Overconfidence and Inaccurate Self-Assessment: Evidence from Student Grade Estimations

Huan Cai, Cornell College
Jie Zheng, Shandong University

Acknowledgement: We sincerely thank the students who participated in our surveys. We are grateful for the support of school administrators, teachers, and staff who facilitated data collection. Special thanks to Tsinghua SEM for financial support and graduate assistance with survey design and implementation by Yunwen He and Yongchun Peng. Their contributions were invaluable to this research. Any errors or omissions are the responsibility of the authors.

A Simple Survey Before & After the Exam

- Each student in 12th grade was asked to estimate:
  1. For the current mock exam of this subject:
     - Your own score: _____
     - Your ordinal rank out of N students in the class: _____
  2. For the Gaokao exam of this subject:
     - Your own score: _____
     - Your ordinal rank out of N students in the class: _____

Motivation

- Moore & Healy (2008): Overconfidence
  - Overestimation: *Score_Error* = (Estimated Score – Actual Score) / Full Score > 0
  - Overplacement: *Percentile_Error* > 0
    - ✓ Murphy & Weinhardt (2020), Yu (2020): *Percentile* = 1 – (*Rank* – 1) / (*N* – 1)
- Kruger & Dunning (1999): Competency↓ → Overplacement↑
- Research questions:

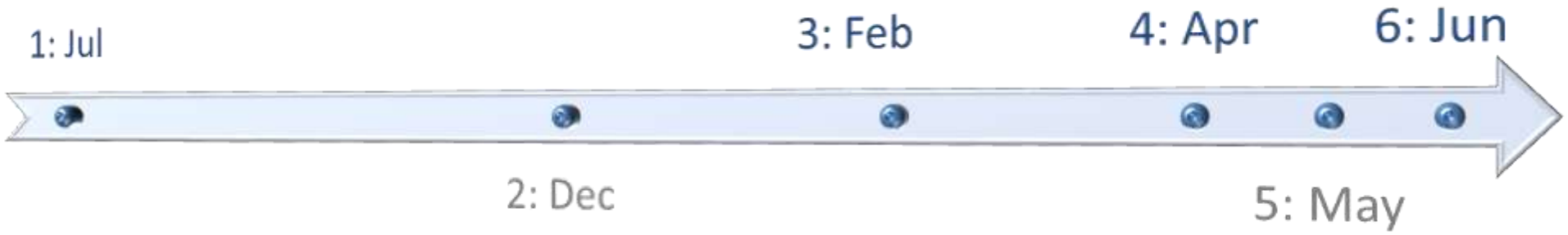
1. What: Overestimation Vs. Overplacement?
2. When: Current Vs. Gaokao exam? Before Vs. After the exam?
   Differences in subjects (e.g. Chinese Vs. Math)? Male Vs. Female?
3. How: Can competency explain overestimation? Or just overplacement?
4. Why? ……

A Series of Surveys in 2 Years

- A senior high school in GD, China, 2022-2023, Gaokao “3+1+2”:
  - 6 subject-categories for survey: *CHI*, *MAT*, *ENG*, *PHY/HIS*, *Sub4*, *Sub6*

No. of Students: School X	2022	% Female	2023	% Female
Science Track (PHY)	711	51%	756	47%
Humanities Track (HIS)	222	74%	214	72%
Total of Both Tracks	933	56%	970	52%

  - Timeline of the surveys: 6 exams in 2022, 4 exams in 2023
  - Each academic year:



Survey Data

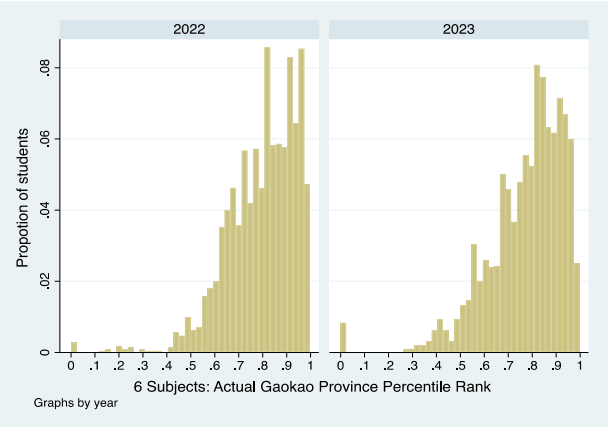
- No. of surveys: $2Surveys(before\ \&\ after) \times 10Exams(for\ 2\ years)$
- Each survey: $2Qs(score\ \&\ rank) \times 6Subjects \times 2Secs(current\ \&\ Gaokao)$

➔2 Final Samples:

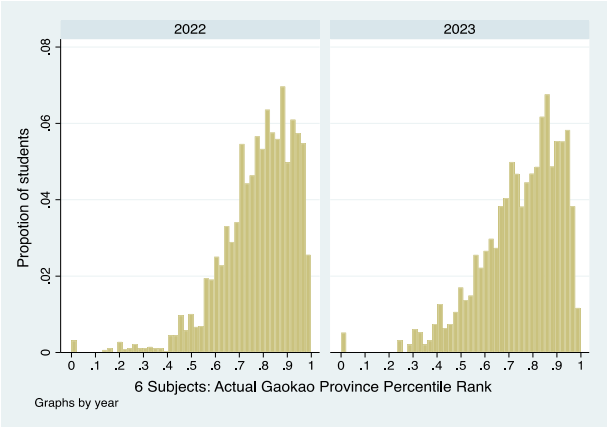
No. of Observations (student-survey-subject)	
Non-Gaokao Sample	18,296
Gaokao Sample	19,256

- A representative survey sample:

