Multidimensional screening of strategic candidates*

Click to access latest version.

Orestis Vravosinos†

October 19, 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 or talent separately. The agent can present evidence of training but not of talent. Although favorable, evidence can make the principal attribute 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 high-training agents without verifying the composite measure. The two forces are complements in inducing these errors.

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, Dilip Abreu, Frank Yang, Gonzalo Cisternas, Guillaume Fréchette, Guillaume Haeringer, Idione Meneghel, Inga Deimen, Johannes Hörner, Laura Doval, Leeat Yariv, Luciano Pomatto, Mariann Ollár, Maren Vairo, Marina Halac, Moritz Meyer-ter-Vehn, Peter Caradonna, Sevgi Yuksel, Susanne Goldlücke, Vasiliki Skreta, Yeon-Koo Che, and audiences at the Twenty-Sixth ACM Conference on Economics and Computation (EC'25), 35th Stony Brook International Conference on Game Theory, 23rd annual International Industrial Organization Conference (IIOC 25), 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 at a cost. Nevertheless, in many cases, 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. For instance, standardized college admission tests pick up a combination of talent and training. Evidence of training is then critical when the evaluator tries to extract information about the candidate's talent through the composite measure.

However, the candidate may strategically withhold evidence of training to make the evaluator attribute the composite measure to talent instead of training. For example, a college applicant may downplay her training and parental support to portray her academic performance and standardized test scores as results of her brilliance. A job candidate might downplay her background to make the employer attribute her achievements and pre-employment test results to talent. An employee may hide how hard she works by working from home rather than the office to make the employer attribute her productivity to talent and promote her. An academic on the job market may strategically withhold certain results she has derived, saving them to answer audience questions later to appear exceptionally adept at thinking on her feet.

The above examples suggest that although presenting all her evidence of training is, in principle, in the candidate's best interest, her incentives to present evidence may be distorted if the evaluator has access to a composite measure of the candidate's training and talent. Thus, there can be a conflict between the two evaluation tools: (i) verifying a composite measure of talent and training 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.

I pursue them in the following principal-agent setting. The agent has a bidimensional type. The first dimension is her *training* and the second is her *talent*.[1] The agent has hard evidence of training. The evidence an agent with higher training has is a proper superset of the evidence an agent with lower training has. Thus, the agent can present evidence to prove any part of her training but cannot prove she is not withholding evidence; that is, she cannot prove her training is not higher than the evidence she has presented suggests. This means that, effectively, the agent can under-report but not over-report training. She cannot unilaterally prove anything about her talent.

The principal's payoff from accepting the agent (i) is non-decreasing in both training and talent and (ii) can be positive or negative. The principal ultimately wants to decide whether to accept or reject the agent (and receive payoff 0 in the latter case). 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 principal then (i) either pays a fixed cost to verify 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 mapping from the agent's type to the (scalar) composite measure is exogenous and increasing in training and talent. The agent wants to get accepted independently of her type.

Whether verification distorts the agent's incentives to present evidence of training depends crucially on the comparison between the principal's (acceptance payoff's) marginal rate of substitution of talent for training (MRS) and the MRS of the composite measure. I say that the composite measure is *talent-biased* if the principal's MRS is higher in absolute value than the composite measure's MRS. This means that the composite measure is more sensitive to talent than talent is valuable to the principal—and conversely, less sensitive to training than training is valuable to the principal. In the opposite case, I say that the composite measure is *training-biased*.

I show that if the composite measure is talent-biased, verification does not create incentives for the agents to withhold evidence, so that 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 mechanism

[1] 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 and preparation), I solve the problem for exogenous training. Section 6.3 endogenizes training.

in the opposite case: when the composite measure is training-biased. In that case, the optimal evaluation scheme never combines evidence and verification in the evaluation of an 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. It accepts some high-training agents—including unworthy agents 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. Among agents who do not meet that level of evidence, (i) it accepts after verification some unworthy agents with high training but low talent, although the payoff from accepting them does not cover the verification cost, and (ii) it rejects some worthy agents with high talent but low training, although the payoff from accepting them would exceed the verification cost.

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). In that case, the composite measure is automatically training-biased. 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 training-biased composite measure.

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 talent-biased—in which case the second force is absent—the cost of verification induces the principal to accept some high-training agents, including unworthy ones, without verification. Similarly, when the composite measure is training-biased, the optimal mechanism makes errors favoring high- over low-training agents even when verification is free. When the two forces are *combined* (i.e., the composite measure is training-biased 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 favoring high- over 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 composite measure is likely talent-biased, so the hiring process is easy. On the other hand, when talent is most valuable, the hiring process is flawed, favoring candidates with high training at the expense of 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 her efficiency or managerial skills. The employer can verify the employee's productivity in the current position. The employer's payoff from promoting the employee is the difference between her productivity in the new position and her productivity in the current position.[2] 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 effort-biased. The optimal promotion scheme thus 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. Training is the candidate's background and education, and talent is her ability and drive not captured by training. Verification amounts to letting a less prestigious employer hire the candidate with the option to poach the candidate later at a cost (additional to the cost of hiring her from the beginning), after observing her performance with that employer. In the optimal mechanism, Ivy-Leaguers—including unworthy ones—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 is more sensitive to talent in the more prestigious position than in the less prestigious one, then the composite measure (i.e., performance in the less prestigious position) is training-biased, so the prestigious employer makes errors also 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 withhold evidence of their socioeconomic background, preparation, and parental support and (ii) standardized test scores reflect talent (relative to background, preparation, and parental support) 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 imposing constraints on it would compromise colleges' ability to screen for talent.

The hiring and college admissions applications illustrate how inequalities can be perpetuated. When standardized tests are training-biased, college applicants with superior access to high-quality education and extensive preparation have an advantage over 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

[2]This normalizes the payoff from rejecting the employee (i.e., keeping her in her current position) to zero.

prestigious institutions.

After a discussion of related literature, section 2 presents the model. Section 3 characterizes the optimal screening mechanism. Section 4 discusses the results, and section 5 discusses applications. Section 6 studies extensions of the model. Section 7 concludes. Proofs are gathered in the Appendix.

Related literature. The urge to withhold favorable evidence of training to make people overestimate one's talent is so fundamental that children also seem to follow it when they eagerly proclaim how little they have studied for an exam. University students have also been found to deliberately hide how hard they study to project an image of "effortless perfection" (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 who may strategically withhold evidence on one of their qualities that both (i) is, in principle, favorable to them and (ii) contains useful information for the evaluator in order to influence how the evaluator interprets a composite measure of their various qualities.

Nevertheless, this paper has connections to several strands of the literature. It contributes to the multidimensional screening literature (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, 2025a) or (ii) that derive sufficient conditions for menus with specific characteristics to be optimal for a multiproduct monopolist (Haghpanah and Hartline, 2021; Yang, 2025b). The solution to the multiproduct monopolist's problem is famously elusive and complex. The optimal mechanism may use lotteries (Manelli and Vincent, 2006), possibly uncountably 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 propose a novel multidimensional screening problem with a remarkably simple solution. Unlike in the monopolist's screening problem, in this model, different agent types have largely aligned preferences over alternative allocations: All types prefer acceptance over rejection. An agent's type affects her preferences only through its effect on her preferences over how acceptance or rejection decisions depend on the composite measure: Agents with a higher composite measure benefit more by mechanisms that reward high composite measures with high acceptance probabilities. This is a fundamental reason why the technical issues in this setting are different and, ultimately, more tractable than those in the monopolist's multidimensional screening problem. Another difference from multidimensional monopolistic screening is that in this paper, agents have hard evidence for one dimension of their type. However, as seen through a comparison of sections 3 and 6.1.3, this feature of the model complicates rather than simplifies the principal's problem.

My analysis does not rely on ironing procedures (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 continuous functional over a convex and compact space of monotone functions. Bauer's maximum principle then implies an extreme point solves the problem.[3] 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 (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 (Townsend, 1979; Gale and Hellwig, 1985; Dunne and Loewenstein, 1995; Ben-Porath et al., 2014; Bizzotto et al., 2020; Erlanson and Kleiner, 2020; Halac and Yared, 2020; Li, 2020; Kattwinkel and Knoepfle, 2023) or the revelation of one dimension of the agent's multidimensional type (Glazer and Rubinstein, 2004; Carroll and Egorov, 2019; Li, 2021).[4] Therefore, the interpretation of the verification result is not influenced by the agent's initial disclosure as in my setting, where the substitutability between the different dimensions of the agent's type 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 and lobbying models (Holmström, 1999; Esteban and Ray, 2006). Still, in these models the main force is the agent's incentives to increase training through costly effort in order to influence the principal's learning (though costless observation of the composite measure) of the agent's talent. Here, I focus on information transmission and verification. I show that if the principal can ask for hard evidence of effort, the signal-jamming problem is mitigated if the composite measure is talent-biased. However, when the composite measure is training-biased, the signal-jamming problem persists even if the principal can ask for evidence of training. In that case, the agent has incentives to withhold evidence, which she 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 (Shin, 1994;

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

[4] The verification technology in my setting nests the case where verification reveals one of the dimensions of the agent's type. If the composite measure is constant in training and increasing in talent (and thus reveals talent exactly), the optimal mechanism is the same as under a talent-biased composite measure. If the composite measure is constant in talent and increasing in training, verification is useless, and the optimal mechanism only asks for evidence.

Dziuda, 2011; Hart et al., 2017), models with signal manipulation (Frankel and Kartik, 2019, 2022; Perez-Richet and Skreta, 2022; Jungbauer and Waldman, 2023; Ball, 2025) or costly lying (Kartik, 2009; Sobel, 2020), and college admissions and standardized testing (Krishna and Tarasov, 2016; Brotherhood et al., 2023; Dessein et al., 2025a,b).

2 The model

There are an agent (she) and a principal (he). The agent is privately informed of her bidimensional type $(e,t) \in [0,1]^2$, which has a full-support density $f : [0,1]^2 \to \mathbb{R}_{++}$.[5] No other assumption is imposed on f; any form of stochastic dependence between e and t is allowed. e is the agent's *training*. The agent has evidence equal to her training. Thus, an agent of type (e,t) can present any level of evidence $e' \in [0,e]$. By presenting evidence e' she proves that her e is at least e'. However, she cannot prove her e is not higher than e' (i.e., that she is not withholding evidence). t is the agent's *talent*, which she cannot unilaterally prove anything about.[6]

Verification. By paying a cost $c \geq 0$, the principal can observe the value of a composite measure of the agent's training 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 set of the principal, which 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. 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 versus 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.

