Bi-dimensional screening with substitutable attributes: distinguishing talent from hard work*

Preliminary and Incomplete. Please do not circulate.

Orestis Vravosinos†

April 23, 2023

Abstract

A certain outcome can be the result of various combinations of an agent's (she) attributes (e.g., high talent and low effort or vice-versa). The agent can costlessly provide evidence to the principal on one of her attributes (effort) but not the other (talent). She can provide partial evidence, thereby understating the former to make the latter appear higher as her outcome (e.g., work product) will be attributed to talent. The principal (he) decides whether to hire/promote the agent by asking each agent type for a certain level of evidence and then verifying some types' outcome (possibly) at a cost.

I show that the principal can achieve the full information benchmark if he mostly values effort and verification is costless. However, when the principal only values talent, he still optimally rewards effort and can never achieve the first-best. He rejects some talented yet not hard-working agents to avoid hiring some hard-working but untalented agents that can disguise their hard work as talent. Last, when verification is costly, he may even hire the most hard-working agents without verifying their outcome.

Keywords: information transmission, disclosure, persuasion
JEL classification codes: D82, D83

*I am grateful to Erik Madsen for his continued guidance and terrifyingly insightful comments. The paper has also significantly benefited from discussions with Debraj Ray. I also thank Agata Farina, Alessandro Bonatti, Alexander Jakobsen, Basil Williams, Bobby Pakzad-Hurson, David Pearce, Dilip Abreu, Gonzalo Cisternas, Idione Meneghel, Inga Deimen, Joseph Root and participants in the NRET seminar and Third Year Paper PhD Mini Conference at NYU for helpful comments and discussions.

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

1 Introduction

People are characterized by multiple (privately observed) attributes that are often substitutes in producing a certain outcome. A chess player with low ability to calculate lines fast and in depth can have the same Elo as a player with a higher ability if the former prepares more intensively, studying chess theory and engine lines. A less talented student can achieve the same grade as a more talented one by solving more practice math problems increasing the probability that similar problems will appear in an exam. An inefficient employee can be as productive as a more efficient one by working longer hours. A researcher or start-up that is going to give a talk to an academic audience or venture capital firm, respectively, can prepare her answers to possible questions to compensate for her difficulty to think on her feet.

In such cases the following problem often arises. Given two desirable substitutable attributes, an agent may have incentives to understate one to manipulate others' beliefs over her other attribute. An employee may understate how long it took her to complete a project if promotion decisions depend mostly on the employer's beliefs over the employee's ability (*i.e.* the rate at which her work hours translate into value to the firm), because in the higher position ability is more important than working long hours. A researcher may hide some things that she has worked on in order to use them to answer the audience's questions.

This way of thinking is so fundamental that even kids seem to follow it. Students often eagerly proclaim that they have not studied hard for an exam not only if they are informed of their substandard performance but also when they have performed exceptionally well. By stressing their low effort (or even understating it) they may be trying to have their high score attributed to their (overstated) brilliance.

However fundamental, to the best of my knowledge this problem has not been examined before. I study it in the following setting. An agent has a bi-dimensional type. The first dimension is her *evidence* (e.g., a researcher's knowledge, an employee's effort, a start-up's market research quality) and the second her *testable attribute* (e.g., a researcher's ability to think fast, an employee's talent, a start-up's founder's entrepreneurial spirit). The value of the agent to the principal is measured by a weighted average of her evidence and testable attribute. The principal ultimately wants to decide whether to hire (or promote/fund) the agent or not. She does so by committing to a direct mechanism that asks each agent type for a certain level of evidence. Given that the agent provides the evidence, the designer may then test the agent's testable attribute by paying a fixed cost.

The test is imperfect in the following sense. If the agent reveals all her evidence in the first step, then the test will reveal the value of her testable attribute. However, if the agent hides some of her evidence, then she can use the withheld evidence in the test, which will make the test result overstate the value of her testable attribute.

I show that the principal can achieve the full-information benchmark if he mostly values the agent's evidence and the test is free. If the principal only values evidence, he can achieve the first-best without the need to test the agent (and without commitment power), since the latter can has no reason to hide any of her evidence and full unraveling arises.

However, the problem becomes more nuanced when the manipulable attribute is the valuable one. This can be the case for example when an academic hiring committee, which mostly values a candidate professor's ability to learn, do research and offer valuable feedback to students and fellow professors fast. I show that even when the principal only values the testable attribute, he still optimally rewards agents for their evidence. Agents with more evidence face lower standards (regarding their testable attribute) when tested. When the test is costly, the principal may even hire agents with a lot of evidence without examining their testable attribute. This happens because he would otherwise need to test many of them to reject only those with a very low testable attribute, since many agents with a low testable attribute (who, if hired, give a negative value to the principal) would still be able to get hired by withholding evidence and imitating more able agents with less evidence. Last, the principal has to reject some agents of positive value with high testable attribute who—because of their low evidence—are easy for less able agents with high evidence to imitate.

After this introduction, section 2 discusses related literature and section 3 presents the model. Section 4 characterizes incentive-compatible mechanisms and then solves the principal's problem. Section 5 applies the model to academic job market talks. Last, section 6 concludes. Proofs are gathered in Appendix A.

2 Related literature

While to the best of my knowledge no other work has studied a similar problem, this paper has links to a few strands of the literature.

First, it connects to the literature of persuasion games, where an informed party discloses verifiable information to a decision maker to influence his actions.[1] When the informed party (e.g., seller) perfectly knows her quality and can costlessly and verifiably disclose it to the DM (e.g., buyer), whose payoff is increasing in quality, full unraveling emerges in equilibrium; every seller type (except possibly the lowest one) discloses her quality (Viscusi, 1978; Grossman, 1981; Milgrom, 1981). In my case, this happens without the need for commitment power when the principal only values the agent's evidence which the agent can fully reveal at no cost.

Second, it has links to the literature of evidence games, where an agent chooses what part of her verifiable evidence to disclose to the principal without being able to prove whether she has disclosed everything or not. Multiple works have studied problems with this feature (e.g., see Shin, 1994; Dziuda, 2011; Hart et al., 2017). They all have several differences to mine. Most importantly, the principal cannot verify some of the agent's attributes after she provides evidence.[2]

The ability of the principal to perform an experiment after an initial information transmission by the agent is considered in Glazer and Rubinstein (2004). However, their model is different from the one I propose in many aspects. Rather than providing verifiable evidence, the agent only sends a cheap talk message. Importantly, the information that the principal obtains from verification is not influenced by the agent's initial disclosure.[3]

[1]See Milgrom (2008) for a review.

[2]There also are other differences. For example, in Hart et al. (2017) the principal does not always prefer more evidence to less and the agent can choose to withhold damaging evidence. In Dziuda (2011) the existence of a behavioral type is central, while in my model all agents are strategic.

[3]Glazer and Rubinstein (2006) examine a model similar in spirit to their previous one, where instead of sending a cheap talk message the agent presents some hard evidence. However, there is no verification by the principal.

3 A model of bi-dimensional screening with substitutable attributes

There are two players, an agent (she) and a principal (he). The agent is privately informed of her bi-dimensional type (e,t) with full support density $f : [0,1]^2 \to \mathbb{R}_+$. Dimension e of the type has an "evidence structure" and I thus refer to e as the agent's *evidence*. That is, an agent of type (e,t) can prove to the principal that her e is at least r for any $r \in [0,e]$ by presenting evidence $r \in [0,e]$. If she reveals $r < e$, we say that she partially reveals her evidence. However, for no $r \in [0,1)$ can she prove that her e is not higher than r; in other words, she cannot prove that she is not withholding evidence. The agent cannot unilaterally prove anything about the second dimension of her type t, which I call her *testable attribute* or *talent*. The principal can test this attribute by paying a cost $c \geq 0$.

3.1 Testing technology

This test is imperfect and works as follows. If after an agent (e,t) has fully revealed her evidence (*i.e.*, she presents evidence $r = e$) the principal tests her t, then the test will reveal her true t. However, this is not the case when $r < e$. When an agent (e,t) with $t < 1$ that has presented evidence $r < e$ is tested, the test returns result $\widehat{t}(e,t,r) \in (t,1)$, higher than her real attribute t because the agent uses the withheld evidence $e - r$ when tested. In other words, an agent (e,t) who has partially revealed evidence r performs in the test like a fully-revealing agent with talent $\widehat{t}(e,t,r)$. For $t = 1$, $\widehat{t}(e,t,r) = 1$ for any e and $r \in [0,e]$. Overall, $\widehat{t} : \{(e,t,r) \in [0,1]^3 : r \leq e\} \to [0,1]$ describes the testing technology. A test performed on an agent of type (e,t) who has revealed evidence $r \in [0,e]$ returns result $\widehat{t}(e,t,r)$, where (i) $\widehat{t}(e,t,e) = t$ for any (e,t) and $\widehat{t}(e,t,r)$ is (ii) increasing in t and (iii) increasing in e and decreasing in r (unless $t = 1$, in which case it is constant in e and r).

$\widehat{t}$ is also assumed to be continuous and satisfy two additional assumptions. First, it is required to satisfy assumption 1. This says that agent (e_3,t) that reveals evidence e_1 will perform at the test like an agent with less evidence $e_2 < e_3$ but testable attribute equal to the test score that type (e_3,t) would achieve after revealing evidence e_2.[4]

[4]Notice that for $t = 1$, $\widehat{t}(e_3,t,e_2) = 1$, so that $\widehat{t}(e_2, \widehat{t}(e_3,t,e_2),e_1) = \widehat{t}(e_3,t,e_1) = 1$ and the assumption is automatically satisfied.

Assumption 1. For any t and any e_1,e_2,e_3 with $0 \le e_1 < e_2 < e_3 \le 1$, it holds that $\widehat{t}(e_2, \widehat{t}(e_3,t,e_2),e_1) = \widehat{t}(e_3,t,e_1)$.

This assumption means that the testing technology has the following crucial property. Agent type (e_3,t) can imitate a fully-revealing type $(e_2, \widehat{t}(e_3,t,e_2))$ with lower evidence $e_2 < e_3$.[5] Then, the set of fully-revealing types $\{(e_1,\widehat{t}(e_2, \widehat{t}(e_3,t,e_2),e_1)) : e_1 < e_2\}$ with even lower evidence $e_1 < e_2$ that can be imitated by type $(e_2, \widehat{t}(e_3,t,e_2))$ is a subset of the types that can be imitated by type (e_3,t).

Second, it should satisfy assumption 2.

Definition 1. A testing technology $\widehat{t}$ has *revelation-neutral ranking* if for any (e,t), (e',t'), r and r' with $r,r' \le e' \le e$, $\text{sgn}\left\{\widehat{t}(e',t',r) - \widehat{t}(e,t,r)\right\} = \text{sgn}\left\{\widehat{t}(e',t',r') - \widehat{t}(e,t,r')\right\}$.

Assumption 2. $\widehat{t}$ has revelation-neutral ranking.

This assumption says that given any pair of agents who both reveal the same level of evidence, the ranking of their test scores does not depend on that common level of evidence revealed. This condition will mean that we can constrain attention to fully-revealing mechanisms (*i.e.* mechanisms where every type is asked to reveal all their evidence).

