Teachers' Ethnic and Gender Biases: Behavioral Evidence from Danish Registry Data

Julian Schuessler
Kim Mannemar Sønderskov (CEPDISC)
Peter Thisted Dinesen (UCL/UniCPH)

February 14 2023
WZB Berlin

DEPARTMENT OF POLITICAL SCIENCE
AARHUS UNIVERSITY

CENTRE FOR THE EXPERIMENTAL-PHILOSOPHICAL STUDY OF DISCRIMINATION
DEPARTMENT OF POLITICAL SCIENCE
AARHUS UNIVERSITY

Section 1

Motivation

Motivation

- Danish 9th grade students receive
  - teacher grade T
  - externally (co-)graded, standardized test grade E

Motivation

- Danish 9th grade students receive
  - teacher grade T
  - externally (co-)graded, standardized test grade E
- Regress $T - E$ on students' gender/migration background
- Effect = teacher bias
- Large literature in econ doing exactly this

Prior findings

- Bias against boys (very robust) (Lavy 2009; Lavy and Megalokonomou 2019; Terrier 2020; Di Liberto, Casula and Pau 2021; Gibbons and Chevalier 2008)

Prior findings

- Bias against boys (very robust) (Lavy 2009; Lavy and Megalokonomou 2019; Terrier 2020; Di Liberto, Casula and Pau 2021; Gibbons and Chevalier 2008)
- Mixed findings on bias based on students' migration background (Alesina et al. 2018; Botelho, Madeira and Rangel 2015; Burgess and Greaves 2013; Gibbons and Chevalier 2008)

Prior findings

- Bias against boys (very robust) (Lavy 2009; Lavy and Megalokonomou 2019; Terrier 2020; Di Liberto, Casula and Pau 2021; Gibbons and Chevalier 2008)
- Mixed findings on bias based on students' migration background (Alesina et al. 2018; Botelho, Madeira and Rangel 2015; Burgess and Greaves 2013; Gibbons and Chevalier 2008)
- Biases correlate with implicit measures of teacher's bias (Alesina et al. 2018; Carlana 2018)

Prior findings

- Bias against boys (very robust) (Lavy 2009; Lavy and Megalokonomou 2019; Terrier 2020; Di Liberto, Casula and Pau 2021; Gibbons and Chevalier 2008)
- Mixed findings on bias based on students' migration background (Alesina et al. 2018; Botelho, Madeira and Rangel 2015; Burgess and Greaves 2013; Gibbons and Chevalier 2008)
- Biases correlate with implicit measures of teacher's bias (Alesina et al. 2018; Carlana 2018)
- Mixed and Null findings on teacher demographics, especially gender (Lavy 2009; Terrier 2020; Lindahl 2016)

Our contribution

- Discuss identification
- Estimate teacher biases using Danish register data, 2002–2019 (> 1 million students)
- Very small average biases (pro-girl, pro-migrant background), large heterogeneity across schools and teachers

Our contribution

- Discuss identification
- Estimate teacher biases using Danish register data, 2002–2019 ($>$ 1 million students)
- Very small average biases (pro-girl, pro-migrant background), large heterogeneity across schools and teachers
- We suggest *contact* as an explanation of biases
- H1: Share of migrants on teacher/school-level moderates bias
- H2: Relative performance of gender/migrant group moderates bias

Section 2

Empirical Framework

Data

- 2002–2019
- (Almost) all Danish 9th-grade students (both public and private schools)
- $N \approx 1,050,000$
- Two math grades:
  - One given by teacher
  - One from standardized, externally & often blindly graded test
- Student gender & (parents') migration background/country of origin

Data, ctd.

