Multidimensional screening of strategic candidates*

Orestis Vravosinos†

September 23, 2026

Abstract

An evaluator can pay to verify the value of a composite measure of a candidate's training and talent before deciding whether to accept her. Although both qualities are valuable, when the composite measure is oversensitive to training, verification can induce the candidate to under-report training to make the evaluator attribute the composite measure to talent instead. The optimal mechanism is forced to make—but minimizes—errors favoring high- over low-training candidates. Failing to account for incentives to under-report leads to extensive under-reporting and excessive acceptance of unworthy high-training candidates. The results have implications for promotions, hiring, poaching, and college admissions.

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

*Previously circulated as "Multidimensional screening with substitutable attributes." I am grateful to Erik Madsen for his continued guidance and terrifyingly insightful comments. The paper has also immensely benefited from discussions with Ben Brooks and 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, Todd Jones, 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.

†University of Chicago; e-mail: orestis.vravosinos@uchicago.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

A candidate's suitability for a position typically depends on multiple valuable qualities, such as education, training, knowledge, intelligence, or adaptability. While the evaluator is often unable to assess each of the candidate's qualities in isolation, he can observe—possibly at a cost—the value of a *composite measure* of the candidate's qualities, without being able to disentangle the individual contribution of each quality to the composite measure. For instance, to decide whether to promote an employee, an employer can monitor the employee's performance in the current position, which depends both on how hardworking and on how efficient the employee is.[1] At the same time, in some—but not all—cases, the candidate has hard evidence on some of her qualities and can choose how much of this evidence to present. For example, if an employer monitors the employee's office hours, the employee will be able to choose how much evidence of hardworkingness to present by deciding how to allocate her work hours between working from the office and working from home.[2]

In principle, the candidate has no incentives to under-report a valuable quality. If she has hard evidence on that quality, full unraveling will emerge without the need for the evaluator to carefully design incentives (Viscusi, 1978; Grossman, 1981; Milgrom, 1981).[3]

[1] Monitoring and evaluating employee performance costs organizations millions annually (Pulakos et al., 2019). The cost is not only direct (e.g., for performance management software) but also in terms of foregone work hours that managers and employees spend on performance evaluation.

[2] Flexibility to work from home is widespread, with employers offering such flexibility to—among other reasons—attract and retain employees (Bloom et al., 2015; McMullin, 2025; Dua et al., 2022; Aksoy et al., 2022; Barrero et al., 2023; Bloom et al., 2024). Even if an employer requires employees to work from the office, imperfect monitoring of work hours can still allow employees discretion with respect to how much evidence of hardworkingness to present (e.g., by choosing how many of her work hours are observed by the employer). Also, employees may still be able to choose how much evidence of hardworkingness to present by deciding how to allocate hours worked beyond the required office hours between staying late in the office and working extra hours from home.

[3] That is, absent a strong negative stochastic dependence between the quality the candidate has evidence on and other valuable qualities.

However, when the evaluator gains access to a composite measure of the candidate's qualities, perverse incentives may arise: The candidate might strategically under-report one of her qualities to make the evaluator attribute the composite measure to another quality instead, thereby overestimating that quality. An employee may hide how hard she works to make the employer attribute her performance to talent and promote her. Indeed, although accounting firms' codes of conduct instruct employees to report time completely and accurately (PwC, 2024; KPMG, 2025b; Deloitte, 2025), there is extensive evidence that auditors consistently under-report their work hours (e.g., by working on personal time without reporting it) to appear more efficient, receive better performance evaluations, and increase their chances for promotion. Under-reporting of work hours negatively affects the firm not only by complicating the assessment of employee efficiency—thereby undermining the firm's personnel evaluation, retention, and promotion processes—but also by making it harder to predict future auditing project duration (Shapeero and Killough, 1993; Smith et al., 1996; Otley and Pierce, 1996; Akers et al., 1998; Sweeney and Pierce, 2006; Mendoza, 2020; Blum et al., 2022). Also, audit hours are reported in public disclosures of audit quality, making it important that work hours are reported completely (Mendoza, 2020; PCAOB, 2024; Deloitte, 2024; Ernst & Young, 2025; KPMG, 2025a; PwC, 2025).

A college applicant may not fully reveal her training and parental support, so that her academic performance and standardized test scores are attributed to her brilliance. 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. A professional might downplay her background to make a potential employer attribute her career trajectory and performance at her current employer to talent. Indeed, there is suggestive evidence that people downplay their privileged background to enhance their career prospects by leading others to attribute their achievements to talent (Jacobs, 2024; Reeves and Friedman, 2024).

The above examples suggest that there can be a conflict between the two evaluation tools: (i) verifying a composite measure of the candidate's qualities and (ii) asking the candidate to reveal information about her qualities—be it with or without evidence. Under what circumstances does the conflict arise? When it does, how does the evaluator *optimally evaluate* the candidate while taking the conflict into account? What happens under the *naive benchmark*, where the evaluator designs the evaluation scheme failing to realize that the agent may under-report one of her qualities to distort the interpretation of the composite measure? 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*.[4] She cannot

[4]The Online Appendix shows how the characterization of the optimal mechanism generalizes to the case where there are m dimensions of training and n dimensions of talent. 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 understate effort and preparation), to isolate the screening aspect of the

unilaterally prove anything about her talent. I study two cases: (i) the case where the agent has hard evidence of training and (ii) the case where she does not. In the former case, the evidence an agent with higher training has is a proper superset of the evidence an agent with lower training has. This means that the agent can under-report but not over-report training. In the latter case, the agent can both under- and over-report training.

The principal's payoff from accepting the agent (i) is non-decreasing in both training and talent and (ii) can be positive or negative. His payoff from rejecting the agent is zero. He ultimately wants to decide whether to accept or reject the agent. He does so by committing to a mechanism that asks the agent to report her training and talent. Conditional on the report, 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.

Whether verification creates incentives to under-report training or talent 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 lower 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*.[5] If the composite measure is talent-biased, verification creates incentives to under-report talent and over-report training—unless incentives are carefully designed. If the composite measure is training-biased, verification creates incentives to under-report training and over-report talent.

Consider first the case where the agent has evidence of training. If the composite measure is talent-biased, verification does not create incentives to withhold evidence. Thus, 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. On the other hand, when the composite measure is training-biased, verification *does* create incentives to withhold evidence. The optimal evaluation scheme never combines evidence and verification in the evaluation of an agent. Rather, it asks for a certain level e^* of evidence only to accept some high-training agents—including unworthy agents who give the principal a negative payoff when accepted—*without* verification. Among agents who do not have that level of evidence, it accepts those whose composite measure meets a threshold s^*. In doing so, (i) it rejects without verification some worthy agents with

interaction, I solve the problem for exogenous training. The Online Appendix discusses endogenous training.

[5] In the model, talent and training-bias are defined more generally in terms of single-crossing conditions.

high talent but low training, although the payoff from accepting them would exceed the verification cost (Type I error), and (ii) 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 (Type II error). Overall, the principal optimally favors high- over low-training agents due to two forces: (i) to save on verification costs by accepting high-training agents without verification and (ii) due to the strategic incentives of agents to withhold evidence.

The two forces are complements in inducing errors favoring high- over low-training agents. Each of the two forces *individually* induces such errors. 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. 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 high-training agents without verification.

There is a trade-off between Type I and Type II errors. The optimal mechanism finely balances the two errors. Instead, the naive benchmark induces the agent to withhold evidence of training, thereby leading to excessive Type II while eliminating Type I errors. As a result, too many agents are accepted after verification. At the same time, in the naive benchmark, the evidence threshold for acceptance without verification is too strict: Too few agents are accepted without verification. Failing to realize that verification creates incentives to withhold evidence, the principal overestimates the effectiveness of verification and thus under-utilizes evidence in evaluation.

The verification cost c mediates the trade-off between the two errors while also shaping the incentives to accept high-training agents without verification. An increase in c tends to directly cause (i) lower standards for acceptance without verification (i.e., e^* tends to decrease), leading to more agents getting accepted without verification, and (ii) stricter standards for acceptance after verification (i.e., s^* tends to increase), leading to fewer agents getting accepted after verification. However, the interaction between these two forces works in the opposite direction. The decrease in e^* tends to relax the standards for acceptance after verification by mitigating the risk that a high composite measure reflects high training but low talent. Therefore, the overall effect is ambiguous. Particularly, an increase in the verification cost can relax both the standards for acceptance without verification and the standards for acceptance after verification.

Consider now the case where the agent does not have evidence of training. Under a training-biased composite measure, the optimal scheme (i) rejects without verification

some worthy agents with low training but high talent (Type I error) and (ii) accepts after verification some unworthy agents with high training but low talent (Type II error).[6] The composite measure required for acceptance is higher than in the case where the agent has evidence. In that case, the principal accepts high-training agents without verification, which mitigates the risk that a high composite measure reflects high training but low talent. Therefore, the standards for acceptance after verification can be relaxed. The naive benchmark induces the agent to under-report training. As a result, while the optimal mechanism finely balances the two errors, the naive benchmark leads to excessive Type II while eliminating Type I errors.

The results have implications for promotions, hiring, poaching, and college admissions. They can explain the extensive under-reporting of work hours by auditors as a consequence of sub-optimally designed incentives that put too much emphasis on efficiency, leading to under-reporting of work hours. The results suggest that accounting firms should relax their emphasis on efficiency to mitigate under-reporting of work hours and the excessive reliance on performance evaluations for promotion decisions. Also, if they can monitor employee work hours, they should not only promote high-performing employees. They should also consider offering a path to promotion for employees who work long hours. Apart from limiting performance evaluation costs, offering this path reduces the risk that an employee who is promoted thanks to high performance is hardworking but untalented. This would allow the firm to relax performance standards for promotions. In designing the promotion path for hard-working employees, the firms should take into account that the effectiveness of performance evaluation is undermined by the incentives to under-report work hours. Failing to take this into consideration would prevent the firms from taking full advantage of work hours monitoring, making them overly rely on performance monitoring instead.

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. The cost can reflect job switching costs or a non-compete clause in the contract offered by the less prestigious firm.

It has been argued that non-compete clauses may make it harder to attract valuable employees (see, e.g., Garmaise, 2011). The results suggest that a non-compete clause, which increases the cost c of poaching, may undermine the less prestigious employer's ability not only to attract but also even to *retain* employees. The direct effect of the non-compete on poaching is indeed negative. However, if job candidates have evidence of training and performance in the less prestigious position is training-biased (compared to performance in the prestigious position), there is an indirect effect working in the

[6]The case of a talent-biased composite measure is equivalent up to relabelling.

opposite direction. By making it costlier to poach, the non-compete clause pushes the prestigious firm to lower its standards for hiring a candidate immediately, without waiting to observe her performance at the less prestigious employer. This, in turn, reduces the risk that a well-performing employee of the less prestigious firm is well-trained but untalented. Therefore, poaching becomes more attractive. Thus, non-compete clauses may be counterproductive in their goal of improving employee retention.

Last, the results have implications for college admissions. In this context, the cost c of incorporating the applicant's standardized test score in the evaluation is likely negligible. If standardized test scores reflect talent (relative to background, preparation, and parental support) less than colleges value talent, the optimal admissions policy requires the same test score from every applicant for admission—regardless of background. As a result, some worthy, talented applicants with low training are rejected, while some unworthy, untalented applicants with high training are accepted. The results can explain the College Board's introduction of a redesigned SAT test in 2016 with the goal to "rein in the intense coaching and tutoring on how to take the test that often gave affluent students an advantage" (Lewin, 2014).

The paper contributes methodologically to the multidimensional screening literature. It proposes a multidimensional screening problem that is relevant for real-world applications and fully characterizes its solution. Tractability is afforded by the fact that (i) the agent has transparent motives as in Lipnowski and Ravid (2020) and (ii) the principal's ultimate decision is binary.[7] At the same time, the model allows for substantial generality in other ways. Namely, it allows (i) for arbitrary stochastic dependence between the different dimensions of the agent's type, (ii) for costly verification of a composite measure of the agent's multidimensional type (which effectively makes the principal's choice non-binary), encompassing settings without verification, as well as settings where the principal can verify one dimension of the agent's type, and (iii) for the agent to have hard evidence on some dimensions of her type but not others.[8]

Related literature. The urge to downplay 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. However, to the best of my knowledge, no prior work has studied the following problem: evaluating people who may strategically understate a valuable quality 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 con-

[7]The Online Appendix shows how the IC characterization generalizes to the case where the principal's ultimate decision lies in a real interval.