[5]The Online Appendix studies the more general case where (e,t) lies in a hypercube.

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

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 must 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 $s \in [0,1]$.[7] 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 $P : [0,1]^2 \times ([0,1] \cup \{\emptyset\}) \to [0,1]$ 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

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

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 Characterization of the optimal mechanism

This section first shows that it is without loss to restrict attention to truthful mechanisms that accept the agent with certainty if she meets a composite measure threshold (section 3.1) and then characterizes this class of mechanisms (section 3.2). Next, it further simplifies the class of candidate mechanisms by showing it is without loss to restrict attention to (fully) deterministic mechanisms and the optimal mechanism does not overspend on verification (section 3.3). Last, it derives the optimal mechanism under free (section 3.4) or costly verification (section 3.5).

[7]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. The condition can also be interpreted as a "fairness" constraint on the mechanism.

3.1 Simplifying the class of mechanisms

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

3.1.1 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).

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

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

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.[8]

3.1.2 Mechanisms that accept the agent with certainty if she meets a composite measure threshold 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 of a truthfully-reporting agent. 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).[9]

[8]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.

[9]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 provide additional incentives

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 truthful mechanism $M' \equiv \langle T',P' \rangle$ with threshold acceptance policy is outcome-equivalent to another truthful mechanism $M \equiv \langle T,P \rangle$ with threshold acceptance policy 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, any optimal mechanism accepts the agent with probability 1 if she passes the threshold. When verification is free, it is still without loss to constrain attention to such mechanisms.

Lemma 1. Given any truthful mechanism M with threshold acceptance policy, there exists a truthful mechanism $M' \equiv \langle T',P' \rangle$ with threshold acceptance policy and $P'_{at}(e,t) = 1$ 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 $T(e,t) > 0$.[10]

Here is the intuition behind this result. 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 threshold is met). But then, 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.

3.2 Incentive-compatible mechanisms

Given what we have seen, we constrain attention to truthful mechanisms that accept the agent with certainty if she meets the composite measure threshold.

Definition 3. A mechanism $M \equiv \langle T,P \rangle$ is simply incentive-compatible (SIC) if it is truthful and

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

for 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.

[10]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.

for every $(e,t) \in [0,1]^2$.

Proposition 1 characterizes these mechanisms. 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 training 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 $\overline{e}(s) \coloneqq \max\{e \in [0,1] : \sigma(e,0) \leq s\}$.[11]

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

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

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

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

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

Figure 1(a) shows the different ways an agent can misreport her type: (1) present all evidence of training but understate talent, (2) present all evidence but overstate talent, betting on the prospect of acceptance without verification, (3) withhold evidence to overstate talent, imitating an agent with the same composite measure, (4) withhold evidence, imitating an agent with lower composite measure, and (5) withhold evidence, imitating an agent with higher composite measure, betting on the prospect of acceptance without verification. Figure 1(b) schematically summarizes conditions (i) and (ii) of Proposition 1.

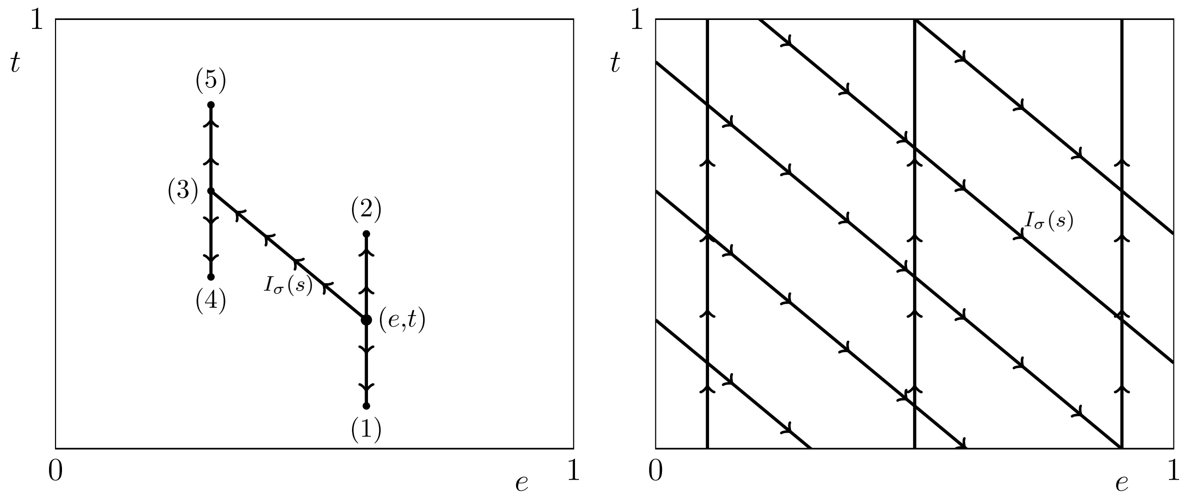
Here is why the conditions of Proposition 1 are necessary and sufficient to preclude all five kinds of deviations. First, condition (i) is necessary and sufficient to rule out deviation (1). Agent (e,t) does not want to present all her evidence but understate her talent to imitate agent (e,t') with $t' < t$, meet the composite measure threshold, and get accepted with probability $\Pi(e,t')$. Second, condition (iii) is necessary and sufficient to rule out deviation (2) by an untalented agent $(e,0)$. Agent $(e,0)$ does not want to overstate her talent, imitating an agent (e,t) and possibly getting accepted without verification. Put differently, among agents with the same level of training 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. Moreover, conditions (i) and (iii) combined rule out deviation (2) by *any* agent (e,t). Combined, they imply that $\Pi(e,t) \geq \Pi(e,0) \geq (1 - T(e,t'))P(e,t',\emptyset)$

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

Figure 1: Possible misreports and incentive compatibility

(a) Five possible ways agent (e,t) can misreport her type

(b) Directions in which $\Pi(e,t)$ is non-decreasing in SIC mechanisms



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 $\Pi(e,t)$. Third, condition (ii) is necessary and sufficient to rule out deviation (3). Agent (e,t) does not want to imitate an agent (e',t') with less training $e' < e$, more talent $t' > t$, and equal composite measure $\sigma(e',t') = \sigma(e,t)$ to get accepted with probability $\Pi(e',t')$ instead of $\Pi(e,t)$. Fourth, conditions (i) and (ii) combined rule out deviation (4). Fifth, conditions (i), (ii), and (iii) combined rule out deviation (5), since they 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.

3.3 Further simplifying the class of mechanisms

Before deriving the optimal mechanism, we show that it is deterministic and does not overspend on verification.

3.3.1 The optimal mechanism does not overspend on verification

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 training, 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 SIC mechanism $M \equiv \langle T,P \rangle$, there exists an SIC 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 SIC 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, 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.[12] Conditions (i) and (ii) of Proposition 1 are also still satisfied since Π has not changed. (e,t)'s composite measure is verified with lower but still high enough 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 SIC 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 training.

3.3.2 The optimal mechanism is deterministic

The principal's objective function is $\int_0^1 \int_0^1 [\Pi(e,t)u(e,t) - cT(e,t)]\, f(e,t)dtde$. Given Lemma 2, any Π that satisfies conditions (i) and (ii) of Proposition 1 can be optimally implemented with T and P that satisfy condition (iii) with equality. Thus, we can substitute $T(e,t) = \Pi(e,t) - \Pi(e,0)$ to write the objective function only in terms of Π:

$$\int_0^1 \int_{\underline{e}(s)}^{\overline{e}(s)} [\Pi(e,\tau(e,s))(u(e,\tau(e,s)) - c) + c\Pi(e,0)]\, f(e,\tau(e,s))deds, \tag{2}$$

where instead of integrating over e and t, we integrate over e and s. The principal's problem amounts to choosing $\Pi(e, \tau(e,s))$ 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. 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. Any extreme Π maps each (e,s) to either 0 or 1.

Lemma 3. There exists an optimal mechanism that is deterministic (i.e., with $\Pi(e,t) \in \{0,1\}$ for all (e,t)).

[12] Notice that because M is SIC, 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.

3.4 Optimal screening under free verification

We are now ready to characterize the optimal mechanism under free verification.

3.4.1 Talent-biased composite measure

Consider first the case where the composite measure is *talent-biased* in the sense that it is more sensitive to talent than talent is valuable to the principal—and conversely, less sensitive to training than training is valuable to the principal. In the linear and Cobb-Douglas specifications (see section 2), σ is talent-biased when $\gamma_\sigma < \gamma_u$. A talent-biased σ can be defined more generally as follows.[13]

Definition 4. σ is talent-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 talent-biased, it is too generous towards those with high talent and low training and too strict towards those with low talent and high training. 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) training.

Clearly, if the principal's payoff from accepting the agent is increasing along iso-composite-measure curves, σ is talent-biased. This is the case if the principal's marginal rate of substitution of talent for training (MRS) is higher (in absolute value) than the composite measure's MRS.

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 talent-biased. 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).

If the principal only values training, σ is automatically talent-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, he can accept every agent with sufficient training to be of positive value. Allowing for the principal to also value talent, Proposition 2 shows that when verification is (i) free and (ii) the composite measure is talent-biased, the principal can still achieve the full information benchmark.

[13] We define a talent-biased composite measure for any verification cost c. The optimal mechanism under costly verification is studied in section 3.5.

Proposition 2. Let $c = 0$, and assume that σ is talent-biased. Then, $\Pi(e,t) = \mathbf{I}(u(e,t) \geq 0)$ is incentive-compatible, so the principal achieves the full information first-best.[14]

Proposition 2 shows that the principal can achieve the first-best, safely disregarding the possibility that the agent will withhold evidence to manipulate the interpretation of the composite measure. The agent's incentive-compatibility constraints are not binding. Figure 3(a) presents the optimal mechanism.

3.4.2 Training-biased composite measure

Consider now the case where the composite measure is *training-biased*. In the linear and Cobb-Douglas specifications (see section 2), σ is training-biased when $\gamma_\sigma > \gamma_u$. A training-biased σ can be defined more generally as follows.

Definition 5. σ is training-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 training-biased. 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. 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 incentive-compatible, 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 training-biased? Proposition 3 describes the optimal mechanism when verification is free and the composite measure is training-biased. In the optimal mechanism, agent (e,t) is accepted if and only if $\sigma(e,t) \geq s^*$.

