly, she may want to withhold evidence to make the principal attribute the composite measure to talent, thereby overestimating her talent.

This problem arises when the composite measure is less sensitive to talent than talent is valuable to the principal. In that case, the optimal mechanism makes errors in favor of high-training and against low-training agents. The errors are due to two forces: (i) the principal's incentive to save on verification costs by accepting high-training agents without verifying their composite measure and (ii) due to the strategic incentives of agents to withhold evidence of training when the principal verifies the value of a composite measure that is under-sensitive to talent. Each of the two forces *individually* causes the principal to optimally favor high- over low-training agents. When the two forces are *combined* (i.e., the composite measure is under-sensitive to talent and verification is costly), the second force reduces the effectiveness of verification, causing the principal to accept even more agents without verification to save on verification costs, thereby exacerbating the errors the principal makes by accepting agents without verification. The two forces are complements in inducing errors in favor of high-training and against low-training agents. Remarkably, these errors arise even when the principal *only* values talent. The principal still optimally rewards evidence of training—making errors in favor of high-training and against low-training agents—even though training 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 withhold evidence of prior training and parental support, 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

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.

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.

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.

Choquet, G. (1954). Theory of capacities. In *Annales de l'institut Fourier*, volume 5, pages 131–295.

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

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.

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.

Gates, B. (2025). *Source Code: My Beginnings*. Random House.

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. (2006). Bundling as an optimal selling mechanism for a multiple-good monopolist. *Journal of Economic Theory*, 127(1):1–35.

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. 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).[46]

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 verification 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 (e',t') (with $e' \leq e$) by

$$\widetilde{P}(e',t';e,t) := (1 - T(e',t'))P(e',t',\emptyset) + T(e',t')\mathbf{I}\left(\sigma(e,t) \geq \sigma(e',t')\right).$$

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

[46] 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.

Step 2: I now show that condition (iii') is necessary for IC by showing the contrapositive.[47] 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

$$\max_{(e',t') \leq (e,1)} [(1 - T(e',t'))P(e',t';\emptyset) + T(e',t')\mathbf{I}\left(\sigma(e,t) \geq \sigma(e',t')\right)] = \Pi(e,t). \tag{5}$$

Assume that conditions (i) and (iii') are satisfied. Then, $\Pi(e,t) \geq \Pi(e,0) \geq (1 - T(e',t'))P(e',t',\emptyset)$ for any (e',t') with $e' \leq e$. Therefore, (5) is equivalent to

$$\max_{(e',t') \in \{(x,y) \in [0,1]^2 : x \leq e \text{ and } \sigma(e,t) \geq \sigma(x,y)\}} [(1 - T(e',t'))P(e',t';\emptyset) + T(e',t')] = \Pi(e,t). \tag{6}$$

Given that $\Pi(e,t)$ is non-decreasing in t (condition (i)), (6) 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,

$$e \in \operatorname*{arg\,max}_{r \in [\underline{e}(\sigma(e,t)),e]} \Pi(r, \tau(r,\sigma(e,t))). \tag{7}$$

Thus, IC is satisfied for every type if and only if for every (e,t), (7) 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 (7) 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 $r \in [\underline{e}(\sigma(e,t)),e]$ for every (e,t) is equivalent to condition (ii).

Step 5: Finally, notice that provided that conditions (i) and (ii) hold, conditions (iii)

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

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),[48] 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 $M' \coloneqq \langle T',P' \rangle$ with[49]

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

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 verification 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). Denote by $\mathcal{P} \subseteq L^1(\{(e,s) \in [0,1]^2 : e \in [\underline{e}(s),\overline{e}(s)]\})$ 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 and compact (e.g., see Yang and Yang, 2025). The objective function (3) is linear (and thus convex) in Π. By the Dominated Convergence Theorem, it is also continuous in Π. By Bauer's 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 (see Theorem 40.1 in Choquet, 1954). **Q.E.D.**

[48]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 $\Pi(e',0) \geq \Pi(e,0)$.

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

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 a pro-e biased composite measure, 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 a pro-e biased composite measure) e_s is non-decreasing in s. Take any $\underline{s},\overline{s} \in [0,1]$ with $\overline{s} > \underline{s}$, and define $S := (e_{\overline{s}},e_{\underline{s}}) \cap [\underline{e}(\overline{s}), \overline{e}(\overline{s})] \cap [\underline{e}(\underline{s}), \overline{e}(\underline{s})]$.