[8]It encompasses settings where there is no hard evidence, as well as settings where the agent has hard evidence on every dimension of her type. The latter case is trivial given the agent's transparent motives and the principal's monotone preferences over every dimension of the agent's type.

tributes to the multidimensional screening literature by proposing a novel multidimensional screening problem. The solution to the monopolist's multidimensional screening problem is elusive and, even when known, complex (Manelli and Vincent, 2006; Daskalakis et al., 2017).[9] On the other hand, I fully characterize the solution to the proposed problem. Unlike in the monopolist's screening problem, in this model, the agent has transparent motives: All types prefer acceptance over rejection. An agent's type affects her preferences over how acceptance or rejection decisions depend on the composite measure: Agents with a higher composite measure benefit more from mechanisms that reward high composite measures with high acceptance probabilities. Due to the agent's transparent motives, the technical issues in this setting are different and, ultimately, more tractable than those in the monopolist's problem. Another difference is that in this paper, I consider the case where agents have hard evidence for some dimensions of their type. As seen through a comparison of sections 3 and 4, this feature of the model complicates rather than simplifies the principal's problem.

This paper also fits into the literature on models with costly verification. A main difference between my model and existing models with costly verification is that in existing work, verification amounts to either the revelation of the agent's one-dimensional type (see, e.g., Townsend, 1979; Gale and Hellwig, 1985; Ben-Porath et al., 2014; Halac and Yared, 2020; Kattwinkel and Knoepfle, 2023) or the revelation of one dimension of the agent's multidimensional type (see, e.g., Glazer and Rubinstein, 2004; Carroll and Egorov, 2019).[10] Therefore, the interpretation of the verification result is independent of the agent's report.

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 to influence the principal's learning (through costless observation of the composite measure) of the agent's talent. Here, I focus on information transmission and costly verification. I show that if the principal can ask for hard evidence of training, 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. The agent has incentives to withhold evidence, which she should be paid information rents to reveal.

[9]Frick et al. (2026) show that the monopolist's problem is in a sense simplified if she has sufficiently precise information about the buyer's valuations.

[10]The verification technology in this paper nests the case where verification reveals one of the dimensions of the agent's type. For example, in the case where the agent has evidence of training, 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.

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; Dziuda, 2011; Hart et al., 2017), countersignaling (Feltovich et al., 2002), and 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).

2 The model

There is an agent (she) and a principal (he). The agent is privately informed of her type $(e,t) \in [0,1]^2$ with full-support density $f : [0,1]^2 \to \mathbb{R}_{++}$. No other assumption is imposed on f; any form of stochastic dependence between e and t is allowed. t is the agent's *talent*, which she cannot unilaterally prove anything about. e is the agent's *training*. Sections 3 and 4 will study two different specifications of the model. In section 3, the agent has evidence equal to her training. Thus, an agent of type (e,t) can present any level of evidence $e' \in [0,e]$. This means that the agent can under-report but not over-report training. It is not important that the evidence itself be a real number. This setting is equivalent to one where (i) each agent has many pieces of evidence, (ii) she can choose which pieces to present, and (iii) when $e > e'$, the pieces of evidence (e,t) has are a proper superset of the pieces of evidence (e',t') has. In section 4, the agent has no evidence, so she can both under-report and over-report training.

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. $\sigma : [0,1]^2 \to [0,1]$ is differentiable and increasing 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 decides 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 a level set of the principal's payoff from acceptance, 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 example. 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, $\sigma(e,t) \coloneqq \gamma_\sigma e + (1-\gamma_\sigma)t$, where $\gamma_\sigma \in (0,1)$ measures how sensitive the composite

measure is to e versus t. No parametric assumptions are imposed on u or σ. For simplicity in depiction, all figures use the linear specification.

The principal's problem. To decide whether to accept the agent, the principal commits to a direct mechanism $M \equiv \langle T,P \rangle$ that specifies: (i) the probability $T(e,t) \in [0,1]$ with which the principal will verify the composite measure if the agent's report is (e,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 when the agent's report is (e,t) and the composite measure is $s \in [0,1]$.[11] 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. 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

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

subject to the agent's incentive compatibility (IC) constraint. If the agent has evidence of training (section 3), the IC constraint is

$$\phi(e,t) \in \underset{(e',t') \in [0,e] \times [0,1]}{\arg\max} \underbrace{\{T(e',t')P(e',t',\sigma(e,t)) + (1 - T(e',t'))P(e',t',\emptyset)\}}_{\text{total probability that } (e,t) \text{ is accepted if she reports } (e',t')}.$$

If she does not have evidence of training (section 4), the IC constraint is

$$\phi(e,t) \in \underset{(e',t') \in [0,1]^2}{\arg\max} \{T(e',t')P(e',t',\sigma(e,t)) + (1 - T(e',t'))P(e',t',\emptyset)\}.$$

Model discussion. In certain cases, the agent's evidence may—much like the composite measure—capture a combination of the agent's qualities. For example, evidence of a college applicant's performance in high school can capture a combination of training and talent. Assume type (e,t) has evidence $\varepsilon(e,t)$, where $\varepsilon(e,t)$ is increasing in e and t. Then,

[11] The condition that $P(e,t,s)$ be non-decreasing in $s \in [0,1]$ can be interpreted as a "fairness" condition or as an incentive-compatibility condition in a model where $\sigma(e,t)$ is the (maximum) composite measure that agent (e,t) can achieve, and the agent can intentionally manipulate her composite measure downwards. The Online Appendix shows that if such manipulation is possible, requiring that $P(e,t,s)$ be non-decreasing in $s \in [0,1]$ is without loss.

we can redefine the agent's type to be $(\widetilde{e},t)$ with $\widetilde{e} \coloneqq \varepsilon(e,t)$, $u(\widetilde{e},t) = u(e(\widetilde{e},t),t)$, and $\sigma(\widetilde{e},t) = \sigma(e(\widetilde{e},t),t)$, where $e(\widetilde{e},t)$ is implicitly defined by $\varepsilon(e(\widetilde{e},t),t) = \widetilde{e}$. Thus, the model captures the case where evidence measures a combination of talent and training, and t is the part of talent not captured by evidence $\widetilde{e}$.

In some cases, some part of training cannot be hidden. For example, if an employee is required to be in the office for some hours each week while the employer (or manager) is there to observe this, part of the employee's work hours is observed. The part of a college applicant's training that is captured by the identity of the high school she attended may not be hidden. If we let the agent's type (e_p,e,t) be distributed over $[0,1]^3$, where e_p is the publicly observed part of training and e is the privately observed one, the analysis of the following sections is valid conditional on e_p.

The Online Appendix generalizes the model to the case of m dimensions of training and n dimensions of talent.

3 The optimal mechanism with evidence of training

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 that it is without loss to restrict attention to deterministic mechanisms and that 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).

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 principal can without loss of optimality restrict attention to truthful mechanisms.

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.[12] The Online Appendix discusses when the outcome induced by the optimal mechanism can be implemented in simpler, non-truthful ways.

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

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

for any (e,t) and some $P_{av} : [0,1]^2 \to [0,1]$. If type (e,t) truthfully reports her type and the composite measure is verified, she is accepted with probability $P_{av}(e,t)$. 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 $P_{av}(e,t)$, the one that satisfies equation (1) minimizes incentives of other types to imitate (e,t).[13]

We can further restrict attention to mechanisms that accept the agent with certainty if she meets the composite measure threshold (i.e., $P_{av}(e,t) = 1$). To see why, denote the total probability with which agent (e,t) is accepted if she truthfully reports her type by $\Pi(e,t) \coloneqq (1 - T(e,t))P(e,t,\emptyset) + T(e,t)P_{av}(e,t)$ 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 $\Pi(e,t) \equiv (1-T(e,t))P(e,t,\emptyset)+T(e,t)P_{av}(e,t)$ and $\Pi'(e,t) \equiv (1 - T'(e,t))P'(e,t,\emptyset) + T'(e,t)P'_{av}(e,t)$.

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.

[12] Essentially, the principal implements a social choice function $g : [0,1]^2 \to [0,1]^2 \times [0,1]^{[0,1]}$, where $g_1(e,t)$ the probability of 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.

[13] Namely, accepting the agent with even higher probability for performing above $\sigma(e,t)$ will result in the same probability of accepting type (e,t) in case of verification and only provide additional incentives 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.

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

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 $T(e,t)P_{av}(e,t)$ of acceptance after verification (provided that the composite measure threshold is met). If the principal pays for verification, he may as well set $P_{av}(e,t) = 1$ 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 for every $(e,t) \in [0,1]^2$, $P(e,t,s) = \mathbf{I}(s \geq \sigma(e,t))$.

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 $\overline{e}(s) \coloneqq \max\{e \in [0,1] : \sigma(e,0) \leq s\}$.[15]

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.

[14] Strictly put, $P_{av}(e,t)$ 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 $P_{av}(e,t)$ does not matter, so we can again set $P_{av}(e,t) = 1$ without loss.

[15] $\underline{e}(s)$ (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(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 $\Pi(e,t) \geq \Pi(e,0) \geq (1 - T(e,t'))P(e,t',\emptyset)$ 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.

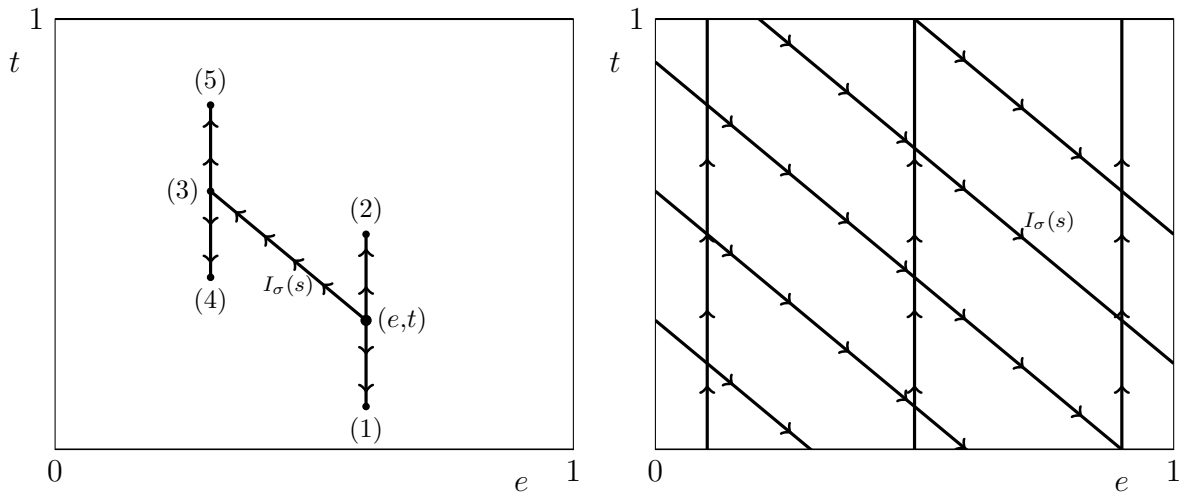
3.3.1 The optimal mechanism does not overspend on verification

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

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



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.[16] 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 attention 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).

[16]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.

3.3.2 The optimal mechanism is deterministic

The principal's objective function is $\int_0^1 \int_0^1 [\Pi(e,t)u(e,t) - cT(e,t)]\, f(e,t)dedt$. 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:

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

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)).

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 specification (see section 2), σ is talent-biased when $\gamma_\sigma < \gamma_u$. A talent-biased σ can be defined more generally as follows.[17]

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,

[17] We define a talent-biased composite measure for any verification cost c. The optimal mechanism under costly verification is studied in section 3.5. In the Online Appendix, a talent-biased composite measure is defined more generally, allowing for m dimensions of training and n dimensions of talent.

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

Proposition 2. Let $c = 0$, and assume that σ is talent-biased. Then, $\Pi(e,t) = \mathbf{I}(u(e,t) > 0)$ is incentive-compatible, so the principal achieves the full information first-best.[18]

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 specification (see section 2), σ is training-biased when $\gamma_\sigma > \gamma_u$. A training-biased σ can be defined more generally as follows.