3.2 Payoffs

Ultimately, the principal wants to choose between two actions: hire or not the agent. He receives payoff $u(e,t) \coloneqq \gamma e + (1-\gamma)t - \underline{q}$ from hiring an agent of type (e,t), where $\gamma \in [0,1]$ 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 for her to be of (positive) value to the principal. If he does not hire an agent, the principal receives payoff normalized to 0. The agent wants to maximize the probability of getting hired.

3.3 The principal's problem

To decide whether to hire the agent, the principal designs (with commitment) a direct mechanism $M \equiv \langle R,T,P \rangle$ that specifies (i) the evidence $R(e,t) \in [0,e]$ that an agent that claims her type to be (e,t) is required to present, (ii) the probability $T(e,t) \in [0,1]$ with

[5] This is to say that agent (e_3,t) that reveals evidence e_2 and gets tested will score $\widehat{t}(e_3,t,e_2)$. Therefore, she cannot be distinguished from a fully-revealing agent $(e_2,\widehat{t}(e_3,t,e_2))$.

which the principal will test an agent's talent when she reports her type to be (e,t), and (iii) the probability $P(e,t,\widetilde{t})$, which should be non-decreasing in $\widetilde{t} \in [0,1]$, with which the principal will hire the agent after the agent has reported her type to be (e,t) and the test has returned result $\widetilde{t} \in [0,1]$.[6] When no test is performed, $\widetilde{t} = \emptyset$ and the agent is hired with probability $P(e,t,\emptyset)$. Overall, the principal chooses an incentive-compatible (IC) mechanism $M \equiv \langle R,T,P \rangle$, where $R : [0,1]^2 \to [0,1]$ with $R(e,t) \leq e$ for every (e,t), $T : [0,1]^2 \to [0,1]$, and $P : [0,1]^2 \times [0,1] \cup \{\emptyset\} \to [0,1]$ with $P(e,t,\widetilde{t})$ non-decreasing in $\widetilde{t} \in [0,1]$, to maximize

$$\int_0^1 \int_0^1 \left[\left(T(e,t)P(e,t,\widehat{t}(e,t,R(e,t)) + (1 - T(e,t))P(e,t,\emptyset) \right) u(e,t) - cT(e,t) \right] f(e,t)dtde.$$

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

$$(e,t) \in \underset{\{(e',t') \in [0,1]^2 : R(e',t') \leq e\}}{\arg\max} \left\{ T(e',t')P(e',t',\widehat{t}(e,t,R(e',t')) + (1 - T(e',t'))P(e',t',\emptyset) \right\}.$$

4 Optimal bi-dimensional screening with substitutable attributes

This section characterizes IC mechanisms and then solves the principal's problem.

4.1 Simplifying the class of candidate IC mechanisms

Given that $P(e,t,\widetilde{t})$ should be non-decreasing in $\widetilde{t} \in [0,1]$, in solving the principal's problem we can constrain attention to mechanisms such that for $\widetilde{t} \in [0,1]$

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

for any (e,t) for some $P_{at} : [0,1]^2 \to [0,1]$, where *at* is a mnemonic for the probability of getting hired *after testing* (given that the threshold test score is met). Notice that if type (e,t) reports truthfully and is tested, she is then hired with probability $P_{at}(e,t)$. To

[6] The condition that $P(e,t,\widetilde{t})$ be non-decreasing in $\widetilde{t} \in [0,1]$ can be understood as an incentive-compatibility condition in a model where $\widehat{t}(e,t,r)$ gives the score that agent type (e,t) can achieve after disclosing evidence r but the agent can intentionally make mistakes in the test.

see why constraining attention to such mechanisms is without loss of optimality, observe that among all mechanisms that (conditional on testing) hire type (e,t) with probability $P_{at}(e,t)$, the one that satisfies (1) minimizes incentives of other types to imitate (e,t).[7]

The total probability with which type (e,t) is hired if she reports (e',t') (with $e' \leq e$) is then equal to

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

and the total probability with which agent (e,t) is hired if she truthfully reports her type is equal to

$$\Pi(e,t) \coloneqq (1-T(e,t))P(e,t,\emptyset) + T(e,t)P_{at}(e,t).$$

Definition 3. A mechanism $M' \equiv \langle R',T',P' \rangle$ is outcome-equivalent to a mechanism $M \equiv \langle R,T,P \rangle$ if for every (e,t), $\Pi(e,t) = \Pi'(e,t)$, where $\Pi(e,t) \equiv (1-T(e,t))P(e,t,\emptyset) + T(e,t)P_{at}(e,t)$ and $\Pi'(e,t) \equiv (1-T'(e,t))P'(e,t,\emptyset) + T'(e,t)P'_{at}(e,t)$.

Lemma 1 shows that when testing is costly, an agent that is tested and passes the test is hired with probability 1 in any optimal mechanism. With costless testing, it is still without loss to constrain attention to mechanisms that satisfy this condition.

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

Given Lemma 1, from now on we constrain attention to mechanisms with $P_{at}(e,t) = 1$ for any (e,t).[8] I will now show that we can further simplify the analysis by constraining attention to fully-revealing mechanisms where agents with zero talent are never tested.

Definition 4. A mechanism $M \equiv \langle R,T,P \rangle$ is fully-revealing if $R(e,t) = e$ for every (e,t).

Lemma 2. Given any IC mechanism M, there exists a fully-revealing IC mechanism $M' \equiv \langle R',T',P' \rangle$ with $T'(e,0) = 0$ for every e that is outcome-equivalent to M and has the same expected testing cost as M.

[7]Namely, rewarding the agent further for performing above $\widehat{t}(e,t,R(e,t))$ will result to the same probability of hiring (e,t) after testing her and only give additional incentives to other agents ti imitate (e,t). Similarly, there is no reason to hire the agent for test scores lower than $\widehat{t}(e,t,R(e,t))$.

[8]For (e,t) with $T(e,t) = 0$ the value of $P_{at}(e,t)$ does not matter.

4.2 Incentive-compatible mechanisms

Thus, we now constrain attention to fully-revealing mechanisms that satisfy $P_{at}(e,t) = 1$ and $T(e,0) = 0$ for any (e,t). Proposition 1 characterizes such IC mechanisms.

Proposition 1. A fully-revealing mechanism $M \equiv \langle R,T,P \rangle$ with $P_{at}(e,t) = 1$ for every (e,t) is IC if and only if

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

(ii) $\Pi(r, \widehat{t}(e,t,r))$ is non-decreasing in r over $r \in [0,e]$ for every (e,t), and

(iii) $P(e',t',\emptyset) \leq \Pi(e,0)$ for every e,e',t' with $e' \leq e$,

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

Lemma 3. For any fully-revealing IC mechanism $M \equiv \langle R,T,P \rangle$ with $P_{at}(e,t) = 1$ and $T(e,0) = 0$ for every (e,t), there exists an IC mechanism $M' \equiv \langle R',T',P' \rangle$ with the same properties and $P'(e,t,\emptyset) = P'(e,0,\emptyset) = \Pi'(e,0)$ for every (e,t), and that is outcome-equivalent to M. Also, for $c > 0$, any optimal mechanism $M \equiv \langle R,T,P \rangle$ satisfies $P(e,t,\emptyset) = P(e,0,\emptyset)$ for any (e,t) with $T(e,t) < 1$.

Remark 4.1. Notice that by Proposition 1, IC also requires that $P(e,0,\emptyset)$ be non-decreasing in e.

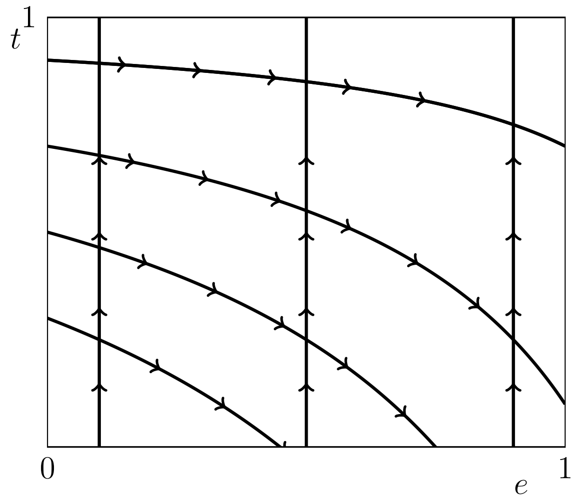
We can now constrain attention to regular mechanisms.

Definition 5. A mechanism $M \equiv \langle R,T,P \rangle$ is regular if it is fully-revealing and satisfies (i) $P_{at}(e,t) = 1$, (ii) $T(e,0) = 0$ and (iii) $P(e,t,\emptyset) = P(e,0,\emptyset)$ for every (e,t).

We will now use assumption 1 to simplify condition (ii) of Proposition 1. Define the function $\tau : \{(x,y) \in [0,1]^2 : \widehat{t}(x,0,0) \leq y\} \to [0,1]$, implicitly given by $\widehat{t}(e,\tau(e,s),0)) = s$, which given evidence e and test score s returns talent level $\tau(e,s)$ such that agent $(e,\tau(e,s))$ will achieve test score s if she reveals no evidence.[9] Define also $\widetilde{e} : [0,1] \to [0,1]$, implicitly given by $\widehat{t}(\widetilde{e}(s),0,0) = \min\{\widehat{t}(1,0,0),s\}$, which given a test score s returns the maximum level of evidence $\widetilde{e}(s)$ that an agent can have and achieve test score $\widetilde{e}(s)$ if she reveals

[9] Notice that $\tau(e,s)$ is decreasing in e (unless $s = 1$, on which case $\tau(e,1) = 1$ constant in e) and increasing in s.

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



no evidence.[10] For each $s \in [0,1]$ then, the set $I_s \coloneqq \{(e,\tau(e,s)) : e \in [0,\widetilde{e}(s)]\}$ is the iso-test-score curve s. It includes all agent types who will achieve the same test score s if they reveal no evidence.[11]

Lemma 4 then shows that condition (ii) is satisfied as long as the probability of getting hired (when every agent type truthfully reports her type) is non-decreasing along the iso-test-score curves. Figure 1 schematically summarizes the IC conditions of Corollary 1.1, which describes regular IC mechanisms. It shows the directions towards which $\Pi(e,t)$ (weakly) increases in IC mechanisms.

Lemma 4. The following statements are equivalent:

(i) $\Pi(r, \widehat{t}(e,t,r))$ is non-decreasing in r for $r \in [0,e]$ for every (e,t),

(ii) $\Pi(r, \tau(r,s))$ is non-decreasing in r for $r \in [0,\widetilde{e}(s)]$ for every $s \in [0,1]$.

Corollary 1.1. A regular mechanism $M \equiv \langle R,T,P \rangle$ is IC if and only if

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

(ii) $\Pi(r, \tau(r,s))$ is non-decreasing in r for $r \in [0,\widetilde{e}(s)]$ for every $s \in [0,1]$,

where $\Pi(e,t) \equiv (1 - T(e,t))P(e,0,\emptyset) + T(e,t)$.