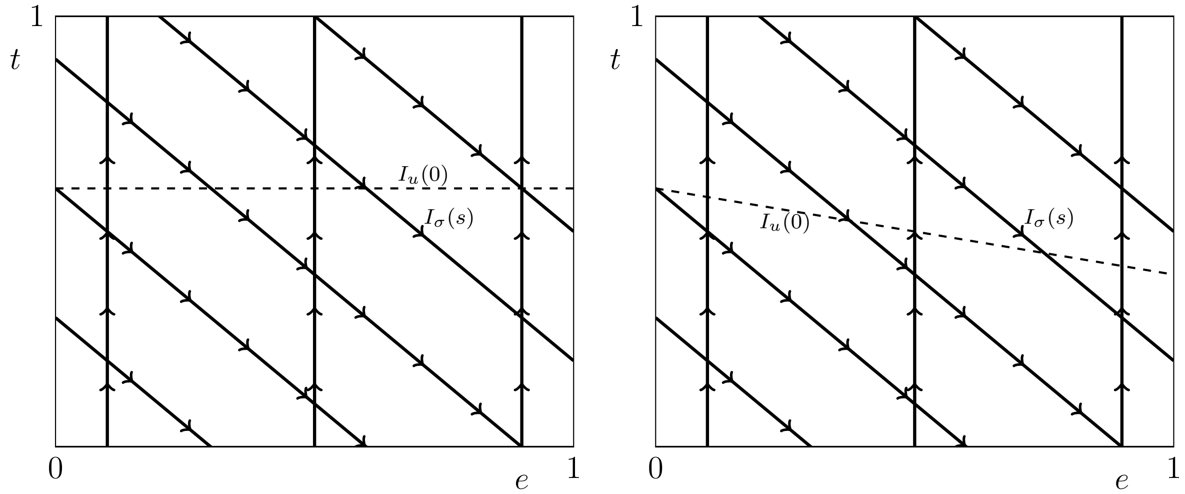
Proposition 3. Let $c = 0$, and assume that σ is training-biased. Then, there exists an optimal mechanism with $\Pi(e,t) = \mathbf{I}(\sigma(e,t) \geq s^*)$.[15]

[14] 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 $P(e,t,\emptyset) = \mathbf{I}(u(e,0) \geq 0)$. Clearly, since verification is free, $T(e,t) = \mathbf{I}(u(e,t) \geq 0)$ and $P(e,t,\emptyset) = 0$ is, for example, also optimal, as is always verifying the composite measure and accepting only the valuable agents.

[15] 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 $P(e,t,\emptyset) = \mathbf{I}(e > \overline{e}(s^*))$. Clearly, since verification is free, $T(e,t) = \mathbf{I}(\sigma(e,t) \geq s^*)$ and $P(e,t,\emptyset) = 0$ is, for example, also optimal, as is always verifying the composite measure and accepting only the agents who pass the threshold s^*.

Figure 2: *Not* achieving the first-best: training-biased composite measure

(a) Only talent valuable to the principal **(b)** Training-biased composite measure



Note: the arrowed lines represent the directions in which $\Pi(e,t)$ is non-decreasing in any SIC mechanism. The dashed lines represent the principal's indifference curve $I_u(0)$.

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 $\widetilde{f}(e,s) := f(e,\tau(e,s))$.[16] The principal effectively chooses a threshold s^* and accepts every agent with composite measure at least as high. 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 training-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 $e' > e$, incentive-compatibility 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 SIC 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) training than the threshold type are accepted (resp. rejected). Next, notice that incentive-compatibility 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 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 by condition (ii) of Proposition 1 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

[16]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 SIC.

automatically satisfied.

3.5 Optimal screening under costly verification

We now turn to characterizing the optimal mechanism under costly verification.

3.5.1 Talent-biased composite measure

Proposition 4 characterizes the optimal mechanism under a talent-biased composite measure, generalizing Proposition 2 by allowing for costly verification (i.e., $c \geq 0$).

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

Every agent with training $e \geq e^*$ is accepted without verification, while agents with training $e < e^*$ 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 optimal mechanism.

An increase in the threshold e^* would lead to: (i) increased verification costs by having 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 with accuracy in acceptance/rejection decisions (i.e., the net effect of (ii) and (iii)).

3.5.2 Training-biased composite measure

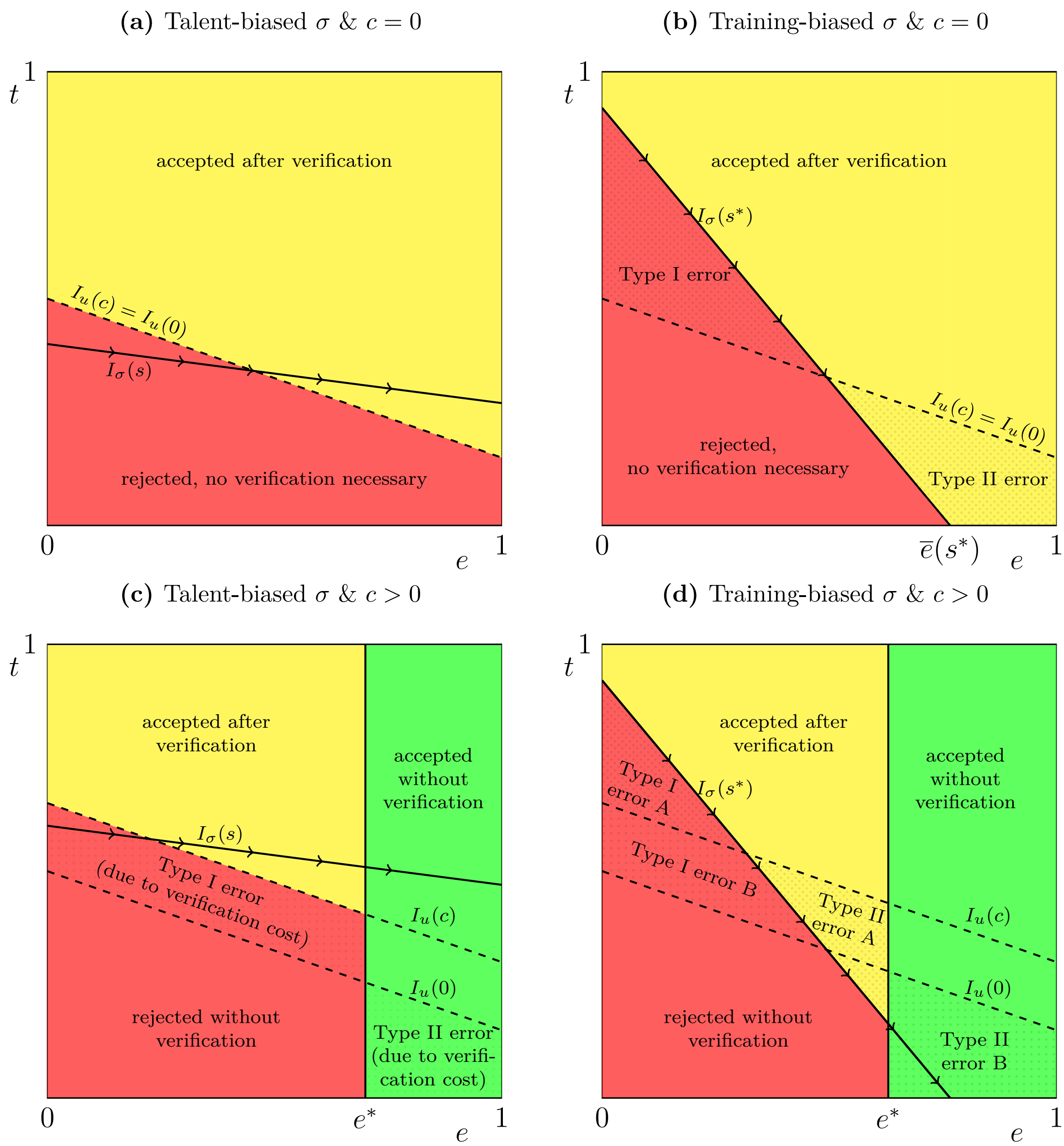
Proposition 5 characterizes the optimal mechanism under a training-biased composite measure, generalizing Proposition 3 by allowing for costly verification (i.e., $c \geq 0$).

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

Every agent with training $e \geq e^*$ is accepted without verification, while agents with training $e < e^*$ are accepted after verification if their composite measure is at least s^*. The remaining agents are rejected without verification. Figure 3(d) presents the optimal mechanism.

The principal makes four types of errors: (i) He rejects without verification some agents whom he would prefer to accept after verification (Type I error A), (ii) he rejects without verification some agents whom he would prefer to accept without verification

Figure 3: The optimal mechanism



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^*.

(Type I error B), (iii) he accepts after verification some agents whom he would prefer to reject without verification (Type II error A), and (iv) he accepts without verification some agents whom he would prefer to reject without verification (Type II error B).

In choosing s^*, he considers the trade-off between Type I error A and Type II error A.[17] An increase in the threshold e^* would lead to: (i) increased verification costs by having additional agents who lie above $I_\sigma(s^*)$ get 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^*).[18] 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 with accuracy in acceptance/rejection decisions (i.e., the net effect of (ii) and (iii)).

4 Discussion

This section discusses (i) the effects of the verification cost and bias of the composite measure on the principal's decision errors, (ii) comparative statics, (iii) the implementation of the optimal mechanism, and (iv) implications of the results for the difficulty of hiring for different types of positions.

4.1 Effects of the verification cost and bias of the composite measure on principal's decision errors

The verification cost and the bias of the composite measure in favor of training induce the principal to make errors favoring high- over low-training agents. Indeed, a comparison of Figures 3(a) and 3(c) shows that the verification cost *alone* gives rise to such errors. A comparison Figures 3(a) and 3(b) shows that the bias of the composite measure in favor of training also gives rise to such errors *even* when verification is free. But how do the verification cost an the bias of the composite measure interact interact in inducing errors favoring high- over low-training agents? Proposition 6 shows that the two forces are *complements*: The bias of the composite measure in favor of training exacerbates the errors due to the verification cost by decreasing the threshold level of evidence required for acceptance without verification, as can be seen through a comparison of Figures 3(c) and 3(d).

[17] 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 $I_u(c)$.

[18] Channel (iii) is not necessarily present, as can be seen in Figure 3(d).

