Optimal Test Design: Screening vs Investment

Vasundhara Mallick

European University Institute

A model of costly screening in the presence of moral hazard and adverse selection.

A Principal faces a heterogeneous pool of candidates, and must decide whether to accept of reject them.

She would like to accept only *high* types, but screening is costly.

Prior to screening, agent's can exert effort to improve their type.

Examples: job market screening, firm facing a quality control or regulatory board, college admissions.

Since screening is costly, Principal might not want to evaluate all candidates.

May want to filter some of them using a first stage test.

Such first stage screening tests are often manipulable.

In this setting of costly screening and investment, what is the Principal-optimal first stage test, in the class of all **manipulable** *tests?*

In this setting of costly screening and investment, what is the Principal-optimal first stage test, in the class of all **manipulable** *tests?*

(Surprising?) answer: the best test may be the one that's always manipulated in equilibrium, and filters out no one.

It **appears** to be useless, but incentivises the maximum investment by the agents to improve their types, leading to the best **pool** of agents.

Benchmark Model

A Principal (P), and a unit mass of agents.

Each agent has two possible types: High (h) or Low (l).

Initial mass of high type agents: $\pi \in (0, 1)$.

Principal has a binary decision: $a \in \{A, R\}$.

- P's payoff from accepting h: 1, from accepting l: -1, from rejecting: 0.
- Agent's payoff: 1 from getting accepted, 0 from rejection.

Before deciding, P can engage in *costly information acquisition*.

Before this screening stage, agents (low type) can *invest* in their type.

Timeline:

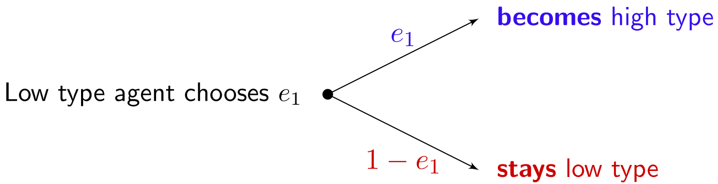
- Agents invest in their type.
- Some low types may become high type; *new pool of candidates*.
- P evaluates this pool, and makes her decision.

Agent's Investment Stage

Agents of low type choose effort $e_1 \in [0, 1]$; cost of effort is $c_1(e_1) = k_1 e_1^2$, $k_1 > 0$.

Agent's Investment Stage

Agents of low type choose effort $e_1 \in [0, 1]$; cost of effort is $c_1(e_1) = k_1 e_1^2$, $k_1 > 0$.



Agent's Investment Stage

Once the result of agents' effort realises, there is a *new* pool of candidates.

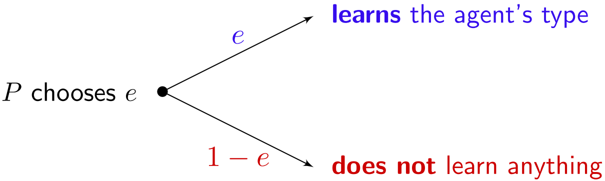
What is the new pool?

- mass $(1 - \pi)$: initially bad.
- After effort: $(1 - \pi)e_1$ high type, $(1 - \pi)(1 - e_1)$ low type.
- Posterior going into the screening phase: $\pi + (1 - \pi)e_1$.

Principal's Evaluation Stage

Once the new pool realises, P exerts costly effort to acquire information about the candidates.

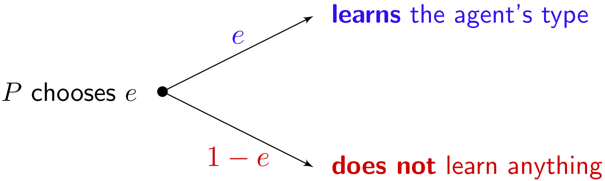
Chooses effort $e \in [0, 1]$. Cost given by $c(e) = ke^2$, where $k > 0$.



Principal's Evaluation Stage

Once the new pool realises, P exerts costly effort to acquire information about the candidates.

Chooses effort $e \in [0, 1]$. Cost given by $c(e) = ke^2$, where $k > 0$.



After the result of effort, makes decision.

- If effort successful, accepts h, rejects l.
- If effort unsuccessful, optimal decision depends on beliefs about pool she is evaluating.

Equilibrium of Benchmark Model

Solution concept: Perfect Bayesian Equilibrium.