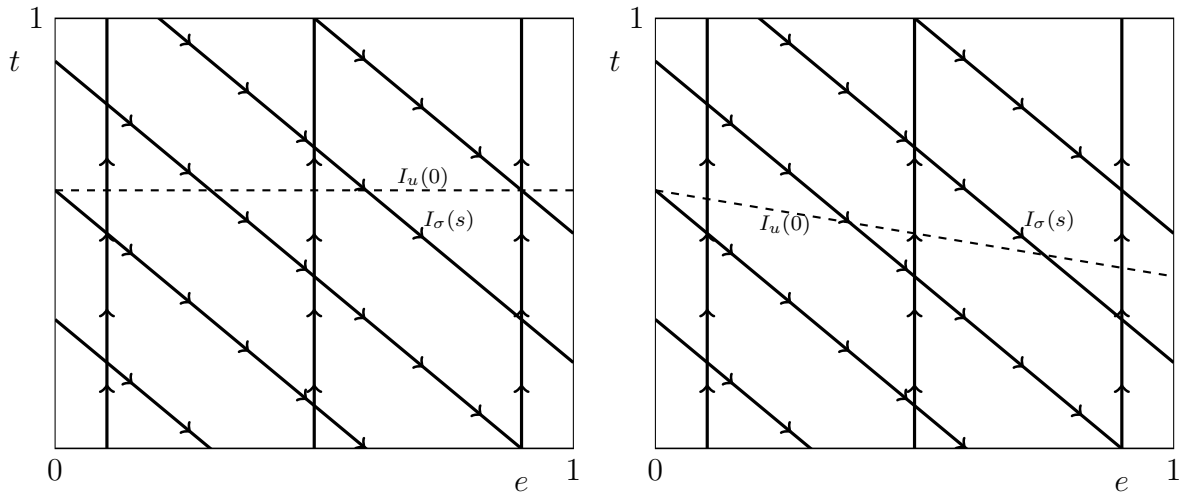
Definition 5. σ is training-biased if for every composite measure $s \in [0,1]$ there exists e_s such that for every (e,t), if $e < e_s$ (resp. $e > e_s$) and $\sigma(e,t) = s$, then $u(e,t) > c$ (resp. $u(e,t) < c$).

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.

[18] 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.

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 $I_u(0)$.

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 $\frac{\partial u(e,t)/\partial t}{\partial u(e,t)/\partial e} > \frac{\partial \sigma(e,t)/\partial t}{\partial \sigma(e,t)/\partial e}$ 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 $\Pi(e,t) = \mathbf{I}(\sigma(e,t) \geq s^*)$.[19]

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).

[19]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^*.

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 (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.[20] 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^*).[21] 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 The optimal mechanism without evidence

This section characterizes the optimal mechanism when the agent does not have evidence. That is, agent (e,t) can report any $(e',t') \in [0,1]^2$. We can still restrict attention to SIC mechanisms, which Proposition 6 characterizes.

Proposition 6. 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),

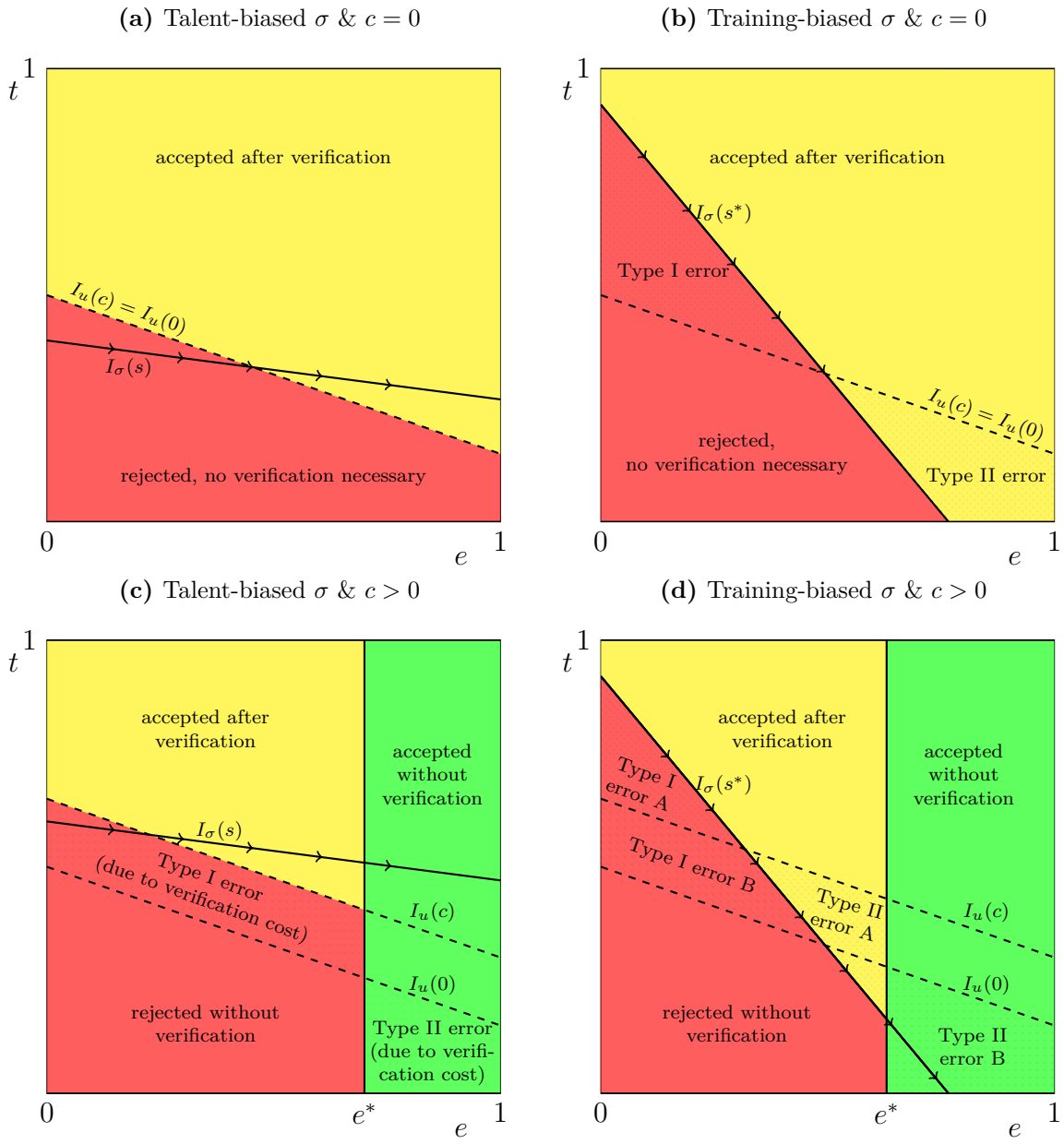
where $\Pi(e,t) \equiv (1-T(e,t))P(e,t,\emptyset) + T(e,t)$.

Condition (i) is identical to the one in Proposition 1, 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 1. 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 1 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

[20] 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 $I_u(c)$.

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

Figure 3: The optimal mechanism when the agent has evidence of training



Note: The dashed line $I_u(c)$ is the principal's iso-utility curve at utility level c. 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 colors denote whether the agent is (i) accepted without verification, (ii) accepted after verification, or (iii) rejected without verification.

$\sigma(e',t') = \sigma(e,t)$ to get accepted with probability $\Pi(e',t')$ instead of $\Pi(e,t)$. The condition is stricter than the one in Proposition 6 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 6 with equality. Thus, $\Pi(e,t) = \Pi(0,0) + T(e,t)$.

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.

The principal's objective function can be written as

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

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 7 describes that extreme optimal mechanism.

Proposition 7. 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.[22] Also, the principal now has to choose 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—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. Figure 4 presents the optimal mechanism.

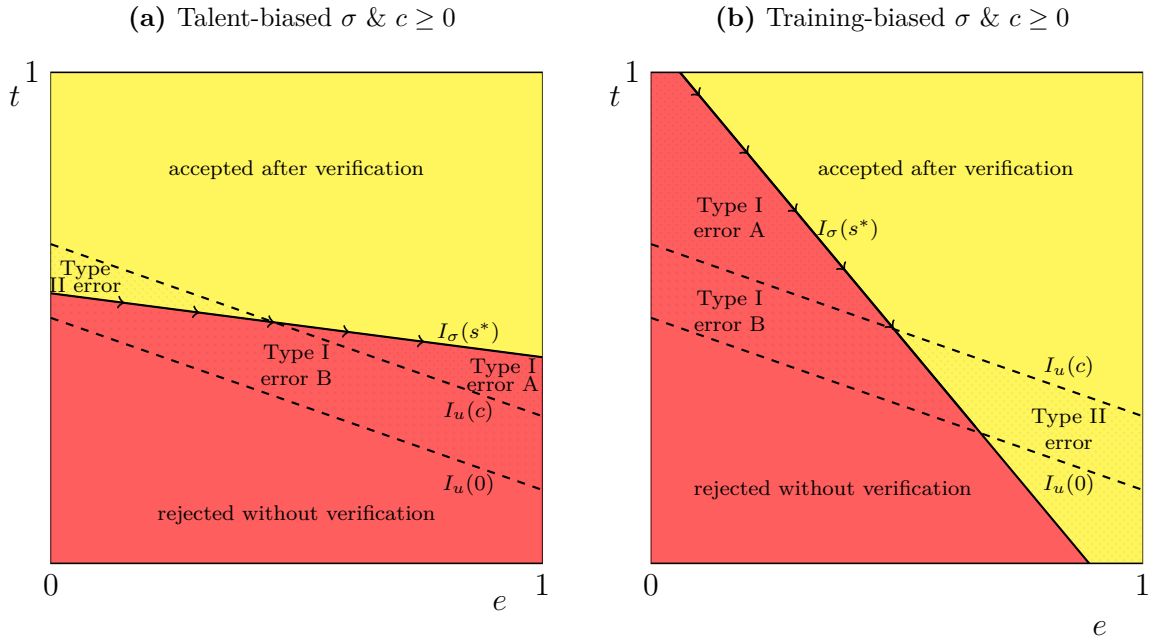
5 A naive benchmark

This section presents a naive benchmark: the outcome implemented by a principal who wrongly believes that whenever he verifies the composite measure, he can correctly interpret it, assuming the agent does not under-report one of her qualities to over-report the other.[23]

[22] 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.

[23] The Online Appendix describes formally how the naive benchmark arises when the principal wrongly believes that the agent never under-reports training or talent.

Figure 4: The optimal mechanism when the agent does *not* have evidence of training



Note: The dashed line $I_u(c)$ is the principal's iso-utility curve at utility level c. The arrowed line represents the iso-composite-measure curve at the optimal level s^*. The colors denote whether the agent is (i) accepted after verification or (ii) rejected without verification.

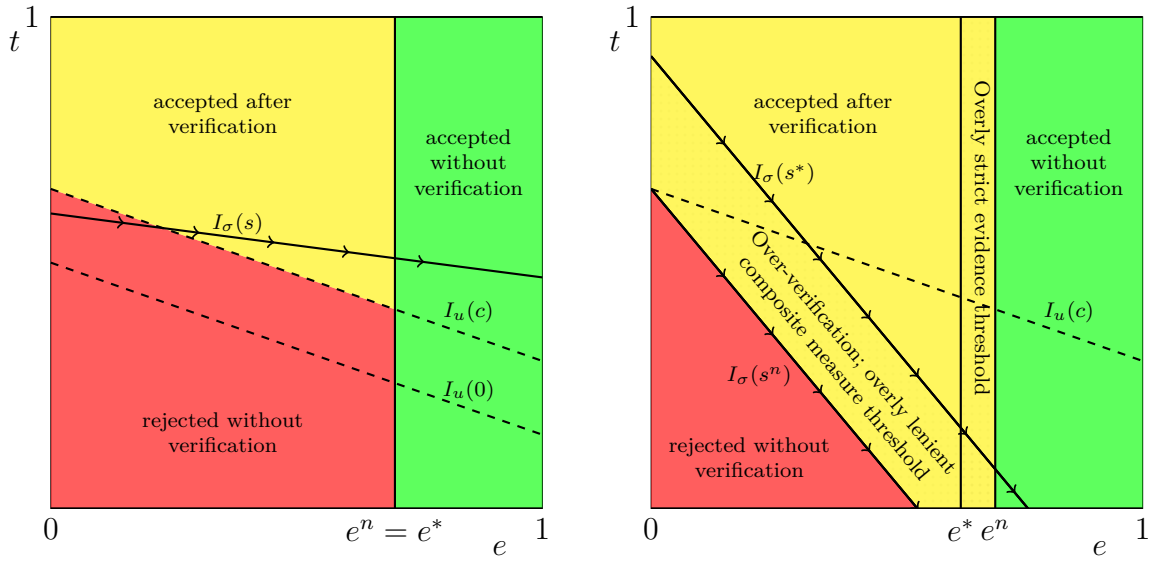
Agent with evidence of training. The outcome that the principal attempts to implement is the one of Proposition 4: Each agent (e,t) is (a) accepted without verification if $e \geq e^n$, (b) accepted after verification if $u(e,t) \geq c$ and $e < e^n$, or (c) rejected without verification if $u(e,t) < c$ and $e < e^n$. e^n is the evidence threshold of the naive benchmark. Assuming that the agent will not withhold evidence, the principal attempts to implement this outcome by specifying that for every e, an agent presenting evidence e is (a) accepted without verification if $e \geq e^n$, (b) accepted after verification if $\sigma(e,t) \geq \widehat{s}(e)$ and $e < e^n$, or (c) rejected without verification if $\sigma(e,t) < \widehat{s}(e)$ and $e < e^n$, where $\widehat{s}(e)$ is such that for every (e,t), $\sigma(e,t) \geq \widehat{s}(e) \iff u(e,t) \geq c$. The trade-offs are the same as with the optimal mechanism under a talent-biased composite measure, so e^n is equal to the optimal threshold e^* of Proposition 4.