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-composite-measure curve $I_\sigma(s)$. That is, any other agent (e,t) on that iso-composite-measure 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.[50] 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

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

The mechanism affects the second term only through e^*. Given e^*, setting $\Pi(e,t) = \mathbf{I}\left(u(e,t) \geq c \text{ or } e \geq e^*\right)$ 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

[50] 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

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

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 Proposition 6 First, notice the following relationship between $v_e^{\text{pro-}t}(e_{min})$ and $v_e^{\text{pro-}e}(e_{min},s_{min})$:

$$\begin{aligned}
v_e^{\text{pro-}t}(e_{min}) =& \int_0^1 (u(e_{min},t) - c)\mathbf{I}(u(e_{min},t) \geq c)f(e_{min},t)dt - \int_0^1 u(e_{min},t)f(e_{min},t)dt \\
=& \int_0^1 [u(e_{min},t)\left(\mathbf{I}(u(e_{min},t) \geq c) - 1\right) - c\mathbf{I}(u(e_{min},t) \geq c)]\, f(e_{min},t)dt \\
=& -\int_0^1 [u(e_{min},t)\mathbf{I}(u(e_{min},t) < c) + c\mathbf{I}(u(e_{min},t) \geq c)]\, f(e_{min},t)dt \\
=& -\int_0^1 [\tilde{u}(e_{min},s)\mathbf{I}(\tilde{u}(e_{min},s) < c) + c\mathbf{I}(\tilde{u}(e_{min},s) \geq c)]\, \tilde{f}(e_{min},s)ds \\
=& -\int_0^1 \left\{\begin{aligned} \tilde{u}(e_{min},s)\, [\mathbf{I}(\tilde{u}(e_{min},s) < c) - \mathbf{I}(s < s_{min})] \\ +c\, [\mathbf{I}(\tilde{u}(e_{min},s) \geq c) - \mathbf{I}(s \geq s_{min})] \end{aligned}\right\} \tilde{f}(e_{min},s)ds \\
& - \int_0^{s_{min}} \tilde{u}(e_{min},s)\tilde{f}(e_{min},s)ds - c\int_{s_{min}}^1 \tilde{f}(e_{min},s)ds \\
=& -\int_0^1 \left\{\begin{aligned} \tilde{u}(e_{min},s)\, [\mathbf{I}(\tilde{u}(e_{min},s) < c) - \mathbf{I}(s < s_{min})] \\ +c\, [\mathbf{I}(\tilde{u}(e_{min},s) \geq c) - \mathbf{I}(s \geq s_{min})] \end{aligned}\right\} \tilde{f}(e_{min},s)ds + v_e^{\text{pro-}e}(e_{min},s_{min}) \\
=& v_e^{\text{pro-}e}(e_{min},s_{min}) - \int_0^1 \left\{\begin{aligned} \tilde{u}(e_{min},s)\, [\mathbf{I}(\tilde{u}(e_{min},s) < c) - \mathbf{I}(s < s_{min})] \\ +c\, [1 - \mathbf{I}(\tilde{u}(e_{min},s) < c) - (1 - \mathbf{I}(s < s_{min}))] \end{aligned}\right\} \tilde{f}(e_{min},s)ds \\
=& v_e^{\text{pro-}e}(e_{min},s_{min}) - \int_0^1 \left(\tilde{u}(e_{min},s) - c\right) [\mathbf{I}(\tilde{u}(e_{min},s) < c) - \mathbf{I}(s < s_{min})]\, \tilde{f}(e_{min},s)ds.
\end{aligned}$$

Now, fix some pro-e biased composite measure σ and take any s^* such that $(e^*,s^*) \in \arg\max_{(e_{min},s_{min})} v^{\text{pro-}e}(e_{min},s_{min})$ for some e^* under σ. Define

$$\begin{aligned}
v(e_{min},\alpha) :=& \alpha v^{\text{pro-}t}(e_{min}) + (1 - \alpha)v^{\text{pro-}e}(e_{min},s^*) \\
=& \alpha \left\{\int_0^1 \int_0^{e_{min}} (u(e,t) - c)\mathbf{I}(u(e,t) \geq c)f(e,t)dedt + \int_0^1 \int_{e_{min}}^1 u(e,t)f(e,t)dedt\right\}
\end{aligned}$$

