

A Few Bad Apples?

Academic Dishonesty, Political Selection, and Institutional Performance in China

John Zhuang Liu

University of Hong Kong

Wenwei Peng

Hong Kong Baptist University

Shaoda Wang

University of Chicago

CESI

August, 2026

Motivation

- ▶ Institutions ultimately act through people (Knack and Keefer, 1997; Bertrand and Schoar, 2003; Caselli and Morelli, 2004; Besley, 2005; Jones and Olken, 2005)
- ▶ We rely on individuals' honesty in key roles, especially when discretion is involved
 - ▶ Public sector: judges be impartial; officials to serve the public interest
 - ▶ Private sector: executives/auditors be honest in financial reporting
- ▶ Recognizing this, societies invest heavily in screening for and cultivating honesty:
 - ▶ Strict honor codes of academic integrity in education

Academic Dishonesty is Taken Seriously



(a) Biden withdrew from the presidential race



(b) Hungarian president resigns

- Academic dishonesty is treated with serious penalties
 - Students expelled; academic degrees revoked
 - National leaders face career setbacks

How Prevalent is Academic Dishonesty? Its Societal Costs?

Despite the perceived importance, limited rigorous evidence on:

1. The prevalence of academic dishonesty (esp. among public officials)
 - ▶ Measurement problem: successful dishonesty is hidden
2. The societal/institutional consequences of placing dishonest individuals in power
 - ▶ Sorting issue: people with different honesty levels sort into different careers/tasks

This Paper

Focusing on the world's largest higher-education & public service system: China

- ▶ Develop a high-stakes, scalable measure of hidden academic dishonesty:
 - plagiarism in graduate dissertations (half-million graduates)
 - validate its strong corr. with dishonesty using a survey: “academic dishonest behavior”
- ▶ Linking dissertation plagiarism to public officials' selection and performance

We study three questions:

1. **Political Selection:** Who enters public service?
 - ▶ Focusing on all positions: tax/customs officials, judges, and others
2. **Institutional Performance:** What happens once dishonest people hold power?
 - ▶ Focusing on the judiciary: performance (140M cases) + quasi-random assignment
3. **Policy:** Can promoting academic integrity in university help ?
 - ▶ Staggered-rollout plagiarism-ceiling: plagiarism/political selection/professional conduct

Roadmap

Motivation

Background and Data

Academic Dishonesty and Political Selection

Impacts of Academic Dishonesty in the Public Service

Effect of Anti-Plagiarism (ceiling) Policy

Conclusion

Higher Education and Academic Integrity in China

- ▶ World's largest higher ed system: 53M Bachelors & 11M Postgrads since 1978
- ▶ Dissertation is mandatory for all graduate degrees
- ▶ Centralized repository: China National Knowledge Infrastructure (CNKI)
 - ▶ State-owned enterprise, central digital archive for all academic work
 - ▶ MOE requires universities to publish graduate theses online since 2010
 - ▶ CNKI: 6 million postgrads. dissertations covering 450 universities (our data source)
- ▶ CNKI also develops a powerful plagiarism detection algorithm:
 - ▶ Widely adopted by universities for degree conferral checks since 2016
 - ▶ Low awareness of plagiarism detection before 2016 (Baidu Search Index, MOE 2016)
 - ▶ Unaccessible to individuals before 2022, only selected university staff have access
- ▶ Unique setting: centralized data + powerful tool → large-scale academic misconduct analysis

China's Public Service System

- ▶ World's largest bureaucracy: >8M civil servants, 70M broader state employees
- ▶ Entry via Civil Service Examinations (CSE):
 - ▶ A merit-based system since 1990s
 - ▶ Highly competitive: national gov $\sim 1.5\%$ admission rate; local gov $\sim 3.2\%$ Harvard
 - ▶ Public disclosure of admitted candidates (our data source)
- ▶ Public officials wield significant discretion: limited checks and transparency
- ▶ Civil servants are the pipeline to political power:
 - ▶ No difference between bureaucrats and politicians in China
 - ▶ Promotion leads to political leadership roles

Wage & Corruption

Data

We link half-million graduates' dissertations (plagiarism) to public sector entry, promotion, and performance:

Data

We link half-million graduates' dissertations (plagiarism) to public sector entry, promotion, and performance:

- ▶ Link academic dishonesty to political selection [Linking Process](#)
 1. 6 million postgraduate dissertations covering 450 universities from CNKI
 2. Comprehensive rosters of public officials from 2014-2022 [Details](#)
 - ▶ 350K admitted civil servants + 190K active judges
 - 60K civil servants + 61K judges; 400K random classmates in private sector
 - 0.52M dissertation plagiarism link to public career entry/advancement

Data

We link half-million graduates' dissertations (plagiarism) to public sector entry, promotion, and performance:

- ▶ Link academic dishonesty to political selection [Linking Process](#)
 1. 6 million postgraduate dissertations covering 450 universities from CNKI
 2. Comprehensive rosters of public officials from 2014-2022 [Details](#)
 - ▶ 350K admitted civil servants + 190K active judges
 - 60K civil servants + 61K judges; 400K random classmates in private sector
 - 0.52M dissertation plagiarism link to public career entry/advancement
- ▶ Link academic dishonesty to institutional performance (judiciary) [Linking Process](#)
 3. Universe of 140m court verdicts from 2014 to 2022 [Details](#)
 - ▶ plaintiff/defendant, rulings/win rate, judicial reasoning, dispute, legal codes, appeal
 4. Universe of business registration records from 1978 to 2023: Tianyancha [Details](#)
 - ▶ Ownership (SOE/private), registered capital, government contractor, etc.
 - Case-level: judge's academic honesty + rulings outcomes

CNKI's Plagiarism-Detection Algorithm

The most widely used plagiarism-detection algorithm in China:

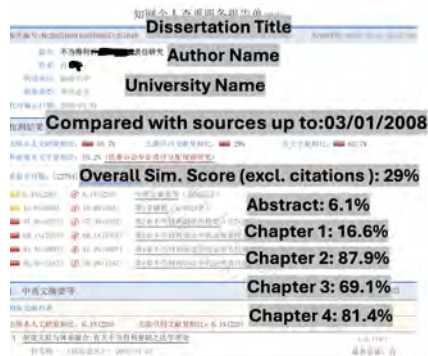
- ▶ Dominant tool
 - ▶ Adopted by most universities for degree conferral checks
 - ▶ Netizens use to scrutinize prominent figures (e.g., Zhang Wenhong, Zhai Tianlin)
- ▶ Distinctive Features
 1. Comprehensive corpus
 - ▶ Journals, books, conference papers, theses, newspapers, patents, standards, encyclopedias, web content, etc
 2. Smart algorithm [Details](#)
 - ▶ Detects paraphrasing, mixed-source weaving, synonym substitution, translation
 - ▶ Filters out legitimate citations (e.g., legal quotations or self-citations)
 3. Temporal control
 - ▶ Compare sources published *before* a specific date: plagiarism at the time of writing

CNKI's Plagiarism-Detection Algorithm

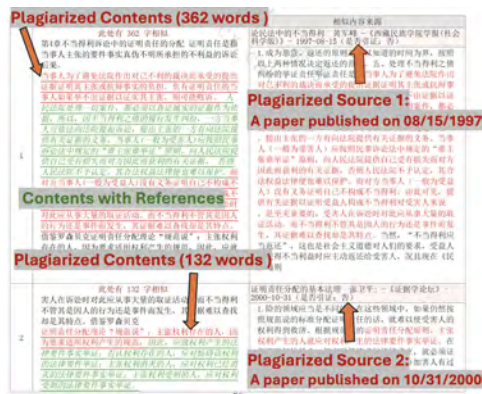
The most widely used plagiarism-detection algorithm in China:

- ▶ Dominant tool
 - ▶ Adopted by most universities for degree conferral checks
 - ▶ Netizens use to scrutinize prominent figures (e.g., Zhang Wenhong, Zhai Tianlin)
- ▶ Distinctive Features
 1. Comprehensive corpus
 - ▶ Journals, books, conference papers, theses, newspapers, patents, standards, encyclopedias, web content, etc
 2. Smart algorithm [Details](#)
 - ▶ Detects paraphrasing, mixed-source weaving, synonym substitution, translation
 - ▶ Filters out legitimate citations (e.g., legal quotations or self-citations)
 3. Temporal control
 - ▶ Compare sources published *before* a specific date: plagiarism at the time of writing
 4. The most expensive one (US\$20 per thesis)

An Example of Plagiarism Record



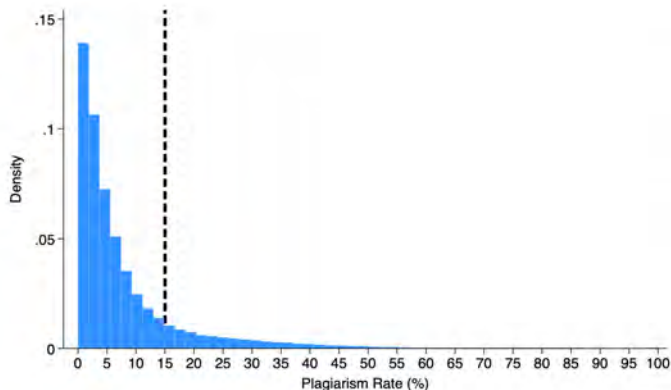
(a) Overall Plagiarism Indicators



(b) Paragraph-level Plagiarism Breakdown

- ▶ Overall similarity score: 0-100 → % content is plagiarized
- ▶ Paragraph-level: background/literature review vs. core analytical/main results sections

Distribution of Plagiarism Scores of Half A Million Dissertations



- ▶ Most students have low plagiarism scores, but also a noticeable long tail
- ▶ **Prevalence:** 14% of students exceed the 15% similarity rate ($1/7$) (a common degree-conferral threshold)

What Does Plagiarism Measure?

What Does Plagiarism Measure?