Under a talent-biased composite measure, $\widehat{s}(e)$ is decreasing (i.e., agents presenting more evidence of training are required to meet a lower composite measure threshold to get accepted), so the naive benchmark does not create incentives to withhold evidence. It coincides with the optimal mechanism. On the other hand, under a training-biased composite measure, $\widehat{s}(e)$ is increasing, so the naive benchmark creates incentives to withhold evidence and therefore does not implement the intended outcome. The outcome it actually implements is the following: Each agent (e,t) is (a) accepted without verification if $e \geq e^n$, (b) accepted after verification if $\sigma(e,t) \geq s^n$ and $e < e^n$, or (c) rejected without verification if $\sigma(e,t) < s^n$ and $e < e^n$, where $s^n := \sigma(e_\ell, t_\ell)$, $e_\ell := \min_e\{e : u(e,1) \geq c\}$,

$t_\ell \coloneqq \min_t\{t : u(e_\ell,t) \geq c\}$. s^n is the effective composite measure threshold of the naive benchmark (i.e., the composite measure threshold that an agent who does under-report training needs to meet to get accepted). Figure 5 presents the outcome implemented by the naive benchmark, as well as the optimal acceptance thresholds.

Figure 5: The naive benchmark when the agent has evidence of training

(a) Talent-biased σ & $c \geq 0$ **(b)** Training-biased σ & $c \geq 0$



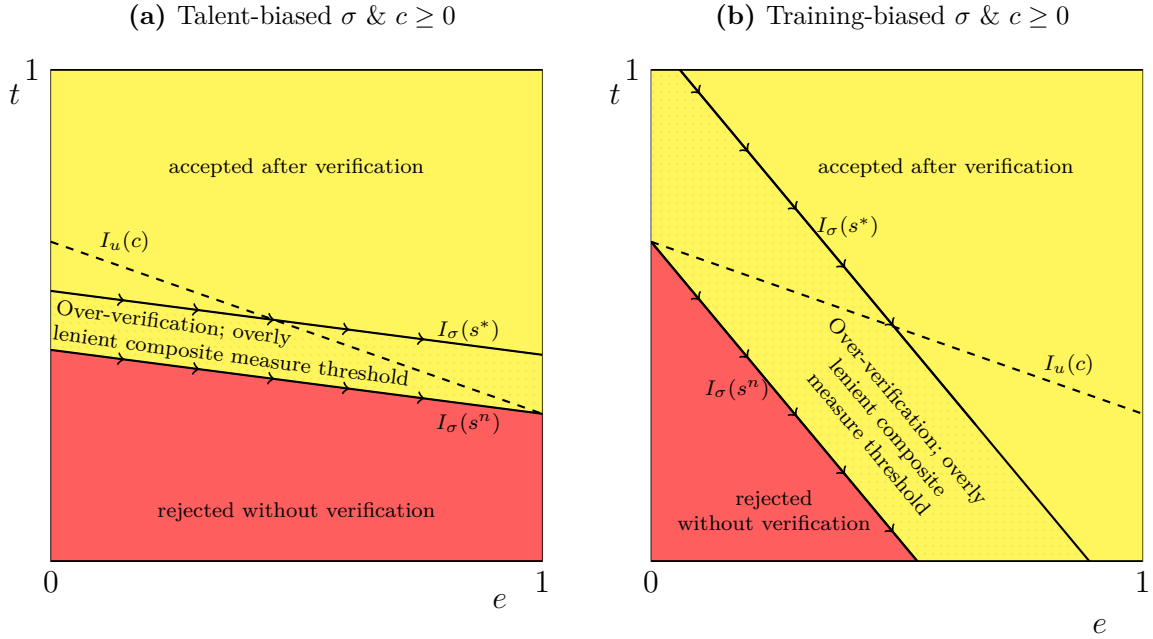
Note: The dashed line $I_u(c)$ is the principal's iso-utility curve at utility level c. The arrowed line represents an iso-composite-measure curve, at an arbitrary level s in panel (a), and at the optimal s^* and naive effective s^n composite measure thresholds in panel (b). The colors denote whether the agent is (i) accepted without verification, (ii) accepted after verification, or (iii) rejected without verification.

Agent without evidence. The outcome that the principal attempts to implement is the following: Each agent (e,t) is (a) accepted after verification if $u(e,t) \geq c$, or (b) rejected without verification if $u(e,t) < c$. He attempts to do so by specifying that for every e, an agent reporting training e is (a) accepted after verification if $\sigma(e,t) \geq \widehat{s}(e)$, where $\widehat{s}(e)$ is such that $\sigma(e,t) \geq \widehat{s}(e) \iff u(e,t) \geq c$, or (b) rejected without verification if $\sigma(e,t) < \widehat{s}(e)$.

Under a talent-biased composite measure, $\widehat{s}(e)$ is decreasing, so the naive benchmark creates incentives to over-report training (and under-report talent). The outcome it implements is the following: Each agent (e,t) is (a) accepted after verification if $\sigma(e,t) \geq s^n$, or (b) rejected without verification if $\sigma(e,t) < s^n$, where $s^n \coloneqq \sigma(e_h, t_h)$, $e_h \coloneqq \max_e\{e : u(e,t) = c \text{ for some } t\}$, t_h is such that $u(e_h,t_h) = c$. s^n is the effective composite measure threshold of the naive benchmark. Analogously, under a training-biased composite measure, $\widehat{s}(e)$ is increasing, so the naive benchmark creates incentives to under-report training (and over-report talent). The outcome it implements is the following: Each agent (e,t)

is (a) accepted after verification if $\sigma(e,t) \geq s^n$, or (b) rejected without verification if $\sigma(e,t) < s^n$, where $s^n \coloneqq \sigma(e_\ell, t_\ell)$, $e_\ell \coloneqq \min_e\{e : u(e,1) \geq c\}$, $t_\ell \coloneqq \min_t\{t : u(e_\ell,t) \geq c\}$. s^n is the effective composite measure threshold of the naive benchmark. Figure 6 presents the outcome implemented by the naive benchmark, as well as the composite measure threshold of the optimal mechanism.

Figure 6: The naive benchmark when the agent does *not* have evidence of training



Note: The dashed line $I_u(c)$ is the principal's iso-utility curve at utility level c. The arrowed line represents an iso-composite-measure curve, at the optimal s^* and naive effective s^n composite measure thresholds. The colors denote whether the agent is accepted after verification or rejected without verification.

6 Discussion and comparisons

This section (i) studies the effects of the verification cost and bias of the composite measure on the principal's decision errors, (ii) compares the optimal mechanism to the naive benchmark, (iii) studies how the optimal composite measure threshold changes depending on whether the agent has evidence of training or not, and (iv) discusses comparative statics of the optimal mechanism.

6.1 Effects of bias and verification cost on decision errors

When the agent has evidence of training, 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. A comparison of Figures 3(a) and 3(c) shows that the verification cost

alone gives rise to such errors. A comparison of 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 and the bias of the composite measure interact in inducing errors favoring high- over low-training agents? Proposition 8 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).

Proposition 8. 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.[24]

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 agent's incentives 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. Thus, fewer agents are accepted without verification under a talent-biased than under a training-biased composite measure.

6.2 Optimal mechanism versus naive benchmark

This section compares the optimal mechanism with the naive benchmark in each case: (i) when the agent has evidence of training and (ii) when she does not.

[24] 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 $e^*_{t-\text{biased}} > e^*_{e-\text{biased}}$.

Agent with evidence of training. As we have seen, under a talent-biased composite measure, the naive benchmark coincides with the optimal mechanism. On the other hand, under a training-biased composite measure, the effective composite measure threshold s^n of the naive benchmark is lower than the optimal one. Instead of finely balancing Type I error A against Type II error A, s^n leads to excessive Type II error A while eliminating Type I error A. The naive benchmark creates incentives for the agent to manipulate the interpretation of the composite measure. As a result, too many agents are accepted after verification. At the same time, from Proposition 8, it follows that the evidence threshold e^n of the naive benchmark is higher than the optimal one. Failing to understand that verification creates incentives to withhold evidence, the principal overestimates the effectiveness of verification. As a result, he under-utilizes evidence in evaluation.

Agent without evidence. Whether the composite measure is training or talent-biased, the effective composite measure threshold s^n of the naive benchmark is lower than the optimal one. Instead of finely balancing Type I against Type II errors, s^n leads to excessive Type II errors while eliminating Type I errors. The naive benchmark creates incentives for the agent to manipulate the interpretation of the composite measure. As a result, too many agents are accepted after verification.

6.3 Effect of evidence on composite measure threshold

When the agent does not have evidence, the principal cannot accept high-training agents without verification. Thus, the risk that an (e,t) agent who is accepted after verification has high training and low talent, delivering payoff $u(e,t) < c$, is high. Therefore, the threshold composite measure required for acceptance is higher in the setting where the agent does not have evidence than in the setting where she does have evidence of training, as can be seen through a comparison of Figures 3(d) and 4(d).

Proposition 9. Take any training-biased composite measure. For any positive optimal composite measure threshold $s^*_{\text{no evidence}}$ when the agent does *not* have evidence of training and any optimal evidence and composite measure thresholds $(e^*_{e-\text{biased}}, s^*_{e-\text{biased}})$ when the agent *has* evidence of training, (i) the composite measure threshold is weakly higher when the agent does *not* have evidence of training: $s^*_{\text{no evidence}} \geq s^*_{e-\text{biased}}$, or (ii) both composite measure thresholds are optimal whether the agent has evidence of training or not.

6.4 Comparative statics

Now I briefly discuss comparative statics for the case where the agent has evidence of training, which are covered in more detail in the Appendix. Analogous results are easy to derive when the agent does not have evidence.

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 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.

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^*.[25] 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. Particularly, both e^* and s^* may decrease in response to an increase in c.

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.[26] 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.

[25] 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.

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

7 Applications

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

7.1 Promotions

An employee's e is her hardworkingness. t is her efficiency, talent, or managerial skills. $\sigma(e,t)$ is her performance in her current position. Depending on the firm and the position, the employee may be able to provide or withhold evidence on e. For example, if we interpret e as the employee's total work hours, the employee can present any amount of evidence $\widehat{e} \in [0,e]$ by choosing to work $\widehat{e}$ hours in the office—so that her employer (or manager) can see her—and $e - \widehat{e}$ hours from home—or in the office without being observed by her employer. The employer can monitor the employee's performance $\sigma(e,t)$ at cost c. 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.[27] 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 means that for the employee's performance 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.