Proposition 6. Take any pair of talent- and training-biased composite measures. For any optimal evidence threshold $e^*_{t-\text{biased}}$ under the talent-biased measure and any optimal evidence and composite measure thresholds ($e^*_{e-\text{biased}}$,$s^*_{e-\text{biased}}$) under the training-biased measure, (i) the evidence threshold is weakly higher under the talent-biased measure: $e^*_{t-\text{biased}} \geq e^*_{e-\text{biased}}$, or (ii) both evidence thresholds are optimal under both measures.[19]

To see why, we can examine Figures 3(c),(d). Both under a talent- and a training-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 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 talent-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 training-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 rather than to the yellow area. Therefore, an increase in e^* is more attractive to the principal under a talent-biased than under a training-biased composite measure.

Simply put, because verification does not distort the incentives of agents to present evidence under a talent-biased composite measure but it *does* distort incentives under a training-biased composite measure, verification is more effective under a talent-biased than under a training-biased composite measure.[20] Thus, fewer agents are optimally accepted without verification under a talent-biased than under a training-biased composite measure.

4.2 Comparative statics

Now I briefly comparative statics, which are covered in more detail in the Appendix.

4.2.1 Talent-biased composite measure

The following comparative statics hold for the case of a talent-biased composite measure. First, an increase in c causes the (combined) magnitude of channels (i) and (iii) discussed in section 3.5.1 to increase without affecting the magnitude of channel (ii). Thus, e^* is non-increasing in c: The more costly verification is, the more high-training agents are

[19] If $e^*_{e-\text{biased}} \in (0,1)$ and the principal's objective function under the talent-biased composite measure is single-peaked in the evidence threshold, then $e^*_{t-\text{biased}} > e^*_{e-\text{biased}}$.

[20] Absent a strong negative dependence between training and talent, evidence of training is indeed good news about the payoff from accepting the agent. Even with a strong negative dependence, absent verification, any incentive-compatible evaluation scheme that utilizes evidence should reward the candidate for presenting it. Therefore, absent verification, presenting evidence is always weakly in the agent's best interest.

accepted without verification. Second, the principal's optimal payoff is non-increasing in c. Third, since the principal's objective function is independent of σ, the optimal mechanism and payoff are the same under any two talent-biased composite measures.

4.2.2 Training-biased composite measure

The following comparative statics hold for the case of a training-biased composite measure. First, an increase in c tends to directly cause (i) e^* to decrease by enhancing the verification cost savings associated with a decrease in e^* and (ii) s^* to increase by enhancing the verification cost savings associated with an increase in s^*.[21] 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, a decrease in e^* tends to cause s^* to decrease by decreasing 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 rendering the net effect ambiguous.

Second, the principal's optimal payoff is non-increasing in c. Third, the optimal payoff is higher under less training-biased composite measures. Take any two training-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 training-biased than σ'), the principal's optimal payoff is higher under σ' than under σ.[22] Fourth, the principal's payoff will tend to increase as training and talent become more positively stochastically dependent. A strong positive stochastic dependence between e and t means that there are not many agents with high (resp. low) talent and low (resp. high) training, 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.

4.3 Implementation of the optimal mechanism

We have so far without loss restricted attention to truthful mechanisms. However, under a training-biased composite measure, the principal can implement the optimal mechanism simply by offering 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

[21]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.

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

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 talent-biased composite measure is not possible. In that case, the principal needs to ask for evidence also from agents whose composite measure he verifies. As can be seen in Figures 3(a),(c), there exist $e_\ell, e_h, t_\ell, t_m, t_h$ such that $e_\ell < e_h$, $t_\ell < t_m < t_h$, $u(e_h,t_m) > c > \max\{u(e_h,t_\ell), u(e_\ell,t_h)\}$, and $\sigma(e_h,t_m) = \sigma(e_\ell,t_h)$. The principal needs to verify (e_h,t_m)'s composite measure to accept her but not (e_h,t_ℓ). He also needs to ask (e_h,t_m) for evidence of training to accept her but not (e_ℓ,t_h).

4.4 Hiring for different types of positions

Comparing the optimal mechanism under a talent-biased composite measure with the optimal mechanism under a training-biased composite measure reveals a stark contrast in the difficulty of hiring different types of employees. When training (e.g., skills and knowledge) provable through hard evidence are most valuable, the hiring process is easy. On the other hand, when talent is most valuable, the hiring process is harder: Hiring errors arise unless the composite measure (e.g., interview or test performance) is very sensitive to talent. If it is not, the hiring process is flawed, favoring unworthy candidates with advanced training over worthy candidates with limited training.

5 Applications

This section examines the implications of the results for hiring for prestigious positions, promotions, college admissions, and academic job market hiring.

5.1 Hiring for prestigious positions

A job candidate's training e is her resume quality. t is her ability and drive not captured by e. An employer wants to decide whether to hire the candidate for a prestigious position. Verification works as follows: The employer has the option to let another employer hire the candidate in some less prestigious position, observe her performance in that position (i.e., the composite measure), and decide whether to poach her. Poaching is costly because it is more expensive to poach the candidate than hire her from the beginning.

In the optimal mechanism, candidates with strong credentials—including unworthy ones—are immediately hired for prestigious positions. On the other hand, talented candidates with weak credentials have to first work in less prestigious positions to prove their worth before landing a prestigious position. Also, if the candidates' performance is more sensitive to talent in the more prestigious position than in the less prestigious one, worthy candidates with low credentials are at a disadvantage also in the poaching stage.

5.2 Promotions

An employee's e is her hardworkingness. t is her efficiency, talent, and managerial skills. $\sigma(e,t)$ is her productivity in her current position. The employee can provide or withhold evidence on e by, for example, choosing how many hours to work in the office and how many to work from home. The employer can verify the employee's productivity $\sigma(e,t)$. If the employee continues to work in her current position, the employer's payoff is $\sigma(e,t)$. If she is promoted, the employer's payoff is $v(e,t)$. The employer's problem is equivalent to the one in section 2 with $u(e,t) \coloneqq v(e,t) - \sigma(e,t)$, as long as the difference $v(e,t) - \sigma(e,t)$ is non-decreasing in e and t.[23] This condition has a natural interpretation: The higher position comes with increased responsibilities that allow the employee's talent and hardworkingness to have a larger impact. The composite measure is training-biased if for every (e,t),

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

This condition also has a natural interpretation: For the employee's productivity in the higher position, the relative importance of talent (relative to hardworkingness) is higher than in the current position. The composite measure is talent-biased if the inequality is reversed.

5.3 College admissions and standardized testing

A college applicant's e is her prior training, preparation, parents' education and professions, and parental support. t is her talent and drive 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.[24]

In the optimal mechanism, if the standardized test is training-biased, admission decisions are flawed at the expense of students with low training, preparation, and parental support to the extent that applicants can withhold evidence of those. Particularly, if colleges only value talent and, thus, try to control for the applicants' unequal backgrounds, the above problem is necessarily present.

[23] $u(e,t)$ could also be defined as $u(e,t) \coloneqq v(e,t) - \sigma(e,t) - \underline{q}$, where $\underline{q}$ is the threshold productivity differential for the promotion to be beneficial to the firm. For example, $\underline{q}$ could be (i) the productivity differential of another employee who could be promoted instead, (ii) the salary raise associated with the promotion, or (iii) the surplus from hiring someone from outside the firm.

[24] In this setting, $c = 0$ and the college does not condition the requirement to submit a test score on the candidate's report. However, the optimal mechanism takes the same form as in the setting we have studied for $c = 0$, where there is no need for the principal to accept some agents who present evidence without verification.

5.4 Academic job market talks

An academic job market candidate's research topic is comprised of a "mass" $b > 1$ of (uncountably infinitely many) problems.[25] $e \in [0,1]$ is the candidate's knowledge, the mass of problems she has solved. 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 sends a cheap talk message about t, the hiring committee may verify the proportion of questions she can answer. Verification amounts to posing to the candidate infinitely many problems randomly sampled from the mass of problems that the candidate has not initially disclosed answers to.[26] 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 the candidate already has answers to (but has not presented) and (ii) the proportion $(b - e)/(b - e')$ of problems sampled from the set of problems 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 we have studied.

6 Extensions and robustness

This section discusses optimal screening under alternative evidence structures (section 6.1), costly composite measure design (section 6.2), endogenous training (section 6.3), and the constraint that $P(e,t,s)$ be non-decreasing in s (section 6.4). The Online Appendix generalizes the characterization of the optimal mechanism to the case of m dimensions of training and n dimensions of talent.

6.1 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).

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

[26] These can be countably infinitely many problems or a mass of (uncountably many) problems smaller than $b - 1$. 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.

6.1.1 Optimal screening when the agent cannot withhold evidence

Assume that training is observed by the principal.[27] Now, 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 two options: (i) accept without verification every agent with training e or (ii) accept an agent with training e after verification if $u(e,t) \geq c$ and reject an agent with training e without verification if $u(e,t) < c$. Denote by $\mathbb{E}_t\left[\min\{u(e,t), c\}|\, e\right]$ the expectation of $\min\{u(e,t), c\}$ conditional on e. 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 superior. Clearly, when $c = 0$, option (ii) is superior. Proposition 7 describes the optimal mechanism.

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

(i) if $\mathbb{E}_t\left[\min\{u(e,t), c\}|\, e\right] > 0$, then every agent with training e is accepted without verification, and

(ii) if $\mathbb{E}_t\left[\min\{u(e,t), c\}|\, e\right] \leq 0$, then an agent with training e is (a) accepted after verification if $u(e,t) \geq c$ and (b) rejected without verification if $u(e,t) < c$.

The optimal mechanism does not depend on σ. Part (ii) of Proposition 7 implies that under a training-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 training) to get accepted, then the composite measure threshold that (e_1,t_1) needs to meet is lower than the composite measure threshold that (e_2,t_2) needs to meet.[28] 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.

Combined with the analysis of the baseline model, these results have implications for affirmative action in college admissions. The baseline model has shown that if (i) college applicants can to a large extent withhold evidence of training and (ii) standardized tests are training-biased, every applicant has to achieve the same test score to get admitted, and affirmative action is not effective. On the other hand, if condition (ii) fails (i.e., standardized tests are talent-biased, that is, adequately sensitive to talent), then college admissions are no longer flawed in favor of applicants from advantaged backgrounds. The results of this section show that if condition (i) fails (i.e., college applicants *cannot* withhold evidence), then college admissions are not flawed—even if standardized tests are training-biased. In that case, applicants from disadvantaged backgrounds will face lower

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

[28] To see this, notice that under a training-biased composite measure, if $u(e_1,t_1) = u(e_2,t_2) = c$ and $e_2 > e_1$, then $\sigma(e_1,t_1) < \sigma(e_2,t_2)$.

test score cutoffs. We conclude that if either of the two conditions fails, affirmative action is effective; in that case, imposing constraints on it would compromise colleges' ability to screen for talent.