A survey of 387 legal professionals applying for a job at an elite research institute

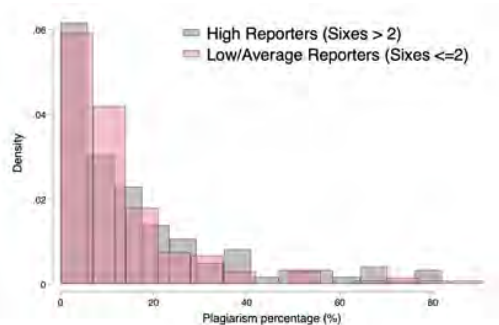
► Voluntary & confidential survey, we collect:

1. Dissertations → plagiarism
2. Personality traits: Big Five, honesty beliefs, and others
3. Dishonesty: incentivized dice game, payoff based on self-reported sixes over 10 rolls

What Does Plagiarism Measure?

A survey of 387 legal professionals applying for a job at an elite research institute

- ▶ Voluntary & confidential survey, we collect:
 1. Dissertations → plagiarism
 2. Personality traits: Big Five, honesty beliefs, and others
 3. Dishonesty: incentivized dice game, payoff based on self-reported sixes over 10 rolls
- ▶ Over-reporting in the dice game show higher plagiarism, especially $> 40\%$



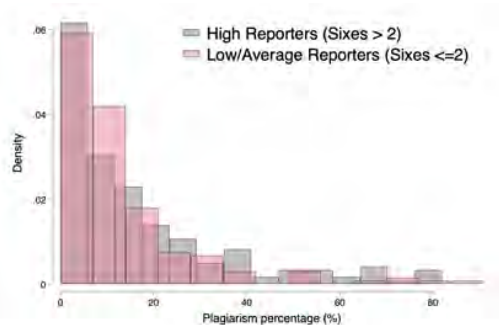
What Does Plagiarism Measure?

A survey of 387 legal professionals applying for a job at an elite research institute

► Voluntary & confidential survey, we collect:

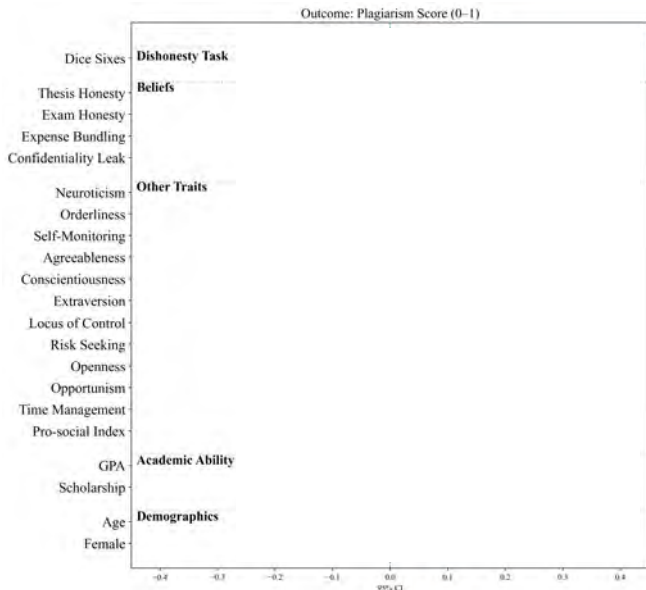
1. Dissertations → plagiarism
2. Personality traits: Big Five, honesty beliefs, and others
3. Dishonesty: incentivized dice game, payoff based on self-reported sixes over 10 rolls

► Over-reporting in the dice game show higher plagiarism, especially $> 40\%$

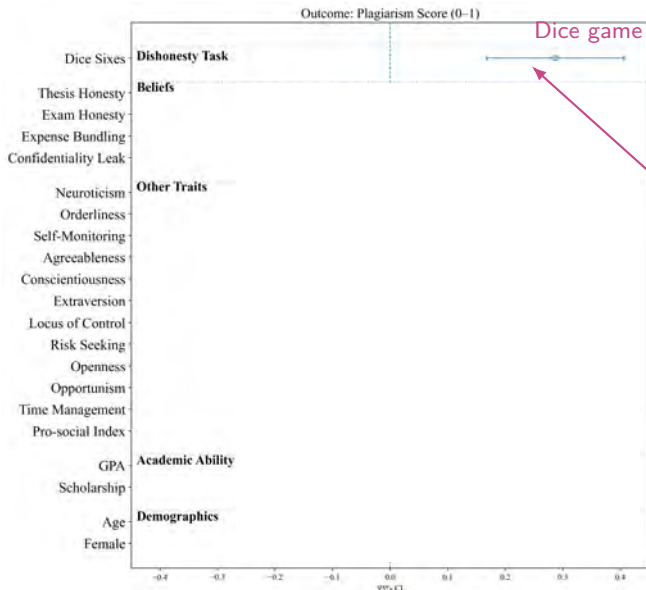


- Many respondents are 5-10 yrs after graduation
- Echoing psychology literature, dishonesty is a stable trait after adulthood

What Does Plagiarism Measure?



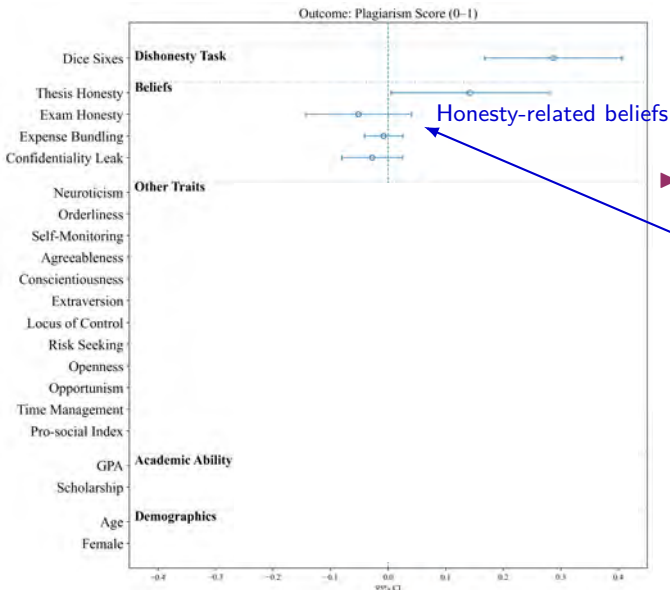
What Does Plagiarism Measure?



4 Findings:

1. ↑ corr. with over-reporting (dice)

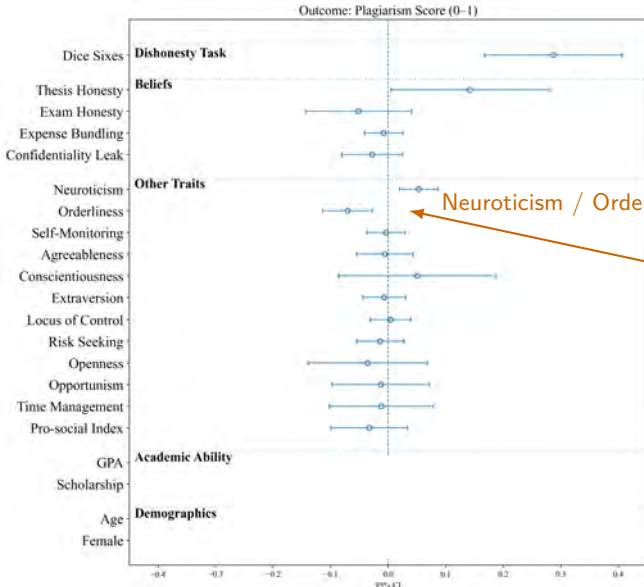
What Does Plagiarism Measure?



► 4 Findings:

1. ↑ corr. with over-reporting (dice)
2. No corr. with stated honesty beliefs

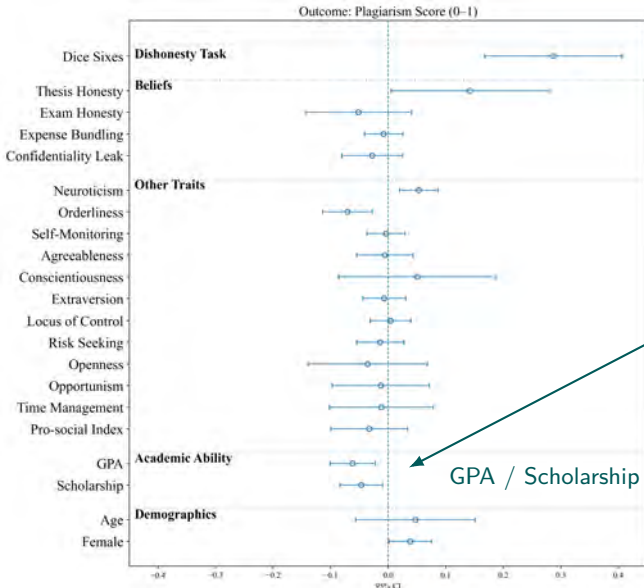
What Does Plagiarism Measure?



► 4 Findings:

1. ↑ corr. with over-reporting (dice)
2. No corr. with stated honesty beliefs
3. Neuroticism/orderliness

What Does Plagiarism Measure?

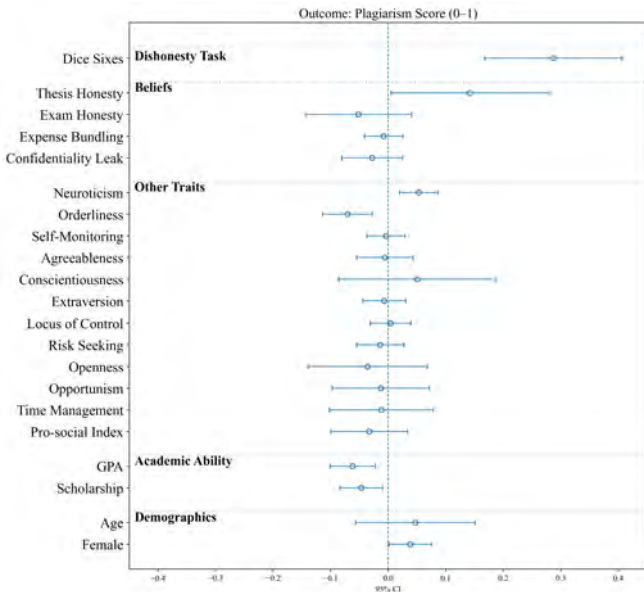


► 4 Findings:

1. ↑ corr. with over-reporting (dice)
2. No corr. with stated honesty beliefs
3. Neuroticism/orderliness
4. ↓ academic ability

GPA / Scholarship / Demographics

What Does Plagiarism Measure?