7.2 College admissions and standardized testing

A college applicant's e is her prior training, preparation, parents' education and professions, and parental support. t is the part of 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 $\sigma(e,t)$. In this setting, $c \approx 0$ provided it is not very costly for the college to incorporate the applicant's test score in the evaluation. Therefore, the optimal admissions rule does not allow for test-optional

[27] $u(e,t)$ could also be defined as $u(e,t) \coloneqq v(e,t) - \sigma(e,t) - \underline{q}$, where $\underline{q}$ is the threshold performance differential for the promotion to be beneficial to the firm. For example, $\underline{q}$ could depend on (i) the performance 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.

admissions, regardless of whether applicants have evidence of e or not.[28]

In the admissions rule, if the standardized test is training-biased, admission decisions are flawed at the expense of low-training applicants to the extent that applicants can under-report training. Particularly, if colleges only value talent and, thus, try to control for the applicants' unequal backgrounds, the above problem is necessarily present. If the standardized test is talent-biased, the efficacy of the optimal admissions rule depends on whether applicants have evidence of training or not. If they do not, admission decisions are flawed at the expense of applicants with high training. If applicants have evidence of training, admission decisions do not favor any applicant.

The case where the standardized test is training-biased seems most relevant.[29] It can explain the College Board's decision to begin providing free test preparation to "level the playing field" (College Board, 2014) as a way of limiting variance in training. Indeed, unobserved heterogeneity in e is crucial for the errors against low-training applicants to arise under a training-biased standardized test. Similarly, Frankel and Kartik (2019) explain this policy as a reduction in the heterogeneity of the applicants' "gaming ability." The presented model can also explain the College Board's redesign of the SAT test with the goal to "rein in the intense coaching and tutoring on how to take the test that often gave affluent students an advantage" (Lewin, 2014). The less training-biased the SAT test, the fewer the errors in college admissions (see section 6.4).

7.3 Hiring and poaching

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 at cost c. Alternatively, the prestigious employer can directly hire the candidate without work experience. Poaching is costly because it is more expensive to poach the candidate than hire her from the beginning (e.g., due to job switching costs or a non-compete clause in the contract offered by the less prestigious firm).[30] Also, the cost c may arise because it is costly to monitor

[28] Given that in our setting, absent the verification cost, the principal does not accept the agent without verification, the model cannot explain test-optional admissions. Indeed, given that a college can decide how to use test scores in its admission decisions, alternative assumptions are needed for explaining test-optional admissions (Dessein et al., 2025b,a).

[29] Remember that when the composite measure is training-biased and $c = 0$, it does not matter whether the agent can over-report e. The optimal mechanism is the same whether she has evidence of e or not. Also, to explain the College Board's decision, it is important that applicants can under-report training.

[30] This application departs somewhat from the model of section 2. The prestigious employer pays the poaching cost c for observing the candidate's performance *and* then poaching her—not for just observing her performance. However, the Online Appendix shows that the results go through under this alternative modeling assumption. Intuitively, this is the case because the principal has no reason to verify an agent (e,t)'s composite measure without accepting her. Verification is useful for preventing agent (e',t') from

the (other firm's) employee's performance.

In the optimal mechanism with evidence of training, high-training candidates—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. 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.

An increase in the poaching cost c can be interpreted as a (strengthening of a) non-compete clause in the contract offered by the less prestigious employer (e.g., an increase in the non-compete buyout amount). The analysis of section 6.4 suggests that non-compete clauses may be counterproductive in their goal of improving employee retention. The direct effect of a non-compete clause in the contract offered by the less prestigious firm is indeed to enhance employee retention. Namely, all else fixed, an increase in c causes s^* to rise (when the optimal mechanism involves a composite measure threshold) and an upward vertical shift of the curve $I_u(c)$. Thus, the set of employees who are poached tends to shrink regardless of (i) whether employees have evidence of training or not and (ii) whether performance in the less prestigious firm is training- or talent-biased.

However, when workers have evidence of training, the non-compete tends to reduce e^*, thereby shrinking the set of workers who are not hired by the prestigious firm in the first place. Overall, when performance in the less prestigious firm is talent-biased, the non-compete makes it harder for the firm to attract (i.e., e^* decreases) but easier to retain employees (i.e., $I_u(c)$ shifts upward). However, when performance in the less prestigious firm is training-biased, the non-compete can make it harder for the firm not only to attract (i.e., e^* decreases) but also to retain employees (i.e., s^* decreases). That is, because as the prestigious firm reduces the training requirements e^* to hire the candidate from the beginning, it also reduces the risk that a well-performing employee of the less prestigious firm is well-trained but untalented (see section 6.4). Therefore, poaching the employee after she has worked in the training-biased position becomes more attractive for the prestigious firm.

7.4 Academic job market talks

An academic job market candidate's research topic is comprised of a "mass" $b > 1$ of (uncountably infinitely many) problems.[31] $e \in [0,1]$ is the candidate's knowledge, the mass

imitating (e,t) when (e,t) is accepted. If (e,t) is rejected, (e',t') has no reason to imitate (e,t) in the first place, so verification of (e,t)'s composite measure is unnecessary.

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

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.[32] 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 studied in section 3.

In deciding whether to hire the candidate, the committee should recognize that the candidate might save some results she has derived to use them in answering questions. Failing to recognize this could effectively make the hiring standards too low by allowing candidates to use results they have already derived to appear adept at thinking on their feet (see section 6.2).

8 Conclusion

Candidates for a position are often evaluated through costly tests or indices that measure a combination of multiple underlying qualities. For example, a performance metric of an employee under consideration for a promotion (or for poaching by another employer) depends not only on "how hard" but also on "how smart" the employee works. Standardized college admission tests measure a combination of talent and preparation. The following problem may then arise when the evaluation relies on such a composite measure: The candidate may understate one of her valuable qualities (e.g., training) to make the evaluator attribute the composite measure to another quality instead (e.g., talent). This can happen even though the evaluator values training, and even if the candidate has hard evidence of training. This paper has shown how the optimal evaluation scheme takes into account the candidate's incentives to understate a valuable quality. It has also studied the consequences of failing to take these incentives into account.

Future work can build on the model developed in this paper to study, for example,

[32] 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.

the problem of designing tests and performance metrics taking into account the perverse incentives that may arise when these tests and performance metrics measure a combination of an agent's qualities. Methodologically, future work can examine how the tools used in this paper can be employed in studying multidimensional screening problems with transparent agent motives, allowing for transfers or a continuous choice variable by the principal.

The results open avenues also for empirical work. They could help explain the limited effectiveness of performance management processes (Pulakos et al., 2019): When performance metrics are biased towards hardworkingness, their informativeness is undermined by the employees' incentives to understate how hard they work. The results could also help us understand the heterogeneity across job positions of the effect of working long hours on promotion chances (Frederiksen et al., 2026): In positions where performance metrics are biased towards hardworkingness, their informativeness is limited. This means that the optimal promotion scheme is more likely to rely on workers proving how hard they work. Further, future work can study whether gender differences in honesty (Graf et al., 2026) affect career progression by giving rise to gender-specific rates of strategic understatement of valuable qualities. Last, the results suggest that hiding one's effort might be more common among younger people, whose talent has not been revealed through successive stages of evaluation in their studies and career. Indeed, there is some evidence pointing in this direction: Hiding effort has been identified as an expression of perfectionistic self-presentation and narcissism (Flett et al., 2016; Beattie et al., 2017), which decrease with age (Weidmann et al., 2023; Orth et al., 2024).

References

Akers, M. D., Horngren, C. T., and Eaton, T. V. (1998). Underreporting chargeable time: A continuing problem for public accounting firms. *Journal of Applied Business Research*, 15(1):13–20.

Aksoy, C. G., Barrero, J. M., Bloom, N., Davis, S. J., Dolls, M., and Zarate, P. (2022). Working from home around the world. *Brookings Papers on Economic Activity*, 2022(2):281–360.

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

Barrero, J. M., Bloom, N., and Davis, S. J. (2023). The evolution of work from home. *Journal of Economic Perspectives*, 37(4):23–49.

Beattie, S., Dempsey, C., Roberts, R., Woodman, T., and Cooke, A. (2017). The

moderating role of narcissism on the reciprocal relationship between self-efficacy and performance. *Sport, Exercise, and Performance Psychology*, 6(2):199–214.

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

Bloom, N., Han, R., and Liang, J. (2024). Hybrid working from home improves retention without damaging performance. *Nature*, 630(8018):920–925.

Bloom, N., Liang, J., Roberts, J., and Ying, Z. J. (2015). Does working from home work? Evidence from a Chinese experiment. *The Quarterly journal of economics*, 130(1):165–218.

Blum, E. S., Hatfield, R. C., and Houston, R. W. (2022). The effect of staff auditor reputation on audit quality enhancing actions. *The Accounting Review*, 97(1):75–97.

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

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

College Board (2014). The college board announces bold plans to expand access to opportunity; redesign of the SAT. Press Release.

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

Deloitte (2024). Audit quality report. Technical report, Deloitte US.

Deloitte (2025). Global principles of business conduct. Technical report, Deloitte.

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. *AEA Papers and Proceedings*, 115:669–75.

Dua, A., Ellingrud, K., Kirschner, P., Kwok, A., Luby, R., Palter, R., and Pemberton, S. (2022). Americans are embracing flexible work—and they want more of it.

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

Ernst & Young (2025). Audit quality report. Technical report, Ernst & Young LLP.

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

Feltovich, N., Harbaugh, R., and To, T. (2002). Too cool for school? Signalling and countersignalling. *RAND Journal of Economics*, pages 630–649.

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.

Frederiksen, A., Kato, T., and Smith, N. (2026). Working hours, top management appointments, and gender: Evidence from linked employer-employee data. *Forthcoming at Journal of Labor Economics*.

Frick, M., Iijima, R., and Ishii, Y. (2026). Multidimensional screening with precise seller information. *Econometrica*, 94(1):35–70.

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

Garmaise, M. J. (2011). Ties that truly bind: Noncompetition agreements, executive compensation, and firm investment. *The Journal of Law, Economics, & Organization*, 27(2):376–425.

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.

Graf, C., Pondorfer, A., and Schulz, J. (2026). Culture and gender differences in honesty. *Forthcoming at The Economic Journal*.

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.

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.

Jacobs, E. (2024). Why playing down a privileged background might be a savvy career move. *Financial Times*.

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.

KPMG (2025a). Audit quality report. Technical report, KPMG International Limited.

KPMG (2025b). Code of Conduct. Technical report, KPMG International Limited.

Lewin, T. (2014). A new SAT aims to realign with schoolwork. *The New York Times*.

Lipnowski, E. and Ravid, D. (2020). Cheap talk with transparent motives. *Econometrica*, 88(4):1631–1660.

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.

McMullin, C. (2025). Work-life balance, talent attraction and retention top reasons to offer flexible work arrangements. *Benefits Quarterly*, 41(4):49–50.

Mendoza, K. I. (2020). Reducing underreporting by aggregating budgeted time. *The Accounting Review*, 95(5):299–319.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 12(2):380–391.

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):643–665.

Otley, D. T. and Pierce, B. J. (1996). The operation of control systems in large audit firms. *Auditing*, 15(2):65–84.

PCAOB (2024). Firm and engagement metrics. Technical report.

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

Pulakos, E. D., Mueller-Hanson, R., and Arad, S. (2019). The evolution of performance management: Searching for value. *Annual Review of Organizational Psychology and Organizational Behavior*, 6:249–271.

PwC (2024). Code of conduct. Technical report, PricewaterhouseCoopers LLP.

PwC (2025). Audit quality report. Technical report, PricewaterhouseCoopers LLP.

Reeves, A. and Friedman, S. (2024). *Born to rule: The making and remaking of the British elite*. Harvard University Press.

Shapeero, M. and Killough, L. N. (1993). Underreporting of chargeable hours in public accounting: An ethical dilemma. *Journal of Managerial Issues*, pages 373–386.

Shin, H. S. (1994). News management and the value of firms. *The RAND Journal of Economics*, 25(1):58–71.

Smith, W. R., Hutton, M. R., and Jordan, C. E. (1996). Underreporting time: Accountants are doing it more and enjoying it less. *The CPA Journal*, 66(10):67–70.

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

Sweeney, B. and Pierce, B. (2006). Good hours, bad hours and auditors' defence mechanisms in audit firms. *Accounting, Auditing & Accountability Journal*, 19(6):858–892.

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

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):1277–1298.

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

Appendix

A Characterization of optimal thresholds

This section discusses how the principal chooses thresholds in the optimal mechanism.