- Agents' effort is optimal given her beliefs about P's screening effort.
- P's effort, and her action if she learns nothing must be sequentially rational given her beliefs about the pool of candidates she is facing.
- Beliefs must be derived using Baye's Rule, wherever possible.

Model with first stage test

Screening is costly, P may want to screen out some candidates using a first stage test.

- Filtering through resumes or aptitude test before conducting interviews.
- SAT in college admissions.
- Quality certification: first stage test to determine if minimum quality standards are met.

These tests are often costless for P, or of negligible cost compared to the eventual costly evaluation.

Model with first stage test

Screening is costly, P may want to screen out some candidates using a first stage test.

- Filtering through resumes or aptitude test before conducting interviews.
- SAT in college admissions.
- Quality certification: first stage test to determine if minimum quality standards are met.

These tests are often costless for P, or of negligible cost compared to the eventual costly evaluation.

Problem: These tests are often manipulable (at a cost).

- resume padding
- gaming the SAT
- Companies can cheat on quality tests for their products, emission tests for cars.

Test

A test is a Blackwell Experiment. For today, restrict attention to a simple class of tests.

Test

A test is a Blackwell Experiment. For today, restrict attention to a simple class of tests.

Pass-Fail tests: high type passes wp1, low type passes with prob $p \in [0, 1]$.

So p is the parameter that characterises the test.

Test

A test is a Blackwell Experiment. For today, restrict attention to a simple class of tests.

Pass-Fail tests: high type passes wp1, low type passes with prob $p \in [0, 1]$.

So p is the parameter that characterises the test.

Those who pass the test are then subject to costly screening by P.

Test

A test is a Blackwell Experiment. For today, restrict attention to a simple class of tests.

Pass-Fail tests: high type passes wp1, low type passes with prob $p \in [0, 1]$.

So p is the parameter that characterises the test.

Those who pass the test are then subject to costly screening by P.

Tests

Test is falsifiable: At a cost c, type l can *falsify* her input; for the test, it is **as if** he is the high type.

Tests

Test is falsifiable: At a cost c, type l can *falsify* her input; for the test, it is **as if** he is the high type.

So, p is the *natural probability* of low type passing:

- if he falsifies, she passes with probability 1.
- If not, he passes with probability p.

Tests

Test is falsifiable: At a cost c, type l can *falsify* her input; for the test, it is **as if** he is the high type.

So, p is the *natural probability* of low type passing:

- if he falsifies, she passes with probability 1.
- If not, he passes with probability p.

Important: By falsifying, the low type does not **become** the high type.

She passes the test with prob 1, but *remains* teh low type for the subsequent costly screening.

Tests

Let β_2 be the probability with which the low type is accepted in second stage screening.

- Payoff from falsifying: $\beta_2 - c$.
- Payoff from not falsifying: $p\beta_2$.
- So low type falsifies if and only if $p \leq \frac{\beta - c}{\beta}$

Of course, β_2 is determined in equilibrium.

New timeline:

- Low type chooses investment effort: may or may not become high type.
- Those who are still low type, choose whether or not to manipulate the test.
- Those who pass are evaluated by P.

New timeline:

- Low type chooses investment effort: may or may not become high type.
- Those who are still low type, choose whether or not to manipulate the test.
- Those who pass are evaluated by P.

As before, solution concept: PBE

Agent's falsification decision is optimal given P's equilibrium screening effort.

Agent's investment effort is optimal given her payoff from the falsification stage.

P's screening effort is optimal given her beliefs about the pool that has passed.

Results

Proposition

A test **always** makes P strictly better off: *even when* all low types falsify it, i.e., it does **no screening** in equilibrium.

Intuition

After investment stage, some pool of high and low types.

If all low types falsify the test, then the test is *uninformative* in equilibrium.

The why is P strictly better off compared to **no-test** equilibrium?

Intuition

After investment stage, some pool of high and low types.

If all low types falsify the test, then the test is *uninformative* in equilibrium.

The why is P strictly better off compared to **no-test** equilibrium?

Encourages higher investment effort by the low type, leading to a better *pool* of agents.

- Suppose $p = 0$.
- High type passes for free, low type must incur a cost c to pass.
- Test is an *extra* costly hurdle for the low type that wasn't there before.
- Increases the relative cost of being the low type.

Results

In fact, a test could be most effective when it does **not** screen in equilibrium.

This is because there is a trade off between screening and investment.

A test that screens is a test that the agent does *not* find optimal to falsify.