Non-Gaokao Sample



School X Full Sample



Gaokao Sample

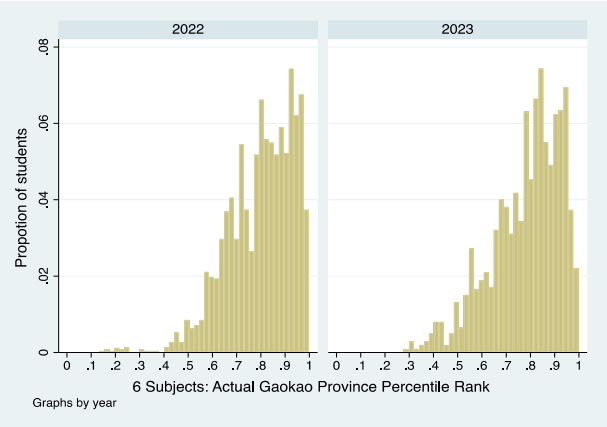
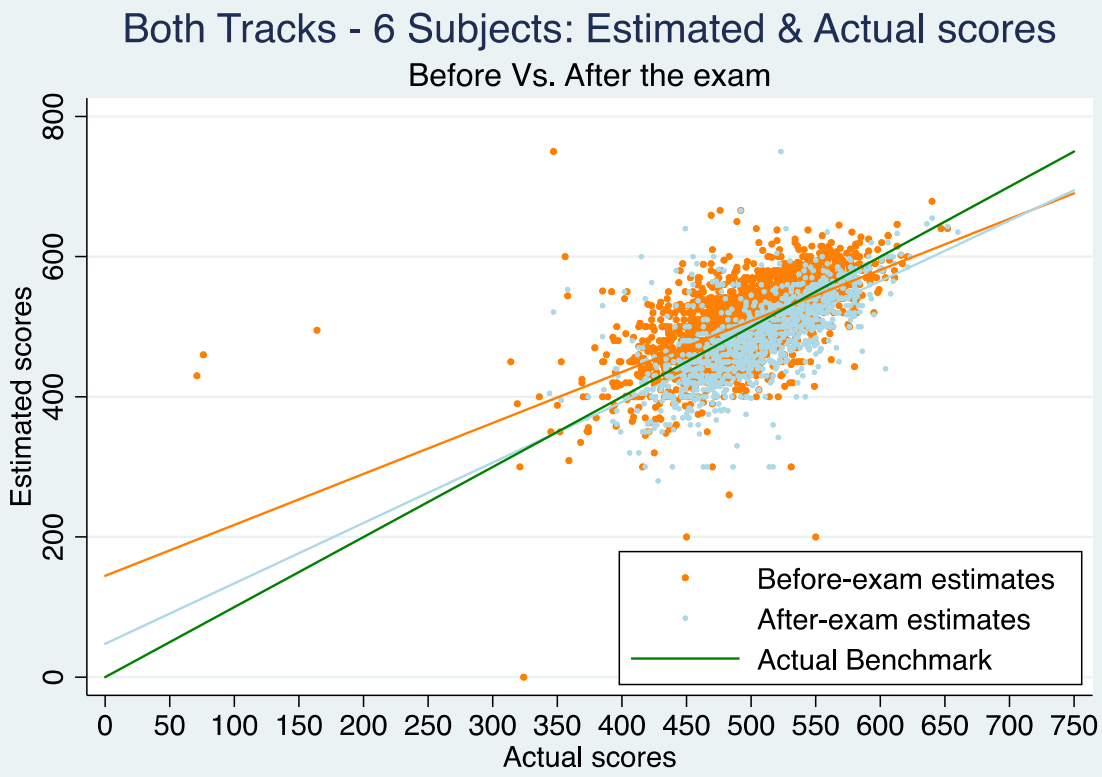


Figure 1-1: Estimated Vs. Actual Scores of *Sub6*

Non-Gaokao Sample

Overestimation!

Gaokao Sample



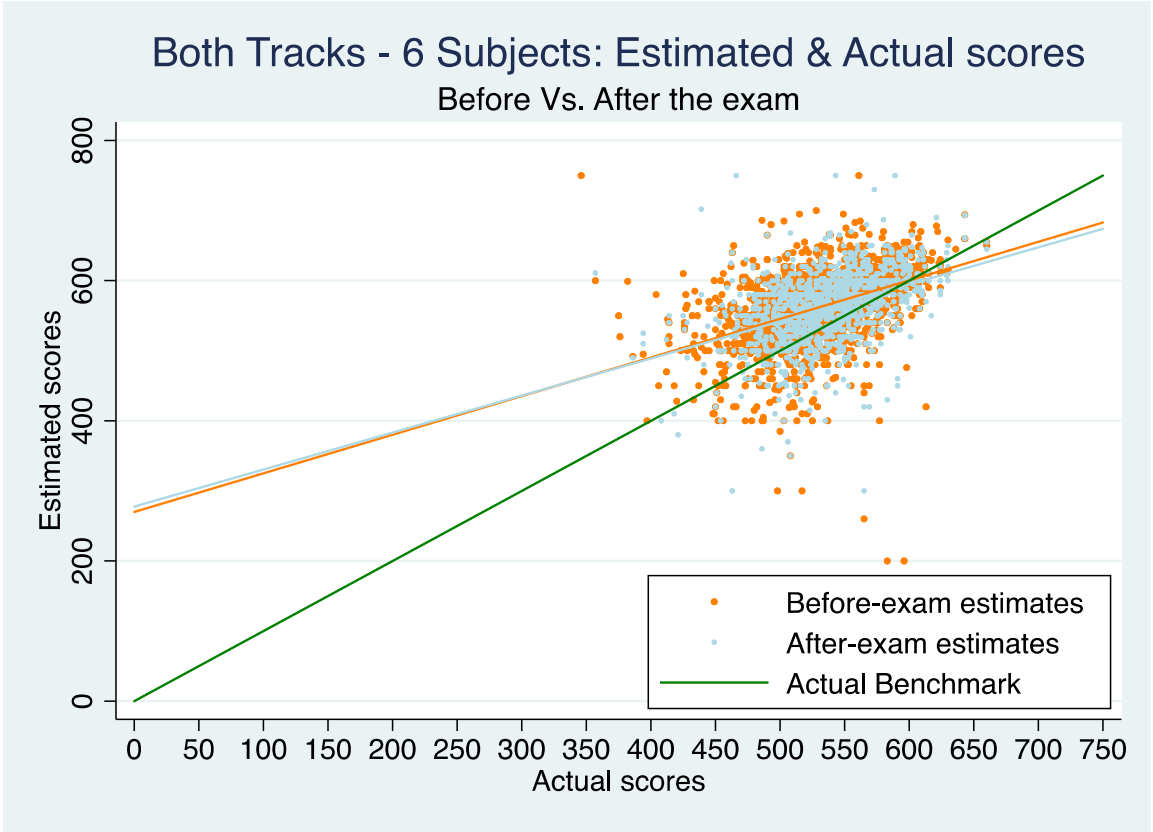
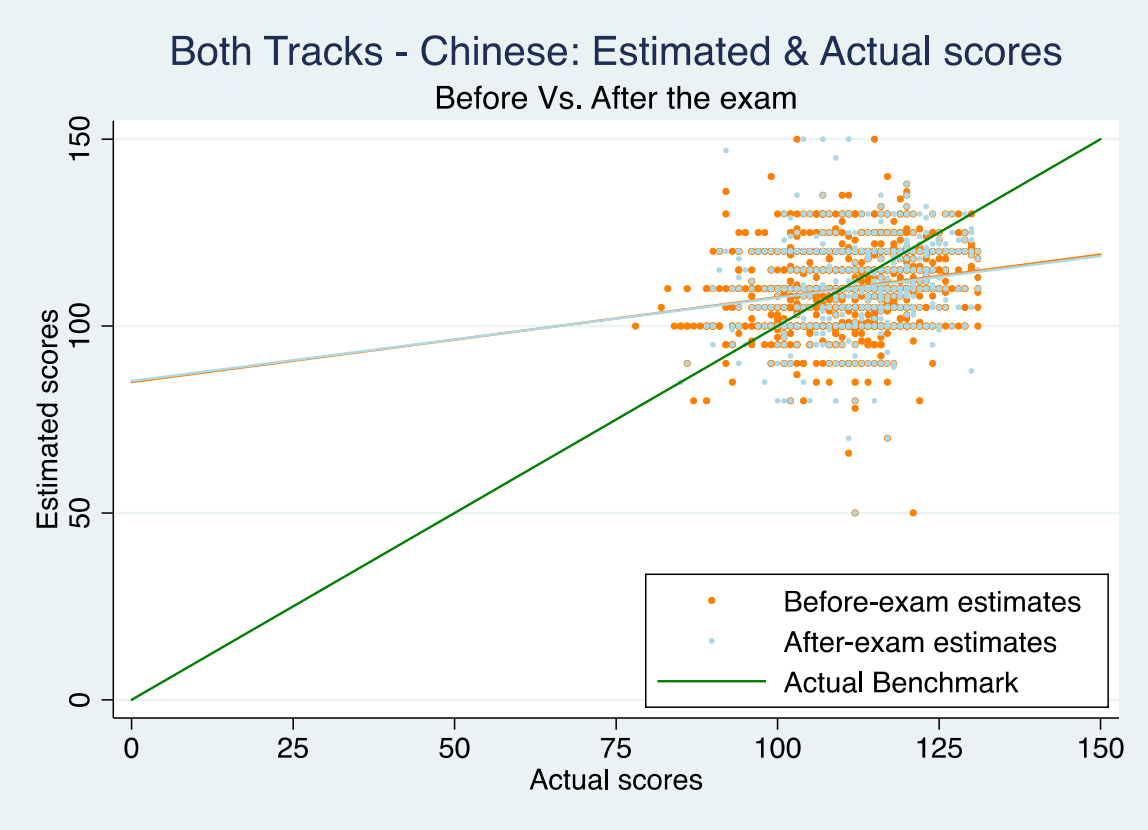