A.1 Talent-biased composite measure; agent with evidence

The principal chooses threshold $e^* \in \arg\max_{e_{min}\in[0,1]} v^{t-\text{biased}}(e_{min})$,[33] 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 $v_e^{t-\text{biased}}(e^*) = -\int_0^1 \min\{u(e^*,t),c\}f(e^*,t)dt = 0$, or equivalently

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

The cross-partial derivative $\partial v_e^{t-\text{biased}}/\partial c$ is non-positive, so by Topkis' monotonicity theorem, the set of optimal evidence thresholds is decreasing in c in the strong set order. If $e^* \in (0,1)$ is unique with the second-order condition of the principal's problem satisfied strictly and verification is used for a positive measure of agents,[34] $\partial v_e^{t-\text{biased}}(e^*)/\partial c = -\int_0^1 \mathbf{I}(u(e^*,t)\geq c)f(e^*,t)dt < 0$, and thus $de^*/dc = -\partial v_e^{t-\text{biased}}(e^*)/\partial c / v_{ee}^{t-\text{biased}}(e^*) < 0$, so e^* is decreasing in c. Also, the principal's optimal payoff is decreasing in c.

[33] 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.

[34] 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; agent with evidence

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

$$v^{e-\text{biased}}(e_{min},s_{min}) \coloneqq \overbrace{\int_{s_{min}}^{1}\int_{\underline{e}(s)}^{\max\{\min\{\overline{e}(s),e_{min}\},\underline{e}(s)\}} (\widetilde{u}(e,s)-c)\,\widetilde{f}(e,s)deds}^{\text{payoff from agents accepted after verification net of verification costs}} + \underbrace{\int_{0}^{1}\int_{\max\{\underline{e}(s),e_{min}\}}^{\max\{\overline{e}(s),e_{min}\}} \widetilde{u}(e,s)\widetilde{f}(e,s)deds}_{\text{payoff from agents accepted without verification}},$$

where $\widetilde{u}(e,s) \coloneqq u(e,\tau(e,s))$ and $\widetilde{f}(e,s) \coloneqq f(e,\tau(e,s))\partial\tau(e,s)/\partial s$.[35]

When $\underline{e}(s^*) < e^*$ (i.e., some agents are accepted after verification) and $e^*,s^* \in (0,1)$,[36] the first-order conditions with respect to e_{min} and s_{min} are, respectively,

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

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

Assume $s^*,e^* \in (0,1)$ 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. $v_{es}^{e-\text{biased}}(e^*,s^*) = v_{se}^{e-\text{biased}}(e^*,s^*) = -(\widetilde{u}(e^*,s^*)-c)\,\widetilde{f}(e^*,s^*) > 0$, and the total derivatives of e^* and s^* with respect to c are:

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

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

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

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

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

A.3 Agent without evidence

Consider the case where the agent cannot present evidence of training (or talent). The principal chooses threshold $s^* \in \arg\max_{s_{min}\in[0,1]} v^{\text{no evidence}}(s_{min})$, where

$$v^{\text{no evidence}}(s_{min}) \coloneqq \overbrace{\mathbf{I}(s_{min}>0)\int_{s_{min}}^1 \int_{\underline{e}(s)}^{\overline{e}(s)} (\widetilde{u}(e,s)-c)\widetilde{f}(e,s)deds}^{\text{payoff from agents accepted after verification net of verification costs}} + \underbrace{\mathbf{I}(s_{min}=0)\int_0^1 \int_{\underline{e}(s)}^{\overline{e}(s)} \widetilde{u}(e,s)\widetilde{f}(e,s)deds}_{\text{payoff from agents accepted without verification}}.$$

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

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

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_{av}. Construct the mechanism $M' \equiv \langle T',P'\rangle$ with (i) $P'_{av}(e,t) = 1$, (ii) $T'(e,t) = T(e,t)P_{av}(e,t) \leq T(e,t)$, and (iii) $P'(e,t,\emptyset) = (1-T(e,t))P(e,t,\emptyset)/(1-T'(e,t))$ for any (e,t).[37] We have then that (a) $T'(e,t)P'_{av}(e,t) = T(e,t)P_{av}(e,t)$, (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 $P_{av}(e,t)<1$. **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

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

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).

[37] 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.

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, $\Pi(e,t) \geq \Pi(e,0) \geq (1-T(e',t'))P(e',t',\emptyset)$ for any (e',t') with $e' \leq e$ and $\Pi(e,t)$ is non-decreasing in t. Therefore, (4) is equivalent to

$$e \in \underset{r\in[\underline{e}(\sigma(e,t)),e]}{\arg\max}\ \Pi(r,\tau(r,\sigma(e,t))). \tag{5}$$

Thus, incentive-compatibility is satisfied for every type if and only if for every (e,t), (5) 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 (5) 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, $r_2 \notin \arg\max_{x\in[\underline{e}(\sigma(e,t)),r_2]} \Pi(x,\tau(x,\sigma(e,t)))$, 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),[38] the second from condition (i), and the third from condition (iii). **Q.E.D.**

[38]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[39]

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

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. 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 (see, e.g., 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

[39]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. $u(e,t) \leq c = 0$) 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 $S \coloneqq (e_{\overline{s}},e_{\underline{s}}) \cap [\underline{e}(\overline{s}),\overline{e}(\overline{s})] \cap [\underline{e}(\underline{s}),\overline{e}(\underline{s})]$.

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

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

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.[40] 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 Lemma 3, there exists an optimal mechanism with $\Pi(e_s,t_s) = \mathbf{I}(s \geq s^*)$ for some $s^* \in [0,1]$.

[40]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

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

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 $\Pi(e,t) = \Pi(e,0) + T(e,t)$. **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

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

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) such that $e < e^*$. 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 $\Pi(e,t) = \Pi(e,0) + T(e,t)$. **Q.E.D.**

Proof of Proposition 6. Denote the total probability with which type (e,t) is accepted if she reports (e',t') (with $e' \leq e$) by

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

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

$$\max_{(e',t') \leq (1,1)} \left[(1-T(e',t'))P(e',t';\emptyset) + T(e',t')\mathbf{I}\left(\sigma(e,t) \geq \sigma(e',t')\right)\right] = \Pi(e,t). \tag{6}$$

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') and $\Pi(e,t)$ is non-decreasing in t. Therefore, (6) is equivalent to

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

Thus, incentive-compatibility is satisfied for every type if and only if for every (e,t), (7) is satisfied. This is true if and only if $\Pi(r,\tau(r,\sigma(e,t)))$ is 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[41]

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

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

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

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 6. 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 6. 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 7. Straightforward and thus omitted.

Proof of Proposition 8. We use the objective functions defined in section A. Fix some training-biased composite measure σ and take any s^* such that for some e^*, $(e^*,s^*) \in \arg\max_{(e_{min},s_{min})} v^{e-\text{biased}}(e_{min},s_{min})$ under σ. Define

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

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 $e^*_{e-\text{biased}} \in \arg\max_{e_{min}} v(e_{min},0) = \arg\max_{e_{min}} v^{e-\text{biased}}(e_{min},s^*)$. We have that

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

so $v(e_{min},\alpha)$ has increasing differences in (e_{min},α), and Topkis' Monotonicity Theorem implies that $\arg\max_{e_{min}} v(e_{min},\alpha)$ is increasing in α in the strong set order. Therefore, (i) $e^*_{t-\text{biased}} \geq e^*_{e-\text{biased}}$ 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 $e^*_{e-\text{biased}} \in \arg\max_{e_{min}} v^{t-\text{biased}}(e_{min})$. 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 9. First, fix some training-biased composite measure σ and notice that for any $s_{min} > 0$, $v^{\text{no evidence}}(s_{min}) = v^{e-\text{biased}}(1,s_{min})$. Now, take any $s^*_{\text{no evidence}} \in \arg\max_{s_{min}} v^{\text{no evidence}}(s_{min}) = \arg\max_{s_{min}} v^{e-\text{biased}}(1,s_{min})$, assuming $s^*_{\text{no evidence}} > 0$, and any $(e^*_{e-\text{biased}},s^*_{e-\text{biased}}) \in \arg\max_{(e_{min},s_{min})} v^{e-\text{biased}}(e_{min},s_{min})$. For every $e \in [0,1]$, define

$\overline{s}(e) \coloneqq \max\{s \in [0,1] : \widetilde{u}(e,s) \leq c\} \cup \{0\}$. Define also

$$\psi(e_{min}) \coloneqq \underset{s_{min} \in [0,\overline{s}(e_{min})]}{\arg\max} v^{e-\text{biased}}(e_{min},s_{min}) = \underset{s_{min} \in [0,1]}{\arg\max}\, v^{e-\text{biased}}(e_{min},s_{min}),$$

where the equality follows because for $s > \overline{s}(e_{min})$, $v_s^{e-\text{biased}}(e_{min},s) < 0$. Notice that $s^*_{\text{no evidence}} \in \psi(1)$ and $s^*_{e-\text{biased}} \in \psi(e^*_{e-\text{biased}})$. Also, for every $e_{min} \in [e^*_{e-\text{biased}},1]$ and $s_{min} \in [0,\overline{s}(e_{min})]$

$$\frac{\partial v^{e-\text{biased}}(e_{min},s_{min})}{\partial s_{min}} = -\int_{\underline{e}(s_{min})}^{\max\{\min\{\overline{e}(s_{min}),e_{min}\},\underline{e}(s_{min})\}} (\widetilde{u}(e,s_{min}) - c)\, \widetilde{f}(e,s_{min}) de,$$

which is non-decreasing in e_{min}, since by definition $\overline{s}(e_{min})$, $\widetilde{u}(e_{min},s_{min}) \leq c$ for every e_{min} and $s_{min} \in [0,\overline{s}(e_{min})]$. Also, $\overline{s}(e_{min})$ is non-decreasing in e_{min} because σ is training-biased. Topkis' Monotonicity Theorem implies that $\psi(e_{min})$ is increasing in e_{min} in the strong set order. **Q.E.D.**

Online Appendix

Multidimensional screening of strategic agents

Orestis Vravosinos

C Implementation of the optimal mechanism

In the paper, we have without loss restricted attention to truthful mechanisms. However, the optimal mechanism can sometimes be implemented in simpler ways.

Implementation with evidence of training. When the composite measure is training-biased and the agent has evidence of training, 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 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).

When the composite measure is talent-biased and the agent has evidence of training, a similarly simple implementation of the optimal mechanism 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_ℓ,e_h,t_ℓ,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).

Implementation without evidence. When the agent does not have evidence, the principal can implement the optimal mechanism by offering the agent a single path to getting accepted: ask the principal to verify your composite measure, and if it is at least s^*, you will be accepted.

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

I study optimal screening when the agent can also present evidence of 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.[42] Given also the results of the baseline model, we conclude that when the agent

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

can provide evidence of training but not of talent, and the principal can only imperfectly try to verify talent 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. When the agent can provide evidence of neither training nor talent, the principal is constrained in his evaluation of the agent as long as the composite measure is not perfectly aligned with the principal's preferences.

E The naive optimal mechanism

This section studies the naive optimal mechanism; that is, the mechanism implemented by the principal if he wrongly believes that the agent is *non-strategic* in the sense that she never understates any of her qualities when in reality she is strategic (i.e., potentially under-reports talent and/or training). In this section, I assume that $e \mapsto \mathbb{E}_t\left[\min\{u(e,t), c\} \middle| e\right]$ does not cross zero from above. This means that if for a certain level of training e, the principal prefers to (i) accept every agent with training e without verification rather than (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 $u(e,t) < c$, then for any higher level of training $e' > e$, the principal still prefers option (i) (i.e., to accept every agent with training e' without verification) over option (ii). Particularly, I assume that there exists e^* such that $\operatorname{sgn}\left\{\mathbb{E}_t\left[\min\{u(e,t), c\} \middle| e\right]\right\} = \operatorname{sgn}\{e - e^*\}$.[43]

This assumption simplifies the exposition by isolating the main force we are interested in. It guarantees that when the naive optimal mechanism is used, the only reason why an agent may under-report one of her qualities is to over-report the other quality by influencing how the principal interprets the composite measure. Without this assumption, an agent faced with the naive optimal mechanism may under-report training to get accepted without verification.

E.1 The optimal mechanism against a non-strategic agent

We need to derive the optimal mechanism when the agent is indeed *non-strategic* (i.e., always truthfully reports her training). From this analysis, it then follows that the outcome implemented by the naive optimal mechanism (i.e., the mechanism when the principal wrongly believes that the agent is *non-strategic* when in reality she is strategic) is as described in section 5.

[43] A sufficient condition is that $\mathbb{E}_t\left[\min\{u(e,t), c\} \middle| e\right]$ 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.