[10] That is, if an agent has evidence $e > \widetilde{e}(s)$, then the test score she will achieve by revealing no evidence will be higher than s even if her true testable attribute is $t = 0$.

[11] By assumption 2, for any $\underline{e} \in [0,\widetilde{e}(s)]$, the subset $\{(e,t) \in I_s : e \geq \underline{e}\}$ of the line includes all types who achieve the same score after revealing evidence $\underline{e}$.

Corollary 1.1 implies that in any IC mechanism, if for some e, $\Pi(e,0) = 1$, then $\Pi(e',0) = 1$ for every $e' \geq e$. Given a mechanism, define $\overline{e} \coloneqq \min\{e \in [0,1] : \Pi(e,0) = 1\}$.[12] Then, $T(e,0) = 0$ implies that for every (e,t)

$$T(e,t) = \begin{cases} \dfrac{\Pi(e,t) - \Pi(e,0)}{1 - \Pi(e,0)} & \text{if } e < \overline{e} \\ 0 & \text{if } e \geq \overline{e}. \end{cases}$$

Thus, the principal's problem amounts to choosing a function $\Pi(e,t)$ (which implies a specific $\overline{e} \in [0,1]$) that satisfies the conditions of Corollary 1.1 to maximize

$$\int_0^{\overline{e}} \int_0^1 \left[\Pi(e,t)u(e,t) - c\frac{\Pi(e,t) - \Pi(e,0)}{1 - \Pi(e,0)}\right] f(e,t)dtde + \int_{\overline{e}}^1 \int_0^1 u(e,t)f(e,t)dtde.$$

4.3 The optimal regular mechanism under free testing

For $c = 0$, this can equivalently be written as

$$\int_0^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \Pi(e,\tau(e,s))u(e,\tau(e,s))f(e,\tau(e,s))deds,$$

where instead of integrating over e and t, we integrate across iso-test-score curves and e.

Proposition 2 then examines when the principal can achieve the full information benchmark.

Proposition 2. Let $c = 0$. Then,

(i) the principal achieves the full information first-best[13] if and only if for every $\underline{e} \in [0,1]$

$$I_{\widehat{t(\underline{e},\underline{t}(\underline{e}),0)};\underline{e}} \subset \left\{(e,t) \in [0,1]^2 : t \geq \underline{t}(e)\right\}$$

where $I_{s;\underline{e}} \coloneqq \{(e,\tau(e,s)) : e \in [\underline{e},\widetilde{e}(s)]\}$ and $\underline{t}(e) \coloneqq \max\{(\underline{q} - \gamma e)/(1-\gamma),0\}$ is such that $t \geq \underline{t}(e) \iff u(e,t) \geq 0$,

(ii) for $\gamma = 0$, the principal does not achieve the first-best,

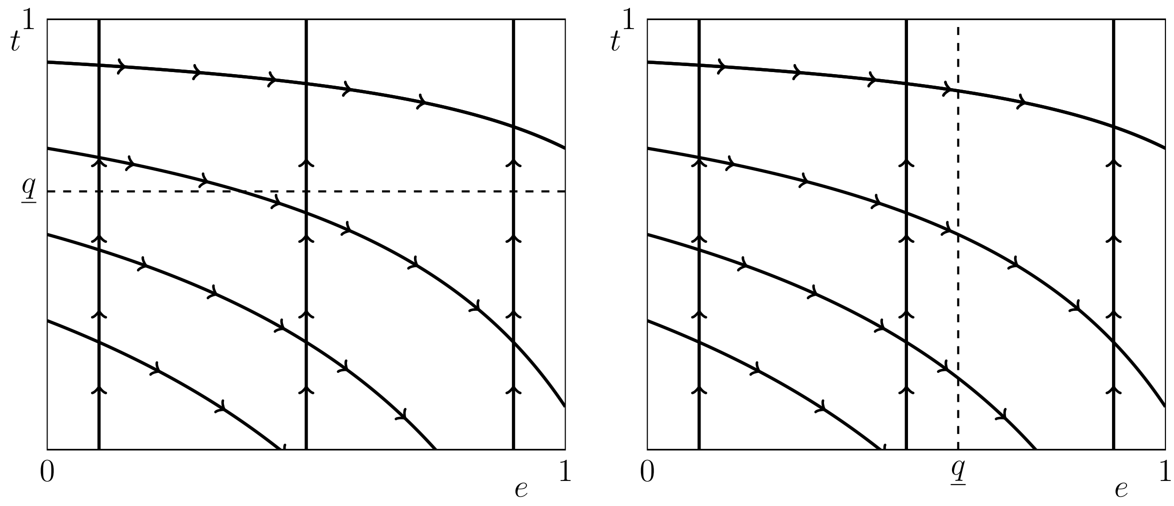
(iii) for $\gamma = 1$, the principal achieves the first-best.

[12]For simplicity in notation I let $\Pi(1,1) = 1$ and assume $\Pi(e,0)$ to be right-continuous in e. This is without loss, since these assumptions affect the hiring decisions for a zero-measure subset of types and make IC constraints most lax.

[13]That is, he hires every agent with $u(e,t) > 0$ and does not hire any agent with $u(e,t) < 0$.

Figure 2: (Not) achieving the first-best

(a) Failure to achieve the first-best: $\gamma = 0$ **(b)** Achieving the first-best: $\gamma = 1$



Note: the arrowed lines represent the direction of (weak) increase in $\Pi(e,t)$ in any IC mechanism.

When the principal only values evidence, full unraveling arises without the need for commitment or testing. However, when the principal only values talent, he cannot achieve the first best. This is because to do so, he would have to only reward talent and hire every agent with talent higher that $\underline{q}$. But then the principal cannot reject agent $(e + \varepsilon, \underline{q} - \delta)$, who can imitate agent $(e - \varepsilon, \underline{q} + \delta)$ for $\delta > 0$ small enough and $\varepsilon > 0$.

Graphically, these can be seen in Figure 2. For $\gamma = 0$, the directions of increase in $\Pi(e,t)$ show that hiring every type with $t > \underline{q}$ implies that types with $t < \underline{q}$ also have to be hired. How significantly valuing talent makes it impossible to achieve the first-best will be studied in more detail in section 5. On the other hand, for $\gamma = 1$, hiring every agent with $e > \underline{q}$ is IC.

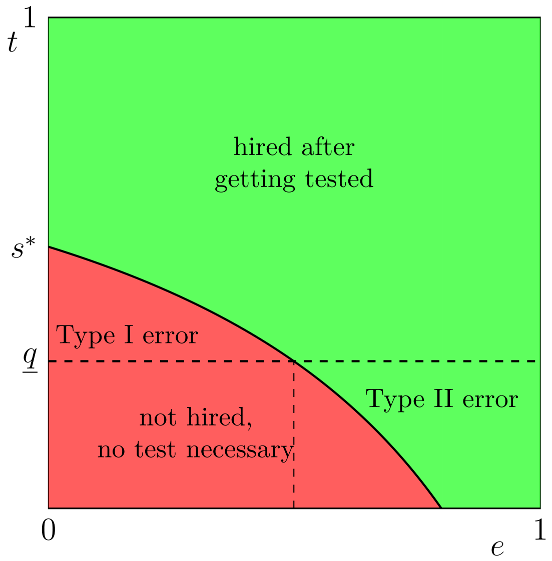
Proposition 3 studies what *can* actually be achieved when $\gamma = 0$.

Proposition 3. Let $c = \gamma = 0$. Then, in the optimal mechanism $\Pi(e,t) = \mathbf{I}(\widehat{t}(e,t,0) \geq s^*)$ for some $s^* \in (0,1)$. That is, agent (e,t) is hired if and only if $\widehat{t}(e,t,0) \geq s^*$.

To see why, notice that for a fixed value of $\Pi(e, \underline{q})$ for every $e \in [0,1]$ (which by IC has to be non-decreasing in e), it is optimal to keep $\Pi(e,t)$ constant along the iso-test-score curves. Then, the proposition follows from the fact that maximizing over any possible non-decreasing $\Pi(e, \underline{q})$ leads to a bang-bang $\Pi(e, \underline{q})$ solution.

The principal effectively chooses a threshold test score s^* and hires every agent that can achieve this score by not revealing any evidence. In choosing this threshold he balances

Figure 3: The optimal mechanism under $\gamma = 0$ and free testing



type I (*i.e.* rejecting agents of talent $t > \underline{q}$) and type II (*i.e.* accepting agents of talent $t < \underline{q}$) errors. This trade-off can be seen in Figure 3.

4.4 The optimal regular mechanism under costly testing

Notice that the optimal mechanism of Proposition 3 turns out to be pure. Restricting attention to pure mechanisms, Proposition 4 then derives the optimal mechanism when $\gamma = 0$ but testing is costly.

Proposition 4. Let $\gamma = 0$ and $c > 0$ and constrain attention to mechanisms with $P(e,t,\emptyset) \in \{0,1\}$ for every (e,t). Then, in the optimal mechanism $\Pi(e,t) = \mathbf{I}(\hat{t}(e,t,0) \geq s^*$ or $e \geq \bar{e})$ and $T(e,t) = \mathbf{I}(\hat{t}(e,t,0) \geq s^*$ and $e < \bar{e})$ for some $s^* \in (0,1)$ and $\bar{e} \in [0,1]$. That is, agent (e,t) is hired either (i) without getting tested if $e \geq \bar{e}$ or (ii) after getting tested if $\hat{t}(e,t,0) \geq s^*$ and $e < \bar{e}$.

The optimal mechanism has a similar structure to the one under free testing with one important difference. Some agents with a lot of evidence are hired without getting tested. This is because the principal would need to test many of the high-evidence types to only manage to reject few of the least talented among them.

5 Application: academic job market hiring

In this section I use the model to analyze academic job market talks and promotion decisions.

5.1 Academic job market hiring

The agent's (*i.e.* candidate) research topic is comprised by a "mass" $b > 1$ of problems. This means that there are uncountably infinitely many problems. $e \in [0,1]$ is the candidate's knowledge, that is, the mass of problems to which she has answers to. t is her acumen, 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. Testing here amounts to posing to the candidate countably infinitely many problems randomly sampled from the mass of problems that the candidate has not initially disclosed answers to.[14]

Then, if she reveals part $r \in [0,e]$ of her knowledge and is tested, she will answer proportion

$$\widehat{t}(e,t,r) := \frac{e - r + (b - e)t}{b - r},$$

of the problems posed to her. This is the sum of (i) the proportion $(e - r)/(b - r)$ of problems sampled from the set of problems that the candidate already has answers to (but has not disclosed them) and (ii) the proportion $(b - e)/(b - r)$ of problems sampled from the set of problems that the candidate does not already have answers to multiplied by the proportion t to which the candidate will find answers on the spot.

Proposition 5 studies when the first-best can be achieved.