- $p \geq \frac{\beta_2 - c}{\beta_2}$.
- Since the test is not falsified, β_2 is actually higher in this equilibrium than in an equilibrium with a test that **is** falsified.
- So, in an equilibrium where screening occurs, post falsification payoff better, yet optimal to not falsify.
- So the payoff from *staying the low* type is higher.
- This discourages effort at the investment stage.

Main Result

Theorem

There exists a k^*, such that when $k \leq k^*$, i.e. P's information acquisition cost is *low*, the optimal test is one that all low types falsify, i.e. one that does no screening in equilibrium. When $k > k^*$, the optimal test involves $p \in (0, 1)$, such that in equilibrium, low types find it optimal to **not** falsify the test, i.e. the test screens out some low types in equilibrium.

Example

An example to illustrate the main forces

First: Equilibrium of benchmark model with no test.

Add test: In an equilibrium where the test is falsified wp1, P strictly better off than with no test.

Consider another test that's not falsified and screens in equilibrium: this leads to less effort by agents and a worse pool of agents.

Example

$\pi = 0.5.$

$k_1 = k = 0.5$, so agent's cost is $c_1(e_1) = 0.5e_1^2$, and P's cost is $c(e) = 0.5e^2$.

Principal's optimization problem:

- Suppose her belief is π' (she believes this is the mass of high types).
- She chooses effort to maximise:

$$e\pi' + (1-e)\max\{\pi' - (1-\pi'), 0\} - 0.5e^2$$

- Suppose $\pi' \geq 0.5$. Then P is maximising

$$e\pi' + (1-e)\{\pi' - (1-\pi')\} - 0.5e^2$$

- Optimal effort $e^* = 1 - \pi'$.

Example (continued)

Agents' optimization problem:

- Suppose P's strategy is to accept wp1 if effort is unsuccessful.
- If this is the case, and the P's effort is e:
  - h gets accepted wp1.
  - l gets accepted wp $1 - e$.
- So, if agent believes that P's effort is e, he is maximising:

$$e_1 + (1 - e_1)(1 - e) - 0.5e_1^2$$

- Optimal effort given e is $e_1^*(e) = e$

Example (continued)

Let e^* and e_1^* be the equilibrium levels of P's and agents' efforts.

Step 1: The pool of high types *after* investment has mass $\pi + (1-\pi)e_1^*$ of high types.

Step 2: In equilibrium, P correctly anticipates agents' effort. So her belief prior to screening: $\pi' = \pi + (1-\pi)e_1^*$

Step 3: In equilibrium, irrespective of e^*, $\pi' \geq 0.5$ always holds, as $\pi = 0.5$.

Step 4: Therefore, $e^* = 1 - \pi' = (1-\pi)(1-e_1)$

Step 5: For agent:$e^* = e_1^*$. Replacing this into the expression for e^*, and since $\pi = 0.5$, we have

$$e^* = 0.5(1-e^*)$$

So, $e^* = e_1^* = \frac{1}{3}$.

Example (continued)

We assumed for the Agent's best response that P's strategy is to accept if effort is unsuccessful.

In equilibrium, P's posterior is $\pi' = 0.5 + 0.5(\frac{1}{3}) > 0.5$.

So P would indeed like to accept wp1 if effort is unsuccessful.

This is the unique equilibrium outcome.

Add test

Let $c = 0.1$.

We will consider two tests:

1. $p = 0$: This test will be falsified wp1 in equilibrium. The equilibrium effort of the low type is $e_1^* = 0.4 > \frac{1}{3}$.
2. $p = 0.9$: In equilibrium, low types will not falsify. So, with prob 0.1, low types do not pass. Equilibrium effort at the investment stage is $e_1^* = 0.3654$. This is lower than 0.4 but higher than investment effort in the no test equilibrium.

$p = 0$

Let e^* and e_1^* be the equilibrium levels of P's and agents' efforts.

Claim: In equilibrium, $e^* \leq 0.5$ must hold.

- $0.5 + 0.5e_1^* \geq 0.5$ the mass of high types after the investment stage. So $1 - \pi' \leq 0.5$ are low types.
- Low types may or may not falsify.
- In equilibrium, $\pi' \geq 0.5$. Since $e^* = 1 - \pi'$, $e^* \leq 0.5$.

$p = 0$

Let e^* and e_1^* be the equilibrium levels of P's and agents' efforts.

Claim: In equilibrium, $e^* \leq 0.5$ must hold.