► 4 Findings:

1. ↑ corr. with over-reporting (dice)
2. No corr. with stated honesty beliefs
3. Neuroticism/orderliness
4. ↓ academic ability

► We refer to it as: academic dishonesty

► Psychology evidence

► Regression table

► Survey design

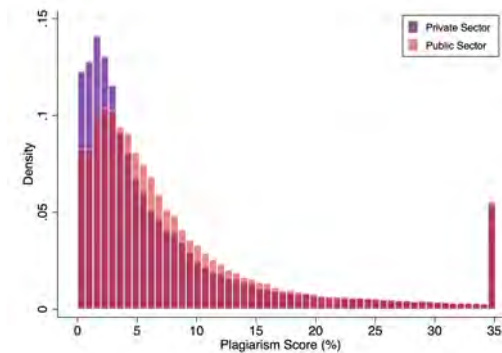
► Questionnaire

► Attrition balance

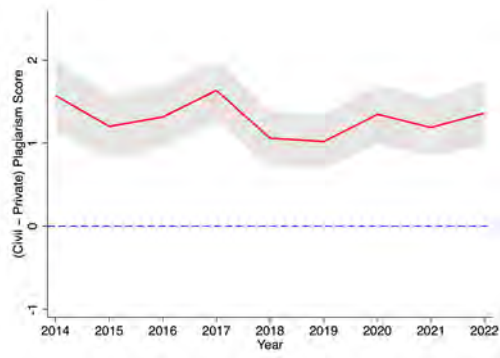
1. Academic Dishonesty and Political Selection

- a. Extensive Margin: Who is more likely to enter the public service? honest vs. dishonest
- b. Intensive Margin: Conditional on entering, do dishonest individuals advance faster?

Descriptive Patterns: Public Officials vs. Private-sector Classmates



(a) Overall Plagiarism Distribution



(b) Plagiarism Gap Over Time

- ▶ Public sector officials have higher plagiarism rates than their private-sector classmates
- ▶ The gap has persisted over the past decade

Are Public Officials More Likely to Plagiarize?

	(1)
	Plagiarism Score
Public Sector	1.201*** (0.060)
% to mean	15.60%
Degree FE	Y
Major FE	Y
Cohort FE	Y
University FE	Y
Individual chara.	Y
Uni-cohort-major	N
Observations	511,021
R-squared	0.234

- Public sector officials: 1.2 pct point/ 15% ↑ plagiarism (vs. private-sector classmates)

Are Public Officials More Likely to Plagiarize?

	(1)	(2)	(3)
	Plagiarism Score	Non-Essential Sections	Essential Sections
Public Sector	1.201*** (0.060)	1.141*** (0.112)	1.261*** (0.125)
% to mean	15.60%	14.82%	16.38%
Degree FE	Y	Y	Y
Major FE	Y	Y	Y
Cohort FE	Y	Y	Y
University FE	Y	Y	Y
Individual chara.	Y	Y	Y
Uni-cohort-major	N	N	N
Observations	511,021	511,021	511,021
R-squared	0.234	0.207	0.245

- ▶ Public sector officials: 1.2 pct point/ 15% ↑ plagiarism (vs. private-sector classmates)
- ▶ Plagiarism, as an academically dishonest behavior, is a predictor for public sector entry

Are Public Officials More Likely to Plagiarize?

	(1)	(2)	(3)	(4)	(5)
	Plagiarism Score	Non-Essential Sections	Essential Sections	Elite Universities	Non-Elite Universities
Public Sector	1.201*** (0.060)	1.141*** (0.112)	1.261*** (0.125)	1.046*** (0.101)	1.387*** (0.134)
% to mean	15.60%	14.82%	16.38%	14.93%	17.23%
Degree FE	Y	Y	Y	Y	Y
Major FE	Y	Y	Y	Y	Y
Cohort FE	Y	Y	Y	Y	Y
University FE	Y	Y	Y	Y	Y
Individual chara.	Y	Y	Y	Y	Y
Uni-cohort-major	N	N	N	N	N
Observations	511,021	511,021	511,021	120,881	390,103
R-squared	0.234	0.207	0.245	0.233	0.196

- ▶ Public sector officials: 1.2 pct point/ 15% ↑ plagiarism (vs. private-sector classmates)
- ▶ Plagiarism, as an academically dishonest behavior, is a predictor for public sector entry

Are Public Officials More Likely to Plagiarize?

	(1)	(2)	(3)	(4)	(5)	(6)
	Plagiarism Score	Non-Essential Sections	Essential Sections	Elite Universities	Non-Elite Universities	Plagiarism Score
Public Sector	1.201*** (0.060)	1.141*** (0.112)	1.261*** (0.125)	1.046*** (0.101)	1.387*** (0.134)	1.312*** (0.098)
% to mean	15.60%	14.82%	16.38%	14.93%	17.23%	17.03%
Degree FE	Y	Y	Y	Y	Y	N
Major FE	Y	Y	Y	Y	Y	N
Cohort FE	Y	Y	Y	Y	Y	N
University FE	Y	Y	Y	Y	Y	N
Individual chara.	Y	Y	Y	Y	Y	Y
Uni-cohort-major	N	N	N	N	N	Y
Observations	511,021	511,021	511,021	120,881	390,103	478,162
R-squared	0.234	0.207	0.245	0.233	0.196	0.261

- ▶ Public sector officials: 1.2 pct point/ 15% ↑ plagiarism (vs. private-sector classmates)
- ▶ Plagiarism, as an academically dishonest behavior, is a predictor for public sector entry

Heterogeneity across Public Positions

	(1)	(2)	(3)	(4)	(5)
	Judges	Tax Officials	Law Enforcement	Custom Officials	Other Admin
Public Sector	1.234*** (0.147)	2.005*** (0.125)	1.120*** (0.125)	1.981*** (0.140)	0.945*** (0.136)
% to mean	16.03%	25.96%	14.52%	24.65%	12.27%
Observations	190,141	305,451	279,551	341,567	401,515
R-squared	0.195	0.207	0.266	0.208	0.214
Degree FE	Y	Y	Y	Y	Y
Major FE	Y	Y	Y	Y	Y
Cohort FE	Y	Y	Y	Y	Y
University FE	Y	Y	Y	Y	Y
Individual chara.	Y	Y	Y	Y	Y

- More salient in the tax and customs system; modest in judges and others

Heterogeneity across Regions

	(1)	(2)	(3)	(4)
	Economic Development		Corruption	
	High GDP	Low GDP	More Corrupt	Less Corrupt
Public Sector	1.079*** (0.083)	1.453*** (0.104)	1.401*** (0.124)	0.955*** (0.078)
% to mean	14%	19%	18%	12%
Degree Type	Y	Y	Y	Y
Major FE	Y	Y	Y	Y
Cohort FE	Y	Y	Y	Y
University FE	Y	Y	Y	Y
Student characteristics	Y	Y	Y	Y
Observations	209,257	191,843	167,579	215,778
R-squared	0.116	0.130	0.104	0.142

- Wider gap in less developed and more corrupt regions

Are Plagiarists Promoted Faster?

	(1)	(2)	(3)	(4)	(5)
	Civil Servants	Tax Officials	Customs Officials	Judges	
	Rank (1-27, 27 = highest)			Rank (1-5, 5 = highest)	
High-plagiarism Official	0.213*** (0.020)	0.401*** (0.034)	0.369*** (0.036)	0.271*** (0.026)	0.223*** (0.025)
% to control mean	9%	15%	12%	9%	7%
Unit-Entry-Year FE	Y	Y	Y	Y	Y
Individual Characteristics	Y	Y	Y	Y	Y
Performance Controls	N	N	N	N	Y
Observations	60,113	10,104	8,925	61,024	61,024
R-squared	0.081	0.084	0.091	0.096	0.105

- ▶ High-plagiarism officials get ahead faster: 7-15% higher rank after the first 5 years
- ▶ Heterogeneity: salient in the tax and customs system; modest in judges and others

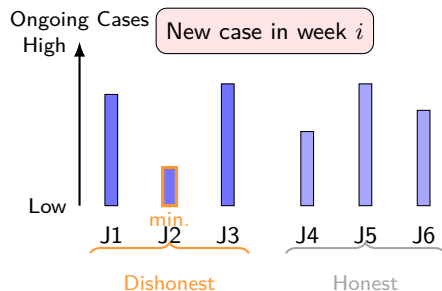
Full Table

2 Any Institutional Consequences When Dishonest Individuals Hold Power?

- ▶ For most public positions, performance is hard (impossible) to measure:
 - ▶ Multi-dimensional, difficult to measure objectively/quantitatively
 - ▶ Vary substantially across individuals
- ▶ Focusing on the judiciary:
 - ▶ Every decision is documented, can be measure quantitatively (win rate, appeal, etc.)
 - ▶ Cases of similar types are routinely/quasi-randomly assigned to different judges

IV for Case Assignment: Fluctuation in Judges' Weekly Caseload

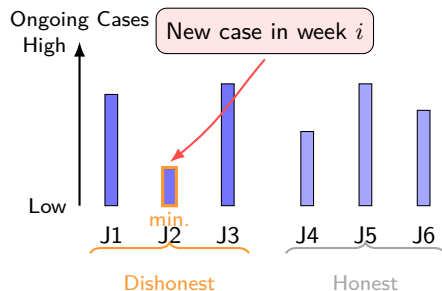
Rule: **equal caseload** across judges



- Instrument: minimum ongoing caseload among high-plagiarism (dishonest) judges in court c , domain d , week t (*scheduling luck*)

IV for Case Assignment: Fluctuation in Judges' Weekly Caseload

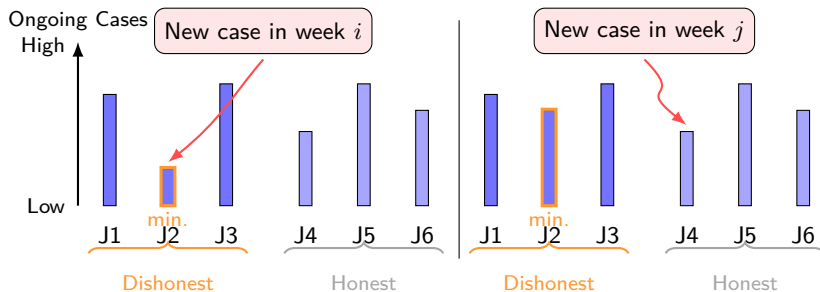
Rule: **equal caseload** across judges



- ▶ Instrument: minimum ongoing caseload among high-plagiarism (dishonest) judges in court c , domain d , week t (*scheduling luck*)
- ▶ If the minimum is low, a new case $\Rightarrow$ a high-plagiarism (dishonest) judge