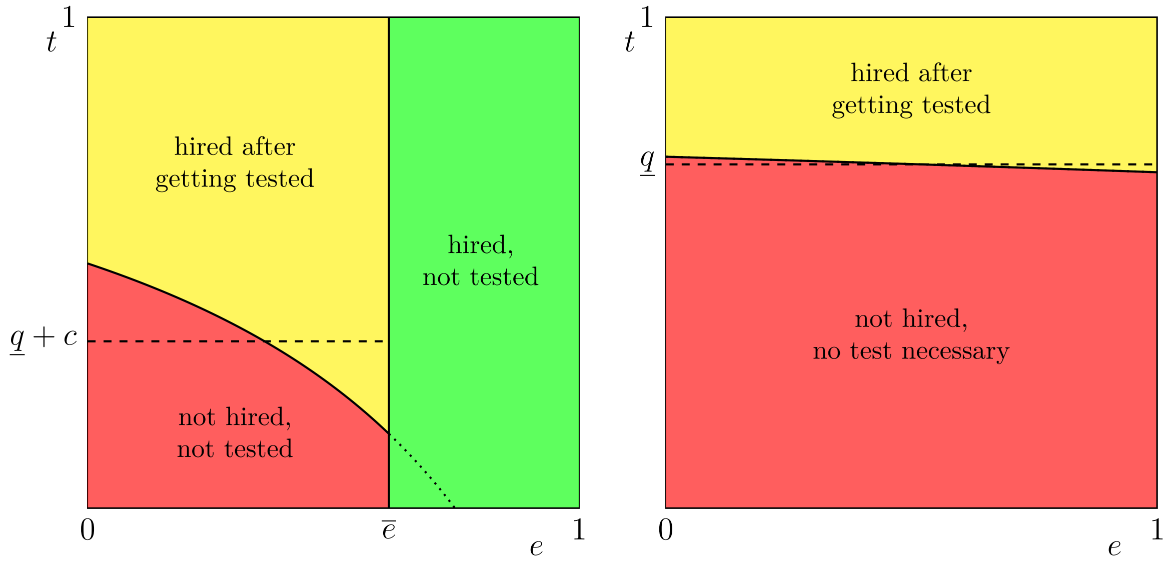
Proposition 5. Let $c = 0$. The principal achieves the first-best outcome if and only if $\gamma(1 + b) \geq 1 + \underline{q}$.

When the principal can test the candidate at no cost, he can achieve the first-best if

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

Figure 4: Optimal academic job market hiring: $\gamma = 0$

(a) $b = 3/2$, $\underline{q} = 3/10$, $c = 1/25$ **(b)** $b = 10$, $\underline{q} = 7/10$, $c = 0$



Note: (e,t) is uniformly distributed.

(i) he mostly cares about knowledge rather than acumen (γ high) and (ii) any candidate's knowledge is limited relative to the universe of problems. (b high).

Figure 4 shows the optimal mechanism when the principal only values acumen.

5.2 Promotions

An employee of efficiency t has exerted effort e. $\tilde{t}(e,t)$ is the employee's productivity, increasing in e (unless $t = 1$, then constant in e) and in t. Testing by the employer then amounts to verifying the employee's productivity $\tilde{t}(e,t)$.

To see how this fits into out framework, define $t^*(e,p)$ implicitly given by $\tilde{t}(e,t^*(e,p)) = p$. That is, $t^*(e,p)$ is the efficiency that an employee who has exerted effort e needs to achieve productivity p.[15] Now, define the test score as $\hat{t}(e,t,r) \coloneqq t^*(r,\tilde{t}(e,t))$, which satisfies all of the assumptions in section 3. Finally, notice that the employer observing the revealed evidence r and productivity $\tilde{t}(e,t)$ is equivalent to him observing r and the test score $\hat{t}(e,t,r)$.

[15] $t^*(e,p)$ is decreasing in e (unless $p = \tilde{t}(0,1)$, then constant in e) and increasing in p.

6 Conclusion

In this paper I have proposed a model where an agent with two attributes (e.g. talent and effort) that are substitutes in producing a certain outcome. For example, an inefficient employee can be as productive as a more efficient one if the former exerts more effort. In such cases, the following problem arises: the agent, who can verifiably disclose (possibly part) of her effort may choose to understate it to make the principal attribute her outcome (e.g. work product) to her talent, thereby manipulating the principal's beliefs over that attribute upwards.

I have shown that the principal can achieve the full information benchmark if he mostly values effort and can verify the agent's outcome at no cost. However, when the principal only values talent, he can never achieve the first-best. He still has to reward effort to the extent of accepting some untalented yet hard-working agents, because they can disguise their hard work as talent. Conversely, he has to reject some talented yet not hard-working agents to limit the number of hard-working but untalented individuals that use their hard work to present themselves as talented. When it is costly to verify an agent's outcome, the principal may even hire the most hard-working agents without verifying their outcome at all.

I have applied this model to the case of job interviews, and particularly academic job market talks. A candidate's performance depends both on how much she has worked on her paper and on the extent to which she can find answers to questions from the audience fast. The complication in that case is that the candidate may have already prepared answers to problems that are not in her slides. She can then use this preparation to answer audience questions as effectively as a candidate that is more able to think on her feet.

The results reveal a stark difference in the difficulty of interviewing and hiring specialized employees. When the employer mostly values knowledge in an employee, the interview and hiring process is easy. The employer can perfectly learn the agent's knowledge at minimal cost. On the other hand, when the employer mostly values acumen, as can be the case when a department hires a research professor, the interview and hiring process is much more difficult.

References

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

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

Glazer, J. and Rubinstein, A. (2006). A Study in the Pragmatics of Persuasion: A Game Theoretical Approach. *Theoretical Economics*, 1(4):395–410.

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.

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

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

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

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

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

A Proofs

Proof of Lemma 1 Take an IC mechanism $M \equiv \langle R,T,P \rangle$. Construct the mechanism $M'\langle R',T',P' \rangle$ with (i) $R'(e,t) = R(e,t)$, (ii) $P'_{at}(e,t) = 1$, (iii) $T'(e,t) = T(e,t)P_{at}(e,t) \leq T(e,t)$ and (iv) $P'(e,t,\emptyset) = (1 - T(e,t))P(e,t,\emptyset)/(1 - T'(e,t))$ for any (e,t).[16]

We have then that (a) $T'(e,t)P'_{at}(e,t) = T(e,t)P_{at}(e,t)$, (b) $(1 - T'(e,t))P'(e,t,\emptyset) = (1 - T(e,t))P(e,t,\emptyset)$ and (c) $\Pi'(e,t) = \Pi(e,t)$ for any (e,t). (a)-(c) combined imply that

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

the problem of every agent type under M' is the same as it was under M'. (c) means that M' is outcome-equivalent to M.

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

Proof of Lemma 2 Let $M \equiv \langle R,T,P \rangle$ be an IC mechanism. Then, construct the mechanism $M' \coloneqq \langle R',T',P' \rangle$ with (i) $R'(e,t) \coloneqq e$ for every (e,t), (ii) $T'(e,0) = 0$ for every e and $T'(e,t) = T(e,t)$ for every (e,t) with $t > 0$, (iii) $P'(e,t,\emptyset) = P(e,t,\emptyset) + (T(e,t) - T'(e,t))\,(1 - P(e,t,\emptyset))$ for every (e,t) and (iv)

$$P'(e,t,\widetilde{t}) \coloneqq \begin{cases} 0 & \text{if } \widetilde{t} < t \\ P_{at}(e,t) & \text{if } \widetilde{t} \geq t \end{cases}$$

for every (e,t) and $\widetilde{t} \in [0,1]$.

If every type reports truthfully, M' hires each agent type with the same probability that M does, so it remains to show that M' is IC.

By IC of M we have that for every (e,t)

$$(e,t) \in \underset{(e',t') \in \{(x,y) \in [0,1]^2 : R(x,y) \leq e\}}{\arg\max} \left\{ \begin{array}{r} (1 - T(e',t'))P(e',t',\emptyset) \\ +T(e',t')\mathbf{I}\left(\widehat{t}(e,t,R(e',t')) \geq \widehat{t}(e',t',R(e',t'))\right) \end{array} \right\}. \quad (2)$$

Since $R'(e,t) \coloneqq e \geq R(e,t)$, for every e we have that

$$\left\{(x,y) \in [0,1]^2 : R'(x,y) \leq e\right\} = \left\{(x,y) \in [0,1]^2 : x \leq e\right\} \subseteq \left\{(x,y) \in [0,1]^2 : R(x,y) \leq e\right\}.$$

By construction, we have that $T(e',t') = T'(e',t')$ and $P(e',t',\emptyset) = P'(e',t',\emptyset)$ for every (e',t') with $t' > 0$. Also, for any (e,t) and any $(e',t') \in \{(x,y) \in [0,1]^2 : R'(x,y) \leq e\}$ with $t' = 0$, $\widehat{t}(e,t,R'(e',t')) \geq t \geq 0 = \widehat{t}(e',t',R'(e',t'))$, which means that (again for $t' = 0$)

$$\begin{aligned} (1 - T'(e',t'))P'(e',t',\emptyset) + T'(e',t')\mathbf{I}\left(\widehat{t}(e,t,R'(e',t')) \geq \widehat{t}(e',t',R'(e',t'))\right) &= \\ (1 - 0)\left[P(e',t',\emptyset) + (T(e',t') - 0)\,(1 - P(e',t',\emptyset))\right] + 0 &= \\ (1 - T(e',t'))P(e',t',\emptyset) + T(e',t')\mathbf{I}\left(\widehat{t}(e,t,R(e',t')) \geq \widehat{t}(e',t',R(e',t'))\right)&, \end{aligned}$$

where in the last line we have used assumption 2.

In words, we have shown three points. First, that under M', any agent type (e,t) has enough information to try to imitate (in the first step of evidence revelation) a subset of types she could try to imitate under M because M' is fully-revealing. Second, for any type (e',t') with $t' > 0$ that she may imitate, the payoff from doing so is the same as it was under M, because $T(e',t') = T'(e',t')$ and $P(e',t',\emptyset) = P'(e',t',\emptyset)$ for every (e',t') with $t' > 0$ and by assumption 2, she will pass the test when imitating (e',t') if and only if she also would under M. Third, for any type (e',t') with $t' = 0$ that she may imitate, the payoff from doing so is the same as it was under M, because $e' \leq e$ (since M' is fully-revealing) and $t' = 0$ imply that she would pass the test under M' and under M she simply gets hired with the same total probability without getting tested.

Combining these observations with (2) and assumption 2, we conclude that M' is IC. **Q.E.D.**

Proof of Proposition 1 *Step 1:* I first show that condition (i) is necessary for IC by showing the contrapositive. Assume that for some e,t_1,t_2 with $t_2 > t_1$, $\Pi(e,t_2) < \Pi(e,t_1)$. Then, IC of type (e,t_2) is violated, since $\widetilde{P}(e,t_1;e,t_2) = \Pi(e,t_1) > \Pi(e,t_2)$, that is, (e,t_2) can imitate (e,t_1) to (reach (e,t_1)'s test score threshold and) get hired with higher probability that she would if she truthfully reported her type.