Figure 2-1: Scores, *CHI* Vs. *MAT* (Gaokao Sample)

Chinese (*CHI*)

Overestimation!

Mathematics (*MAT*)



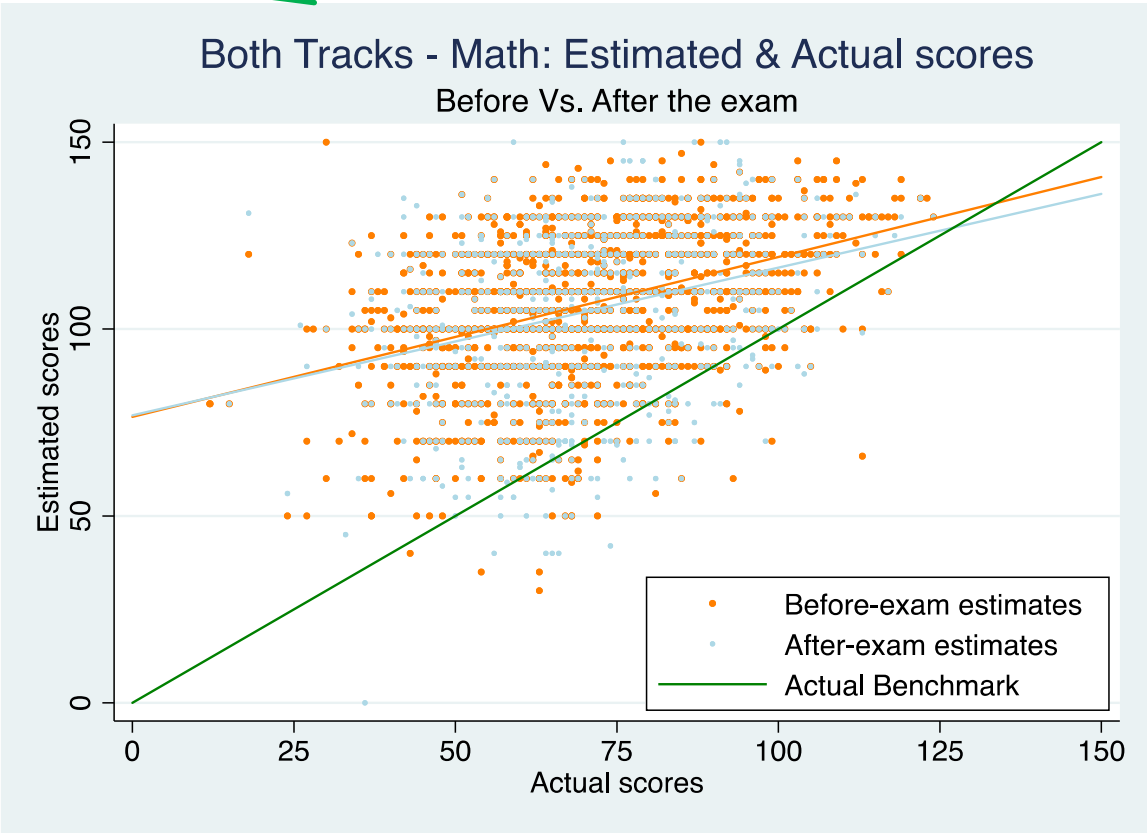
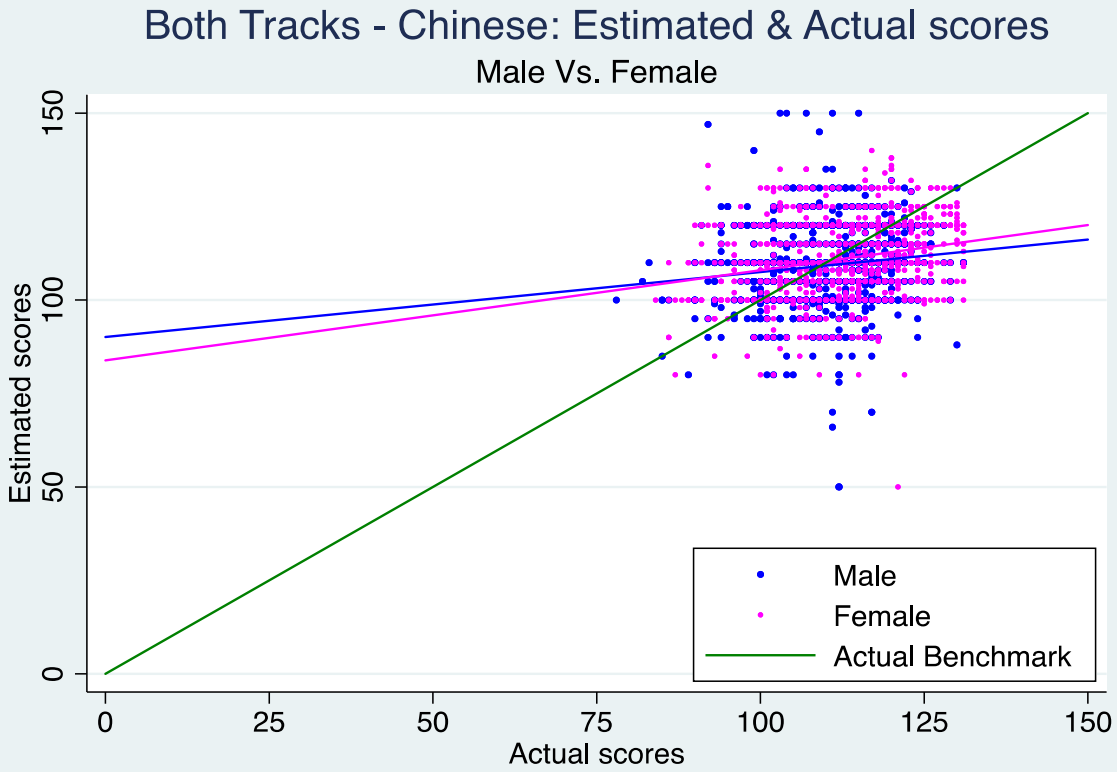


Figure 2-3: Male Vs. Female, Scores, *CHI* Vs. *MAT* (Gaokao)

Chinese (*CHI*)

Overestimation!