IV for Case Assignment: Fluctuation in Judges' Weekly Caseload

Rule: **equal caseload** across judges



- Instrument: minimum ongoing caseload among high-plagiarism (dishonest) judges in court c , domain d , week t (*scheduling luck*)
- If the minimum is low, a new case $\Rightarrow$ a high-plagiarism (dishonest) judge
- If the minimum is high, a new case $\Rightarrow$ a low-plagiarism (honest) judge

IV: Weekly Caseload of the Least Busy Plagiarizing Judge

$$PlagiarizingJudge_{icdt} = \alpha \cdot CaseLoad_{cdt} + \lambda_{cd} + \gamma_t + \epsilon_{icdt}, \quad (1)$$

$$Ruling_{icdt} = \widehat{PlagiarizingJudge}_{icdt} + \lambda_{cd} + \gamma_t + \epsilon_{icdt}, \quad (2)$$

- ▶ $CaseLoad_{cdt}$: lowest caseload among high-plagiarism judges in court c , domain d , week t
- ▶ $PlagiarizingJudge_{icdt} = 1$ if case i is assigned to a high-plagiarism (dishonest) judge
 - ▶ $>$ Cross-section median
 - ▶ Robust to other cutoffs
- ▶ $Ruling_{icdt}$: case outcomes (rulings, etc.)
- ▶ Court-by-domain fixed effects λ_{cd} and weekly fixed effects γ_t .
- ▶ SEs are clustered at the court level.

Plagiarizing Judges, Rulings Favoritism, and Quality

	(1)	(2)	(3)	(4)	(5)	(6)
	Win Rate (Favoritism)			Quality		
	Larger Firms	SOEs	Gov't	Appeal Rate	Discretionary Codes	Reasoning Length
High-plagiarism Judge	0.041*** (0.011)	0.056*** (0.015)	0.045*** (0.017)	0.024*** (0.006)	0.068*** (0.011)	-36.985*** (9.475)
F-stat	225	144	105	201	201	201
% of Mean	12%	15%	10%	18%	19%	7%
Observations	1,759,415	150,251	251,356	1,735,195	1,735,195	1,735,195
R^2	0.197	0.291	0.157	0.161	0.131	0.271
Fixed Effects	Year-Week, Court-Domain, Cohort, University					

- ▶ Caseload strongly predicts case assignment: F-stat 105-225
- ▶ High-plagiarism judges render more:
 - ▶ preferential rulings: 10-15% ↑ win rates for powerful connected parties
 - ▶ low-quality rulings: 18% ↑ appeal, 19% ↑ discretionary, and 7% ↓ reasoning length

Balance in other dimensions

Verdict

Win Rate Measure

Possible Explanations for the “Questionable” Rulings

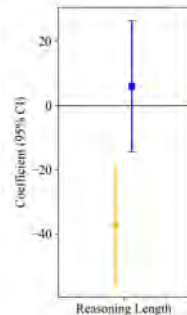
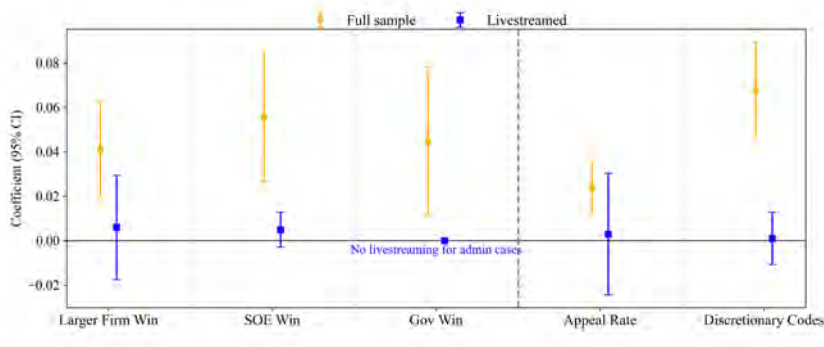
- ▶ How to understand these patterns of preferential low-quality rulings?
- ▶ Not necessarily about misconduct/corruption, plagiarizing judges:
 1. different preferences
 2. try to correct the existing favoritism towards less powerful connected litigants
 3. less competent

Suggestive Evidence: Corruption Investigations

	(1)	(2)
	Corruption Investigation (=1)	
High-plagiarism Judge	0.007*** (0.002)	0.006*** (0.002)
% to outcome mean	17%	17%
Judge Characteristics	N	Y
Entry Year FE	Y	Y
University FE	Y	Y
Court FE	Y	Y
Observations	59,392	59,392
R-squared	0.107	0.094

- ▶ High-plagiarism judges are 0.6% more likely to be investigated (17% ↑ in the mean)
- ▶ 78% investigations involve bribery, abuse of power, or case-related favoritism
- ▶ Suggesting preferential rulings are likely linked to actual misconduct

Suggestive Evidence: Livestreamed Trials



- ▶ In **livestreamed cases**, the high-plagiarism judge effect disappears
- ▶ Suggesting not about judges' incompetence but strategic choice
- ▶ Transparency can be a powerful constraint on misconduct

2.2. Do High-Plagiarism (Dishonest) Judges Generate Spillover Effects?

- ▶ We've shown that high-plagiarism judges directly undermines institutional performance
- ▶ Next, we ask whether their behavior also *spreads to others*.
 1. Spillovers to peer judges
 2. Spillovers (complementarity) between judges and lawyers

Spillover to Peer Judges

	(1)	(2)	(3)
	Larger Firm's Win Rate	SOE's Win Rate	Gov's Win Rate
Mentored by High-plagiarism Senior	0.015** (0.009)	0.021** (0.011)	0.019* (0.010)
Junior Judge's Plagiarism	0.026*** (0.008)	0.024*** (0.007)	0.030*** (0.008)
First Stage F-statistics	85	81	130
Observations	370,457	60,9845	102,067
R-squared	0.156	0.145	0.156
Year-Week FE	Y	Y	Y
Dispute Type	Y	Y	Y
Court-Domain FE	Y	Y	Y

Newly appointed judges will be mentored by a senior judge for 1-2 years

- ▶ Mentorship by high-plagiarism seniors → rulings favoring powerful/connected
- ▶ Magnitude is around half of their own dishonesty
- ▶ Assimilation rather than selection, as no systematic sorting on probity

No sorting

Spillover (Complementarity) between Judges and Lawyers

	(1)	(2)	(3)
	Client's (Lawyer's) Win Rate		
	Contract Dispute	Non-Contract Dispute	Administrative Disputes
Dishonest Judge $\times$ Dishonest Lawyer	0.023** (0.009)	0.021** (0.011)	0.016* (0.010)
Dishonest Judge	-0.013* (0.008)	-0.009 (0.011)	-0.010 (0.008)
Dishonest Lawyer	0.007 (0.009)	0.009 (0.014)	-0.006 (0.010)
% to Outcome mean	8%	6%	5%
First Stage F-statistics	106	96	88
Observations	111,546	61,667	74,081
Year-Week FE	Y	Y	Y
Litigant Type	Y	Y	Y
Court-Domain FE	Y	Y	Y

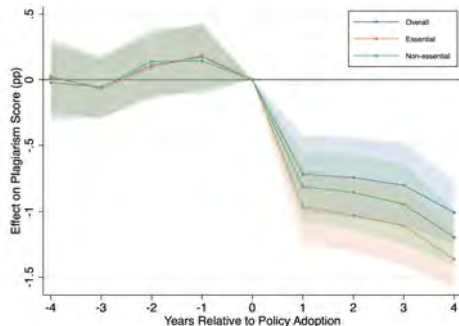
- ▶ Dishonest lawyers achieve 5-8% $\uparrow$ win rate when randomly assigned to dishonest judges
- ▶ No such effect or even reverse for honest lawyers
- ▶ Dishonest judges systematically empower dishonest lawyers

3. Can Promoting Academic Integrity in Universities Help?

A high-stakes anti-plagiarism policy in university

- ▶ Set similarity ceilings for dissertations (tied to degree conferral) since 2013
- ▶ Exploiting the staggered-rollout, we examine:
 1. Plagiarism in dissertations
 2. Selection into the public service
 3. Long-run professional conducts

University Plagiarism-ceiling Policy: Sanity Check



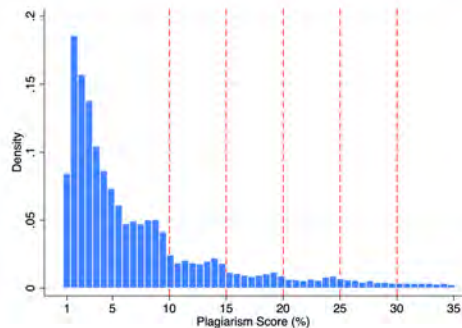
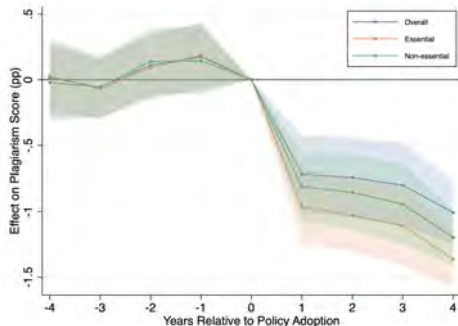
- Ceiling policy leads to a 1 pct point decline in plagiarism rates (12% ↓ pre-reform)

Regression Table

Policy

Bunching Before vs. After

University Plagiarism-ceiling Policy: Sanity Check



- ▶ Ceiling policy leads to a 1 pct point decline in plagiarism rates (12% ↓ pre-reform)
- ▶ Bunching below threshold: strategic responses, not necessarily improved honesty

Regression Table

Policy

Bunching Before vs. After

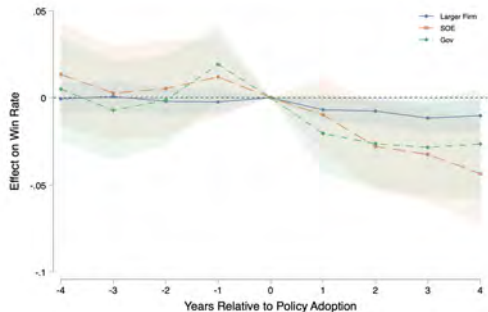
University Plagiarism-ceiling Policy and Political Selection