Step 2: I now show that condition (iii) is necessary for IC by showing the contrapositive.[17] Assume that for some e,e',t with $e' \geq e$, $P(e,t,\emptyset) > \Pi(e',0)$. Then, IC of type $(e',0)$ is violated, since $\widetilde{P}(e,t;e',0) \geq P(e,t,\emptyset) > \Pi(e',0)$, that is, $(e',0)$ can imitate (e,t) to get hired with higher probability that she would if she truthfully reported her type (even if she cannot achieve (e,t)'s test score threshold).

Step 3: I finally show that provided that (i) and (iii) are satisfied, (ii) is necessary and sufficient for IC. IC of type (e,t) is satisfied if and only if

$$\max_{(e',t')\in\{(x,y)\in[0,1]^2:e'\leq e\}} \left[(1-T(e',t'))P(e',t';\emptyset) + T(e',t')\mathbf{I}\left(\widehat{t}(e,t,e') \geq t'\right)\right] = \Pi(e,t). \quad (3)$$

Assume that conditions (i) and (iii) are satisfied. Then, $\Pi(e,t) \geq \Pi(e,0) \geq P(e',t',\emptyset)$ for any (e',t') with $e' \leq e$. Therefore, (3) is equivalent to

$$\max_{(e',t')\in\left\{(x,y)\in[0,1]^2:x\leq e \text{ and } \widehat{t}(e,y,x)\geq y\right\}} [(1-T(e',t'))P(e',t';\emptyset) + T(e',t')] = \Pi(e,t). \quad (4)$$

[17] That $P(e',0,\emptyset) = \Pi(e,0)$ follows from $T(e,0) = 0$.

Given that $\Pi(e,t)$ is non-decreasing in t (condition (i)), (4) can equivalently be written as

$$\max_{e' \leq e} \Big[(1 - T(e', \widehat{t}(e,t,e')))P_{nt}(e',\widehat{t}(e,t,e')) + T(e',\widehat{t}(e,t,e'))\Big] = \Pi(e,t)$$

or equivalently,

$$e \in \arg\max_{e' \in [0,e]} \Pi(e', \widehat{t}(e,t,e')). \tag{5}$$

Thus, IC is satisfied for every type if and only if for every (e,t), (5) is satisfied. This is true if and only if $\Pi(r, \widehat{t}(e,t,r))$ is non-decreasing in r for $r \in [0,e]$ 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), $\Pi(r, \widehat{t}(e,t,r))$ is *not* non-decreasing in r for $r \in [0,e]$. That is, for some (e,t) there exist r_1, r_2 with $0 \leq r_1 < r_2 \leq e$ such that $\Pi(r_2, \widehat{t}(e,t,r_2)) < \Pi(r_1, \widehat{t}(e,t,r_1))$.

Define $t' := \widehat{t}(e,t,r_2) \in [e,1]$ (with $t' = t$ if and only if $t = 1$ or $e = r_2$). It follows then that

$$\Pi(r_2, \widehat{t}(r_2,t',r_2)) = \Pi(r_2, t') = \Pi(r_2, \widehat{t}(e,t,r_2)) < \Pi(r_1, \widehat{t}(e,t,r_1)) = \Pi(r_1, \widehat{t}(r_2,t',r_1)), \tag{6}$$

where the last equality follows since by Assumption 1, $\widehat{t}(r_2,\widehat{t}(e,t,r_2),r_1)) = \widehat{t}(e,t,r_1)$, where $t' \equiv \widehat{t}(e,t,r_2)$. But then it follows that

$$r_2 \notin \arg\max_{x \in [0,r_2]} \Pi(x, \widehat{t}(r_2,t',x)),$$

since $r_1 \in [0,r_2]$ and $\Pi(r_2, \widehat{t}(r_2,t',r_2)) < \Pi(r_1, \widehat{t}(r_2,t',r_1))$ by (6). Namely, IC of type (r_2, t') is violated, as she prefers to imitate type $(r_1,\widehat{t}(r_2,t',r_1))$. **Q.E.D.**

Proof of Lemma 4 The direction $(i) \Rightarrow (ii)$ is immediate. It remains to prove the opposite direction of the implication.

Step 1: I first show that if for some $(\overline{e},t)$, $\Pi(r, \widehat{t}(\overline{e},t,r))$ is non-decreasing in $r \in [0,\overline{e}]$, then $\Pi(r, \widehat{t}(\underline{e},\widehat{t}(\overline{e},t,\underline{e}),r))$ is non-decreasing in $r \in [0,\underline{e}]$ for every $\underline{e} \in [0,\overline{e}]$. To see this, take some arbitrary $(\overline{e},t)$ and r_2,r_1 with $\overline{e} > r_2 > r_1 \geq 0$ and assume that $\Pi(r_2, \widehat{t}(\overline{e},t,r_2)) \geq$

$\Pi(r_1,\hat{t}(\bar{e},t,r_1))$. It follows then that for any $\underline{e} \in [r_2,\bar{e}]$

$$\Pi(r_2,\hat{t}(\underline{e},\hat{t}(\bar{e},t,\underline{e}),r_2)) = \Pi(r_2,\hat{t}(\bar{e},t,r_2)) \geq \Pi(r_1,\hat{t}(\bar{e},t,r_1)) = \Pi(r_1,\hat{t}(\underline{e},\hat{t}(\bar{e},t,\underline{e}),r_1)).$$

where the equalities follow because by assumption 1, $\hat{t}(\underline{e},\hat{t}(\bar{e},t,\underline{e}),r_2) = \hat{t}(\bar{e},t,r_2)$ and $\hat{t}(\underline{e},\hat{t}(\bar{e},t,\underline{e}),r_1) = \hat{t}(\bar{e},t,r_1)$.

Clearly written proof in progress. **Q.E.D.**

Proof of Lemma 3 Take any fully-revealing IC mechanism $M \equiv \langle R,T,P \rangle$ with $P_{at}(e,t) = 1$ and $T(e,0) = 0$ for every (e,t). Part (iii) of Proposition 1 combined with $T(e,0) = 0$ imply that $P(e,0,\emptyset) = \Pi(e,0) \geq P(e,t,\emptyset)$ for any (e,t).

Then, construct the mechanism $M' := \langle R',T',P' \rangle$ with $R'(e,t) := e$, $P'(e,t,\emptyset) = P(e,0,\emptyset)$ and[18]

$$T'(e,t) := 1 - \frac{(1-T(e,t))\,(1-P(e,t,\emptyset))}{1-P(e,0,\emptyset)} = 1 - \frac{(1-T(e,t))\,(1-P(e,t,\emptyset))}{1-\Pi(e,0)} \leq T(e,t),$$

where the second equality follows from $T(e,0) = 0$ and the inequality from $\Pi(e,0) \geq P(e,t,\emptyset)$, which follows from part (iii) Proposition 1 combined with the fact that M is IC.

By construction we have that $\Pi'(e,t) = \Pi(e,t)$ for every (e,t), so M' satisfies conditions (i) and (ii) of Proposition 1. By IC of the original mechanism M (combined with part (iii) of Proposition 1) and given the construction of M', we also have that $\Pi'(e,0) \geq P'(e,t,\emptyset)$ for every (e,t). Thus, M' also satisfies condition (iii) of Proposition 1. Therefore, M' is IC.

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

Clearly written proofs of the remaining results are in progress. A major part of the proofs is Lemma 5, which is used to show the existence of a bang-bang solution.

Lemma 5. Let $f : [0,1] \to \mathbb{R}$ be a bounded and continuous function. Then the problem of choosing a non-decreasing $g : [0,1] \to [0,1]$ to maximize $\int_0^1 g(x)f(x)dx$ has a bang-bang solution (*i.e.* a solution of the form $g(x) = \mathbb{1}(x \geq \underline{x})$ for some $\underline{x} \in [0,1]$).

[18] For e such that $P(e,0,\emptyset) = 1$, set $T'(e,t) = 0$ for every $t \in [0,1]$.

Proof of Lemma 5 Take any non-decreasing g. First, notice that since g has to be non-decreasing, it will have at most countably many discontinuity points, so given also that f is continuous, gf has at most countably many discontinuity points. This, combined with the fact that gf is bounded implies that gf is indeed integrable.

Also, there is a sequence of non-decreasing step functions $\{g_n\}_{n=1}^{\infty}$ that converges pointwise to g almost everywhere. Then, by the Dominated Convergence Theorem, $\lim_{n\to\infty} \int_0^1 g_n(x)f(x)dx = \int_0^1 g(x)f(x)dx$.

Further, for any function g_n of the sequence, there exists a function h_n of the form $h_n(x) = \mathbb{1}(x \geq \underline{x}_n)$ for some $\underline{x}_n \in [0,1]$ such that[19]

$$\int_0^1 g_n(x)f(x)dx \leq \int_0^1 h_n(x)f(x)dx.$$

h_n has a convergent subsequence, so wlog let h be the limit of h_n. We have then that

$$\int_0^1 g(x)f(x)dx = \lim_{n\to\infty} \int_0^1 g_n(x)f(x)dx \leq \lim_{n\to\infty} \int_0^1 h_n(x)f(x)dx = \int_0^1 h(x)f(x)dx.$$

We conclude that for any non-decreasing function g, there exists a function $h(x) = \mathbf{I}(x \geq \underline{x})$ for some $\underline{x} \in [0,1]$ such that $\int_0^1 g(x)f(x)dx \leq \int_0^1 h(x)f(x)dx$, so the problem has a bang-bang solution. **Q.E.D.**

[19]To see this, take a function $g_n(k) \equiv \sum_{i=1}^{m} t_i \mathbf{I}(x \in [x_{i-1},x_i))$ with $0 = x_0 < x_1 < \cdots < x_m = 1$ and $t_1 < t_2 < \cdots < t_m$. Then, construct g_n' to be equal to g_n except for $x \in [0,x_1]$; for such x, (i) set $t_1' = 0$ if $\int_{x_0}^{x_1} f(x)dx \leq 0$, (ii) otherwise set $t_1' = t_2$ (or set $t_1' = 1$ if there is no t_2). This process defines an operator Φ on a step function that reduces the step function's number of steps by one and makes $\int_0^1 \Phi(g_n)(x)f(x)dx \geq \int_0^1 g_n(x)f(x)dx$. Applying this operator m times to g_n gives a function $\Phi^m(g_n)$ of the form $\Phi^m(g_n)(x) = \mathbb{1}(x \geq \underline{x}_n)$ that makes the objective function at least as high as g_n.