Mathematics (*MAT*)



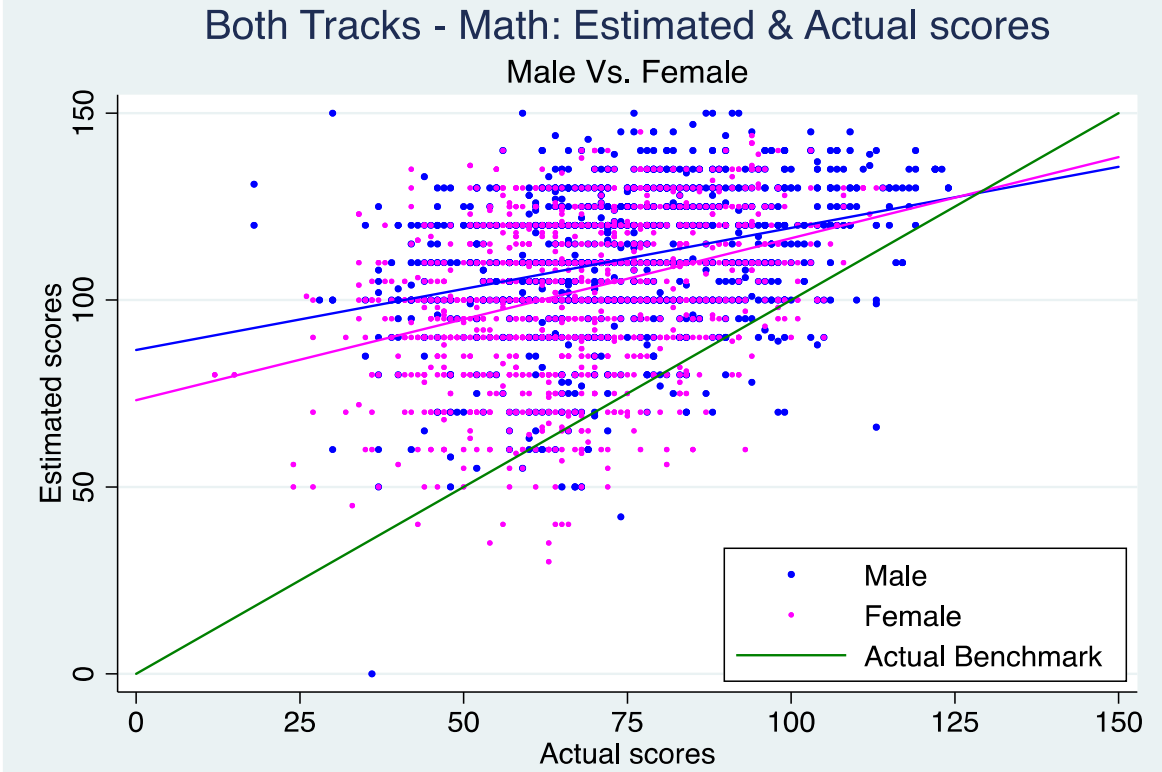
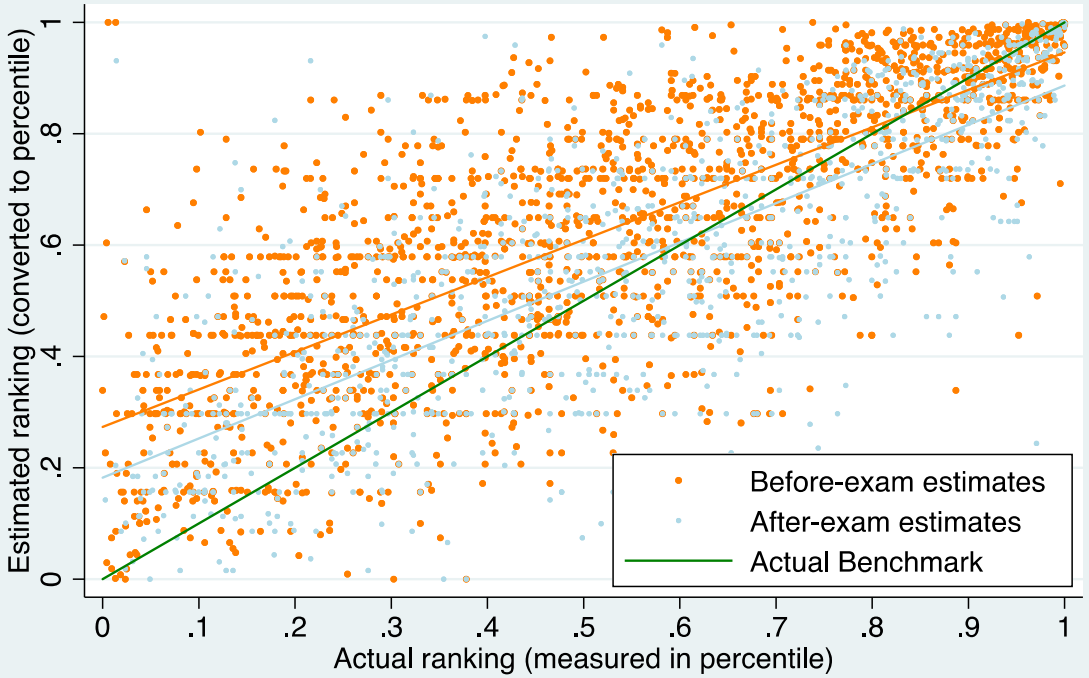