	(1)	(2)	(3)	(4)
	Plagiarism Score	Non-Essential Sec.	Essential Sec.	"Surprised" Cohort
Public x Treated	0.041 (0.093)	0.031 (0.104)	0.047 (0.105)	0.006 (0.855)
Treated Graduates	-0.906*** (0.131)	-0.811*** (0.097)	-1.026*** (0.103)	-1.201*** (0.133)
Public Sector	1.077*** (0.110)	1.062*** (0.104)	1.128*** (0.096)	1.108*** (0.104)
Fixed Effects	Degree, Major, Cohort, University FEs			
Observations	307,981	307,981	307,981	230,815
R-squared	0.193	0.206	0.215	0.167

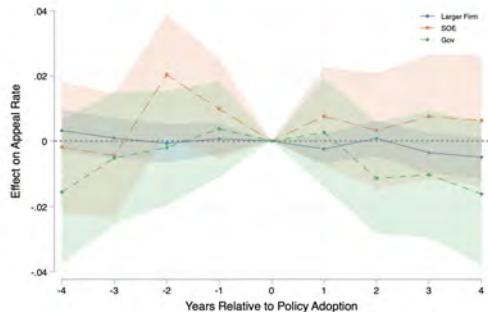
- ▶ Null effect on political selection, even for cohorts admitted before, graduated after
- ▶ Plagiarists still more likely to enter the public sector

Bunching vs. Complier

Can Plagiarism-ceiling Policy Foster Ethical Behaviors?



(a) Reduction in Favoritism towards Powerful Litigants



(b) Reduction in Low-quality Rulings

- ▶ Compare rulings of judges graduated before and after
- ▶ Modest ↓ in favoritism/low-quality rulings ($\approx 15\%$ baseline dishonest judge effect)
- ▶ Policy isn't purely symbolic $\rightarrow$ marginally reduced misconduct

Conclusions

- ▶ What we do:
 - ▶ A high-stakes, scalable measure of academic dishonesty: dissertation plagiarism
 - ▶ Link half-million graduate dissertation plagiarism checks to public officials' selection and performance
- ▶ What we find:
 - ▶ Adverse political selection on academic honesty
→ plagiarist more likely to enter & advance in the public sector
 - ▶ Placing dishonest actors in public positions can harm institutional performance
 - ▶ Dishonesty spread via professional networks, shaping social norms
 - ▶ Enforcing academic integrity in schools marginally decreases professional misconduct

Thank You!

Preview of Main Findings

1. Adverse political selection in all positions

- ▶ Individuals entering the public service: 15% ↑ in plagiarism
- ▶ Plagiarists are promoted faster

2. Dishonesty harms institutional performance

- ▶ Dishonest judges favor powerful/connected: ↑ win rate(big firm/SOE/Govt)
- ▶ Low quality: more discretionary codes; higher appeal; corruption ↑
- ▶ Null effect in livestreamed trials → strategic choice rather than incompetence
- ▶ Negative spillover to peer judges and between judges and lawyers

3. University's anti-plagiarism (ceiling) policy has modest impacts

- ▶ plagiarism ↓, adverse selection remains, marginal ↑ in professional behavior

Contributions to the Literature

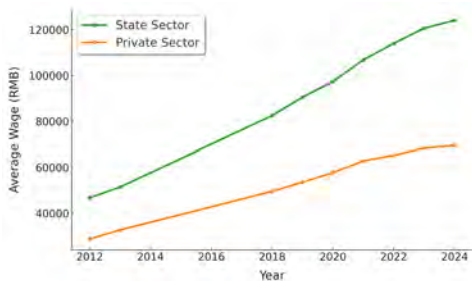
1. (Dis)honesty in political selection and its consequence

- ▶ Lab/survey experiment honesty measures and its correlation with preference for public jobs and real-world behavior
(Fischbacher & Föllmi-Heusi 2013; Abeler et al. 2014; Hanna & Wang 2017; Barfort et al., 2019; Olsen et al., 2019; Fisman et al., 2025)
- ▶ Our paper:
 - provide the scalable, real-world measure of hidden academic dishonesty: linking half-million dissertation plagiarism to political selection (entry+promotion)
 - causally estimate impacts on professional misconduct in high-stakes settings

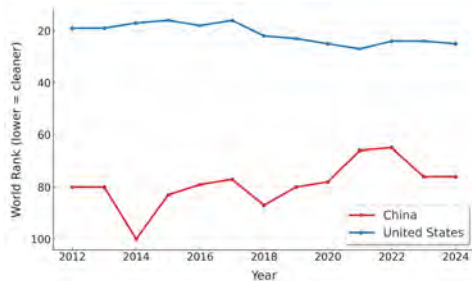
2. Corruption and development

- ▶ Corruption undermines development: institutions, monitoring, auditing, etc. (Shleifer & Vishny 1993; Mauro 1995; Fisman & Svensson, 2007; Olken 2007; Ferraz & Finan 2011; Colonnelli & Prem, 2022)
- ▶ Our paper:
 - individual virtue matters — dishonest actors behave unethically/corruptly
 - dishonesty spreads via peers, while mitigated by transparency

Wage Gap & Corruption



(a) Higher pay and widening gap



(b) High corruption: 80th in the world

Background

Low-plagiarism → Honesty or Sophistication?

Are students with lower plagiarism rates simply more **sophisticated** at gaming the system (e.g., using re-phrasing tools) rather than genuinely more honest?

Evidence suggests sophistication is not the main driver:

1. Bunching Analysis:

- ▶ Strategic agents typically bunch at enforcement thresholds.
- ▶ We observe **no bunching** in the pre-enforcement period for either public officials or private-sector students. [Bunching](#)

2. Technological Availability:

- ▶ Results are robust to the *early period* of the sample (before 2016) [Early Period Sample](#)
- ▶ During this time, detection tools were scarce and student awareness of the algorithm was negligible

[Back to Distribution](#)

[Back to Selection](#)

Bunching Before and After Plagiarism-check Policy

./Figures/no_bunching_before.jpg

(a) Before

./Figures/bunching_after.jpg

(b) After

University Plagiarism-ceiling Policy and Political Selection

	(1)	(2)
	Public Official	
Bunching Complier	0.0435*** (0.0105)	0.0435*** (0.0105)
Comfortable Complier	-0.056*** (0.007)	-0.056*** (0.007)
Observations	74,753	74,753
R-squared	0.174	0.174
College FE	Y	Y
Major FE	Y	Y
Cohort FE	Y	Y
Controls	N	Y

- Marginal/bunching compliers are more likely to become public officials, while comfortable compliers are less likely, relative to those who fall between

Summary of Spillover Results

1. Spillovers from senior judges to junior judges via mentorship
 - ▶ Mentorship by high-plagiarism seniors → rulings favoring powerful/connected
 - ▶ Assimilation rather than selection, as no systematic sorting on plagiarism
2. Spillovers (complementarity) between judges and lawyers
 - ▶ Dishonest lawyers achieve 5-8% ↑ win rate when randomly assigned to dishonest judges
 - ▶ Dishonest judges systematically empower dishonest lawyers

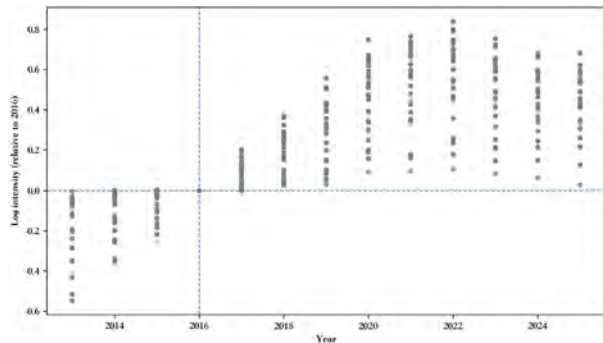
[Back](#)

[Go To Policy Part](#)

Access to CNKI's Plagiarism System

- ▶ Before 2022, students could not access CNKI's detection tool directly
- ▶ Only universities held institutional accounts
- ▶ The process was fully managed by the university (library or graduate office); students could not self-screen or pre-test their drafts.
 - ▶ Some universities may delegate this to thesis supervisors
 - ▶ A sizable underground market: students sometimes paid intermediaries or had university staff run checks privately (300 USD per thesis, around half of one year's tuition at top universities)
- ▶ Therefore, most authors lacked access to the detection system
- ▶ Our measure is less likely to be contaminated by manipulation/strategic responses at a large scale

Low Awareness of Plagiarism Detection Before 2016



(a) Baidu Search Index on "Plagiarism Detection/ "

- ▶ Before 2016, public interest in "plagiarism detection" was consistently low
- ▶ In 2016 the MOE released an Order: "Measures for Preventing and Handling Academic Misconduct in Higher Education Institutions"
- ▶ This regulation triggers an immediate and sustained jump in search activity

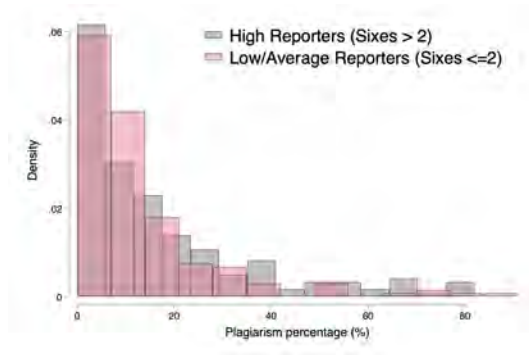
Low-Awareness Period (Pre-2016) Robustness

	(1)	(2)	(3)
Panel A Selection into Public Sectors			
	Overall Plagiarism Score	Non-Essential Sections	Essential Sections
Public Sector	1.325*** (0.084)	1.204*** (0.121)	1.314*** (0.145)
Fixed Effects/Controls	University, Major, Cohort, Student characteristics		
Panel B Favoritism in Rulings			
	Larger Firms	SOEs	Gov's
High-plagiarism Judge	0.036*** (0.012)	0.052* (0.032)	0.041** (0.017)
Fixed Effects	Court-domain, Week, Cohort, University		
Panel C Low Quality of Rulings			
	Appeal Rate	Discretionary Codes	Judicial Reasoning Length
High-plagiarism Judge	0.023*** (0.009)	0.054*** (0.015)	-30.985*** (10.056)
Fixed Effects	Court-domain, Week, Cohort, University		