For a fixed $\overline{e}$

$$\begin{aligned}
&V\left(\lambda\Pi+(1-\lambda)\Pi'\right)\\
=&\int_0^{\overline{e}}\int_0^1\left[\begin{array}{r}\left(\lambda\Pi(e,t)+(1-\lambda)\Pi'(e,t)\right)u(e,t)\\ -c\dfrac{(\lambda\Pi(e,t)+(1-\lambda)\Pi'(e,t))-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))}{1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))}\end{array}\right]f(e,t)dtde\\
&+\int_{\overline{e}}^1\int_0^1 u(e,t)f(e,t)dtde.
\end{aligned}$$

while

$$\begin{aligned}
&\lambda V(\Pi)+(1-\lambda)V(\Pi')\\
=&\int_0^{\overline{e}}\int_0^1\left[\begin{array}{r}\left(\lambda\Pi+(1-\lambda)\Pi'\right)u(e,t)-c\dfrac{\lambda\Pi(e,t)-\lambda\Pi(e,0)}{1-\Pi(e,0)}\\ -c\dfrac{(1-\lambda)\Pi'(e,t)-(1-\lambda)\Pi'(e,0)}{1-\Pi'(e,0)}\end{array}\right]f(e,t)dtde\\
&+\int_{\overline{e}}^1\int_0^1 u(e,t)f(e,t)dtde.
\end{aligned}$$

so that

$$\begin{aligned}
&\lambda V(\Pi)+(1-\lambda)V(\Pi')-V\left(\lambda\Pi+(1-\lambda)\Pi'\right)\\
=&c\int_0^{\overline{e}}\int_0^1\left[\frac{(\lambda\Pi(e,t)+(1-\lambda)\Pi'(e,t))-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))}{1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))}\right]f(e,t)dtde\\
&-c\int_0^{\overline{e}}\int_0^1\left[\frac{\lambda\Pi(e,t)-\lambda\Pi(e,0)}{1-\Pi(e,0)}+\frac{(1-\lambda)\Pi'(e,t)-(1-\lambda)\Pi'(e,0)}{1-\Pi'(e,0)}\right]f(e,t)dtde\\
\propto&\int_0^{\overline{e}}\int_0^1\left[\frac{(\lambda\Pi(e,t)+(1-\lambda)\Pi'(e,t))-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))}{1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))}\right]f(e,t)dtde\\
&-\int_0^{\overline{e}}\int_0^1\left[\frac{\lambda\Pi(e,t)-\lambda\Pi(e,0)}{1-\Pi(e,0)}+\frac{(1-\lambda)\Pi'(e,t)-(1-\lambda)\Pi'(e,0)}{1-\Pi'(e,0)}\right]f(e,t)dtde\\
=&\int_0^{\overline{e}}\int_0^1\frac{\left\{\begin{array}{r}\{\lambda\left[\Pi(e,t)-\Pi(e,0)\right]+(1-\lambda)\left[\Pi'(e,t)-\Pi'(e,0)\right]\}\left[1-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right]\\ -\lambda\left[\Pi(e,t)-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right]\left[1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\\ -(1-\lambda)\left[\Pi'(e,t)-\Pi'(e,0)\right]\left[1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\left[1-\Pi(e,0)\right]\end{array}\right\}}{\left[1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\left[1-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right]}f(e,t)dtde\\
=&\int_0^{\overline{e}}\int_0^1\frac{\left\{\begin{array}{r}-\lambda\left[\Pi(e,t)-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right]\left[\Pi(e,0)-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\\ -(1-\lambda)\left[\Pi'(e,t)-\Pi'(e,0)\right]\left[\Pi'(e,0)-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\left[1-\Pi(e,0)\right]\end{array}\right\}}{\left[1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\left[1-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right]}f(e,t)dtde
\end{aligned}$$

$$
\begin{aligned}
&=\int_0^{\overline{e}}\int_0^1 \frac{\begin{Bmatrix} -\lambda(1-\lambda)\left[\Pi(e,t)-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right]\left[\Pi(e,0)-\Pi'(e,0)\right] \\ -\lambda(1-\lambda)\left[\Pi'(e,t)-\Pi'(e,0)\right]\left[\Pi'(e,0)-\Pi(e,0)\right]\left[1-\Pi(e,0)\right] \end{Bmatrix}}{\left[1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\left[1-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right]} f(e,t)dtde \\
&=\lambda(1-\lambda)\int_0^{\overline{e}}\int_0^1 \frac{\left[\Pi'(e,0)-\Pi(e,0)\right]\begin{Bmatrix} \left[\Pi(e,t)-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right] \\ -\left[\Pi'(e,t)-\Pi'(e,0)\right]\left[1-\Pi(e,0)\right] \end{Bmatrix}}{\left[1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\left[1-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right]} f(e,t)dtde \\
&=\lambda(1-\lambda)\int_0^{\overline{e}}\int_0^1 \frac{\left[\Pi'(e,0)-\Pi(e,0)\right]\begin{Bmatrix} \Pi(e,t)-\Pi(e,0)-\Pi'(e,0)\Pi(e,t) \\ -\left[\Pi'(e,t)-\Pi'(e,0)-\Pi(e,0)\Pi'(e,t)\right] \end{Bmatrix}}{\left[1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\left[1-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right]} f(e,t)dtde \\
&\geq\lambda(1-\lambda)\int_0^{\overline{e}}\int_0^1 \frac{\left[\Pi'(e,0)-\Pi(e,0)\right]\begin{Bmatrix} \Pi(e,t)-\Pi(e,0)-\Pi(e,0)\Pi(e,t) \\ -\left[\Pi'(e,t)-\Pi(e,0)-\Pi(e,0)\Pi'(e,t)\right] \end{Bmatrix}}{\left[1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\left[1-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right]} f(e,t)dtde \\
&=\lambda(1-\lambda)\int_0^{\overline{e}}\int_0^1 \frac{\left[\Pi'(e,0)-\Pi(e,0)\right]\left[\Pi(e,t)-\Pi'(e,t)\right]\left[1-\Pi(e,0)\right]}{\left[1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\left[1-\Pi(e,0)\right]\left[1-\Pi'(e,0)\right]} f(e,t)dtde \\
&=\lambda(1-\lambda)\int_0^{\overline{e}}\int_0^1 \frac{\left[\Pi'(e,0)-\Pi(e,0)\right]\left[\Pi(e,t)-\Pi'(e,t)\right]}{\left[1-(\lambda\Pi(e,0)+(1-\lambda)\Pi'(e,0))\right]\left[1-\Pi'(e,0)\right]} f(e,t)dtde
\end{aligned}
$$

Fix some non-decreasing $\Pi(e,0)$ and consider the optimal choice of $\Pi(e,t)$ (for (e,t) with $t > 0$) given that $\Pi(e,0)$. The second and third term in (**??**) are fixed given $\Pi(e,0)$, so we need to choose $\Pi(e,t)$ to maximize the first term, which can be rewritten as

$$\int_0^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \Pi(e,\tau(e,s)) \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)} \right) f(e,\tau(e,s)) de ds \tag{7}$$

Let $\gamma = 0$ and define the non-decreasing function $\underline{t} : [0,\overline{e}) \to [\underline{q} + c, +\infty)$, given by $\underline{t}(e) \coloneqq \underline{q} + c/(1 - \Pi(e,0))$, which for any (e,t) with $e < \overline{e}$ satisfies

$$u(e,t) \overset{(resp.<)}{>} \frac{c}{1-\Pi(e,0)} \Longleftrightarrow t \overset{(resp.<)}{>} \underline{t}(e).$$

Without loss, let $\Pi(0,0) = 0$, which implies $\underline{t}(0) = \underline{q} + c$. To deal with possible jumps in $\underline{t}$, define the correspondence $\underline{t}_f : [0,\overline{e}) \rightrightarrows [\underline{q}+c, +\infty)$ given by $\underline{t}_f(e) \coloneqq [\lim_{x\uparrow e} \underline{t}(x), \lim_{x\downarrow e} \underline{t}(x)]$, where gaps due to jumps in $\underline{t}$ are "filled in."[20] Denote by $\mathcal{G} \coloneqq \{(x,y) \in [0,\overline{e}) \times [\underline{q} + c, 1] : y \in \underline{t}_f(x)\}$ the graph of $\underline{t}_f$ in the unit square.

Fix the value of $\Pi(e,t)$ for every $(e,t) \in \widetilde{\mathcal{G}} \coloneqq \mathcal{G} \cup \{0\} \times [0, \underline{q} + c]$. Without loss, let $\Pi(e,t) = 0$ for every $(e,t) \in \{0\} \times [0, \underline{q}+c]$. Define the function $(e^*,t^*) : [0,\overline{s}) \to [0,\overline{e}) \times [0,1]$ implicitly given by

$$(e^*(s),t^*(s)) \in \widetilde{\mathcal{G}} \cap I_s$$

where $\overline{s} \coloneqq \widehat{t}(e', \lim_{e'\uparrow\overline{e}} \underline{t}(e'),0)$. That is, the function returns the (unique) type $(e^*(s),t^*(s))$ that lies on the intersection of $\widetilde{\mathcal{G}}$ and iso-test-score curve s. Notice that choosing the value of $\Pi(e,t)$ for every $(e,t) \in \widetilde{\mathcal{G}}$ is the same as choosing $\Pi(e^*(s),t^*(s))$ for every $s \in [0,\overline{s})$. To satisfy IC, $\Pi(e^*(s),t^*(s))$ must be non-decreasing in s and satisfy $\Pi(e^*(s),t^*(s)) \geq \Pi(e^*(s),0)$ for every s.

Now, define the sets $G^{\uparrow} \coloneqq \{(x,y) \in [0,1]^2 : y > \underline{t}_f(x)\}$ and $G^{\downarrow} \coloneqq \{(x,y) \in [0,1]^2 : y < \underline{t}_f(x)\}$.

For every $(e,t) \in G^{\uparrow}$, the principal wants to make $\Pi(e,t)$ as high as possible. Given Lemma 4, condition (ii) of Corollary 1.1 requires that $\Pi(e^*(s),t^*(s)) \geq \Pi(e, \tau(e,s))$ for every $e \leq e^*(s)$. If we only impose this condition, then the highest we can make $\Pi(e,\tau(e,s))$

[20]If $\Pi(\cdot,0)$ is continuous, so that $\underline{t}$ also is, then $\underline{t} = \underline{t}_f$.

for $(e,\tau(e,s)) \in G^{\uparrow}$ is[21]

$$\Pi(e,\tau(e,s)) = \begin{cases} 1 & \text{if } s \geq \overline{s} \text{ and } (e,\tau(e,s)) \in G^{\uparrow} \\ \Pi\left(e^*(s),t^*(s)\right) & \text{if } s < \overline{s} \text{ and } (e,\tau(e,s)) \in G^{\uparrow}, \end{cases}$$

That is, every type on I_s with evidence less than the type$(e^*(s),t^*(s)) \in \widetilde{\mathcal{G}}$ is hired with the same probability as $(e^*(s),t^*(s))$. If s is so high that no type on $\widetilde{\mathcal{G}}$ is on I_s, then every type on I_s is hired with probability 1. This Π also automatically satisfies condition (i) of Corollary 1.1 because $\Pi(e^*(s),t^*(s))$ is non-decreasing in s and satisfies $\Pi(e^*(s),t^*(s)) \geq \Pi(e^*(s),0)$.[22] We conclude that after an IC $\Pi(e,t)$ is chosen for $(e,t) \in \widetilde{\mathcal{G}}$ and $(e,t) \in [0,1] \times \{0\}$, the optimal $\Pi(e,\tau(e,s))$ is constant along the part of each isotest-score curve I_s with evidence less than $e^*(s)$.