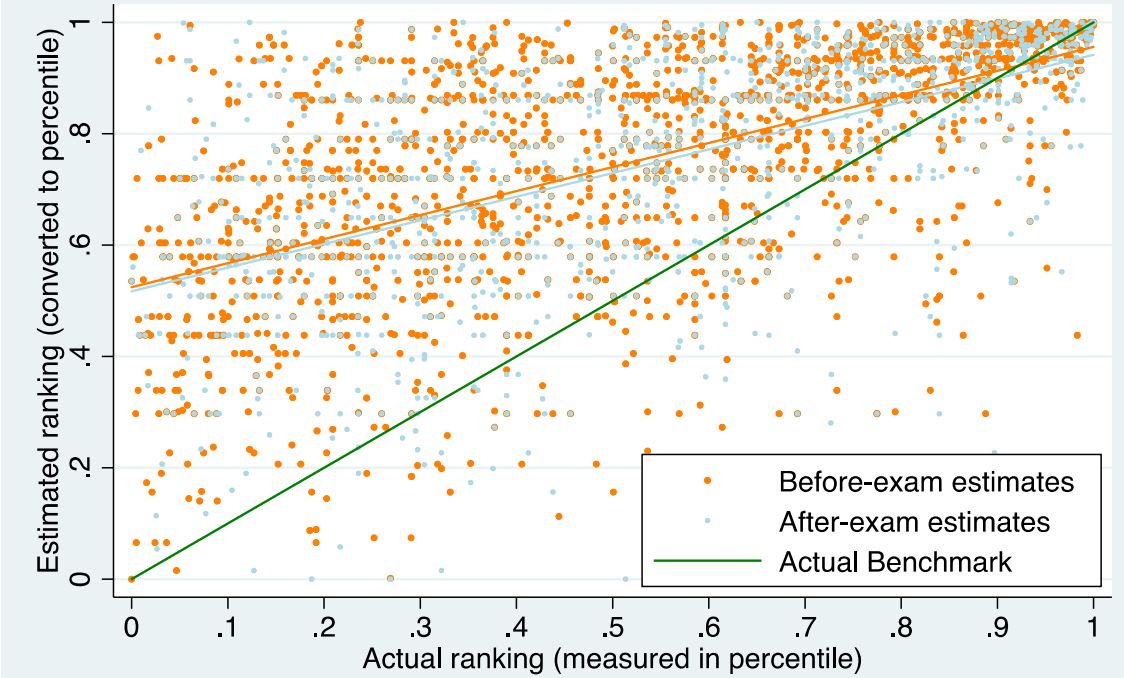
Figure 1-2: Estimated Vs. Actual Ranks (in %) of *Sub6*

Non-Gaokao Sample

Overplacement!

Gaokao Sample





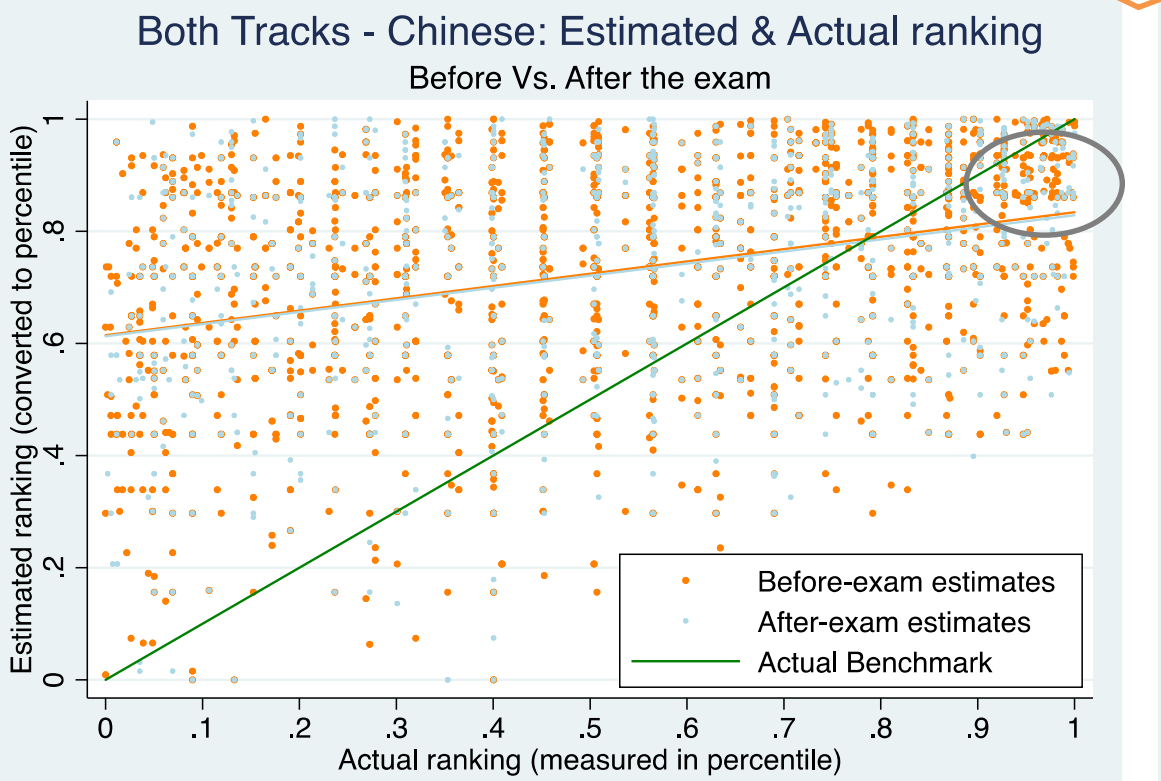
Murphy & Weinhardt (2020), Yu (2020): $Percentile = 1 - (Rank - 1) / (N - 1)$

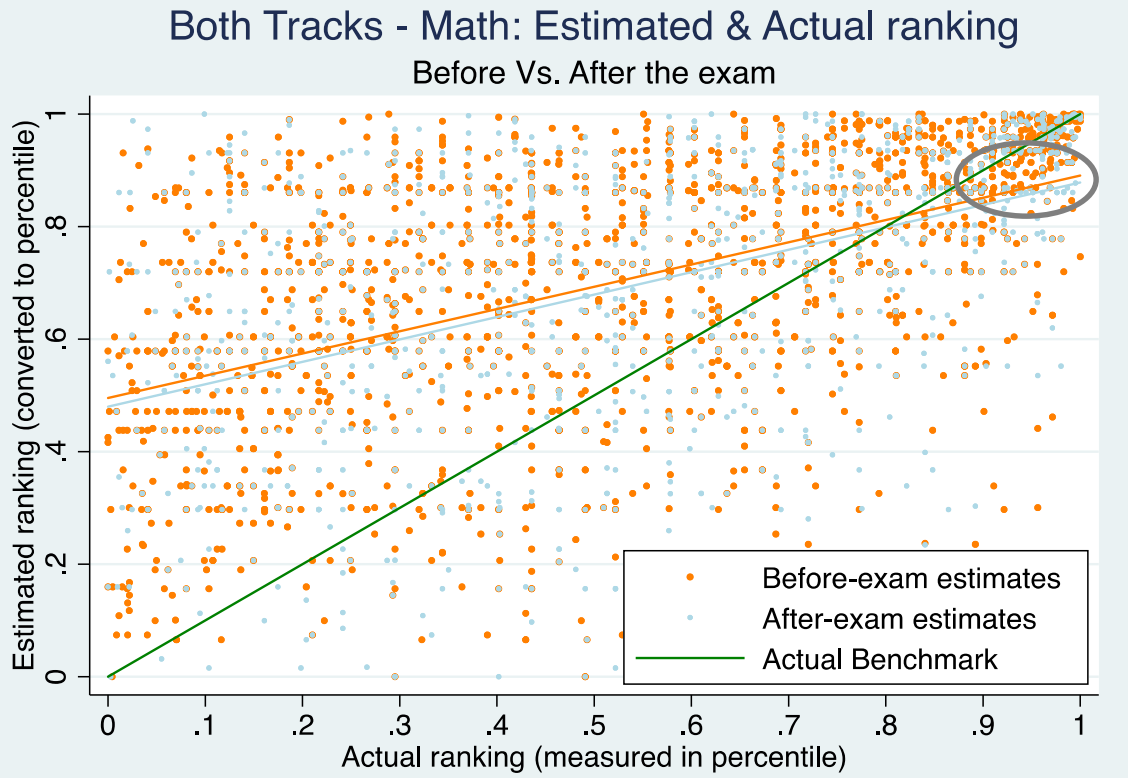
Figure 2-2: Ranks (in %), *CHI* Vs. *MAT* (Gaokao)