[Back to Selection](#)
[Back to Sophistication](#)

Plagiarism and Dice Game Over-reporting

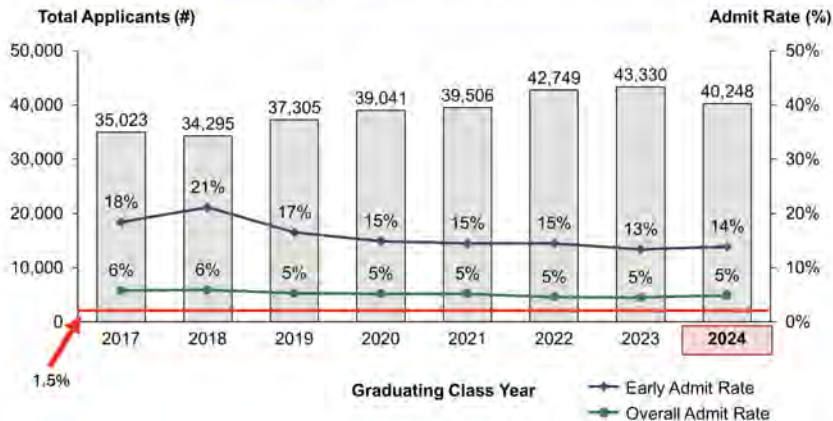
Over-reporting in the dice game shows higher plagiarism, especially $> 40\%$



[Back](#)

National Civil Service Exam: 1.5% Admission Rate

Harvard University: Cumulative Admission Statistics



Roster of Admitted Civil Servants

深圳海关2021年考录公务员拟录用人员公示

© 2005 Blackwell Publishing Ltd *Journal of Internal Medicine* 258: 105–112

—台享到

(原標題：深圳南山2021年度擬錄用公務員公示公告)

根据2013年度中央机关及其直属机构考试录用公务员工作有关要求,经过笔试、面试、体检和考察等程序,确定邵坤等22人为深圳海关拟录公务员,现予以公示。公示期间如有问题,请向深圳海关反映。

会审时间：2003年6月14日和8月8日。（含个工作日）

監督電話：0755-84398368

联系地址：广东省深圳市南山区深圳大道2888号悦洲

附註

復利利率2024年複利利率2024年複利利率

序号	犯罪嫌疑人姓名	性别	年龄	身份证号	学历	单位职务	工作经历	备注
23	犯罪嫌疑人王某某(女) (1985-03-15)	女	34	360101198503150000	本科 (硕士)	中国政法大学	无	犯罪嫌疑人
24	犯罪嫌疑人王某某(女) (1988-07-22)	女	31	360101198807220000	硕士 (博士)	复旦大学	无	犯罪嫌疑人
25	犯罪嫌疑人王某某(男) (1982-05-10)	男	37	360101198205100000	本科 (硕士)	清华大学	无	犯罪嫌疑人
26	犯罪嫌疑人王某某(女) (1990-09-03)	女	29	360101199009030000	本科	浙江工业大学	无	犯罪嫌疑人
27	犯罪嫌疑人王某某(女) (1987-11-18)	女	32	360101198711180000	本科	南京理工大学	无	犯罪嫌疑人
28	犯罪嫌疑人王某某(女) (1992-04-25)	女	27	360101199204250000	本科	中山大学	无	犯罪嫌疑人
29	犯罪嫌疑人王某某(女) (1989-12-01)	女	30	360101198912010000	本科	南方科技大学	无	犯罪嫌疑人

- An example of a roster for the Shenzhen Customs office, posted in 2021
- Individual-level information: name, gender, college, degree, specific work position

Example: Paraphrasing

Original

Integrity is the cornerstone of social order; without it, even the best-designed institutions fail to function.

Paraphrased Version

The normal operation of society depends on honesty and trust; even with well-established systems, a lack of integrity prevents them from working effectively.

Algorithm Detection

CNKI detects overlapping core concepts and reordered syntax

Detected similarity: 80%.

Example: Mixed-Source Weaving

Source Sentences

A:

B:

Source A: Civil servants hold substantial discretion in policymaking.

Source B: Excessive discretion can lead to abuse of power and corruption.

Woven Version

Public officials often exercise significant discretion in decisions, but when unchecked, this discretion may result in abuse of authority or even corruption.

Algorithm Detection

CNKI links phrase fragments to multiple sources, identifying “mosaic plagiarism.”

Detected similarity: 70%.

Example: Synonym Substitution

Original

Universities should establish a scientific system for academic evaluation to improve research quality.

Synonym-Substituted Version

Colleges need to develop a rational mechanism for research assessment to enhance the level of scholarship.

Algorithm Detection

Despite synonym swaps (" → ", " → "), the system's word-vector semantics still flag high similarity

Detected similarity: 75%.

Balance in Case Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Monetary Stake (log)	Big Firm PLTF	Big Firm DFDT	SOE PLTF	SOE DFDT	Top Lawyer PLTF	Top Lawyer DFDT
High-plagiarism JD	-0.064 (0.213)	0.013 (0.030)	-0.031 (0.056)	0.002 (0.041)	-0.014 (0.032)	0.015 (0.062)	0.026 (0.052)
Weekly FE	Y	Y	Y	Y	Y	Y	Y
Cohort FE	Y	Y	Y	Y	Y	Y	Y
University	Y	Y	Y	Y	Y	Y	Y
Court-Domain	Y	Y	Y	Y	Y	Y	Y
Obs.	1,709,415	1,691,502	1,698,134	1,704,145	1,702,167	1,700,240	1,696,228
R-squared	0.241	0.211	0.168	0.193	0.152	0.168	0.193

- Balanced in case/litigant/lawyer characteristics
- No selective assignment of particular cases to dishonest judges

Back

What Does Plagiarism Measure?

A survey of 443 legal professionals applying for a job at an elite research institute

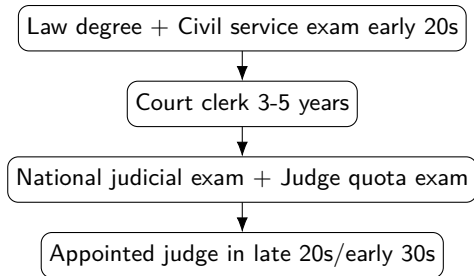
- ▶ Voluntary & confidential survey (387/443, 87%), we collect:
 1. Dissertations → plagiarism
 2. Personality traits: Big Five, honesty beliefs, and others
 3. Dishonesty: incentivized dice game, payoff based on self-reported sixes over 10 rolls
- ▶ Key Findings:
 1. Over-reporting in the dice game strongly predicts plagiarism
 2. Uncorrelated with stated honesty beliefs
 3. Some correlation with personality: neuroticism/orderliness- (coefficient is 1/3 of dice game), no correlation with opportunism/ambition/risk tolerance/pro-social
- ▶ Plagiarism: a strong proxy for dishonesty, capturing other dimensions
→ we interpret it as “academic dishonesty”

Court case assignment in China

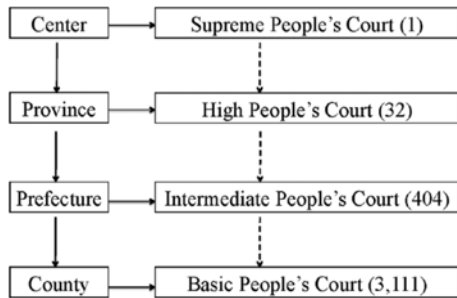
- ▶ Rule: ensure equal caseload across judges within the same domain and court
- ▶ Two types of assignment mechanisms:
 - ▶ Computer assignment
 - ▶ Human decision by court leaders

▶ Back

China's Judicial System



Panel A. Judges' Career Path



Panel B. Court Structure

- ▶ Judges are career officials holding civil service ranks, following civil-law systems
- ▶ Starts young, gaining experience on the bench (vs. experienced practitioners in common-law)
- ▶ Civil law tradition, no jury or binding precedent, granting significant power to judges
- ▶ Most cases start at the county level courts, with appeals moving up

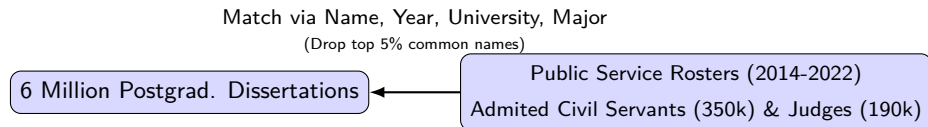
Data: linking academic dishonesty to political selection

6 Million Postgrad. Dissertations

Public Service Rosters (2014-2022)
Admitted Civil Servants (350k) & Judges (190k)

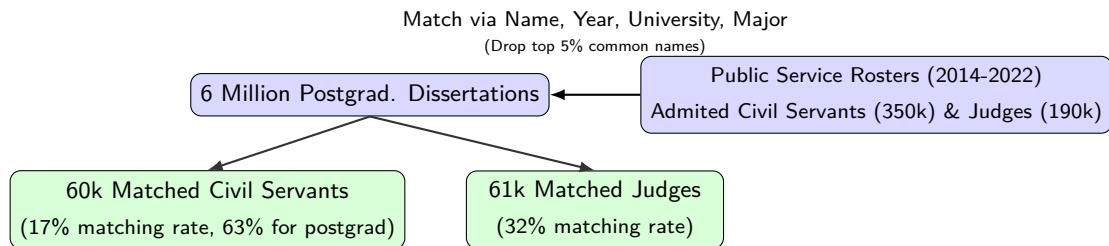
Data

Data: linking academic dishonesty to political selection



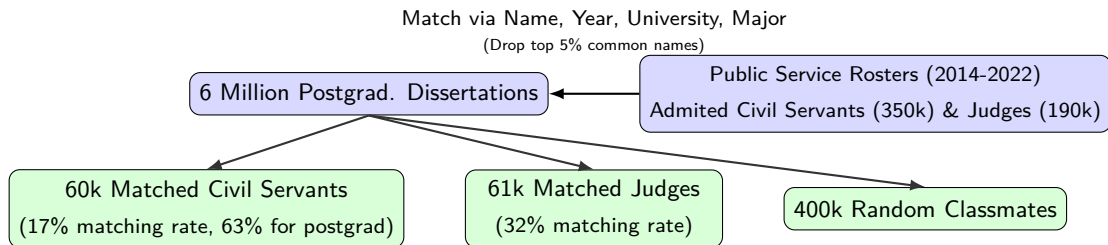
- ▶ Link the universe of 6M dissertations to public service rosters:
 - ▶ Strict 1-to-1 match on name, year, university, and major, dropping top 5% common names