Comparison with baseline model. Assume now that $e \mapsto \mathbb{E}_t\left[\min\{u(e,t), c\} \mid e\right]$ does not cross zero from above. This means that if for a certain level of training e, the principal prefers to accept every agent with training e without verification rather than pay verification costs to reject some unworthy agents with training e, then for any higher level of training $e' > e$, the principal still prefers to accept every agent with training e' without verification.[29] 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 talent-biased (see Proposition 4).

Corollary 7.1. Assume that there exists e^* such that $\operatorname{sgn}\{\mathbb{E}_t\left[\min\{u(e,t), c\} \mid e\right]\} = \operatorname{sgn}\{e - \widehat{e}\}$. In the optimal mechanism, $\Pi(e,t) = \mathbf{I}(u(e,t) \geq c \text{ or } e \geq e^*)$, $T(e,t) = \mathbf{I}(u(e,t) \geq c \text{ and } e < e^*)$, and $P(e,t,\emptyset) = \mathbf{I}(e \geq e^*)$.

This means that as long as the composite measure is talent-biased, the ability of the agent to withhold evidence does not constrain the principal's ability to screen—unless the principal switches from case (i) for some training level e to case (ii) for some training level $e' > e$ in Proposition 7, which is not possible if the agent can withhold evidence.

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

Assume that the agent can also present evidence on talent. That is, agent (e,t) can report any $(e',t') \leq (e,t)$. Then, the principal can achieve the full information first-best without verification, inducing every agent to present all her evidence on both e and t.[30] Given also the results of the baseline model, we conclude that when the agent cannot provide evidence of talent and the principal can only imperfectly verify it through a training-biased composite measure, he is constrained in his evaluation of the agent by the agent's incentives to withhold evidence of training. The constraint vanishes if the agent can also provide evidence of talent or if the composite measure is talent-biased.

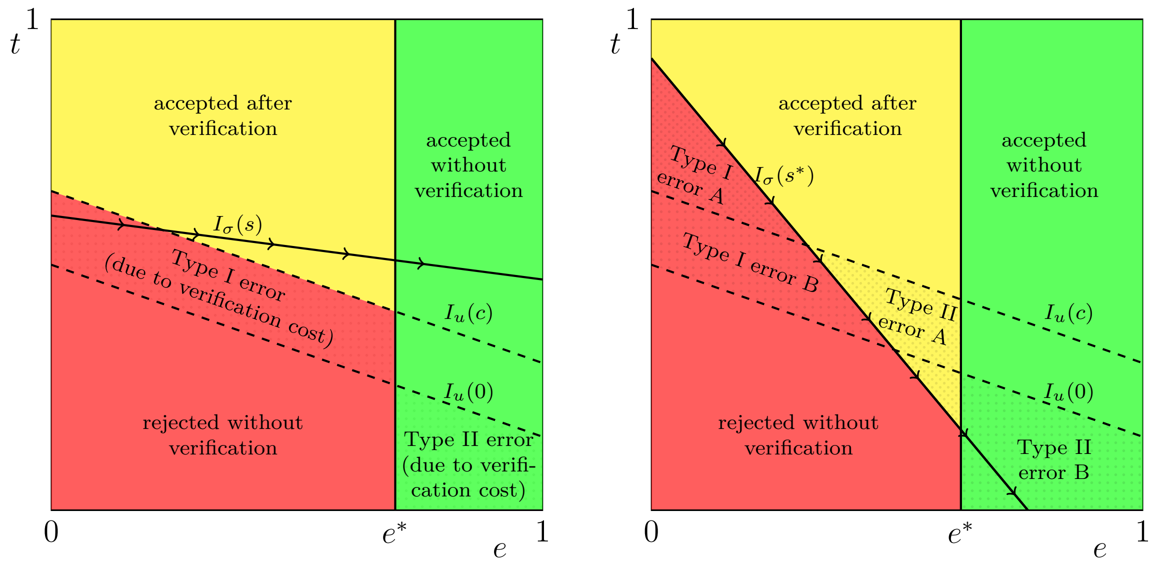
The results suggest that hiding one's effort might be more common among younger people. If a person's talent is revealed as she advances through successive stages of evaluation in her career, then senior professionals should have weaker incentives to hide their effort compared to students and early-career professionals. Indeed, university students

[29] A sufficient condition is that $\mathbb{E}_t\left[\min\{u(e,t), c\} \mid e\right]$ be non-decreasing in e, which is satisfied as long as t does not stochastically depend on e "too negatively." For example, it is sufficient that for any $e' > e$, the distribution of t conditional on e' first-order stochastically dominates the distribution of t conditional on e.

[30] The conclusion is the same if t is observed at no cost by the principal and the agent can present evidence on e.

Figure 4: The optimal mechanism with private or public evidence

(a) Talent-biased σ or publicly observed e **(b)** Training-biased σ and privately observed e



Note: It is assumed that $e \mapsto \mathbb{E}_t\left[\min\{u(e,t),c\} \,|\, e\right]$ does not cross zero from above.

have a desire to project "effortless perfection" by deliberately hiding how hard they study (Travers et al., 2015; Casale et al., 2016). At the same time, hiding effort has been identified as a unique expression of perfectionistic self-presentation (Flett et al., 2016), and narcissism, which is closely linked to perfectionistic self-presentation, decreases with age (Weidmann et al., 2023; Orth et al., 2024).

6.1.3 Optimal screening when the agent cannot present evidence

Assume that the agent can present evidence on neither training nor talent. That is, agent (e,t) can report any $(e',t') \in [0,1]^2$. We can still restrict attention to SIC mechanisms, which Proposition 8 characterizes.

Proposition 8. A mechanism $M \equiv \langle T,P \rangle$ is SIC 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 the agent has evidence on e. 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 *lowest* training does not have incentives to overstate her talent or training 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 training to potentially get accepted without verification. Thus, the need for verification is enhanced due to the agents inability to present evidence on e. Last, condition (ii) ensures that an agent (e,t) does not want to imitate an agent (e',t') with the same 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 training. This nullifies the advantage that agents with high training have (over agents with the same composite measure but lower training) when they can present evidence.

Lemma 4 shows that we can constrain attention to mechanisms that satisfy condition (iii) of Proposition 8 with equality. Thus, the probability of acceptance without verification is the same for every agent type.

Lemma 4. Given any SIC mechanism $M \equiv \langle T,P \rangle$, there exists an SIC 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 4, $\Pi(e,t) = \Pi(0,0) + T(e,t)$. Thus, the principal's objective function can be written as

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

which is linear in Π, so by Bauer's maximum principle, there exists an extreme Π (among all $\Pi(e, \tau(e,s))$ 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 of training limits the set of SIC mechanisms, thereby decreasing the principal's optimal payoff.[31] Also, the principal now has to choose

[31] In more detail, when the composite measure is training-biased, if no $e^* \in \{0,1\}$ is optimal when the agent can present evidence (see Proposition 5), then the principal's payoff is lower when the agent cannot present evidence. When the composite measure is talent-biased, if $e^* = 0$ is not optimal when the agent can present evidence (see Proposition 4), then the principal's payoff is lower when the agent cannot present evidence.

s^* trading-off Type I and Type II errors even when σ is talent-biased. Talent-biased composite measures are not inherently better than training-biased ones when the agent cannot present evidence. Unlike in the baseline model—where all talent-biased composite measures are equally effective (see section 4.2)—when the agent cannot present evidence, the more closely the composite measure aligns with the principal's preferences, the higher the principal's optimal payoff is, regardless of whether the composite measure is talent- or training-biased.

6.2 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 5.1), the employee's production function in the less prestigious position is not chosen by the employer hiring for the prestigious position. In promotion decisions (section 5.2), the employee's production function in the current position depends on her current job description and responsibilities, which may mostly reflect the firm's operating needs rather than support the employer's promotion decisions. In college admissions (section 5.3), a college usually cannot choose the content of the standardized test.

However, in some cases (e.g., hiring decisions where verification amounts to tests and interviews), the principal may be able to choose how the composite measure depends on the agent's type. 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. 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 $\sigma^* \in \Sigma$ from the set Σ of conceivable composite measure functions. The solution to the first step is the one we have already described.[32]

As shown in section 4.2, as long as the composite measure is training-biased, there are gains from making it more sensitive to talent. On the other hand, all talent-biased composite measures are as effective as a composite measure that is exactly aligned with the principal's preferences. Therefore, if composite measures more sensitive to talent are more expensive to design, the principal will want to make the composite measure at most as sensitive to talent as his preferences are. However, as shown in section 6.1.3,

[32] That is, assuming that Σ contains only talent- and training-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 the verification cost c is the same for all composite measures.

when agents cannot present evidence, the principal always gains from finely calibrating the composite measure to make it align with his preferences—regardless of whether the composite measure is talent- or training-biased.

6.3 Endogenous training

Our characterization of the optimal mechanism applies even if training is endogenous, as long as the principal cannot influence the agent's training choice by committing to a mechanism. Indeed, in hiring decisions (section 5.1), an individual employer has limited labor market power to influence a candidate's effort to obtain credentials. Similarly, in college admissions (section 5.3), a single college has little influence over applicants' preparation for standardized tests.[33]

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 obtain training.[34] The cost of effort x is $C_t(x)$. Training 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^*; that is, (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 "training-taking" principals chooses thresholds (e^*,s^*).

While a detailed analysis of endogenous training 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 if 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, training and talent may be stochastically dependent in complicated ways.

6.4 When $P(e,t,s)$ may decrease with s

To see how the principal may be able to do better if he is not required to reward higher composite measures with weakly higher acceptance probabilities, consider the case where verification is free and the composite measure is training-biased. $\Pi(e,t) =$

[33] In other settings, the principal may be able to affect the agent's training by committing to a mechanism *before* the agent obtains training. For example, in promotion decisions (section 5.2), the employer may 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 (i) 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 and (ii) on whether there are better tools (e.g., performance-based bonuses) for incentivizing the employee to exert effort.