Chinese (*CHI*)

Overplacement!

Mathematics (*MAT*)

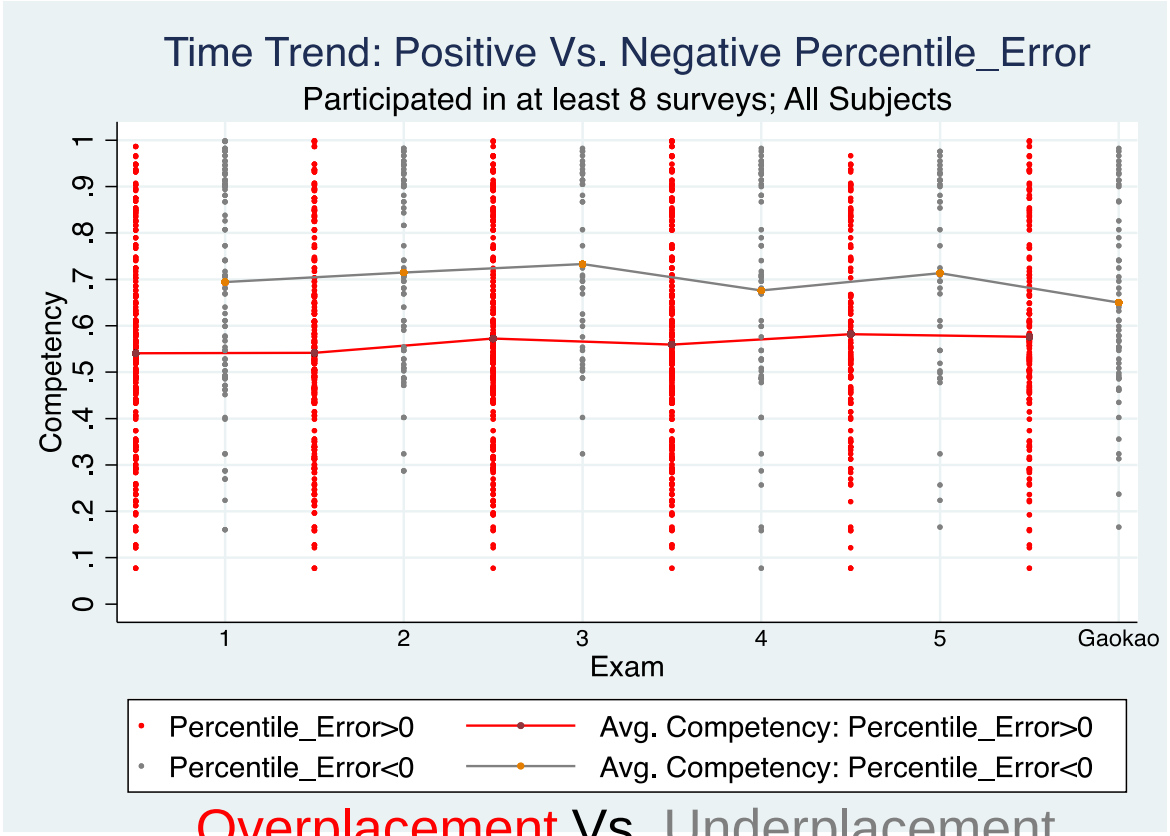




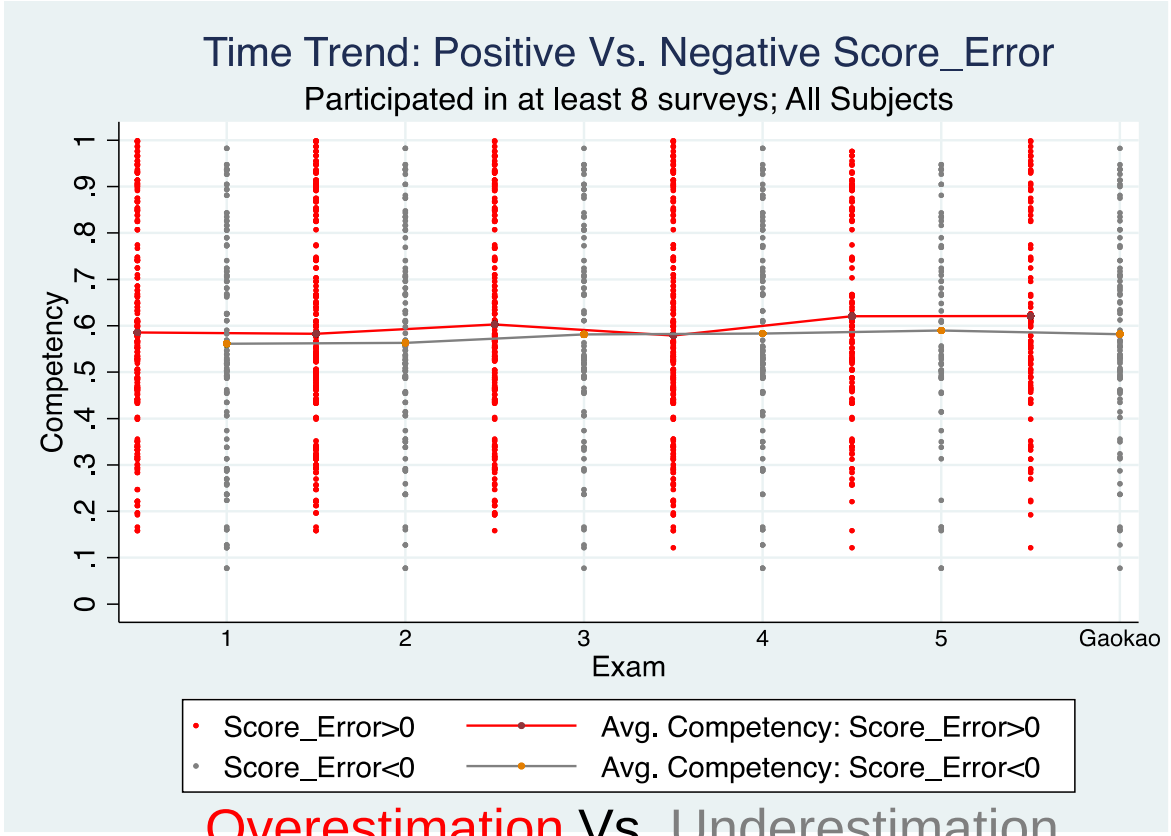
Kruger & Dunning (1999) Effect?