$$+(1-\alpha)\left\{\begin{aligned}&\int_{s^*}^1\int_{\underline{e}(s)}^{\max\{\min\{\overline{e}(s),e_{min}\},\underline{e}(s)\}}\left(\widetilde{u}(e,s)-c\right)\widetilde{f}(e,s)deds\\&\qquad+\int_0^1\int_{\max\{\underline{e}(s),e_{min}\}}^{\max\{\overline{e}(s),e_{min}\}}\widetilde{u}(e,s)\widetilde{f}(e,s)deds\end{aligned}\right\}.$$

where $\widetilde{u}_\alpha(e,s)\equiv u(e,\tau(e,s))$ and $\widetilde{f}(e,s)\equiv f(e,\tau(e,s))$ are defined under σ. Take any $e^*_{\text{pro-}t}\in\arg\max_{e_{min}}v(e_{min},1)=\arg\max_{e_{min}}v^{\text{pro-}t}(e_{min})$ and $e^*_{\text{pro-}e}\in\arg\max_{e_{min}}v(e_{min},0)=\arg\max_{e_{min}}v^{\text{pro-}e}(e_{min},s^*)$. Given the relationship between $v_e^{\text{pro-}t}(e_{min})$ and $v_e^{\text{pro-}e}(e_{min},s_{min})$, we have that

$$\begin{aligned}\frac{\partial^2 v(e_{min},\alpha)}{\partial e_{min}\partial\alpha}&=v_e^{\text{pro-}t}(e_{min})-v_e^{\text{pro-}e}(e_{min},s^*)\\&=-\int_0^1\left(\widetilde{u}(e_{min},s)-c\right)\left[\mathbf{I}(\widetilde{u}(e_{min},s)<c)-\mathbf{I}(s<s^*)\right]\widetilde{f}(e_{min},s)ds\geq 0,\end{aligned}$$

so $v(e_{min},\alpha)$ has increasing differences in (e_{min},α), and Topkis' Monotonicity Theorem implies that $\arg\max_{e_{min}}v(e_{min},\alpha)$ is increasing in α in the strong set order. Therefore, (i) $e^*_{\text{pro-}t}\geq e^*_{\text{pro-}e}$ or (ii) $e^*_{\text{pro-}t}\in\arg\max_{e_{min}}v^{\text{pro-}e}(e_{min},s^*)$ and $e^*_{\text{pro-}e}\in v^{\text{pro-}t}(e_{min})$. We conclude that for any $e^*_{\text{pro-}t}\in\arg\max_{e_{min}}v^{\text{pro-}t}(e_{min})$ and $(e^*_{\text{pro-}e},s^*_{\text{pro-}e})\in\arg\max_{(e_{min},s_{min})}v^{\text{pro-}e}(e_{min},s_{min})$, (i) $e^*_{\text{pro-}t}\geq e^*_{\text{pro-}e}$ or (ii) $e^*_{\text{pro-}t}\in\arg\max_{e_{min}}v^{\text{pro-}e}(e_{min},s^*_{\text{pro-}e})$ and $e^*_{\text{pro-}e}\in v^{\text{pro-}t}(e_{min})$. **Q.E.D.**

Proof of Proposition 7 and Corollary 7.1 Denote by $u_{\text{accept}}(e):=\int_0^1 u(e,t)f(e,t)dt/\int_0^1 f(e,t)dt$ the expected payoff from accepting without verification every agent with evidence e, and by $u_{\text{verify}}(e):=\int_0^1(u(e,t)-c)\mathbf{I}(u(e,t)\geq c)f(e,t)dt/\int_0^1 f(e,t)dt$ the expected payoff from accepting after verification every agent with evidence e who gives payoff at least c.

$$\begin{aligned}u_{\text{accept}}(e)-u_{\text{verify}}(e)&=\frac{\int_0^1 u(e,t)f(e,t)dt-\int_0^1(u(e,t)-c)\mathbf{I}(u(e,t)\geq c)f(e,t)dt}{\int_0^1 f(e,t)dt}\\&=\frac{\int_0^1\left[u(e,t)\mathbf{I}(u(e,t)<c)+c\mathbf{I}(u(e,t)\geq c)\right]f(e,t)dt}{\int_0^1 f(e,t)dt}\\&=\frac{\int_0^1\min\{u(e,t),c\}f(e,t)dt}{\int_0^1 f(e,t)dt}=\mathbb{E}_t\left[\min\{u(e,t),c\}\middle|\,e\right],\end{aligned}$$

and the results follow. **Q.E.D.**

Proof of Proposition 8 Denote the total probability with which type (e,t) is accepted if she reports (e',t') (with $e'\leq e$) by

$$\widetilde{P}(e',t';e,t):=(1-T(e',t'))P(e',t',\emptyset)+T(e',t')\mathbf{I}\left(\sigma(e,t)\geq\sigma(e',t')\right).$$