Data: linking academic dishonesty to political selection



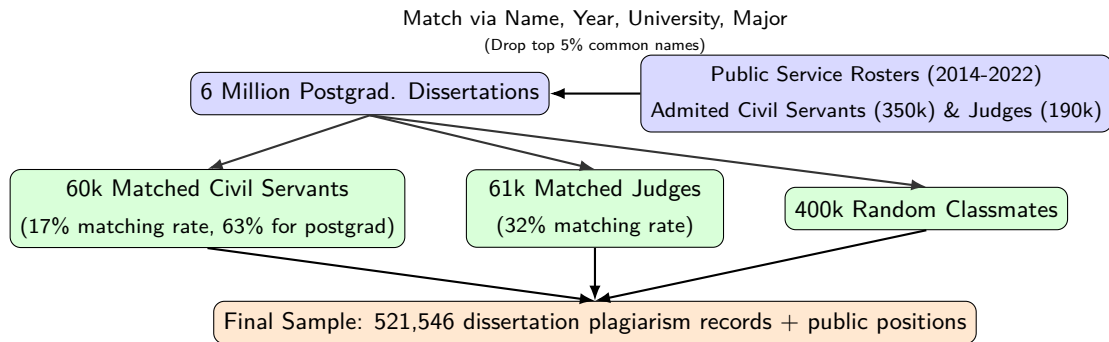
- ▶ Link the universe of 6M dissertations to public service rosters:
 - ▶ Strict 1-to-1 match on name, year, university, and major, dropping top 5% common names
- ▶ Public sector sample: 60k admitted civil servants + 61k judges between 2014-2022

Data: linking academic dishonesty to political selection



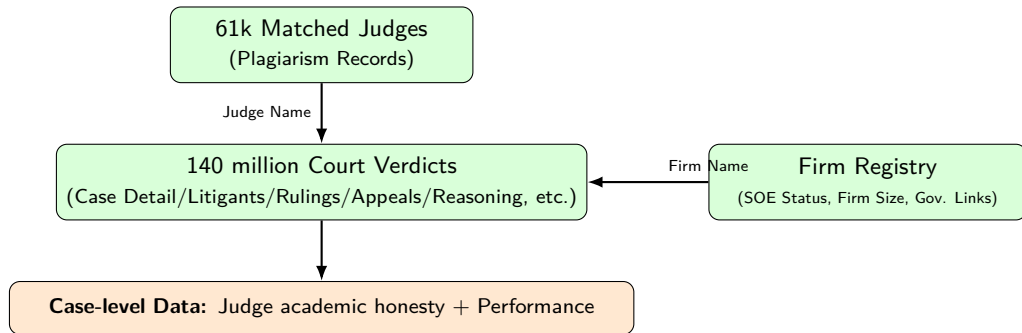
- ▶ Link the universe of 6M dissertations to public service rosters:
 - ▶ Strict 1-to-1 match on name, year, university, and major, dropping top 5% common names
- ▶ Public sector sample: 60k admitted civil servants + 61k judges between 2014-2022
- ▶ Private sector control group: random sample of 400k classmates

Data: linking academic dishonesty to political selection



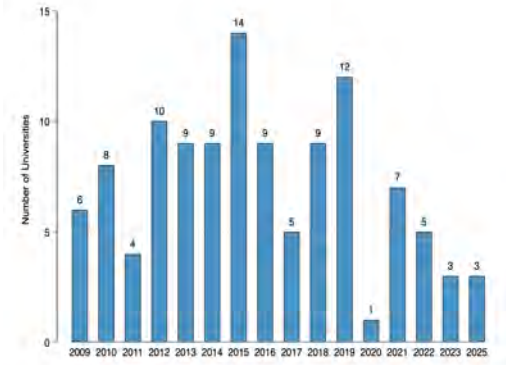
- ▶ Link the universe of 6M dissertations to public service rosters:
 - ▶ Strict 1-to-1 match on name, year, university, and major, dropping top 5% common names
- ▶ Public sector sample: 60k admitted civil servants + 61k judges between 2014-2022
- ▶ Private sector control group: random sample of 400k classmates

Data: linking academic dishonesty to institutional performance

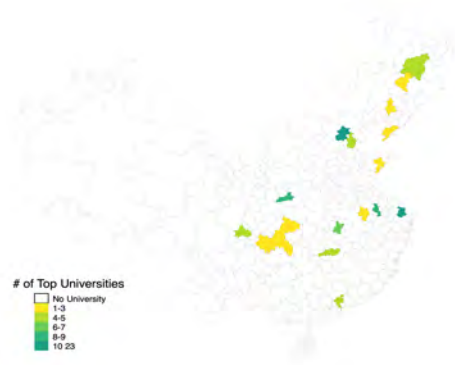


- ▶ Link 61k judges (with plagiarism records) to 140M court cases + firm registry:
 - case-level data: judges' academic honesty with performance outcomes (favoritism and quality)

Universities' Anti-Plagiarism Policy in China



(a) Staggered Rollout of Policies Over Time



(b) Geographic Distribution of Universities with Enforcement

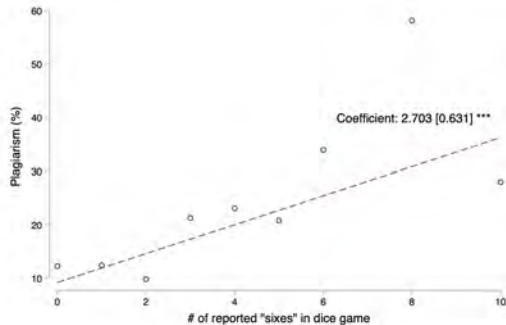
- The staggered rollout of plagiarism detection tools across Chinese universities from 2009-2022

Selection on Livestreaming Cases

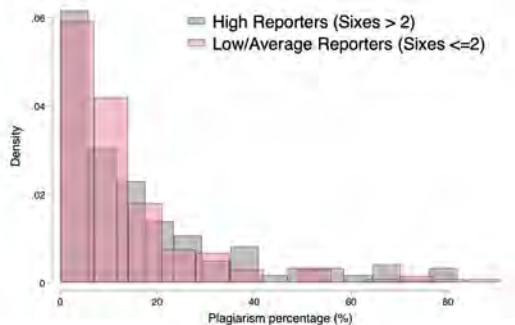
	(1)	(2)	(3)	(4)	(5)	(6)
	Dishonest Judges	Monetary Stake (log)	Contract Dispute (=1)	Involving SOE (=1)	Involving Large Firms (=1)	Case Duration (days, log)
Livestreaming	-0.027 (0.105)	-0.008 (0.243)	0.035 (0.026)	-0.013 (0.010)	0.019 (0.024)	0.004 (0.006)
Year-Week FE	Y	Y	Y	Y	Y	Y
Court FE	Y	Y	Y	Y	Y	Y
Observations	1,709,415	1,709,415	1,709,415	1,709,415	1,709,415	1,709,415
R-squared	0.104	0.091	0.126	0.141	0.172	0.118

- ▶ Principle rule: all the cases should be livestreamed except for: juvenile/family, state/company secrets
- ▶ Coverage varies with technical capacity and budget; empirically we do *not* see systematic selection on key observables

Descriptive evidence: Plagiarism, beliefs, and dishonest behavior



(a) Plagiarism vs. reported sixes



(b) Plagiarism distribution by high vs. low reporters

Survey design

- ▶ Online survey collaborated with an elite legal research institute; sample: experienced legal professionals.
- ▶ Application required uploading master/PhD dissertations (used to measure plagiarism).
- ▶ Immediately after application: optional personality questionnaire; **explicitly unrelated to hiring**.¹
- ▶ Measures: honesty beliefs, incentivized dice-game dishonesty, Big-5, orderliness, locus of control, self-monitoring, risk, demographics, GPA/scholarship.

Back

¹Used only for academic research by an independent team; disclosed to applicants.

Dice game for honesty measurement

1. To thank you for completing this survey, we will issue you a cash reward through a method of your choice. The reward amount will be determined by the outcome of dice rolls. You may use the virtual dice provided on this page, your own physical dice, or any online dice-rolling tool (like a WeChat emoji) to complete this part.
2. Reward Rules:
 - ▶ First 9 rolls: For each "6" you roll, you will receive a 5 RMB cash reward.
9 "6" 5
 - ▶ The 10th (final) roll: If you roll a "6", you will receive a 20 RMB cash reward.
10 "6" 20
3. To ensure effective anonymity, the system will never record or monitor your individual dice roll results. The final bonus depends entirely on the results you report below.

Questionnaire: Honesty-related Belief (Thesis Writing)

Regarding students using ghostwriting platforms for theses, which view is closest to yours?

- ▶ A. This is academic misconduct, violates integrity, and must be firmly resisted.
A.
- ▶ B. Not encouraged, but many are led to it by real-world pressures and ads.
B.
- ▶ C. If it's just a formality, this is more of a way to cope with the system.
C.
- ▶ D. Who writes it is just a 'formality'; the key is to graduate successfully.
D. “ ”

Back

Questionnaire: Opportunism

1. Regarding unexpected opportunities, which is your preference?
 - ▶ A. Carefully consider if the opportunity complies with norms or rules.
A. .
 - ▶ B. Tend to seize opportunities that advance the project.
B. .

Back

Attrition balance test

- 443 invited; 387 completed (87% response rate). No observable differences vs. non-respondents across age, gender, top university/city, scholarship, GPA, degree, or employment.