Figure 3-1: Estimate Errors & Competency (Gaokao)

Percentile_Error **Kruger-Dunning** *Score_Error*



Overplacement Vs. Underplacement



Overestimation Vs. Underestimation

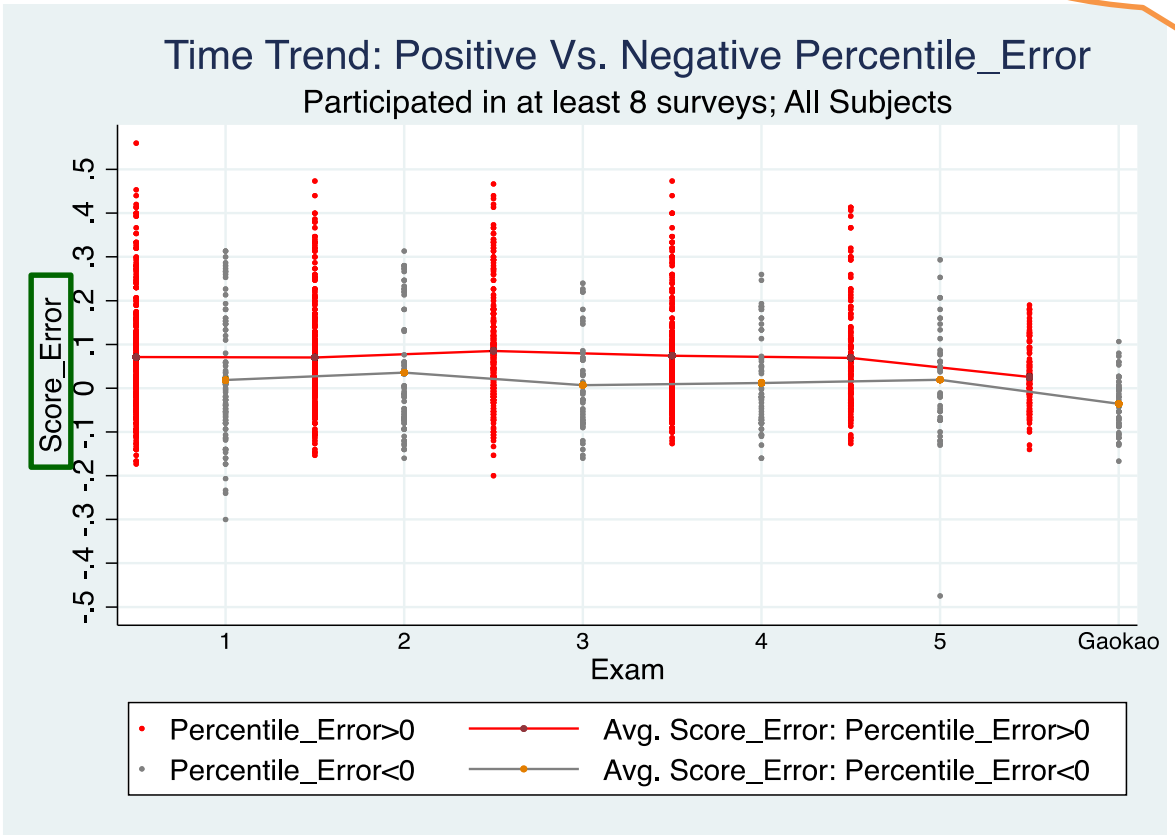
$$Competency_{i,s} = \frac{1}{n_{i,s}} \sum_{j=1}^{j=n_{i,s}} Percentile_{i,j,s}$$

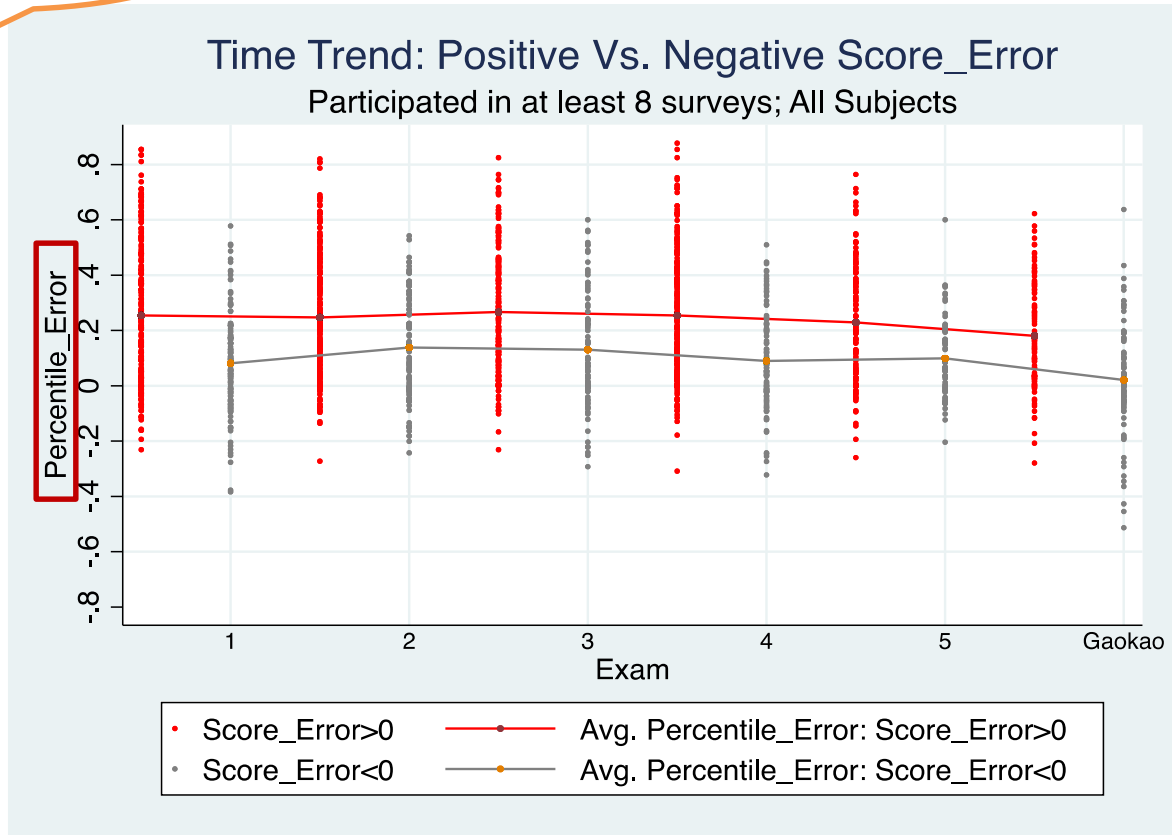
Figure 3-2: *Percentile_Error* Vs. *Score_Error* (Gaokao)

Percentile_Error

+ correlated

Score_Error





Overplacement Vs. Underplacement

Overestimation Vs. Underestimation

Summary: Thank you!