Similarly, for every $(e,t) \in G^{\downarrow}$, the principal wants to make $\Pi(e,t)$ as low as possible. However, now condition (i) of Corollary 1.1 is not automatically satisfied for this subset of types if condition (ii) is satisfied. For each e and s, the lowest we can make $\Pi(e,\tau(e,s))$ (for $\Pi(e,\tau(e,s))$ to be IC given $\Pi(e,0)$ and $\Pi\left(e^*(s), t^*(s)\right)$) is[23]

$$\begin{aligned} \Pi(e,\tau(e,s)) &= \max\left\{\Pi\left(e^*(s), t^*(s)\right), \Pi(e,0)\right\} \\ &= \begin{cases} \Pi(e,0) & \text{if } s \leq \underline{q} + c \text{ and } (e,\tau(e,s)) \in G^{\downarrow} \\ \max\left\{\Pi\left(e^*(s), t^*(s)\right), \Pi(e,0)\right\} & \text{if } s > \underline{q} + c \text{ and } (e,\tau(e,s)) \in G^{\downarrow}, \end{cases} \end{aligned}$$

where the second equality follows because without loss we have set $\Pi\left(e^*(s), t^*(s)\right) = 0$ for $s \in [0, \underline{q} + c]$. Setting the value of $\Pi(e,t)$ on $G^{\downarrow}$ this way also automatically globally condition (i) of Corollary 1.1.

We conclude the following. Fixing the values of $\Pi(e,t)$ for $(e,t) \in \widetilde{\mathcal{G}}$ and $(e,t) \in [0,1] \times \{0\}$, so that $\Pi(e,t)$ is IC in this subset of types, we have found the optimal value of $\Pi(e,t)$ for every other (e,t). We have thereby reduced the problem to choosing non-decreasing $\Pi(e,0)$ for every $e \in [0,1]$ and $\Pi(e^*(s),t^*(s))$ non-decreasing in s for every

[21]The following is a *necessary* bound to satisfy IC for every agent type. With Π set equal to that bound, we can however see that IC for every type is indeed automatically satisfied. Also, because $\tau(e')$ is monotone in e', its left limit exists.

[22]It also does not impose any additional constraints for IC (that are not already imposed by the values of $\Pi(e,t)$ for $(e,t) \in \widetilde{\mathcal{G}}$) on the remaining subset of types $G^{\downarrow}$.

[23]The following is a *necessary* bound to satisfy IC for every agent type. With Π set equal to that bound, we can however see that IC for every type is indeed automatically satisfied.

$s \in [0,\overline{s})$ with $\Pi(e^*(s),t^*(s)) \geq \Pi(e^*(s),0)$ to maximize

$$\begin{aligned}
&\int_0^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \max\left\{\Pi\left(e^*(s),t^*(s)\right),\Pi(e,0)\right\}\left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))deds \\
&\quad + c\int_0^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \frac{\Pi(e,0)}{1-\Pi(e,0)} f(e,\tau(e,s))deds + \int_0^1 \int_{\overline{e}}^1 u(e,t)f(e,t)dedt \\
&= \int_0^1 \Pi\left(e^*(s),t^*(s)\right) \int_0^{e^*(s)} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))deds \\
&\quad + \int_0^1 \int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}} \max\left\{\Pi\left(e^*(s),t^*(s)\right),\Pi(e,0)\right\}\left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))deds \\
&\quad + c\int_0^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \frac{\Pi(e,0)}{1-\Pi(e,0)} f(e,\tau(e,s))deds + \int_0^1 \int_{\overline{e}}^1 u(e,t)f(e,t)dedt
\end{aligned}$$

or equivalently

$$\begin{aligned}
&\int_0^{s^*} \Pi\left(e^*(s),t^*(s)\right) \int_0^{e^*(s)} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))deds \\
&+ \int_0^{s^*} \int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}} \max\left\{\Pi\left(e^*(s),t^*(s)\right),\Pi(e,0)\right\}\left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))deds \\
&+ \int_{s^*}^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))deds \\
&+ c\int_0^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \frac{\Pi(e,0)}{1-\Pi(e,0)} f(e,\tau(e,s))deds + \int_0^1 \int_{\overline{e}}^1 u(e,t)f(e,t)dedt
\end{aligned}$$

where $s^* := \min\{s \in [0,1] : \Pi\left(e^*(s),t^*(s)\right) = 1\} \leq \overline{s}$.

Solve instead the problem of maximizing

$$\begin{aligned}
&\int_0^{s^*} \Pi\left(e^*(s),t^*(s)\right) \int_0^{e^*(s)} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))deds \\
&\quad + \int_0^{s^*} \Pi\left(e^*(s),t^*(s)\right) \int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))deds \\
&\quad + \int_{s^*}^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))deds \\
&\quad + c\int_0^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \frac{\Pi(e,0)}{1-\Pi(e,0)} f(e,\tau(e,s))deds + \int_0^1 \int_{\overline{e}}^1 u(e,t)f(e,t)dedt \\
&= \int_0^{s^*} \Pi\left(e^*(s),t^*(s)\right) \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))deds \\
&\quad + \int_{s^*}^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))deds \\
&\quad + c\int_0^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \frac{\Pi(e,0)}{1-\Pi(e,0)} f(e,\tau(e,s))deds + \int_0^1 \int_{\overline{e}}^1 u(e,t)f(e,t)dedt,
\end{aligned}$$

where we have relaxed the IC constraint for types in $G^{\uparrow}$. For any fixed $\Pi(e,0)$ and s^*, the set of non-decreasing functions $\Pi(e^*(s),t^*(s))$ that dominate $\Pi(e^*(s),0)$ (*i.e.* $\Pi(e^*(s),t^*(s)) \geq \Pi(e^*(s),0)$ for every $s \in [0, s^*)$) is a convex set. Also, the objective function of the relaxed problem is convex (specifically, linear) in $\Pi(e^*(s),t^*(s))$, so by Bauer's maximum principle an extreme point (of the set of admissible $\Pi(e^*(s),t^*(s))$) maximizes the objective function. By Theorem 1 in Yang and Zentefis (2023), an extreme point has the form:

$$\Pi(e^*(s),t^*(s)) = \begin{cases} 1 & \text{if } s \in [s^*, 1] \\ \lim_{x\uparrow \overline{s}_n} \Pi(e^*(x),0) & \text{if } s \in \cup_{n=1}^{\infty}[\underline{s}_n,\overline{s}_n) \\ \Pi(e^*(s),0) & \text{if } s \notin \cup_{n=1}^{\infty}[\underline{s}_n,\overline{s}_n) \\ 0 & \text{if } s \in [0, \underline{q}+c), \end{cases}$$

where $\{[\underline{s}_n,\overline{s}_n)\}_{n=1}^{\infty}$ is a countable collection of (disjoint) intervals with $\underline{s}_1 = \underline{q}+c$ and $\lim_{n\to\infty} \overline{s}_n = s^*$. The objective function then becomes

$$\begin{aligned} &\int_{\underline{q}+c}^{s^*} \Pi\left(e^*(s), t^*(s)\right) \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s)) deds \\ &+ \int_{s^*}^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s)) deds \\ &+ c\int_0^1 \int_0^{\min\{\overline{e},\widetilde{e}(s)\}} \frac{\Pi(e,0)}{1-\Pi(e,0)} f(e,\tau(e,s)) deds + \int_0^1 \int_{\overline{e}}^1 u(e,t) f(e,t) dedt, \end{aligned}$$

or equivalently

$$\int_0^{\overline{e}} \int_{\widehat{t}(e,0,0)}^1 \max\left\{\Pi\left(e^*(s),\tau(e^*(s),s)\right),\Pi(e,0)\right\}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))dsde$$
$$+c\int_0^{\overline{e}} \frac{\Pi(e,0)}{1-\Pi(e,0)}\int_{\widehat{t}(e,0,0)}^1 f(e,\tau(e,s))dsde+\int_{\overline{e}}^1 \int_{\widehat{t}(e,0,0)}^1 u(e,\tau(e,s))f(e,\tau(e,s))dsde$$

or equivalently

$$\int_0^1 \Pi(e,0)\int_{\widehat{t}(e,0,0)}^1 u(e,\tau(e,s))f(e,\tau(e,s))dsde$$
$$+\int_0^{\overline{e}} \int_{\widehat{t}(e,0,0)}^1 \max\left\{\Pi\left(e^*(s),\tau(e^*(s),s)\right)-\Pi(e,0),0\right\}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))dsde.$$

The problem can equivalently be seen as choosing a non-decreasing function $\Pi(e,0)$ for every $e \in [0,1]$ and a non-decreasing $g : [0,1] \to [0,1]$ to be understood as a candidate value of $\Pi(e^*(s), \tau(e^*(s),s))$ for every $s \in [0,1]$ to maximize[24]

$$\int_0^{\overline{e}} \int_{\widehat{t}(e,0,0)}^1 \max\left\{\max\{g(s), \Pi(e^*(s),0)\}, \Pi(e,0)\right\} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))dsde$$
$$+c\int_0^{\overline{e}} \frac{\Pi(e,0)}{1-\Pi(e,0)} \int_{\widehat{t}(e,0,0)}^1 f(e,\tau(e,s))dsde + \int_{\overline{e}}^1 \int_{\widehat{t}(e,0,0)}^1 u(e,\tau(e,s))f(e,\tau(e,s))dsde$$

or equivalently

$$\int_0^{\overline{e}} \int_{\widehat{t}(e,0,0)}^1 \max\left\{g(s), \Pi(e^*(s),0), \Pi(e,0)\right\} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s))dsde$$
$$+c\int_0^{\overline{e}} \frac{\Pi(e,0)}{1-\Pi(e,0)} \int_{\widehat{t}(e,0,0)}^1 f(e,\tau(e,s))dsde + \int_{\overline{e}}^1 \int_{\widehat{t}(e,0,0)}^1 u(e,\tau(e,s))f(e,\tau(e,s))dsde.$$

Only the first term depends on g.

[24]Of course, the subset of types (e,t) to which the function f will give values to will depend on $\Pi(e,0)$ (since e^* depends on $\Pi(e,0)$) but whatever this subset is, the values will be given by f.