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

$$\max_{(e',t')\leq(1,1)} \left[(1-T(e',t'))P(e',t';\emptyset) + T(e',t')\mathbf{I}\left(\sigma(e,t) \geq \sigma(e',t')\right)\right] = \Pi(e,t). \tag{8}$$

Assume that conditions (i) and (iii') are satisfied. Then, $\Pi(e,t) \geq \Pi(e,0) \geq (1-T(e',t'))P(e',t',\emptyset)$ for any (e',t'). Therefore, (8) is equivalent to

$$\max_{(e',t')\in\{(x,y)\in[0,1]^2:\sigma(e,t)\geq\sigma(x,y)\}} \left[(1-T(e',t'))P(e',t';\emptyset) + T(e',t')\right] = \Pi(e,t). \tag{9}$$

Given that $\Pi(e,t)$ is non-decreasing in t (condition (i)), (9) can equivalently be written as

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

or equivalently,

$$e \in \operatorname*{arg\,max}_{r\in[\underline{e}(\sigma(e,t)),\overline{e}(\sigma(e,t))]} \Pi(r,\tau(r,\sigma(e,t))). \tag{10}$$

Thus, IC is satisfied for every type if and only if for every (e,t), (10) 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 (10) 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 5 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 $M' := \langle T',P' \rangle$ with[51]

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

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 (i) and (ii) of Proposition 8. 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 8. Therefore, M' is IC.

Last, to see why the second part is true, notice that for $c > 0$, M' saves on verification 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.**

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

Online Appendix

Multidimensional screening with substitutable attributes and costly verification

Orestis Vravosinos

B $(m+n)$-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 $(e_1,e_2,\dots,e_m,t_1,t_2,\dots,t_n)$ with full-support density $f : [0,1]^{m+n} \to \mathbb{R}_{++}$. $(e_1,e_2,\dots,e_m)$ are different dimensions of evidence and $(t_1,t_2,\dots,t_m)$ are different dimensions of talent. The agent can present any combination of evidence $\boldsymbol{r} \in [\boldsymbol{0},\boldsymbol{e}]$. The composite measure $\sigma : [0,1]^{m+n} \to [0,1]$ 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 threshold acceptance policies after verification are without loss.

Lemma 6 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 6. 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 7 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 verification and acceptance probabilities.

Lemma 7. 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 $\sigma(\boldsymbol{e},\boldsymbol{t}) = \sigma(\boldsymbol{e},\boldsymbol{t}')$.

Here is the intuition behind this result. The only reason to verify an agent's composite measure before accepting her—rather than accept her without verification—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 6, 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. $T(\boldsymbol{e}',\boldsymbol{t}')$). Among all agents in $X(\boldsymbol{e}',s)$, $(\boldsymbol{e},\boldsymbol{t})$ will want to imitate the one with the highest probability of acceptance without verification. 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 $T(\boldsymbol{e}',\boldsymbol{t}') > \inf_{(e'',t'')\in X(e',s)} T(\boldsymbol{e}'',\boldsymbol{t}'')$ (and thus $P(\boldsymbol{e}'',\boldsymbol{t}',\emptyset) < \sup_{(e'',t'')\in X(e',s)} P(\boldsymbol{e}',\boldsymbol{t}'',\emptyset)$) keeping Π fixed. We conclude that among agents with the same evidence and composite measure, there is no point in verifying the composite measure of some agents with higher probability than others, as doing so does not reduce incentives of others to misreport their type and makes the principal incur higher than necessary verification 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 $\sigma(\boldsymbol{e},\boldsymbol{t}) = \sigma(\boldsymbol{e},\boldsymbol{t}')$.[52] In other words, the principal can constrain attention to mechanisms that ask agents only for evidence and a 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 10 generalizes the IC characterization of Proposition 1 to the case of $(m + n)$-dimensional screening.

Proposition 10. 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 $\boldsymbol{e} \in \{\boldsymbol{e} \in [0,1]^m : \sigma(\boldsymbol{e},\boldsymbol{0}) \le s \le \sigma(\boldsymbol{e},\boldsymbol{1})\}$ for every $s \in [0,1]$, and

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

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

The conditions are analogous to those of Proposition 1. Notice that there are no IC conditions on the comparison between the values of T, P, or Π for agent types $(\boldsymbol{e},\boldsymbol{t})$ and $(\boldsymbol{e}',\boldsymbol{t}')$ such that $\boldsymbol{e} \not\geq \boldsymbol{e}'$ and $\boldsymbol{e} \not\leq \boldsymbol{e}'$. That is, because neither agent type has the evidence to imitate the other.