E.1.1 The optimal mechanism against a non-strategic agent with evidence of training

Since the agent has evidence of training and never under-reports, e is effectively public information. 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 $\mathbb{E}_t\left[\min\{u(e,t), c\} \middle| e\right]$ the expectation of $\min\{u(e,t), c\}$ conditional on e. When $\mathbb{E}_t\left[\min\{u(e,t), c\} \middle| e\right] > 0$, option (i) delivers a higher payoff to the principal, while when $\mathbb{E}_t\left[\min\{u(e,t), c\} \middle| e\right] < 0$, option (ii) is superior. Clearly, when $c = 0$, option (ii) is superior.

Claim 3 describes the optimal mechanism.

Claim 3. In the optimal mechanism against a non-strategic agent with evidence of training, $\Pi(e,t) = \mathbf{I}(u(e,t) \geq c \text{ or } e \geq e^*)$, $T(e,t) = \mathbf{I}(u(e,t) \geq c \text{ and } e < e^*)$, and $P(e,t,\emptyset) = \mathbf{I}(e \geq e^*)$.

E.1.2 The optimal mechanism against a non-strategic agent without evidence

Compared to the case where the agent has evidence of training, now the principal has the additional constraint that he cannot switch from option (i) to option (ii) as e increases as that would cause agents to over-report training. Claim 4 describes the optimal mechanism.

Claim 4. In the optimal mechanism against a non-strategic agent without evidence, either

(i) every agent is accepted without verification, or

(ii) $\Pi(e,t) = T(e,t) = \mathbf{I}(u(e,t) \geq c)$; that is, each agent (e,t) is (a) accepted after verification if $u(e,t) \geq c$ or (b) rejected without verification if $u(e,t) < c$.

F 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 7.3), 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 7.1), 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 7.2), 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 he 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.[44]

As shown in section 6.4, as long as the agent has evidence of training and 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 4, 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.

G 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 7.3), an individual employer has limited labor market power to influence a candidate's effort to obtain credentials. Similarly, in college admissions (section 7.2), a single college has little influence over applicants' preparation for standardized tests.[45]

Let the agent's talent t follow a distribution with density g and support [0,1]. Taking

[44]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.

[45]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 7.1), 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.

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.[46] 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.

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

We now generalize the results allowing for multiple dimensions of training and talent. I focus on the more challenging case where the agent has evidence of training. The results for the case where the agent does not have evidence of training also generalize.

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 $(t_1,t_2,\ldots,t_n)$ 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 $\boldsymbol{e},\boldsymbol{t},\boldsymbol{t}'$

[46] Under a talent-biased composite measure, there is only an evidence threshold. When the agent has no evidence of training, there is only a composite measure threshold.

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 $T(\boldsymbol{e}',\boldsymbol{t}') > \inf_{(\boldsymbol{e}'',\boldsymbol{t}'') \in X(\boldsymbol{e}',s)} T(\boldsymbol{e}'',\boldsymbol{t}'')$ (and thus $P(\boldsymbol{e}'',\boldsymbol{t}',\emptyset) < \sup_{(\boldsymbol{e}'',\boldsymbol{t}'') \in X(\boldsymbol{e}',s)} P(\boldsymbol{e}',\boldsymbol{t}'',\emptyset)$) keeping Π fixed. 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 $\sigma(\boldsymbol{e},\boldsymbol{t}) = \sigma(\boldsymbol{e},\boldsymbol{t}')$.[47] 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 $\boldsymbol{e} \in \{\boldsymbol{e} \in [0,1]^m : \sigma(\boldsymbol{e},\boldsymbol{0}) \leq s \leq \sigma(\boldsymbol{e},\boldsymbol{1})\}$ for every $s \in [0,1]$, and

[47] 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.

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

where $\Pi(\boldsymbol{e},s) \coloneqq (1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset) + T(\boldsymbol{e},s)$ is the probability with which 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 $(1-T'(\boldsymbol{e},s))P'(\boldsymbol{e},s,\emptyset) = \Pi'(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))$ for every $(\boldsymbol{e},s)$, $s \in [\sigma(\boldsymbol{e},\mathbf{0}),\sigma(\boldsymbol{e},\mathbf{1})]$ that is outcome-equivalent to M and has at most as high verification costs as M. For $c > 0$, if also $\Pi(\boldsymbol{e},s) > \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))$ for a positive measure of $(\boldsymbol{e},s)$'s, then M' has lower verification costs than M.

Define $\widetilde{f}(\boldsymbol{e},s) \coloneqq \int_{\boldsymbol{t} \in [0,1]^n} \mathbf{I}(\sigma(\boldsymbol{e},\boldsymbol{t}) = s) f(\boldsymbol{e},\boldsymbol{t}) d\boldsymbol{t}$, the probability density of agents with training $\boldsymbol{e}$ and composite measure s, and $\widetilde{u}(\boldsymbol{e},s) \coloneqq \mathbb{E}_{\mathbf{t}}[u(\boldsymbol{e},\boldsymbol{t}) | \sigma(\boldsymbol{e},\boldsymbol{t}) = s] = \int_{\mathbf{t} \in [0,1]^n} u(\mathbf{e},\mathbf{t}) \mathbf{I}(\sigma(\mathbf{e},\mathbf{t}) = s) f(\mathbf{e},\mathbf{t}) d\mathbf{t} / \widetilde{f}(\boldsymbol{e},s)$, 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.[48] 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

$$\int_0^1 \cdots \int_0^1 \int_{\sigma(\boldsymbol{e},\mathbf{0})}^{\sigma(\boldsymbol{e},\mathbf{1})} [\Pi(\boldsymbol{e},s)(\widetilde{u}(\boldsymbol{e},s) - c) + c\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))] \widetilde{f}(\boldsymbol{e},s) ds de_1 \cdots de_m, \tag{8}$$

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.

H.1 Talent-biased composite measure

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

[48] 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, $I_u(\overline{u}) \coloneqq \{(\boldsymbol{e},s) \in [0,1]^{m+1} : \widetilde{u}(\boldsymbol{e},s) = \overline{u}\}$, 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.

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 $\boldsymbol{e}' \in E^*$).[49]

H.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 $\widetilde{u}(\boldsymbol{e},s) \geq c \geq \widetilde{u}(\boldsymbol{e}',s)$ 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 $[0,1]^m$.[50]

I Verification cost paid only after acceptance

In this section, I show that the analysis of sections 3 and 4 goes through if we assume that the principal needs to pay the verification cost only when the value of the composite measure is verified *and* the agent is accepted. The principal now maximizes

$$\int_0^1 \int_0^1 \left\{ \begin{bmatrix} T(\phi(e,t))P(\phi(e,t),\sigma(e,t)) \\ +[1-T(\phi(e,t))]P(\phi(e,t),\emptyset) \end{bmatrix} u(e,t) - cT(\phi(e,t))P(\phi(e,t),\sigma(e,t)) \right\} f(e,t)dtde$$

under the same IC constraints as in sections 3 and 4.

[49] 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 $\widetilde{u}(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) > c$.

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

The optimal mechanism with evidence of training. For the same reasons as in section 3, truthful mechanisms with

$$P(e,t,s) = \begin{cases} 0 & \text{if } s < \sigma(e,t) \\ P_{av}(e,t) & \text{if } s \geq \sigma(e,t) \end{cases} \tag{9}$$

for any (e,t) and some $P_{av} : [0,1]^2 \to [0,1]$ are still without loss. Also, Lemma 1 still applies but now, because the verification cost is paid only after acceptance, there are no verification cost savings from setting $P_{av}(e,t) = 1$.

Lemma 9. 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 $P'_{av}(e,t) = 1$ for every (e,t) that is outcome-equivalent to M and has the same verification costs as M.

Therefore, the principal's objective function can be written as $\int_0^1 \int_0^1 [\Pi(e,t)u(e,t) - cT(e,t)] f(e,t)dtde$, as in section 3. Also, since the IC constraint is the same, Proposition 1 still characterizes SIC mechanisms, and the analysis of section 3 (and, thus, also of section 6) goes through.

The optimal mechanism without evidence. Similarly, when the agent does not have evidence, we can restrict attention to SIC mechanisms, which Proposition 6 characterizes, so the principal's objective function can be written as $\int_0^1 \int_0^1 [\Pi(e,t)u(e,t) - cT(e,t)] f(e,t)dtde$, as in section 4. Therefore, the analysis of section 4 goes through.

J Incentive compatibility with composite measure manipulation

In this section, I allow the agent to downward manipulate her composite measure and the principal to choose acceptance probability $P(e,t,s)$ that is not required to be non-decreasing in $s \in [0,1]$. The agent can choose her composite measure $s \in [0, \sigma(e,t)]$. By paying cost $c \geq 0$, the principal can observe s. The principal designs a mechanism $M \equiv \langle T, P \rangle$, where $T : [0,1]^2 \to [0,1]$ and $P : [0,1]^2 \times ([0,1] \cup \{\emptyset\}) \to [0,1]$,[51] and (breaking the agent's indifferences in his favor) agent response rules $\phi : [0,1]^2 \to [0,1]^2$ and $\chi : [0,1]^2 \to [0, \sigma(1,1)]$ to maximize

$$\int_0^1 \int_0^1 \left\{ \begin{bmatrix} T(\phi(e,t))P(\phi(e,t), \chi(e,t)) \\ +[1 - T(\phi(e,t))]P(\phi(e,t), \emptyset) \end{bmatrix} u(e,t) - cT(\phi(e,t)) \right\} f(e,t)dtde$$

[51] Notice that the message space $[0,1]^2$ is already rich enough. There is no gain from asking the agent to also report her composite measure s.

subject to the agent's IC constraint. If the agent has evidence of training, the IC constraint is

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

If he does not have evidence of training, the IC constraint is

$$(\phi(e,t),\chi(e,t)) \in \underset{(e',t',s)\in[0,1]^2\times[0,\sigma(e,t)]}{\arg\max} \{T(e',t')P(e',t',s) + (1 - T(e',t'))P(e',t',\emptyset)\}.$$

Lemma 10 shows that the principal can without loss restrict attention to truthful mechanisms where non agent manipulates her composite measure and for every (e,t), $P(e,t,s)$ is non-decreasing in s. Particularly, he can constrain attention to mechanisms where $P(e,t,s)$ is a threshold acceptance policy.

Lemma 10. Take any mechanism $M \equiv \langle T,P\rangle$ and agent response rules $\phi : [0,1]^2 \to [0,1]^2$ and $\chi : [0,1]^2 \to [0,1]$, and construct mechanism $M' \equiv \langle T',P'\rangle$ as follows:

$$\begin{aligned} P'(e,t,s) &= \begin{cases} P(\phi(e,t),\chi(e,t)) & \text{if } s \geq \sigma(e,t), \\ 0 & \text{if } s < \sigma(e,t), \end{cases} \\ P'(e,t,\emptyset) &= P(\phi(e,t),\emptyset), \text{ and} \\ T'(e,t) &= T(\phi(e,t)) \end{aligned}$$

for every $(e,t,s) \in [0,1]^3$. If $\phi : [0,1]^2 \to [0,1]^2$ and $\chi : [0,1]^2 \to [0,1]$ are IC under mechanism M, then M' truthfully implements the same outcome as M. Namely, for every (e,t),

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

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 *and* the agent cannot manipulate her composite measure, consider the case where verification is free and the composite measure is training-biased. $\Pi(e,t) = \mathbf{I}(\mathbb{E}[u(x,y)|\sigma(x,y) = \sigma(e,t)] \geq 0)$ 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.

K Incentive compatibility with continuous allocation space

Consider the case where the principal's ultimate choice lies in a closed real interval and the agent has evidence of training. The agent's type $(e,t) \in [0,1]^2$ has a full-support density $f : [0,1]^2 \to \mathbb{R}_{++}$. By paying a cost $c \geq 0$, the principal can observe the value of $\sigma(e,t) \in [0,1]$. Ultimately, the principal must choose an action $p \in [0,1]$. The principal's Bernoulli payoff is $u(e,t,p)$. The agent's Bernoulli payoff is increasing in p; we normalize it to be equal to p.