- $0.5 + 0.5e_1^* \geq 0.5$ the mass of high types after the investment stage. So $1 - \pi' \leq 0.5$ are low types.
- Low types may or may not falsify.
- In equilibrium, $\pi' \geq 0.5$. Since $e^* = 1 - \pi'$, $e^* \leq 0.5$.

Claim 2: In equilibrium, this test must be falsified wp1:

- Don't falsify: 0
- Falsify: $(1 - e^*) - c$, where $c = 0.1$, and $1 - e^* \geq 0.5$.

$p = 0$

Principal's optimization problem same as before.

$e^* = 1 - \pi'$, except that now, π' is the prob of high type, conditional on passing the test.

$p = 0$

Principal's optimization problem same as before.

$e^* = 1 - \pi'$, except that now, π' is the prob of high type, conditional on passing the test.

Agents optimization problem:

- After the investment stage, if still low type, payoff is $(1 - e^*) - 0.1$.
- So, at the investment stage, chooses e_1 to maximise:

$$e_1 + (1 - e_1)((1 - e^*) - 0.1) - 0.5e_1^2$$

- So, $e_1^* = 0.1 + e^*$.

$p = 0$

Plugging back into the Principal's optimization problem,

$$
\begin{aligned}
e^* &= 1 - \pi' \\
&= 1 - \{0.5 + 0.5e_1^*\} \\
&= 1 - \{0.5 + 0.5(0.1 + e^*)\} \\
&\implies e^* = 0.3
\end{aligned}
$$

So, $e_1^* = e^* + 0.1 = 0.4$.

P''s payoff is higher than no test case.

$p = 0.9$

Let e^* and e_1^* be the equilibrium levels of P's and agents' efforts.

Suppose, $0.9(1 - e^*) > (1 - e^*) - 0.1$, so no falsification occurs in equilibrium.

So P is not evaluating everyone, some types are screened out by the test.

$\pi' = \frac{0.5+0.5e_1^*}{0.5+0.5e_1^*+0.5(1-e_1^*)(0.9)}$.

$e^* = 1 - \pi'$.

$p = 0.9$

Agents optimization problem:

- After the investment stage, if still low type, payoff is $0.9(1 - e^*)$.
- So, at the investment stage, chooses e_1 to maximise:

$$e_1 + (1 - e_1)(0.9(1 - e^*)) - 0.5e_1^2$$

- So, $e_1^* = 1 - 0.9(1 - e^*)$.

$p = 0.9$

Recall that $e^* = 1 - \pi'$, with $\pi' = \frac{0.5+0.5e_1^*}{0.5+0.5e_1^*+0.5(1-e_1^*)(0.9)}$.

Solving for e^*, we have $e^* = 0.294916$.

So, $e_1^* = 1 - 0.9(1 - e^*) = 0.3654244$.

$0.9(1 - e^*)$ is indeed strictly greater than $(1 - e^*) - 0.1$, so the Agent does not want to falsify.

P strictly worse off compared to $p = 0$ test.

Intuition

With $p = 0.9$, screening happens in equilibrium; agents do not falsify.

This is even though P's effort with the $p = 0.9$ test is 0.294916, which is strictly lower than with the $p = 0$ test, where it is 0.3.

So, the post falsification payoff $(1 - e^*)$, is **better** in the $p = 0.9$ equilibrium.

Yet agent chooses not to falsify; her *natural* probability of passing is high, from not falsifying, she gets something higher than $(1 - e^*) - c$.

So, folding back to the investment stage, in the $p = 0.9$ test, the agent's payoff from being the low type is strictly higher than her payoff from being the low type with the $p = 0$ test.

So, her incentive to invest into becoming the high type is lower.

P benefits from screening, but loses because the overall pool of candidates is worse.

Intuitively, because candidates are screened out, the pool that passes has a higher posterior than the $p = 0$ case, even though mass of high types is lower.

So equilibrium e^* is lower, and therefore Agent's falsification payoff is $(1 - e^*) - c$, is higher than the $p = 0$ case.

Since falsification is being prevented, after the investment stage, low types get something even higher than $(1 - e^*) - c$.

Conclusion

Investment effort is highest, and therefore the pool of candidates is best for the $p = 0$ test, which is falsifies wp1.

This holds in general, screening always comes at the cost of investment.

When k is low, P always wants the best pool, even if she has to evaluate all.

When k is high, screening effect dominates, she would rather some agents are filtered out at the first stage.