

Hit or Miss? Test Taking Behavior in Multiple Choice Exams

By Pelin Akyol, James Key & Kala Krishna

Discussant: Seda Ertac
Koc University

- Aims to understand skipping behavior in multiple choice test-taking
- Estimation of two structural parameters: “ability” and risk aversion
- Data on scores, coupled with several demographic variables (gender, test experience, GPA)
- Estimates model parameters via final scores.

- Ideally, we would have number of correct answers, number of wrong answers and number of skipped questions.
- For policy: Would design an experiment where a random group of test takers can guess for free, another group is punished a little, another group punished more etc.
- In the absence of this, the paper uses scores to understand skipping behavior, making use of the fact that certain scores can be generated by a limited number of incorrectly answered-skipped question combinations.
- Builds a structural model and estimates the model parameters that generate the observed score patterns.
- This approach allows for counterfactual experiments: What if women had the risk aversion parameter of men? What if there were no penalty for guessing? A higher penalty?

Method

Form bins based on predicted ÖSS scores

Estimate beta and c, allowing them to differ between different groups based on gender, test experience, and predicted scores

Estimation results show that:

- Females are more risk-averse
- Low-scorers are more risk-tolerant

Comments

- Interesting because it is a mechanism design problem—policy implications
- Effect of preferences/non-cognitive skills on achievement and gender differences is important
- Counterfactual analyses reported in the paper provide interesting policy-relevant experiments

Questions

Why not use the math/science tests as well?

Selection into test types may reveal interesting patterns (e.g. gender and risk-aversion)

Self-confidence differences vs. risk aversion?

- What we mean by ability here actually involves self-confidence. Is there any way to incorporate confidence separately into the model? Does it make a difference?

Related to Assumptions...

1. Performance/skipping on one question has no bearing on the next question. Reasons to think otherwise: does how many you've done affect how R-A you are on the next one? Loss aversion-type utility?
 2. Skipping due to time constraints: you may be also allocating time across questions. => May be more prevalent in math/science.
- What do we expect if these assumptions are relaxed? What would be the expected direction of the change in results?

Implications for Policy

- If the policymaker observes the whole set of individual answers, what might be different?
- Teaching students the benefits/costs of skipping (rationality/knowledge of payoff structures may be different across ability levels)
- Interventions to change preferences in childhood