[34] Under a talent-biased composite measure, there is only an evidence threshold.

$\mathbf{I}(\mathbb{E}[u(x,y)|\sigma(x,y) = \sigma(e,t)] \geq 0)$ is incentive-compatible and outperforms the optimal mechanism $\Pi(e,t) = \mathbf{I}(\sigma(e,t) \geq s^*)$ of the baseline model if $\mathbb{E}[u(x,y)|\sigma(x,y) = s]) < 0$ for a positive measure of $s > s^*$. When this is the case, the principal's payoff is not single-peaked in the composite measure threshold.

7 Conclusion

Candidates for a position are often evaluated through costly tests or indices that measure a combination of multiple underlying attributes. For example, pre-employment tests, such as those administered during quantitative finance job interviews measure a combination of talent and preparation. The same is true for standardized college admission tests. A productivity metric of an employee under consideration for a promotion depends not only on "how hard" but also on "how smart" the employee works. An academic job market candidate's effectiveness in responding to questions depends both on how much she has worked on her paper and on how fast she can think on her feet.

The following problem may then arise when the evaluation of a candidate relies on a composite measure: Even if the candidate has hard evidence on one of her valuable qualities (call this quality *training*), she may strategically withhold that evidence to make the evaluator attribute the composite measure to another quality instead (call this quality *talent*). This can happen even though the evaluator values training, which means evidence of training is otherwise favorable to the candidate.

This paper has shown that such perverse incentives arise when the composite measure is training-biased (i.e., over-sensitive to training and under-sensitive to talent compared to how much the evaluator values each in a candidate). In that case, the optimal mechanism makes errors favoring high- over low-training candidates. The errors are due to two forces: (i) the strategic incentives of candidates to withhold evidence of training when the evaluator looks at their composite measure and (ii) the evaluator's incentive to save on evaluation costs by accepting high-training candidates without incurring the cost of gauging their composite measure. On top of this, the two forces are complements in inducing errors favoring high- over low-training candidates: The second force compromises the informativeness of the composite measure, exacerbating the extent to which the evaluator accepts high-training candidates without gauging their composite measure.

The results illustrate how inequalities can be perpetuated. When standardized tests are training-biased and college applicants have considerable room to withhold evidence of training and parental support, affirmative action (i.e., screening for talent taking into account college applicants' unequal backgrounds) has limited effectiveness: College applicants from advantaged backgrounds are favored over better applicants from modest backgrounds. Upon graduation, prestigious employers immediately hire Ivy-Leaguers, thereby avoiding potentially having to poach them later at a higher cost. Graduates of less prestigious institutions need to go through less prestigious employers to prove their worth before landing a prestigious position. They are still at a disadvantage when trying to transition to a more prestigious position if their performance is more sensitive to talent in the more prestigious position (if poached) than in the less prestigious one.

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.

Brotherhood, L., Herskovic, B., and Ramos, J. (2023). Income-based affirmative action in college admissions. *The Economic Journal*, 133(653):1810–1845.

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.

Dessein, W., Frankel, A., and Kartik, N. (2025a). Test-optional admissions. *American Economic Review*, 115(9):3130–70.

Dessein, W., Frankel, A., and Kartik, N. (2025b). The test-optional puzzle.

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.

Esteban, J. and Ray, D. (2006). Inequality, lobbying, and resource allocation. *American Economic Review*, 96(1):257–279.

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.

Jungbauer, T. and Waldman, M. (2023). Self-reported signaling. *American Economic Journal: Microeconomics*, 15(3):78–117.

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.

Krishna, K. and Tarasov, A. (2016). Affirmative action: One size does not fit all. *American Economic Journal: Microeconomics*, 8(2):215–252.

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.

Orth, U., Krauss, S., and Back, M. D. (2024). Development of narcissism across the life span: A meta-analytic review of longitudinal studies. *Psychological Bulletin*, 150(6):643.

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.

Weidmann, R., Chopik, W. J., Ackerman, R. A., Allroggen, M., Bianchi, E. C., Brecheen, C., Campbell, W. K., Gerlach, T. M., Geukes, K., Grijalva, E., et al. (2023). Age and gender differences in narcissism: A comprehensive study across eight measures and over 250,000 participants. *Journal of Personality and Social Psychology*, 124(6):1277.

Yang, F. (2025a). Costly multidimensional screening. *Forthcoming at Review of Economic Studies.*

Yang, F. (2025b). Nested bundling. *American Economic Review*, 115(9):2970–3013.

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.

Appendix

A Characterization of optimal thresholds

This section discusses how the principal chooses (i) the evidence threshold under a talent-biased composite measure and (ii) the evidence and composite measure thresholds under a training-biased composite measure.

A.1 Talent-biased composite measure

The principal chooses threshold $e^* \in \arg\max_{e_{min}\in[0,1]} v^{t-\text{biased}}(e_{min})$,[35] where

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

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

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

or equivalently

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

We now briefly discuss some comparative statics. Notice that the cross-partial derivative $\partial v_e^{t-\text{biased}}/\partial c$ of the objective function is non-positive, so by Topkis' monotonicity theorem, the set of optimal evidence thresholds is decreasing in c in the strong set order.

Stronger results than those discussed in the text can be derived under some additional conditions. 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.[36] Particularly, $\partial v_e^{t-\text{biased}}(e)/\partial c = -\int_0^1 \mathbf{I}(u(e^*,t)\geq c)f(e^*,t)dt < 0$, and by the Implicit Function Theorem $de^*/dc = -\partial v_e^{t-\text{biased}}(e)/\partial c|_{e=e^*}/v_{ee}^{t-\text{biased}}(e^*) < 0$, so e^* is decreasing in c. Second, the principal's optimal payoff is decreasing in c.

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

[36]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 training, in which case he does not use verification.

A.2 Training-biased composite measure

The principal chooses thresholds $(e^*,s^*) \in \arg\max_{(e_{min},s_{min})\in[0,1]^2} v^{e-\text{biased}}(e_{min},s_{min})$, where

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

where $\tilde{u}(e,s) \equiv u(e,\tau(e,s))$ and $\tilde{f}(e,s) \equiv f(e,\tau(e,s))$.[37]

When $\underline{e}(s^*) < e^*$ (i.e., some agents are accepted after verification) and $e^*,s^* \in (0,1)$,[38] the first-order conditions with respect to e_{min} and s_{min} are, respectively,

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

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

We now describe comparative statics of e^* and s^* with respect to c. 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^*), which is by assumption negative definite. Particularly, $v_{ee}^{e-\text{biased}}(e^*,s^*)$, $v_{ss}^{e-\text{biased}}(e^*,s^*) < 0$, and $\det(J(e^*,s^*)) > 0$. Also, $v_{es}^{e-\text{biased}}(e^*,s^*) = v_{se}^{e-\text{biased}}(e^*,s^*) = -\left(\tilde{u}(e^*,s^*)-c\right)\tilde{f}(e^*,s^*) > 0$. The total derivatives of $e^*$and s^* with respect to c are:

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

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

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

where $v_{ec}^{e-\text{biased}}(e^*,s^*) = -\int_{s^*}^1 \tilde{f}(e^*,s)ds < 0$ and $v_{sc}^{e-\text{biased}}(e^*,s^*) = \int_{\underline{e}(s^*)}^{e^*} \tilde{f}(e,s^*)de > 0$ are the partial derivatives of $v_e^{e-\text{biased}}$ and $v_s^{e-\text{biased}}$ with respect to c.

B Proofs

Proof of Lemma 1 Take a truthful mechanism $M \equiv \langle T,P \rangle$ with $P(e,t,s)$ given by 1 for some P_{at}. 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).[39] 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 truthful. (c) means that M' is outcome-equivalent to M. Last, 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') to say that $(1 - T(e,t))P(e,t,\emptyset) \leq \Pi(e',0)$ for every e,t,e' with $e \leq e'$, which is stronger than condition (iii).

Step 1: I first show that condition (i) is necessary for incentive-compatibility 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, incentive-compatibility 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 incentive-compatibility by showing the contrapositive. Assume that for some e,e',t with $e' \geq e$, $(1-T(e,t))P(e,t,\emptyset) > \Pi(e',0)$. Then, incentive-compatibility 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 non-decreasing in r over $r \in [\underline{e}(\sigma(e,t)),e]$ for every (e,t) is necessary and sufficient for SIC. Incentive-compatibility of type (e,t) is satisfied if and only if

$$\max_{(e',t') \leq (e,1)} \left[(1 - T(e',t'))P(e',t';\emptyset) + T(e',t')\mathbf{I}\left(\sigma(e,t) \geq \sigma(e',t')\right)\right] = \Pi(e,t). \tag{4}$$

Assume that conditions (i) and (iii') are satisfied. Then, $\Pi(e,t) \geq \Pi(e,0) \geq (1 -$

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

$T(e',t'))P(e',t',\emptyset)$ for any (e',t') with $e' \leq e$. Therefore, (4) is equivalent to

$$\max_{(e',t')\in\{(x,y)\in[0,1]^2 : x\leq e \text{ and } \sigma(e,t)\geq\sigma(x,y)\}} [(1-T(e',t'))P(e',t';\emptyset)+T(e',t')] = \Pi(e,t). \tag{5}$$

Given that $\Pi(e,t)$ is non-decreasing in t (condition (i)), (5) 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,

$$e \in \operatorname*{arg\,max}_{r\in[\underline{e}(\sigma(e,t)),e]} \Pi(r,\tau(r,\sigma(e,t))). \tag{6}$$

Thus, incentive-compatibility is satisfied for every type if and only if for every (e,t), (6) 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 (6) to hold for every (e,t) is immediate. I show necessity by showing the contrapositive. Assume that 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,

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

since $(r_2,\tau(r_2,\sigma(e,t)))$ 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) 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),[40] the second from condition (i), and the third from condition (iii). **Q.E.D.**

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

Proof of Lemma 2 Take any SIC mechanism $M \equiv \langle T,P \rangle$. Then, construct the mechanism $M' \coloneqq \langle T',P' \rangle$ with[41]

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

for every (e,t), where the inequalities follow from $\Pi(e,0) \geq (1 - T(e,t))P(e,t,\emptyset)$, which is implied by condition (iii) of Proposition 1. 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. 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 SIC. Last, 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 Lemma 3 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 (2) 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) in its domain (see Theorem 40.1 in Choquet, 1954). **Q.E.D.**

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

Condition (i): 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, for any (e,t), $\Pi(e,t) = 1$ implies that $u(e,t) > 0$, which in turn implies $u(e,t') > 0$ for every $t' \geq t$ given that $u(e,t)$ is non-decreasing in t.