Lemma 8 generalizes Lemma 2, showing that we can constrain attention to mechanisms that satisfy condition (iii) of Proposition 10 with equality.

Lemma 8. 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 $(\boldsymbol{e},s)$, $s \in [\sigma(\boldsymbol{e},\boldsymbol{0}),\sigma(\boldsymbol{e},\boldsymbol{1})]$

[52]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 threshold acceptance policies after verification.

that is outcome-equivalent to M and has at most as high verification costs as M. For $c > 0$, if also $\Pi(\boldsymbol{e},s) > \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))$ for a positive measure of $(\boldsymbol{e},s)$'s, then M' has lower verification costs than M.

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. $\widetilde{u}(\boldsymbol{e},s)$ is assumed to be increasing in s.[53] 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 8, condition (iii) of Proposition 10 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

$$\int_0^1 \cdots \int_0^1 \int_{\sigma(\boldsymbol{e},\mathbf{0})}^{\sigma(\boldsymbol{e},\mathbf{1})} [\Pi(\boldsymbol{e},s)(\widetilde{u}(\boldsymbol{e},s) - c) + c\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))]\, \widetilde{f}(\boldsymbol{e},s)dsde_1 \cdots de_m, \qquad (11)$$

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 9. There exists an optimal deterministic mechanism.

B.1 Composite measure biased in favor of talent

The definition of a pro-t biased composite measure 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 11 derives the optimal mechanism under a pro-t composite measure.

Proposition 11. If σ is pro-t biased, then there exists an optimal mechanism with $\Pi(\boldsymbol{e},s) = \mathbf{I}(\widetilde{u}(\boldsymbol{e},s) \geq c \text{ or } \boldsymbol{e} \in E^*)$ and $T(\boldsymbol{e},s) = \mathbf{I}(\widetilde{u}(\boldsymbol{e},s) \geq c \text{ and } \boldsymbol{e} \notin E^*)$ 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 $\boldsymbol{e}' \in E^*$).[54]

[53] For $(\boldsymbol{e},s)$ such that $s = \sigma(\boldsymbol{e},\mathbf{0})$, $\widetilde{u}(\boldsymbol{e},s) \equiv u(\boldsymbol{e},\mathbf{0})$. $\widetilde{u}(\boldsymbol{e},s)$ being increasing in s guarantees that the indifference sets of the principal, $I_u(\overline{u}) := \{(\boldsymbol{e},s) \in [0,1]^{m+1} : \widetilde{u}(\boldsymbol{e},s) = \overline{u}\}$, are m-dimensional, as assumed in the case of $m = n = 1$. The results are still true with $\widetilde{u}(\boldsymbol{e},s)$ non-decreasing in s, which would somewhat complicate the proofs.

[54] Clearly, if $c = 0$, $E^* = \emptyset$ without loss. If $c > 0$, $E^* \supset \{\boldsymbol{e} \in [0,1]^m : \widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) \geq c\}$. This has to be true, because among the agents who are accepted, an agent's composite measure should be verified 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, $(\boldsymbol{e},\mathbf{0})$'s composite measure should not be verified if $\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) > c$.

B.2 Composite measure biased in favor of evidence

The definition of a pro-e biased composite measure 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 $s \in [\max\{\sigma(\boldsymbol{e},\mathbf{0}),\sigma(\boldsymbol{e}',\mathbf{0})\}, \min\{\sigma(\boldsymbol{e},\mathbf{1}),\sigma(\boldsymbol{e}',\mathbf{1})\}]$, 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 12 derives the optimal mechanism under a pro-e biased composite measure.

Proposition 12. 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 $[0,1]^m$.[55]

C Proofs of results in Appendix B

Proof of Lemma 6 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 7 Take any IC mechanism M. Construct the mechanism $M' \equiv \langle T',P' \rangle$ with[56]

$$T'(\boldsymbol{e},\boldsymbol{t}) := \inf_{\boldsymbol{t}' \text{ s.t. } \sigma(\boldsymbol{e},\boldsymbol{t}')=\sigma(\boldsymbol{e},\boldsymbol{t})} T(\boldsymbol{e},\boldsymbol{t}) \leq T(\boldsymbol{e},\boldsymbol{t}), \quad \text{and}$$

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

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

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

[55] If $c > 0$, then $E^* \supset \{\boldsymbol{e} \in [0,1]^m : \tilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) > s^*\}$, because among the agents who are accepted, an agent's composite measure should be verified only if this will prevent others from imitating her.