Characteristic	Respondents	Non-Resp.	Diff.	p-val
Age	29.195	29.341	-0.146	0.691
Female	0.767	0.762	0.005	0.886
Top Univ. (D)	0.214	0.207	0.007	0.835
Top City (D)	0.651	0.661	-0.010	0.798
Scholarship (D)	0.214	0.176	0.038	0.252
GPA (1=High)	0.209	0.181	0.028	0.396
Master+ (D)	0.735	0.757	-0.022	0.554
Employed (D)	0.521	0.535	-0.014	0.743

Dependent Variable: Dissertation Plagiarism Score (0-1)

Dice Game "sixes"	0.270*** (0.063)	0.264*** (0.064)	0.303*** (0.062)	0.261*** (0.062)	0.287*** (0.061)	0.263*** (0.093)
<i>B. Honesty beliefs</i>						
Belief: Thesis Ghostwriter		0.141** (0.070)			0.142** (0.070)	0.114 (0.125)
Belief: Exam Cheating		-0.038 (0.048)			-0.052 (0.047)	-0.057 (0.067)
Situational: Expense		-0.001 (0.017)			-0.008 (0.017)	0.020 (0.028)
Situational: Info Leak		-0.016 (0.027)			-0.028 (0.027)	0.040 (0.049)
<i>C. Personality traits</i>						
Conscientiousness Index			0.094 (0.071)		0.051 (0.070)	0.102 (0.109)
Extraversion			-0.010 (0.020)		-0.007 (0.019)	0.022 (0.032)
Agreeableness			0.000 (0.025)		-0.006 (0.025)	-0.045 (0.041)
Neuroticism			0.049*** (0.017)		0.053*** (0.017)	0.027 (0.029)
Openness			-0.016 (0.054)		-0.036 (0.053)	0.012 (0.090)
Locus of Control (External)			-0.003 (0.018)		0.004 (0.018)	0.019 (0.032)
Self-Monitoring			-0.006 (0.017)		-0.004 (0.017)	-0.038 (0.027)
Risk Seeking			-0.008 (0.021)		-0.014 (0.021)	0.031 (0.034)
Orderliness			-0.093*** (0.022)		-0.071*** (0.022)	-0.102*** (0.033)
Opportunism Index			-0.005 (0.043)		-0.013 (0.043)	-0.068 (0.071)
Time Management Index			-0.038 (0.047)		-0.012 (0.046)	-0.093 (0.074)
Pro-social Index			-0.041 (0.034)		-0.033 (0.034)	0.035 (0.058)
<i>D. Academic ability</i>						
Scholarship				-0.050** (0.019)	-0.047** (0.019)	-0.054* (0.031)
GPA				-0.068*** (0.020)	-0.062*** (0.020)	-0.035 (0.034)
<i>E. Demographics</i>						
Age				0.033 (0.052)	0.047 (0.053)	0.063 (0.160)
Female				0.031 (0.030)	0.038* (0.030)	0.016 (0.030)

Is honesty a stable lifetime trait?

- ▶ Honesty is highly stable in adulthood (age 30-60)

→ An individual's honesty is highly stable, with stability increasing/peaking in adulthood [2].

- ▶ Malleable, not fixed in early-age (age 15-30)

→ develops significantly through adolescence and early adulthood. [1]

- ▶ Dishonesty can be self-reinforcing:

→ A key neuroscience study found that the brain's negative emotional response to lying (in the amygdala) adapts and decreases with repetition, making it psychologically easier to engage in larger acts of dishonesty over time [3].

Key Citations



Ashton, M. C., Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, 11(2), 150-166.



Roberts, B. W., DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, 126(1), 3-25.



Garrett, N., Lazzaro, S. C., Ariely, D., Sharot, T. (2016). The brain adapts to dishonesty. *Nature Neuroscience*, 19(12), 1727-1732.

An Example of Court Judgment

河南省平顶山市新华区人民法院
民事判决书

(2015)新民初字第773号

Plaintiff 原告平顶山市恒泰昌物资贸易有限公司, 住所地平顶山市。
法定代表人张孟霞, 总经理。

Defendant 被告平顶山市纳绿装饰工程有限公司, 住所地平顶山市。
法定代表人周鹤生, 总经理。

原告平顶山市恒泰昌物资贸易有限公司(以下简称恒泰昌公司)与被告平顶山市纳绿装饰工程有限公司(以下简称纳绿公司)承揽合同纠纷一案, 本院于2015年5月5日受理后, 于2015年9月10日依法组成合议庭公开开庭进行了审理。原告恒泰昌公司委托诉讼代理人刘岱, 河南金豫律师事务所律师。被告纳绿公司委托诉讼代理人周鹤生, 到庭参加诉讼。本案现已审理终结。

Rulings 根据《中华人民共和国合同法》第二十九条之规定, 判决如下:

一、平顶山市纳绿装饰工程有限公司自判决生效之日起支付平顶山市恒泰昌物资贸易有限公司227149.8元及违约金, 违约金即227149.8元的利息损失的30%(利息自2013年2月1日起计算至平顶山市纳绿装饰工程有限公司支付之日止, 按银行同期贷款基准利率的四倍计算);

案件受理费5729元, 由平顶山市纳绿装饰工程有限公司负担。

如不服本判决, 可在判决书送达之日起十五日内, 向本院递交上诉状, 并按对方当事人的人数提出副本, 上诉于河南省平顶山市中级人民法院。

审判长 沈卫东
审判员 顾世法
人民陪审员 姚秋娟

The court fee of 5,729 CNY shall be borne by the defendant.

[Back to results](#)

[Back to data](#)

Key Judicial Outcomes

- ▶ Win rates: for (SOE vs private), (gov't vs citizen), (large vs small firms)
 - ▶ Civil practice rule: the court fees shall be paid by the losing party
 - ▶ Applies to both a partial win/lose and a full win/lose
 - ▶ If the court supports x% of your standing, then you pays 1-x% of the court fees.

$$WinRate_i = 1 - \frac{CourtFee_i}{CourtFee_i + CourtFee_j}$$

- ▶ Quality proxies: appeals; discretionary codes; reasoning length; processing time.

[Back to results](#)

[Back to data](#)

Balance in Other Case Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Monetary Stake (log)	Big Firm PLTF	Big Firm DFDT	SOE PLTF	SOE DFDT	Top Lawyer PLTF	Top Lawyer DFDT
Dishonest JD	-0.064 (0.213)	0.013 (0.030)	-0.031 (0.056)	0.002 (0.041)	-0.014 (0.032)	0.015 (0.062)	0.026 (0.052)
Weekly FE	Y	Y	Y	Y	Y	Y	Y
Cohort FE	Y	Y	Y	Y	Y	Y	Y
University	Y	Y	Y	Y	Y	Y	Y
Court-Domain	Y	Y	Y	Y	Y	Y	Y
Obs.	1,709,415	1,691,502	1,698,134	1,704,145	1,702,167	1,700,240	1,696,228
R-squared	0.241	0.211	0.168	0.193	0.152	0.168	0.193

- Balanced in other case/litigant/lawyer characteristics, results are unlikely to be driven by selective assignment of particular cases to dishonest judges

Back

No sorting on plagiarism

(1)	
	Junior Judge's Plagiarism Rate
Senior Judge's Plagiarism Rate	0.018 (0.012)
Senior Judge Expertise	Y
Junior Judges' Cohort FE	Y
Junior Judges' University FE	Y
Court FE	Y
Observations	46,368
R-squared	0.213

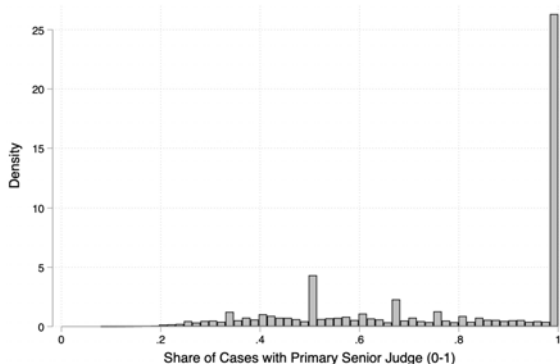
- No systematic sorting on plagiarism rate → assimilation rather than selection

(Dis)integrity in Private Sectors



- Evergrande case: the audit firm complicit in systematic overstatement of revenue

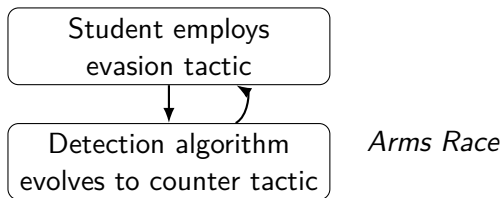
Junior-Senior Mentorship Pairing Stability



- Distribution of share of cases with primary senior judge
- A value of 1: a junior judge worked exclusively with one senior judge for all cases
- The mean of this measure is 0.75, suggesting fixed arrangement

Appendix: The Tech of Plagiarism Detection (Pre-LLM Era)

- ▶ The effectiveness of our measure relies on the sophistication of the detection tools
- ▶ In the pre-LLM era (before ~2020), the field was a constant "arms race" between detection algorithms and students trying to evade them.
- ▶ The core technology involved comparing a submitted text against a massive database of existing works.



[See Evasion Tactics](#) [Countermeasures](#)

[<< Back to Summary](#)

Appendix: Evasion Tactics and Algorithmic Countermeasures

The evolution of detection was a direct response to increasingly sophisticated student methods.

Common Student "Tricks"

- ▶ **Paraphrasing:** Rewriting sentences to change structure while keeping the original meaning.
- ▶ **Synonym Substitution:** Using a thesaurus to replace keywords.
- ▶ **"Mosaic" Plagiarism:** Weaving together phrases from multiple sources without citation.

Algorithmic Countermeasures

- ▶ **Fingerprinting:** Breaks text into overlapping n-grams (phrases of n-words) and compares their hash values. Effective against minor word changes.
- ▶ **Vector Space Models:** Represent documents as vectors based on word frequency. Can detect similarity in topic and keyword usage, good at catching paraphrasing.

[<< Back to Summary](#)

Appendix: Supporting the Monotonicity Assumption

- ▶ **The Assumption:** Making a dishonest judge more available will never *cause* a case that would have gone to them to be reassigned to an honest one (no "defiers").
- ▶ **The Argument:** This is institutionally plausible. The alternative is nonsensical.
- ▶ **Empirical Support:** We check the first-stage relationship across key subgroups.
- ▶ **Finding:** The instrument strongly and positively predicts assignment to a dishonest judge across all major subgroups (by court tier, case type, etc.). This consistency makes the existence of a significant group of hidden "defiers" highly implausible.