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

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

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

$e < e_s$, and since $\sigma(e,\tau(e,s)) = s$, σ being talent-biased implies that $u(e,\tau(e,s)) \leq 0$, a contradiction to $\Pi(e,\tau(e,s)) = \mathbf{I}\left(u(e,\tau(e,s)) > 0\right) = 1$. **Q.E.D.**

Proof of Proposition 3 *Step 1:* In definition 5 of a training-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 may not be 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 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 $e_{\underline{s}} \leq e_{\overline{s}}$.

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 $e_{\underline{s}} \leq e_{\overline{s}}$. Similarly, if $\overline{e}(\underline{s}) \leq e_{\overline{s}}$, then $e_{\underline{s}} \leq \overline{e}(\underline{s}) \leq e_{\overline{s}}$, so $e_{\underline{s}} \leq e_{\overline{s}}$.

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 training-biased and $e^* < e_{\underline{s}}$, it follows that $u(e^*,t^*) > 0$. Similarly, since (i) σ is training-biased, (ii) $e^* > e_{\overline{s}}$, and (iii) $e^* \in [\underline{e}(\overline{s}), \overline{e}(\overline{s})]$, 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 mechanism in three parts. First, we fix an arbitrary "partial" SIC mechanism $s \mapsto \Pi(e_s,t_s)$ for every $s \in [0,1]$. Then, we complete that partial SIC mechanism (i.e., we assign a value to $\Pi(e,t)$ for every (e,t) for which $\Pi(e,t)$ has not been assigned a value in the first step), so that the complete mechanism is SIC 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 SIC mechanism.[42] Given that e_s is non-decreasing in s, by Proposition 1, the values of $\Pi(e_s,t_s)$ are part of some SIC mechanism only if $\Pi(e_s,t_s)$ is non-decreasing in s. Therefore, by Proposition 3, there exists an optimal mechanism with $\Pi(e_s,t_s) = \mathbf{I}(s \geq s^*)$ for some $s^* \in [0,1]$.

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

Step 4: It follows then that for the complete mechanism to be SIC, it must be that (i) $\Pi(e,t) = 1$ for every (e,t) such that $e > e_s$ and $\sigma(e,t) \geq s^*$ and (ii) $\Pi(e,t) = 0$ for every (e,t) such that $e < e_s$ and $\sigma(e,t) < s^*$. 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 incentive-compatibility constraint, it is optimal to set (i) $\Pi(e,t) = 1$ for every (e,t) such that $e < e_s$ and $\sigma(e,t) \geq s^*$ and (ii) $\Pi(e,t) = 0$ for every (e,t) such that $e > e_s$ and $\sigma(e,t) < s^*$. **Q.E.D.**

Proof of Proposition 4 By conditions (i) and (ii) of Proposition 1, any SIC mechanism has $\Pi(e,0)$ non-decreasing in e. Thus, given Lemma 3, there exists an optimal mechanism with $\Pi(e,0) = 1$ for every $e \geq e^*$ for some $e^* \in [0,1]$. The objective function (2) then becomes

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

The mechanism affects the second term only through e^*. Given e^*, setting $\Pi(e,t) = \mathbf{I}\left(u(e,t) \geq c \text{ or } e \geq e^*\right)$ maximizes the first term and—given that σ is talent-biased—makes the mechanism SIC, since it satisfies conditions (i) and (ii) of Proposition 1. T and P are backed out from Lemma 3. **Q.E.D.**

Proof of Proposition 5 By conditions (i) and (ii) of Proposition 1, any SIC mechanism has $\Pi(e,0)$ non-decreasing in e. Thus, given Lemma 3, there exists an optimal mechanism with $\Pi(e,0) = 1$ for every $e \geq e^*$ for some $e^* \in [0,1]$. The objective function (2) then becomes

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

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

Proof of Proposition 6 We will use the objective function defined in section A. First, notice the following relationship between $v_e^{t-\text{biased}}(e_{min})$ and $v_e^{e-\text{biased}}(e_{min},s_{min})$:

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

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

$$v(e_{min},\alpha) :=\alpha v^{t-\text{biased}}(e_{min}) + (1-\alpha)v^{e-\text{biased}}(e_{min},s^*).$$

Take any $e^*_{t-\text{biased}} \in \arg\max_{e_{min}} v(e_{min},1) = \arg\max_{e_{min}} v^{t-\text{biased}}(e_{min})$ and $e^*_{e-\text{biased}} \in \arg\max_{e_{min}} v(e_{min},0) = \arg\max_{e_{min}} v^{e-\text{biased}}(e_{min},s^*)$. Given the relationship between $v_e^{t-\text{biased}}(e_{min})$ and $v_e^{e-\text{biased}}(e_{min},s_{min})$, we have that

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

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) $e^*_{t-\text{biased}} \geq e^*_{e-\text{biased}}$ or (ii) $e^*_{t-\text{biased}} \in \arg\max_{e_{min}} v^{e-\text{biased}}(e_{min},s^*)$ and $e^*_{e-\text{biased}} \in v^{t-\text{biased}}(e_{min})$. We conclude that for any $e^*_{t-\text{biased}} \in \arg\max_{e_{min}} v^{t-\text{biased}}(e_{min})$ and $(e^*_{e-\text{biased}},s^*_{e-\text{biased}}) \in \arg\max_{(e_{min},s_{min})} v^{e-\text{biased}}(e_{min},s_{min})$, (i) $e^*_{t-\text{biased}} \geq e^*_{e-\text{biased}}$ or (ii) $e^*_{t-\text{biased}} \in \arg\max_{e_{min}} v^{e-\text{biased}}(e_{min},s^*_{e-\text{biased}})$ and $e^*_{e-\text{biased}} \in v^{t-\text{biased}}(e_{min})$. Notice also that if $e^*_{e-\text{biased}} \in (0,1)$, then $v_e^{t-\text{biased}}(e^*_{e-\text{biased}}) > 0$, and thus, if also $v^{t-\text{biased}}(e_{min})$ is single-peaked in e_{min}, the inequality is strict: $e^*_{t-\text{biased}} > e^*_{e-\text{biased}}$. **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 training 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 training e who gives payoff at least c.

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

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 incentive-compatibility 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 incentive-compatibility by showing the contrapositive. Assume that for some e,e',t, $(1-T(e,t))P(e,t,\emptyset) > \Pi(e',0)$. Then, incentive-compatibility 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 incentive-compatibility. Incentive-compatibility of type (e,t) is satisfied if and only if

$$\max_{(e',t')\leq(1,1)} \left[(1-T(e',t'))P(e',t';\emptyset) + T(e',t')\mathbf{I}\left(\sigma(e,t) \geq \sigma(e',t')\right)\right] = \Pi(e,t). \tag{7}$$

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, (7) is equivalent to

$$\max_{(e',t')\in\{(x,y)\in[0,1]^2:\sigma(e,t)\geq\sigma(x,y)\}} \left[(1-T(e',t'))P(e',t';\emptyset) + T(e',t')\right] = \Pi(e,t). \tag{8}$$

Given that $\Pi(e,t)$ is non-decreasing in t (condition (i)), (8) 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,

$$e \in \operatorname*{arg\,max}_{r\in[\underline{e}(\sigma(e,t)),\overline{e}(\sigma(e,t))]} \Pi(r,\tau(r,\sigma(e,t))). \tag{9}$$

Thus, incentive-compatibility is satisfied for every type if and only if for every (e,t), (9) 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).

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

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

Proof of Lemma 4 Take any SIC mechanism $M \equiv \langle T,P\rangle$. Condition (iii) of Proposition 1 says that $\Pi(0,0) \geq (1-T(e,t))P(e,t,\emptyset)$ for any (e,t). Then, construct the mechanism $M' \coloneqq \langle T',P'\rangle$ with[43]

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

for every (e,t), where the inequalities follow from $\Pi(0,0) \geq (1-T(e,t))P(e,t,\emptyset)$. By construction, $\Pi'(e,t) = \Pi(e,t)$ for every (e,t), so M' satisfies conditions (i) and (ii) of Proposition 8. Also, 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 SIC. Last, 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.**

Proof of Proposition 9 Straightforward and thus omitted.

[43] 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 of strategic agents

Orestis Vravosinos

C $(m+n)$-dimensional screening of strategic agents

We now generalize the results allowing for multiple dimensions of training and talent. Let the agent's type be $(e_1,e_2,\ldots,e_m,t_1,t_2,\ldots,t_n)$ with full-support density $f:[0,1]^{m+n}\to\mathbb{R}_{++}$. $(e_1,e_2,\ldots,e_m)$ are different dimensions of training and $(t_1,t_2,\ldots,t_m)$ are different dimensions of talent. The agent can present any combination of evidence $\boldsymbol{e}'\in[\mathbf{0},\boldsymbol{e}]$. 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 in the bidimensional type case that truthful mechanisms that accept the agent with certainty if she meets the appropriate composite measure threshold are without loss.

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

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

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

Lemma 6. Given any SIC mechanism M, there exists an SIC 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 5, Π is the same for every agent in $X(\boldsymbol{e},s)$, so if $\sigma(\boldsymbol{e},\boldsymbol{t})\geq s$, then agent $(\boldsymbol{e},\boldsymbol{t})$'s payoff from imitating an agent in $X(\boldsymbol{e}',s)$ does not depend on which particular agent she chooses to imitate. If, on the other hand, $\sigma(\boldsymbol{e},\boldsymbol{t}) < s$, agent $(\boldsymbol{e},\boldsymbol{t})$'s payoff from imitating an agent $(\boldsymbol{e}',\boldsymbol{t}') \in X(\boldsymbol{e}',s)$ is increasing (resp. decreasing) in $P(\boldsymbol{e}',\boldsymbol{t}',\emptyset)$ (resp. $T(\boldsymbol{e}',\boldsymbol{t}')$). 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 $T(\boldsymbol{e}',\boldsymbol{t}') > \inf_{(e'',\boldsymbol{t}'')\in X(e',s)} T(\boldsymbol{e}'',\boldsymbol{t}'')$ (and thus $P(\boldsymbol{e}'',\boldsymbol{t}',\emptyset) < \sup_{(e'',\boldsymbol{t}'')\in X(e',s)} P(\boldsymbol{e}',\boldsymbol{t}'',\emptyset))$ keeping Π fixed. Therefore, among agents with the same training 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 leads to 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 $\sigma(\boldsymbol{e},\boldsymbol{t}) = \sigma(\boldsymbol{e},\boldsymbol{t}')$.[44] 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 SIC characterization of Proposition 1 to the case of $(m+n)$-dimensional screening.