[56] 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 $P'(\boldsymbol{e},\boldsymbol{t},\emptyset) = P(\boldsymbol{e},\boldsymbol{t},\emptyset)$.

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

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 6 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 $(\boldsymbol{e}',\boldsymbol{t}') \leq (\boldsymbol{e},\mathbf{1})$, 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 verification costs. **Q.E.D.**

Proof of Proposition 10 The proof proceeds like the proof of Proposition 1 and is thus omitted. **Q.E.D.**

Proof of Lemma 8 The proof proceeds like the proof of Lemma 2.

Take any IC mechanism $M \equiv \langle T,P\rangle$. Condition (iii) of Proposition 10 says that $\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) \geq (1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset)$ for any $(\boldsymbol{e},s)$. Then, construct the mechanism $M' := \langle T',P'\rangle$ with[57]

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

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

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 $(\boldsymbol{e},s)$

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

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 verification costs compared to M if there exists (a positive measure of) $(\boldsymbol{e},s)$ with $P(\boldsymbol{e},s,\emptyset)(1 - T(\boldsymbol{e},s)) < \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))$, since $T'(\boldsymbol{e},s) < T(\boldsymbol{e},s)$ for such $(\boldsymbol{e},s)$. **Q.E.D.**

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

Proof of Proposition 11 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 10 combined imply that E^* is an upper set of $[0,1]^m$. To see this, take any $\boldsymbol{e} \in E^*$ 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.[58]

Also, condition (i) of Proposition 10 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 (11) can be written as

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

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 $\Pi(\boldsymbol{e},s)=\mathbf{I}(\tilde{u}(\boldsymbol{e},s)>c)$ for every $\boldsymbol{e} \notin E^*$ maximizes the second term. It is also IC.

To show this, we first prove that $\Pi(\boldsymbol{e},s)=\mathbf{I}(\tilde{u}(\boldsymbol{e},s)>c \text{ or } \boldsymbol{e} \in E^*)$ satisfies condition (i) of Proposition 10. 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 $\tilde{u}(\boldsymbol{e},s')<c$ and $\boldsymbol{e} \notin E^*$. Since $\tilde{u}(\boldsymbol{e},s)$ is increasing in s, $\tilde{u}(\boldsymbol{e},s)<\tilde{u}(\boldsymbol{e},s')<c$. Therefore, $\Pi(\boldsymbol{e},s)=0$.

It remains to show that $\Pi(\boldsymbol{e},s)=\mathbf{I}(\tilde{u}(\boldsymbol{e},s)\geq c \text{ or } \boldsymbol{e} \in E^*)$ satisfies condition (ii) of Proposition 10. 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 $\tilde{u}(\boldsymbol{e}',s)<c$ 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 $\tilde{u}(\boldsymbol{e},s)<c$. We will show this by contradiction. Assume that $\tilde{u}(\boldsymbol{e},s)\geq c$. Then, we have that $\tilde{u}(\boldsymbol{e},s)\geq c>\tilde{u}(\boldsymbol{e}',s)$, which, given that σ is pro-t biased, implies that $\boldsymbol{e}' \ngeq \boldsymbol{e}$, a contradiction. **Q.E.D.**

Proof of Proposition 12 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 10 combined imply that E^* is an upper set of $[0,1]^m$.

Also, condition (i) of Proposition 10 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 (11) can be written as

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

The first term depends on the mechanism M only through E^*. The second term depends

[58] 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.

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 $\boldsymbol{e}' \ngeq \boldsymbol{e}$, then $s' < s$. Let $\boldsymbol{e}' \ngeq \boldsymbol{e}$:

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 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 the composite measure being pro-e biased 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 the composite measure being pro-e 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 10, 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 10 requires that $\Pi(\boldsymbol{e},s) \geq \Pi(\boldsymbol{e}^*,s) = 1$, which implies $\Pi(\boldsymbol{e},s) = 1$.

Case 1b: If $\tilde{u}(\boldsymbol{e}',s) > c$ and $s < s^*$, then $\tilde{u}(\boldsymbol{e},s) > c = \tilde{u}(\boldsymbol{e}^*,s)$, so because σ is pro-e biased, $\boldsymbol{e}^* \geq \boldsymbol{e}$, and thus IC condition (ii) of Proposition 10 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 10, 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.**