- Students nested in schools
- We also match students to teachers
  - Only public schools (2/3)
  - 2014–2019; many teachers only have 1–2 9th grade classes

Analytical Strategy

$$\begin{tikzcd} X \arrow[r] \arrow[rr, bend right] & A \arrow[r] & T \end{tikzcd}$$

- Student group X impacts on true ability (true grade) A which impacts on teacher grade T

Analytical Strategy

$$\begin{tikzcd} X \arrow[r] \arrow[rr, bend right] & A \arrow[r] & T \end{tikzcd}$$

- Student group X impacts on true ability (true grade) A which impacts on teacher grade T
- Plus possible direct effect of X on T: bias/discrimination

Analytical Strategy

$$\begin{tikzcd} X \arrow[r] \arrow[rr, bend right] & A \arrow[r] & T \end{tikzcd}$$

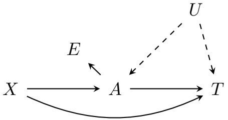
- Student group X impacts on true ability (true grade) A which impacts on teacher grade T
- Plus possible direct effect of X on T: bias/discrimination
- Bias can be estimated by regression of T on X and A

Analytical Strategy

$$\begin{tikzcd} X \arrow[r] \arrow[rr, bend right] & A \arrow[r] & T \end{tikzcd}$$

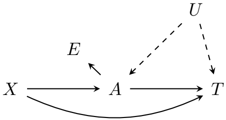
- Student group X impacts on true ability (true grade) A which impacts on teacher grade T
- Plus possible direct effect of X on T: bias/discrimination
- Bias can be estimated by regression of T on X and A
- Note that $X \rightarrow A$ may also contain bias, but cannot be disentangled in our data

Analytical Strategy



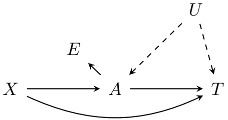
- A unobserved, external test E is proxy

Analytical Strategy



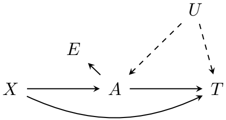
- A unobserved, external test E is proxy
- U unmeasured confounders: e.g., teacher competence, teacher "taste" for bias, parental involvement...

Analytical Strategy



- A unobserved, external test E is proxy
- U unmeasured confounders: e.g., teacher competence, teacher "taste" for bias, parental involvement...
- Without further assumptions, cannot estimate direct effect/bias

Analytical Strategy



- A unobserved, external test E is proxy
- U unmeasured confounders: e.g., teacher competence, teacher "taste" for bias, parental involvement...
- Without further assumptions, cannot estimate direct effect/bias
- Assume grades are a linear function of X, A and U
- And assume A translates equally into E and T
  - By law, E and T should measure same ability

Analytical Strategy

- Idea: Subtract test grade E from teacher grade T (Lavy 2009)
- Teacher deviations from test grade should not be a consequence of ability, but of teacher discretion/bias
- $T - E = \alpha + \beta X + \epsilon$
- β = direct effect of group characteristic on teacher grade, net of ability differences
- Also look at T and E as separate regression outcomes
- We standardize outcomes

Independent Variables

- Student gender, 1G/2G migration background, region of origin
- Teacher gender, migration background, age
- On teacher-/school-level:
  - Share of students with 1G/2G migration background
  - Performance of female/1G/2G students, relative to boys/no migration background

Section 3

Results

Results: Gender / Migration Background in Math

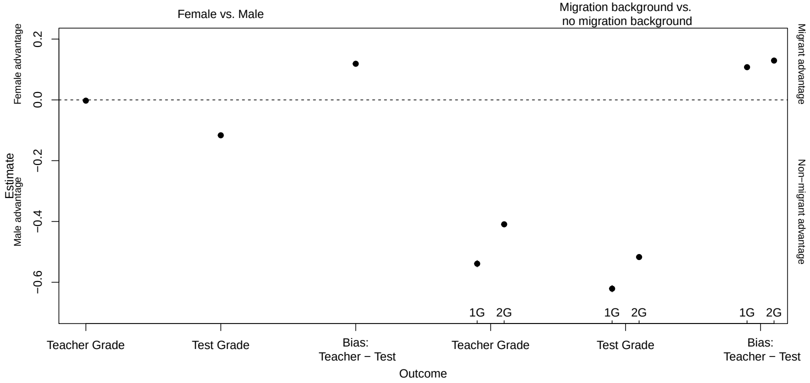
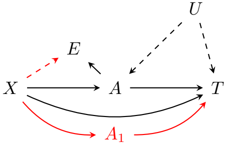


Figure: Results for math

Results: Gender / Migration Background

- For female and migrant students, teacher bias is positive
  - Are teachers “compensating” groups?
- No relevant interactions between gender and migration background
- Biases are relatively small (about $0.05 - 0.1$ standard deviations)
- Teacher FE do not change the results, even though they explain 16% of the variance in $T - E$
- What explains biases?