Proposition 10. A mechanism $M \equiv \langle T,P\rangle$ is SIC 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 an agent is accepted if she truthfully reports her training $\boldsymbol{e}$ and composite measure s.

The conditions are analogous to those of Proposition 1. There are no incentive-compatibility 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 $\boldsymbol{e} \not\leq \boldsymbol{e}'$, because neither agent type has the evidence to imitate the other.

Lemma 7 generalizes Lemma 2, showing that we can constrain attention to mechanisms that satisfy condition (iii) of Proposition 10 with equality.

Lemma 7. Given any SIC mechanism $M \equiv \langle T,P\rangle$, there exists an SIC mechanism $M' \equiv \langle T',P'\rangle$ with $(1-T'(\boldsymbol{e},s))P'(\boldsymbol{e},s,\emptyset) = \Pi'(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0}))$ for every $(\boldsymbol{e},s)$, $s \in [\sigma(\boldsymbol{e},\boldsymbol{0}),\sigma(\boldsymbol{e},\boldsymbol{1})]$ that is outcome-equivalent to M and has at most as high verification costs as M. For

[44] 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 mechanisms that accept the agent with certainty if she meets the appropriate composite measure threshold.

$c > 0$, 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 training $\boldsymbol{e}$ 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 training $\boldsymbol{e}$ and composite measure s. $\widetilde{u}(\boldsymbol{e},s)$ is assumed to be increasing in s.[45] The principal's objective function is $\int_0^1 \cdots \int_0^1 \int_{\sigma(\boldsymbol{e},\mathbf{0})}^{\sigma(\boldsymbol{e},\mathbf{1})} [\Pi(\boldsymbol{e},s)\widetilde{u}(\boldsymbol{e},s) - cT(\boldsymbol{e},s)]\, \widetilde{f}(\boldsymbol{e},s) ds\, de_1 \cdots de_m$. By Lemma 7, condition (iii) of Proposition 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

$$\int_0^1 \cdots \int_0^1 \int_{\sigma(\boldsymbol{e},\mathbf{0})}^{\sigma(\boldsymbol{e},\mathbf{1})} [\Pi(\boldsymbol{e},s)(\widetilde{u}(\boldsymbol{e},s) - c) + c\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))]\, \widetilde{f}(\boldsymbol{e},s) ds de_1 \cdots de_m, \tag{10}$$

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

Lemma 8. There exists an optimal mechanism that is deterministic.

C.1 Talent-biased composite measure

The definition of a talent-biased composite measure generalizes to the case of $(m+n)$-dimensional screening as follows.

Definition 6. σ is talent-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 talent-biased composite measure.

Proposition 11. If σ is talent-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 $\boldsymbol{e}' \in E^*$).[46]

[45] 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 can also be derived with $\widetilde{u}(\boldsymbol{e},s)$ non-decreasing in s, which would somewhat complicate the proofs.

[46] 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 $(\boldsymbol{e},\mathbf{0})$'s. Therefore, $(\boldsymbol{e},\mathbf{0})$'s composite measure should not be verified if $\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) > c$.

C.2 Training-biased composite measure

The definition of a training-biased composite measure generalizes to the case of $(m+n)$-dimensional screening as follows.

Definition 7. σ is training-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 training-biased composite measure.

Proposition 12. If σ is training-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 $[0,1]^m$.[47]

D Proofs of results in Appendix C

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

Proof of Lemma 6 Take any SIC mechanism M. Construct the mechanism $M' \equiv \langle T',P' \rangle$ with[48]

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

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 SIC, 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 $(\boldsymbol{e},\boldsymbol{t})$, training $\boldsymbol{e}' \leq \boldsymbol{e}$, and talent $\boldsymbol{t}'$. It holds that

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

[47] 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.

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

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

where (i) the first equality follows by construction of M', (ii) the inequality because M is SIC, (iii) the second equality by definition of Π, (iv) the third equality by Lemma 5 and M being SIC, which together imply that $\Pi(\boldsymbol{e}',\tilde{\boldsymbol{t}}) = \Pi(\boldsymbol{e}',\boldsymbol{t}')$ for every $\tilde{\boldsymbol{t}}$ such that $\sigma(\boldsymbol{e}',\tilde{\boldsymbol{t}}) = s$, (v) the fifth inequality by construction of M', and the final equality by definition of Π'. We have thus shown that for any agent $(\boldsymbol{e},\boldsymbol{t})$, $\Pi'(\boldsymbol{e},\boldsymbol{t}) \geq (1 - T'(\boldsymbol{e}',\boldsymbol{t}'))P'(\boldsymbol{e}',\boldsymbol{t}',\emptyset)$ for every $(\boldsymbol{e}',\boldsymbol{t}') \leq (\boldsymbol{e},\boldsymbol{1})$, so under mechanism M', no agent has incentives to imitate an agent with higher composite measure than her own. For $c > 0$, M' also minimizes 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 7 The proof proceeds like the proof of Lemma 2.

Take any SIC mechanism $M \equiv \langle T,P\rangle$. Condition (iii) of Proposition 10 says that $\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) \geq (1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset)$ for any $(\boldsymbol{e},s)$. Then, construct the mechanism $M' := \langle T',P'\rangle$ with[49]

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

for every $(\boldsymbol{e},s)$, where the inequalities follow from $\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) \geq (1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset)$. By construction, $\Pi'(\boldsymbol{e},s) = \Pi(\boldsymbol{e},s)$ for every $(\boldsymbol{e},s)$, so M' satisfies conditions (i) and (ii) of Proposition 1. Also, for every $(\boldsymbol{e},s)$

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

so M' also satisfies condition (iii) of Proposition 1. Therefore, M' is SIC. Last, for $c > 0$, M' saves on verification costs compared to M if there exists (a positive measure of) $(\boldsymbol{e},s)$ with $P(\boldsymbol{e},s,\emptyset)(1 - T(\boldsymbol{e},s)) < \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0}))$, since $T'(\boldsymbol{e},s) < T(\boldsymbol{e},s)$ for such $(\boldsymbol{e},s)$. **Q.E.D.**

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 $E^* := \{\boldsymbol{e} \in [0,1]^m : \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\boldsymbol{0})) = 1\}$

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

(so $\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))=0$ for every $\boldsymbol{e}\notin E^*$). Given that M is SIC, 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.[50]

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 (10) can be written as

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

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 incentive-compatible.

To show this, we first prove that $\Pi(\boldsymbol{e},s)=\mathbf{I}(\tilde{u}(\boldsymbol{e},s)>c$ 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 $\Pi(\boldsymbol{e},s)=\mathbf{I}(\tilde{u}(\boldsymbol{e},s)\geq c$ 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 talent-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 SIC, 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 (10) can be written as

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

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 $\boldsymbol{e}\notin E^*$.

Take any $(\boldsymbol{e},s),(\boldsymbol{e}',s')\in I_{\tilde{u}}(c)\setminus E^*$ with $s\neq s'$. That $(\boldsymbol{e},s),(\boldsymbol{e}',s')\in I_{\tilde{u}}(c)$ means that

[50] 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 training $\boldsymbol{e}$ and composite measure $\sigma(\boldsymbol{e}',\mathbf{0})$) but the inequalities still follow if we use conditions (i) and (ii) iteratively.

$\widetilde{u}(\boldsymbol{e},s) = \widetilde{u}(\boldsymbol{e}',s') = c$. 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 training-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 training-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, $\widetilde{u}(\boldsymbol{e},s') < c = \widetilde{u}(\boldsymbol{e},s)$ and $s' < s$, because $s' \neq s$ and by assumption, $\widetilde{u}(\boldsymbol{e},s)$ is increasing in 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 increasing in s, it follows that $\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{1})) \geq \widetilde{u}(\boldsymbol{e},s) = c$. Thus, we have $\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{1})) \geq c = \widetilde{u}(\boldsymbol{e}',s') > \widetilde{u}(\boldsymbol{e}',\sigma(\boldsymbol{e},\mathbf{1}))$, so the composite measure being training-biased 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 increasing 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, $\sigma(\boldsymbol{e}'',\mathbf{0}) = \sigma(\boldsymbol{e},\mathbf{1})$. Given that σ is training-biased, $\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{1})) \geq c \geq \widetilde{u}(\boldsymbol{e}'',\sigma(\boldsymbol{e},\mathbf{1}))$ with one inequality strict implies that $\boldsymbol{e}'' \geq \boldsymbol{e}$, which combined with $\boldsymbol{e}'' \leq \boldsymbol{e}'$ implies $\boldsymbol{e}' \geq \boldsymbol{e}$, a contradiction. Thus, Case 3b is impossible.

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 $(\boldsymbol{e},s) \notin I_{\widetilde{u}}(c) \cup E^*$.

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 training-biased, $\boldsymbol{e} \geq \boldsymbol{e}^*$, and thus 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 training-biased, $\boldsymbol{e}^* \geq \boldsymbol{e}$, and thus 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 incentive-compatibility.

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 incentive-compatibility.

Case 2: If $\tilde{u}(\boldsymbol{e}',s) < c$ for every $\boldsymbol{e}'$ such that $s \in [\sigma(\boldsymbol{e}',\mathbf{0}),\sigma(\boldsymbol{e}',\mathbf{1})]$, 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 incentive-compatibility.

Case 3: If $\tilde{u}(\boldsymbol{e}',s) > c$ for every $\boldsymbol{e}'$ such that $s \in [\sigma(\boldsymbol{e}',\mathbf{0}),\sigma(\boldsymbol{e}',\mathbf{1})]$, 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 incentive-compatibility.

Putting all the above cases together, we get that for $(\boldsymbol{e},s) \notin I_{\tilde{u}}(c) \cup E^*$, $\Pi(\boldsymbol{e},s) = \mathbf{I}(s \geq s^*)$. 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 SIC. Therefore, by solving a relaxed problem ignoring the incentive-compatibility constraints in cases 1c, 1d, 2, and 3, we have also solved the original problem. **Q.E.D.**