- Evidence of Overconfidence:
  - ✓ Overplacement > Overestimation; stronger for Gaokao sample
  - ✓ Overplacement for all subjects, but not always overestimation (e.g. *CHI* Vs. *MAT*)
- Evidence of Kruger-Dunning Effects:
  - ➢ Overplacement significantly negatively associated with competency
  - ➢ Overestimation is positively correlated with overplacement
- Differences in Subjects & in Gender
  - ➢ The puzzle of Mathematics?
- Future suggestions are very welcome!

Huan Cai: hcai@cornellcollege.edu
Jie Zheng: zhengjie@sdu.edu.cn

Summary Statistics

Non-Gaokao Sample

Gaokao Sample

Track	Science (PHY)				Humanities (HIS)				Science (PHY)				Humanities (HIS)			
Student Self-Reported Data	mean	sd	p50	count	mean	sd	p50	count	mean	sd	p50	count	mean	sd	p50	count
Score_Error																
CHI (150)	0.2%	6.4%	0.0%	2,251	-0.6%	5.8%	-0.7%	801	-1.1%	8.0%	-1.3%	2,371	-0.8%	7.1%	-1.3%	842
ENG (150)	1.6%	8.0%	1.3%	2,251	0.1%	7.6%	0.0%	801	-3.0%	9.6%	-3.3%	2,371	-2.9%	9.0%	-3.3%	842
MAT (150)	6.2%	10.5%	4.7%	2,252	4.0%	10.1%	2.7%	801	24.2%	13.8%	26.0%	2,372	21.8%	12.9%	22.7%	842
PHY/HIS (100)	7.7%	12.1%	7.0%	2,253	3.3%	9.6%	3.0%	801	5.1%	12.4%	5.0%	2,373	10.3%	9.4%	11.0%	842
Sub4 (550)	3.6%	7.4%	3.1%	2,233	1.7%	6.3%	1.1%	798	6.2%	8.7%	6.2%	2,349	6.8%	6.8%	6.7%	838
Sub6 (750)	0.1%	7.5%	0.0%	2,253	-1.3%	6.2%	-1.1%	801	3.8%	8.0%	4.0%	2,372	4.1%	5.8%	4.1%	842
Percentile_Error																
CHI	4.8%	28.6%	4.2%	2,251	9.3%	26.8%	10.1%	801	19.5%	29.1%	16.6%	2,371	22.2%	28.6%	21.1%	842
ENG	2.3%	20.6%	1.5%	2,251	4.4%	21.5%	4.0%	801	18.2%	23.8%	16.8%	2,371	18.2%	23.1%	17.6%	842
MAT	5.0%	21.4%	3.8%	2,252	-2.9%	21.9%	-2.4%	801	19.0%	25.9%	16.1%	2,372	9.8%	25.5%	7.2%	842
PHY/HIS	6.7%	22.5%	5.2%	2,253	8.9%	26.0%	6.4%	801	21.6%	25.9%	20.3%	2,373	21.2%	26.8%	18.9%	842
Sub4	6.3%	19.0%	5.3%	2,229	5.9%	18.0%	4.2%	794	20.6%	23.5%	18.4%	2,346	18.0%	22.1%	15.8%	836
Sub6	6.9%	18.8%	5.9%	2,253	6.8%	17.9%	5.3%	801	21.9%	23.8%	20.2%	2,372	19.6%	22.8%	16.9%	842

$$Score\_Error_{i,j,k,s} = (Est\_Score_{i,j,k,s} - Score_{i,j,s}) / Full\_Score_{s}$$

$$Percentile\_Error_{i,j,k,s} = Est\_Percentile_{i,j,k,s} - Percentile_{i,j,s}$$

The Kruger-Dunning Effects: Overplacement

	Non-Gaokao Sample				Gaokao Sample			
Dep. Variable: Percentile_Error	**No Fixed Effects**		**All 3 Fixed Effects**		**No Fixed Effects**		**All 3 Fixed Effects**	
Competency	-0.227***	-0.213***	-0.354***	-0.349***	-0.472***	-0.473***	-0.599***	-0.586***
	(0.012)	(0.016)	(0.014)	(0.017)	(0.017)	(0.023)	(0.019)	(0.022)
Competency * Before		-0.015		-0.008		0.002		-0.021*
		(0.017)		(0.016)		(0.019)		(0.012)
Before		0.079***		0.075***		0.008		0.018**
		(0.011)		(0.010)		(0.013)		(0.008)
Constant	0.173***	0.121***	0.240***	0.193***	0.446***	0.441***	0.513***	0.502***
	(0.007)	(0.010)	(0.007)	(0.010)	(0.012)	(0.016)	(0.010)	(0.012)
Observations	18,288	18,288	18,288	18,288	19,251	19,251	19,251	19,251
R-squared	0.060	0.083	0.400	0.417	0.197	0.198	0.633	0.633

$$Competency_{i,s} = \frac{1}{n_{i,s}} \sum_{j=1}^{j=n_{i,s}} Percentile_{i,j,s}$$

The Kruger-Dunning Effects: Overestimation?

	Non-Gaokao Sample				Gaokao Sample			
Dep. Variable: Score_Error	**No Fixed Effects**		**All 3 Fixed Effects**		**No Fixed Effects**		**All 3 Fixed Effects**	
Competency	-0.036***	-0.028***	-0.025***	-0.014**	-0.087***	-0.083***	-0.067***	-0.068***
	(0.005)	(0.007)	(0.005)	(0.006)	(0.008)	(0.011)	(0.008)	(0.009)
Competency * Before		-0.008		-0.017**		-0.005		0.002
		(0.008)		(0.007)		(0.009)		(0.005)
Before		0.050***		0.050***		0.011**		0.003
		(0.005)		(0.005)		(0.005)		(0.003)
Constant	0.046***	0.013***	0.040***	0.008**	0.106***	0.099***	0.096***	0.094***
	(0.003)	(0.004)	(0.002)	(0.004)	(0.005)	(0.007)	(0.004)	(0.005)
Observations	18,296	18,296	18,296	18,296	19,256	19,256	19,256	19,256
R-squared	0.009	0.068	0.422	0.458	0.025	0.026	0.691	0.691

$$Competency_{i,s} = \frac{1}{n_{i,s}} \sum_{j=1}^{j=n_{i,s}} Percentile_{i,j,s}$$