Further Challenges

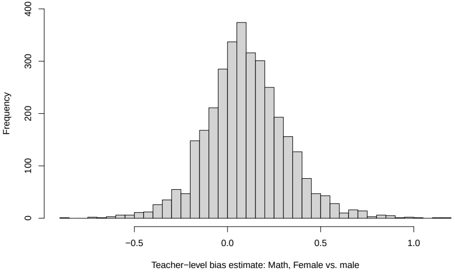


- Standardized test E may suffer from bias (gender differences in risk aversion; language skills of migrants)
- Teachers may grade based on classroom behavior/non-cognitive skills A_1
  - Contra legal regulations, but possible
- We derive a sensitivity analysis; inferences mostly robust (Schuessler et al. 2023)
- Constant bias in average effects does not impact on moderation analysis

Teacher-level analyses

- For each teacher, compute biases/performance across students taught by that teacher
- Relatively small sample (only public schools; many teachers only teach 1–2 classes)
- N $\approx$ 5,000; 2014–2019
- Variance across teachers cannot be explained by unobserved group differences
- However, variance across teachers could be due to small sample noise

Results: Gender biases over teachers



- 30% of teachers have a bias against female students
- 36% of teacher biases are > 0.2 standard deviations

Results: Teacher demographics

	Math Gender	Math 1G	Math 2G
(Intercept)	0.11***	0.06	0.09**
	(0.02)	(0.04)	(0.03)
Female	−0.01	−0.01	−0.01
	(0.01)	(0.02)	(0.01)
Migration Background: 1G	−0.01	−0.01	0.01
	(0.02)	(0.05)	(0.04)
Migration Background: 2G	−0.02	−0.01	0.07
	(0.04)	(0.11)	
Age	−0.02	0.00	0.00
	(0.00)	(0.00)	(0.00)
Num. obs.	5,307	5,307	

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

Table: Teacher-level regressions.

Bringing in Contact Theory

- Teacher demographics do not correlate with bias
- Implicit bias measures not available. Do they “explain” bias?
- Long-standing research strand in soc psychology suggests contact to out-group diminishes biases (Allport 1954, Pettigrew 1998, Levy Paluck et al. 2019)
- Sustained cooperative interactions are thought to reduce fear, build trust, reduce stereotypes (Nathan & Sands 2023)
- H1: The more students with a migration background are assigned to teacher, the more positive the bias

Updating Specific Stereotypes

- Classic contact theory: Contact reduces fear/anxiety; increases knowledge
- What kind of knowledge? About academic ability
- Suppose (some) teachers update group stereotypes based on group's performance
- Suppose (some) teachers have a preference to equalize academic outcomes
- H2: The worse a gender/migrant group assigned to a teacher performs on E, the more positive the bias

Results: Teacher Contact / Information

	Math
Female	0.30^{***}
	(0.01)
Migration Background: 1G	-0.04
	(0.04)
Migration Background: 2G	0.44^{***}
	(0.05)
1G×1G Share	0.58
	(0.63)
1G×2G Share	-0.33
	(0.76)
1G×1G Performance	-0.39^{***}
	(0.03)
2G×1G Performance	0.05
	(0.03)
Num. obs.	150,612

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table: Student-level regressions. Standard errors clustered on teachers.