The problem can equivalently be seen as choosing a non-decreasing function $f : [0,1] \to [0,1]$ to be understood as a candidate value of $\Pi(e^*(s), \tau(e^*(s),s))$ for every $s \in [0,1]$.[25] and a non-decreasing $\Pi(e,0)$ (that does not have to satisfy $\Pi(e^*(s),0) \leq f(s)$) to maximize

$$\int_0^1 \Pi(e,0) \int_{\widehat{t}(e,0,0)}^1 u(e,\tau(e,s)) f(e,\tau(e,s)) ds de$$
$$+ \int_0^{\overline{e}} \int_{\widehat{t}(e,0,0)}^1 \max\left\{\max\{f(s), \Pi(e^*(s))\} - \Pi(e,0), 0\right\} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s)) ds de$$

or equivalently

$$\int_0^1 \Pi(e,0) \int_{\widehat{t}(e,0,0)}^1 u(e,\tau(e,s)) f(e,\tau(e,s)) ds de$$
$$+ \int_0^{\overline{e}} \int_{\widehat{t}(e,0,0)}^1 \max\left\{f(s) - \Pi(e,0), \Pi(e^*(s)) - \Pi(e,0), 0\right\} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s)) ds de$$

or equivalently

$$\int_{\overline{e}}^1 \int_{\widehat{t}(e,0,0)}^1 u(e,\tau(e,s)) f(e,\tau(e,s)) ds de + c \int_0^{\overline{e}} \frac{\Pi(e,0)}{1-\Pi(e,0)} \int_{\widehat{t}(e,0,0)}^1 f(e,\tau(e,s)) ds de$$
$$+ \int_0^{\overline{e}} \int_{\widehat{t}(e,0,0)}^1 \max\left\{f(s), \Pi(e^*(s)), \Pi(e,0)\right\} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s)) ds de$$

or equivalently

$$\int_{\overline{e}}^1 \int_{\widehat{t}(e,0,0)}^1 u(e,\tau(e,s)) f(e,\tau(e,s)) ds de + c \int_0^{\overline{e}} \frac{\Pi(e,0)}{1-\Pi(e,0)} \int_{\widehat{t}(e,0,0)}^1 f(e,\tau(e,s)) ds de$$
$$+ \int_{\widehat{t}(e,0,0)}^1 \int_0^{\min\{e^*(s),\overline{e}\}} \max\left\{f(s), \Pi(e^*(s))\right\} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s)) ds de$$
$$+ \int_{\widehat{t}(e,0,0)}^1 \int_{\min\{e^*(s),\overline{e}\}}^{\overline{e}} \max\left\{f(s), \Pi(e,0)\right\} \left(u(e,\tau(e,s)) - \frac{c}{1-\Pi(e,0)}\right) f(e,\tau(e,s)) ds de$$

[25] Of course, the subset of types (e,t) to which the function f will give values to will depend on $\Pi(e,0)$ (since e^* depends on $\Pi(e,0)$) but whatever this subset is, the values will be given by f.

The objective function becomes

$$\int_0^{\underline{q}+c}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+\int_{\underline{q}+c}^{1}\int_0^{e^*(s)}\Pi\left(e^*(s),\tau(e^*(s),s)\right)\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+\int_{\underline{q}+c}^{1}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\max\left\{\Pi\left(e^*(s),\tau(e^*(s),s)\right),\Pi(e,0)\right\}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+c\int_0^1\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds+\int_{\overline{e}}^1\int_0^1 u(e,t)f(e,t)dtde,$$

where we have used

$$\int_0^{\overline{e}}\frac{\Pi(e,0)}{1-\Pi(e,0)}\int_0^1 f(e,t)dtde=\int_0^1\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds.$$

The problem can equivalently be seen as choosing a non-decreasing function $f:[0,1]\to[0,1]$ to be understood as a candidate value of $\Pi(e^*(s),\tau(e^*(s),s))$ for every $s\in[0,1]$.[26] and a non-decreasing $\Pi(e,0)$ (that does not have to satisfy $\Pi(e^*(s),0)\leq f(s)$) to maximize

$$\int_0^{\underline{q}+c}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+\int_{\underline{q}+c}^{1}\int_0^{e^*(s)}\max\left\{\Pi(e^*(s),0),f(s)\right\}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+\int_{\underline{q}+c}^{1}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\max\left\{\max\left\{\Pi(e^*(s),0),f(s)\right\},\Pi(e,0)\right\}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+c\int_0^1\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds+\int_{\overline{e}}^1\int_0^1 u(e,t)f(e,t)dtde,$$

where notice that whenever $\Pi(e^*(s),0)>f(s)$, $\Pi(e^*(s),\min\underline{t}_f(e^*(s)))$ is set equal to $\Pi(e^*(s),0)$, so that $\Pi(e^*(s),\min\underline{t}_f(e^*(s)))\geq\Pi(e^*(s),0)$ is satisfied. We have thus build in the constraint $\Pi(e,\min\underline{t}_f(e))\geq\Pi(e,0)$ in the objective function.

The objective function can equivalently be written as

$$\int_0^{\underline{q}+c}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)u(e,\tau(e,s))f(e,\tau(e,s))deds$$
$$+\int_{\underline{q}+c}^{1}\int_0^{e^*(s)}\max\left\{\Pi(e^*(s),0),f(s)\right\}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$

[26]Of course, the subset of types (e,t) to which the function f will give values to will depend on $\Pi(e,0)$ (since e^* depends on $\Pi(e,0)$) but whatever this subset is, the values will be given by f.

$$+\int_{\underline{q}+c}^{1}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\max\left\{f(s),\Pi(e,0)\right\}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+c\int_{\underline{q}+c}^{1}\int_{0}^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds+\int_{\overline{e}}^{1}\int_{0}^{1}u(e,t)f(e,t)dtde,$$

where we have used $\Pi(e,0)\geq\Pi(e^*(s),0)$ for any $e\in[e^*(s),\min\{\overline{e},\widetilde{e}(s)\}]$, or equivalently

$$\int_{0}^{\underline{q}+c}\int_{0}^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)u(e,\tau(e,s))f(e,\tau(e,s))deds$$
$$+\int_{\underline{q}+c}^{1}\int_{0}^{e^*(s)}\max\left\{\Pi(e^*(s),0),f(s)\right\}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+\int_{\underline{q}+c}^{1}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\max\left\{f(s),\Pi(e,0)\right\}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+c\int_{\underline{q}+c}^{1}\int_{0}^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds+\int_{\overline{e}}^{1}\int_{0}^{1}u(e,t)f(e,t)dtde,$$

or equivalently

$$\int_0^{\underline{q}+c}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+\int_{\underline{q}+c}^{1}\Pi\left(e^*(s),\tau(e^*(s),s)\right)\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+\int_{\underline{q}+c}^{1}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\max\left\{\Pi(e,0)-\Pi\left(e^*(s),\tau(e^*(s),s)\right),0\right\}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+c\int_0^{1}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds+\int_{\overline{e}}^{1}\int_0^{1}u(e,t)f(e,t)dtde.$$

or equivalently

$$\int_0^{\underline{q}+c}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)u(e,\tau(e,s))f(e,\tau(e,s))deds$$
$$+\int_{\underline{q}+c}^{1}\Pi\left(e^*(s),\tau(e^*(s),s)\right)\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+\int_{\underline{q}+c}^{1}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\max\left\{\Pi(e,0)-\Pi\left(e^*(s),\tau(e^*(s),s)\right),0\right\}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+c\int_{\underline{q}+c}^{1}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds+\int_{\overline{e}}^{1}\int_0^{1}u(e,t)f(e,t)dtde.$$

The third term is always non-positive because $u(e,\tau(e,s))\leq c/(1-\Pi(e,0))$ for every s and $e>e^*(s)$. Ignore it for the time being and let us maximize

$$\int_{\underline{q}+c}^{1}\Pi\left(e^*(s),\tau(e^*(s),s)\right)\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+c\int_0^{\overline{e}}\frac{\Pi(e,0)}{1-\Pi(e,0)}\int_0^{1}f(e,t)dtde+\int_{\overline{e}}^{1}\int_0^{1}u(e,t)f(e,t)dtde$$

instead, relaxing the problem by ignoring also the constraint $\Pi\left(e^*(s),\tau(e^*(s),s)\right)\geq\Pi(e^*(s),0)$. By Lemma 5, given any $\Pi(e,0)$, the problem of choosing $\Pi\left(e^*(s),\tau(e^*(s),s)\right)$ has a bang-bang solution $\Pi\left(e^*(s),\tau(e^*(s),s)\right)=\mathbf{I}(s\geq s^*)$ for some $s^*\in[\underline{q}+c,1]$. This means that $\Pi(e,t)=1$ for every (e,t) with $\widehat{t}(e,t,0)\geq s^*$, and particularly $T(e,t)=1$ for every (e,t) with $\widehat{t}(e,t,0)\geq s^*$ and $e<\overline{e}$. The original objective function then becomes

The original objective function then becomes

$$\int_0^{\underline{q}+c}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$\int_{s^*}^{1}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$

$$+\int_{\underline{q}+c}^{s^*}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)\left(u(e,\tau(e,s))-\frac{c}{1-\Pi(e,0)}\right)f(e,\tau(e,s))deds$$
$$+c\int_0^{\overline{e}}\frac{\Pi(e,0)}{1-\Pi(e,0)}\int_0^1 f(e,t)dtde+\int_{\overline{e}}^1\int_0^1 u(e,t)f(e,t)dtde$$

or equivalently

$$\begin{aligned}
&\int_{\overline{e}}^1\int_0^1 u(e,t)f(e,t)dtde+\int_0^{\underline{q}+c}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)u(e,\tau(e,s))f(e,\tau(e,s))deds\\
&\int_{s^*}^1\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}u(e,\tau(e,s))f(e,\tau(e,s))deds+\int_{\underline{q}+c}^{s^*}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)u(e,\tau(e,s))f(e,\tau(e,s))deds\\
&-c\int_0^{\underline{q}+c}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds\\
&-c\int_{\underline{q}+c}^{s^*}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds\\
&-c\int_{s^*}^1\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{1}{1-\Pi(e,0)}f(e,\tau(e,s))deds+c\int_0^{\overline{e}}\frac{\Pi(e,0)}{1-\Pi(e,0)}\int_0^1 f(e,t)dtde.
\end{aligned}$$

The first four terms can be rewritten as

$$\begin{aligned}
&\int_0^1\int_{\overline{e}}^{\max\{\overline{e},\widetilde{e}(s)\}}u(e,\tau(e,s))f(e,\tau(e,s))deds+\int_0^{\underline{q}+c}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)u(e,\tau(e,s))f(e,\tau(e,s))deds\\
&\int_{s^*}^1\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}u(e,\tau(e,s))f(e,\tau(e,s))deds+\int_{\underline{q}+c}^{s^*}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)u(e,\tau(e,s))f(e,\tau(e,s))deds\\
=&\int_0^1\int_{\overline{e}}^{\widetilde{e}(s)}\Pi(e,0)u(e,\tau(e,s))f(e,\tau(e,s))deds+\int_0^{\underline{q}+c}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)u(e,\tau(e,s))f(e,\tau(e,s))deds\\
&\int_{s^*}^1\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}u(e,\tau(e,s))f(e,\tau(e,s))deds+\int_{\underline{q}+c}^{s^*}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\Pi(e,0)u(e,\tau(e,s))f(e,\tau(e,s))deds
\end{aligned}$$

The last four terms can be rewritten as

$$\begin{aligned}
&c\left[\begin{aligned}-\int_{\underline{q}+c}^{s^*}\int_{e^*(s)}^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds-\int_0^{\underline{q}+c}\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds\\
\int_0^1\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds-\int_{s^*}^1\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}\frac{1}{1-\Pi(e,0)}f(e,\tau(e,s))deds\end{aligned}\right]\\
=&c\left[\int_{\underline{q}+c}^{s^*}\int_0^{e^*(s)}\frac{\Pi(e,0)}{1-\Pi(e,0)}f(e,\tau(e,s))deds-\int_{s^*}^1\int_0^{\min\{\overline{e},\widetilde{e}(s)\}}f(e,\tau(e,s))deds\right]
\end{aligned}$$