The principal commits to a direct mechanism $M \equiv \langle T,P \rangle$ that specifies: (i) the probability $T(e,t) \in [0,1]$ of verification and (ii) the action $P(e,t,s)$, which must be non-decreasing in $s \in [0,1]$, that the principal will choose after the agent has presented evidence e and sent cheap talk message t, and the composite measure is $s \in [0,1]$. If the composite measure is not verified, $s = \emptyset$ and the principal chooses action $P(e,t,\emptyset)$. Overall, the principal designs mechanism $M \equiv \langle T,P \rangle$ and an agent response rule $\phi : [0,1]^2 \to [0,1]^2$ to maximize

$$\int_0^1 \int_0^1 \left\{ \begin{bmatrix} T(\phi(e,t))u(e,t,P(\phi(e,t),\sigma(e,t))) \\ +[1-T(\phi(e,t))]u(e,t,P(\phi(e,t),\emptyset)) \end{bmatrix} - cT(\phi(e,t)) \right\} f(e,t)dtde$$

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

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

For the same reasons as in the baseline model, it is without loss to restrict attention to SIC mechanisms.

Definition 8. A mechanism $M \equiv \langle T,P \rangle$ is simply incentive-compatible (SIC) if it is truthful and

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

for every $(e,t) \in [0,1]^2$ and some $P_{av} : [0,1]^2 \to [0,1]$.

Denote the expected action taken by the principal when agent (e,t) truthfully reports her type by $\Pi(e,t) \coloneqq (1 - T(e,t))P(e,t,\emptyset) + T(e,t)P_{av}(e,t)$. Proposition 13 characterizes SIC mechanisms.

Proposition 13. 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 $(e,t) \in [0,1]^2$.

Similarly, it is easy to see how, in the case where the principal's ultimate choice lies in a closed real interval and the agent does not have evidence, the IC characterization of Propositon 6 generalizes.

L Proofs of results in the Online Appendix

Proof of Claim 3. Denote by $u_{\text{accept}}(e) \coloneqq \int_0^1 u(e,t)f(e,t)dt / \int_0^1 f(e,t)dt$ the expected payoff from accepting without verification every agent with training e, and by $u_{\text{verify}}(e) \coloneqq \int_0^1 (u(e,t)-c)\mathbf{I}(u(e,t)\geq c)f(e,t)dt / \int_0^1 f(e,t)dt$ the expected payoff from accepting after verification every agent with training e who gives payoff at least c.

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

Without assuming that $e \mapsto \mathbb{E}_t\left[\min\{u(e,t),c\}\middle|\, e\right]$ does not cross zero from above, the optimal mechanism specifies that for each level of training $e \in [0,1]$,

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

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

and the result follows. **Q.E.D.**

Proof of Lemma 5. Take any two agents $(\boldsymbol{e},\boldsymbol{t})$ and $(\boldsymbol{e},\boldsymbol{t}')$ with $\sigma(\boldsymbol{e},\boldsymbol{t}) = \sigma(\boldsymbol{e},\boldsymbol{t}')$. $(\boldsymbol{e},\boldsymbol{t})$'s 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[52]

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

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

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

for every $(\boldsymbol{e},\boldsymbol{t})$. Then, $\Pi'(\boldsymbol{e},\boldsymbol{t}) = (1 - T'(\boldsymbol{e},\boldsymbol{t}))P'(\boldsymbol{e},\boldsymbol{t},\emptyset) + T'(\boldsymbol{e},\boldsymbol{t}) = \Pi(\boldsymbol{e},\boldsymbol{t})$ for every $(\boldsymbol{e},\boldsymbol{t})$, where the second equality follows by construction of M'. Thus, M' is outcome-equivalent to M. Given that M is 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

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

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 $(\boldsymbol{e}',\boldsymbol{t}') \leq (\boldsymbol{e},\mathbf{1})$, so under mechanism M', no agent has incentives to imitate an agent with higher composite measure than her own. For $c > 0$, M' also minimizes verification costs. **Q.E.D.**

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

Proof of Lemma 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 $\Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) \geq (1 - T(\boldsymbol{e},s))P(\boldsymbol{e},s,\emptyset)$ for any $(\boldsymbol{e},s)$. Then, construct the mechanism $M' \coloneqq \langle T',P'\rangle$ with[53]

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

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

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

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

so M' also satisfies condition (iii) of Proposition 1. Therefore, M' is SIC. Last, for $c>0$, M' saves on verification costs compared to M if there exists (a positive measure of) $(\boldsymbol{e},s)$ with $P(\boldsymbol{e},s,\emptyset)(1-T(\boldsymbol{e},s)) < \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))$, 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 $E^* \coloneqq \{\boldsymbol{e}\in[0,1]^m : \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0}))=1\}$ (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 $[0,1]^m$. 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.[54]

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 (8) can be written as

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

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

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

It remains to show that $\Pi(\boldsymbol{e},s) = \mathbf{I}(\widetilde{u}(\boldsymbol{e},s)\geq c \text{ or } \boldsymbol{e}\in E^*)$ satisfies condition (ii) of Proposition 10. Take any $\boldsymbol{e},\boldsymbol{e}',s$ with $\boldsymbol{e}'\geq\boldsymbol{e}$. We need to show that $\Pi(\boldsymbol{e}',s)=0$ implies $\Pi(\boldsymbol{e},s)=0$. If $\Pi(\boldsymbol{e}',s)=0$, then $\widetilde{u}(\boldsymbol{e}',s)<c$ 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 $\widetilde{u}(\boldsymbol{e},s)<c$. We will show this by contradiction. Assume that $\widetilde{u}(\boldsymbol{e},s)\geq c$. Then, we have that $\widetilde{u}(\boldsymbol{e},s)\geq c>\widetilde{u}(\boldsymbol{e}',s)$, which, given that σ is talent-biased, implies that $\boldsymbol{e}'\not\geq\boldsymbol{e}$, a contradiction. **Q.E.D.**

[54] 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.

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 $E^* \coloneqq \{\boldsymbol{e} \in [0,1]^m : \Pi(\boldsymbol{e},\sigma(\boldsymbol{e},\mathbf{0})) = 1\}$ (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 (8) can be written as

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

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

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

Case 1: if $s' \in [\sigma(\boldsymbol{e},\mathbf{0}),\sigma(\boldsymbol{e},\mathbf{1})]$, then σ being 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 $\boldsymbol{e}' \geq \boldsymbol{e}''$, $\arg\min_{\boldsymbol{t}:\sigma(\boldsymbol{e}',\boldsymbol{t})=s'} u(\boldsymbol{e}',\boldsymbol{t}) \geq \mathbf{0}$, 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 $\boldsymbol{e}' \not\geq \boldsymbol{e}$, 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 $\boldsymbol{e}^* \notin E^*$ 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 $\widetilde{u}(\boldsymbol{e},s) > c$ and $s < s^*$, then $\widetilde{u}(\boldsymbol{e},s) > c = \widetilde{u}(\boldsymbol{e}^*,s)$, so because σ is 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 $\widetilde{u}(\boldsymbol{e},s) < c$ 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 $\widetilde{u}(\boldsymbol{e},s) > c$ 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 $\widetilde{u}(\boldsymbol{e}',s) < c$ for every $\boldsymbol{e}' \notin E^*$ 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 $\widetilde{u}(\boldsymbol{e}',s) > c$ for every $\boldsymbol{e}' \notin E^*$ 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 $(\boldsymbol{e},s) \notin I_{\widetilde{u}}(c) \cup E^*$, $\Pi(\boldsymbol{e},s) = \mathbf{I}(s \geq s^*)$. 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.**

Proof of Lemma 9. Take a truthful mechanism $M \equiv \langle T,P \rangle$ with $P(e,t,s)$ given by (1) for some P_{av}. Construct the mechanism $M' \equiv \langle T',P' \rangle$ with (i) $P'_{av}(e,t) = 1$, (ii) $T'(e,t) = T(e,t)P_{av}(e,t)$, and (iii) $P'(e,t,\emptyset) = (1 - T(e,t))P(e,t,\emptyset)/(1 - T'(e,t))$ for any (e,t). It follows as in the proof of Lemma 1 that M' is truthful and outcome-equivalent to M. Also, $T'(e,t)P'_{av}(e,t) = T(e,t)P_{av}(e,t)$ for every (e,t), so M' has the same verification costs as M. **Q.E.D.**

Proof of Lemma 10. I prove the result for the case where the agent has evidence of training. The proof for case where she does not have evidence proceeds in a similar fashion.

Under mechanism M', if every agent (e,t) reports truthfully and does not manipulate her composite measure, the outcome is the same as under mechanism M. It remains to show that M' is truthful.

We show the contrapositive. Assume that M' is not truthful. That is, there exists (e,t), (e',t') with $e' \leq e$, and $s \in [0, \sigma(e,t)]$ such that

$$T'(e',t')P'(e',t',s) + (1 - T'(e',t'))P'(e',t',\emptyset) > T'(e,t)P'(e,t,\sigma(e,t)) + (1 - T'(e,t))P'(e,t,\emptyset).$$

Given this, if $s \geq \sigma(e',t')$, we have that

$$\begin{aligned}
&T(\phi(e',t'))P(\phi(e',t'),\chi(e',t')) + (1 - T(\phi(e',t')))P(\phi(e',t'),\emptyset) \\
=&T'(e',t')P'(e',t',s) + (1 - T'(e',t'))P'(e',t',\emptyset) \\
>&T'(e,t)P'(e,t,\sigma(e,t)) + (1 - T'(e,t))P'(e,t,\emptyset) \\
=&T(\phi(e,t))P(\phi(e,t),\chi(e,t)) + (1 - T(\phi(e,t)))P(\phi(e,t),\emptyset),
\end{aligned}$$

where the equalities follow by construction of M'. This means that under mechanism M, agent (e,t) prefers to report $\phi(e',t')$ and choose composite measure $\chi(e',t')$ instead of $\phi(e,t)$ and $\chi(e,t)$, respectively. Also, she is able to do so, since $\phi(e',t') \leq (e',1) \leq (e,1)$ and $\chi(e',t') \leq \sigma(e',t') \leq s \leq \sigma(e,t)$. Therefore, ϕ and χ are not IC under M.

Similarly, if $s < \sigma(e',t')$, we have that

$$\begin{aligned}
&T(\phi(e',t'))P(\phi(e',t'),\chi(e,t)) + (1 - T(\phi(e',t')))P(\phi(e',t'),\emptyset) \\
\geq&(1 - T(\phi(e',t')))P(\phi(e',t'),\emptyset) \\
=&T'(e',t')P'(e',t',s) + (1 - T'(e',t'))P'(e',t',\emptyset) \\
>&T'(e,t)P'(e,t,\sigma(e,t)) + (1 - T'(e,t))P'(e,t,\emptyset) \\
=&T(\phi(e,t))P(\phi(e,t),\chi(e,t)) + (1 - T(\phi(e,t)))P(\phi(e,t),\emptyset),
\end{aligned}$$

This means that under mechanism M, agent (e,t) prefers to report $\phi(e',t')$ instead of $\phi(e,t)$,[55] and she is able to do so. Therefore, ϕ and χ are not IC under M. **Q.E.D.**

Proof of Proposition 13. Denote the expected action taken by the principal when (e,t) reports (e',t') (with $e' \leq e$) by

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

[55] If she reports $\phi(e',t')$, she gets accepted with higher probability than if she reports $\phi(e,t)$ regardless of what composite measure she chooses. For example, composite measure $\chi(e,t)$ works.

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

$$\max_{(e',t') \leq (e,1)} [(1-T(e',t'))P(e',t';\emptyset) + T(e',t')P_{av}(e',t')\mathbf{I}(\sigma(e,t) \geq \sigma(e',t'))] = \Pi(e,t). \quad (10)$$

Assume that conditions (i) and (iii') are satisfied. Then, $\Pi(e,t) \geq \Pi(e,0) \geq (1-T(e',t'))P(e',t',\emptyset)$ for any (e',t') with $e' \leq e$. Therefore, (10) is equivalent to

$$\max_{(e',t') \in \{(x,y) \in [0,1]^2 : x \leq e \text{ and } \sigma(e,t) \geq \sigma(x,y)\}} [(1-T(e',t'))P(e',t';\emptyset) + T(e',t')P_{av}(e',t')] = \Pi(e,t). \quad (11)$$

Given that $\Pi(e,t)$ is non-decreasing in t (condition (i)), (11) can equivalently be written as

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

or equivalently,

$$e \in \arg\max_{r \in [\underline{e}(\sigma(e,t)),e]} \Pi(r,\tau(r,\sigma(e,t))). \quad (12)$$

Thus, incentive-compatibility is satisfied for every type if and only if for every (e,t), (12) 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).

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. **Q.E.D.**