Further results

- Contact (H1) consistently fails
- Information/performance (H2) robustly moderates bias
- 1G performance only moderates 1G bias; 2G performance only moderates 2G bias

Further results

- Contact (H1) consistently fails
- Information/performance (H2) robustly moderates bias
- 1G performance only moderates 1G bias; 2G performance only moderates 2G bias
- Gender performance moderates gender bias
- Migrant performance does not moderate gender bias (placebo test)
- Moderation also holds when subsetting on MENA- or non-MENA immigrants or with regional-background FE
- Interacting with performance on both teacher- and school-level, teacher-level moderation unchanged, school-level moderation zero

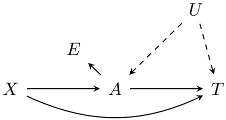
Section 4

Summary

Summary

- Small teacher biases in favor of girls & students with a migration background
  - Replicates results from other countries
- Large heterogeneity across teacher (& schools)
- Simple contact theory prediction fails
- Rather, it seems teacher-specific information about group-performance moderate bias
- Compensation: The worse the performance, the more positive the bias
- Could be a Scandinavian phenomenon: Strong equality norm

Analytical Strategy



$$A = \alpha_A + \beta_1 X + U + \epsilon_A,$$

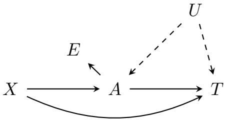
$$T = \alpha_T + \beta_2 X + \beta_3 A + U + \epsilon_T,$$

$$E = \alpha_E + \beta_4 A + \epsilon_E.$$

β_2 is effect of interest. Substitute for A to obtain

$$T - E = \alpha_{T-E} + \beta_2 X + \beta_1(\beta_3 - \beta_4)X + (\beta_3 - \beta_4)U + (\epsilon_T - \epsilon_E).$$

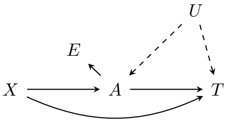
Sensitivity Analysis



$$T - E = \alpha_{T-E} + \beta_2 X + \beta_1(\beta_3 - \beta_4)X + (\beta_3 - \beta_4)U + (\epsilon_T - \epsilon_E).$$

- $\beta_4 < 1$ because of noise in test
- $\beta_3 < 1$ because of 1) noise 2) teachers being biased towards the midpoint of the scale
- Hard to assess whether $\beta_4 < \beta_3$ or $\beta_4 > \beta_3$

Sensitivity Analysis



$$T - E = \alpha_{T-E} + \beta_2 X + \beta_1(\beta_3 - \beta_4)X + (\beta_3 - \beta_4)U + (\epsilon_T - \epsilon_E).$$

$$\beta_2 = \hat{\beta}_2 + \beta_1\beta_4 - \beta_1\beta_3.$$

$\beta_1\beta_4$ is identified from the regression of E on X. $\beta_3, \beta_4 > 0$, so $\text{sgn}(\beta_1\beta_3) = \text{sgn}(\beta_1\beta_4)$

Sensitivity Analysis

$$\beta_2 = \hat{\beta}_2 + \beta_1\beta_4 - \beta_1\beta_3.$$

- Gender: $\hat{\beta}_2 \approx 0.1$ and $\beta_1\beta_4 \approx -0.1$.

Sensitivity Analysis

$$\beta_2 = \hat{\beta}_2 + \beta_1\beta_4 - \beta_1\beta_3.$$

- Gender: $\hat{\beta}_2 \approx 0.1$ and $\beta_1\beta_4 \approx -0.1$.
  - Then $\beta_2 > 0 \implies$ qualitative inference robust

Sensitivity Analysis

$$\beta_2 = \hat{\beta}_2 + \beta_1\beta_4 - \beta_1\beta_3.$$

- Gender: $\hat{\beta}_2 \approx 0.1$ and $\beta_1\beta_4 \approx -0.1$.
  - Then $\beta_2 > 0 \implies$ qualitative inference robust
- 1G/2G: $\hat{\beta}_2 \approx 0.1$ and $\beta_1\beta_4 \approx -0.5$.

Sensitivity Analysis

$$\beta_2 = \hat{\beta}_2 + \beta_1\beta_4 - \beta_1\beta_3.$$

- Gender: $\hat{\beta}_2 \approx 0.1$ and $\beta_1\beta_4 \approx -0.1$.
  - Then $\beta_2 > 0 \implies$ qualitative inference robust
- 1G/2G: $\hat{\beta}_2 \approx 0.1$ and $\beta_1\beta_4 \approx -0.5$.
  - If $\beta_3 > \beta_4$ then $\beta_2 > 0 \implies$ qualitative inference robust