Changing Stereotypes

Eliana La Ferrara

Invernizzi Chair in Development Economics

Bocconi University, Milan

Alberto Alesina Lecture

NES - November 12, 2020

Motivation

- Contemporary societies increasingly diverse along ethnic lines → public debate on inter-group prejudice and stereotypes
- Stereotypes serve a cognitive role, as mental representations of differences between groups that allow for faster processing of information
- But they also induce distortions in social behavior, e.g., through over-simplified judgement and discrimination
- Consequences: unequal access to economic, social or political opportunities; self-confirming aspiration traps; inter-group conflict

Motivation

Q: Can stereotypes be changed, and how?

Two approaches

1. Through social interaction
    - Exposure to roommates of different race in South Africa
2. By making people aware of their own stereoypes
    - “Revelation” experiment with teachers in Italian middle schools

I. Interaction, stereotypes and performance: Evidence from South Africa

Lucia Corno
Catholic University of Milan

Eliana La Ferrara
Bocconi

Justine Burns
University of Cape Town

Research question

- Can stereotypes be changed through social interaction?
- What are the effects on outcomes we care about (e.g., productivity)? Are they mediated by prejudice?

Policy intervention

- Exposure to roommates of different race in South Africa
- Apartheid led to stereotyping & marginalization of blacks

Outcomes

- Stereotypes (Implicit Association Test)
- Explicit attitudes (survey)
- Pro-social behavior (survey + games)
- Academic performance (admin.)

Conceptual framework

Contact hypothesis (Allport, 1954)

Interaction w/ other group → reduction of negative stereotypes if: (i) equal status; (ii) common goals; (iii) interdependent / work cooperatively; (iv) authorities support inter-group contact

"Negative contact hp" (Paolini et al. 2010, Barlow et al. 2012): heightened salience of difference in preferences

Channels

- Information updating
  - Baseline beliefs → interaction → updated beliefs
- Perspective taking/empathy
  - Change in "taste" for interaction

Background and Data

Institutional setting

University of Cape Town (UCT)

- Enrolls 4,000-5,000 freshmen/year; ≈50% live on campus
- Admission policy: Admission Point Score (based on high school grades) + diversity

Residence allocation policy, 1st year

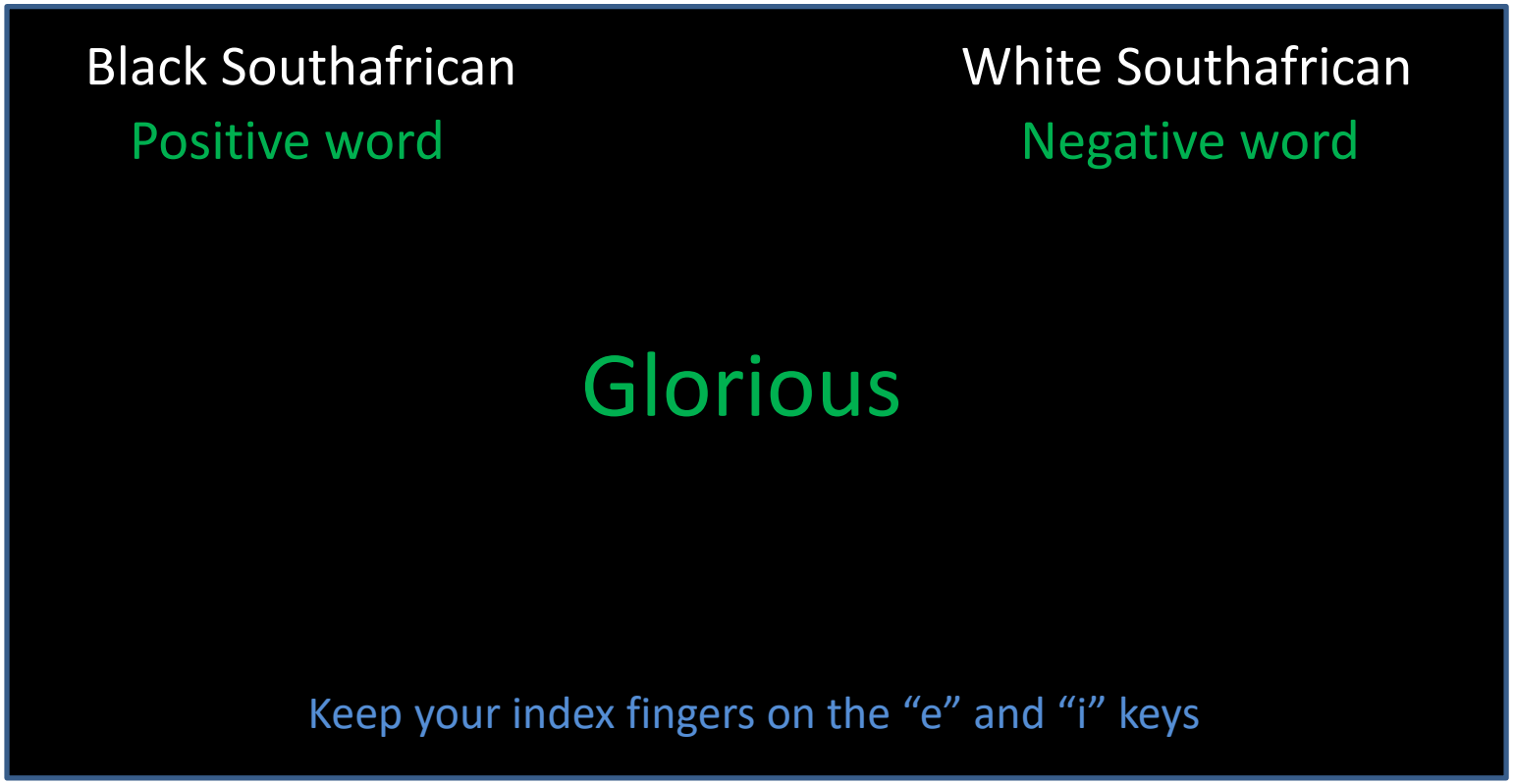
- Assignment to residences is random
- Allocation to rooms within residence (single or double) done by Warden → 8 residences randomize room assignment → our sample: freshmen who joined in 2012

2 survey rounds

- Feb 2012: 625 freshmen, ≈70% of universe in double rooms
- Sept 2012: 508 out of 625 → 19% attrition
  Attrition uncorrelated w/ treatment (mixed room), w/ baseline IAT & w/ interaction Treatment*IAT, Race*IAT

Implicit Association Test (IAT)

- Experimental method from social psychology (Greenwald and Banaji, 1995)
- Pair two concepts in rapid categorization task. Speed in associating: mental process perceives pair as less common



Implicit Association Test (IAT)

Advantages of IAT

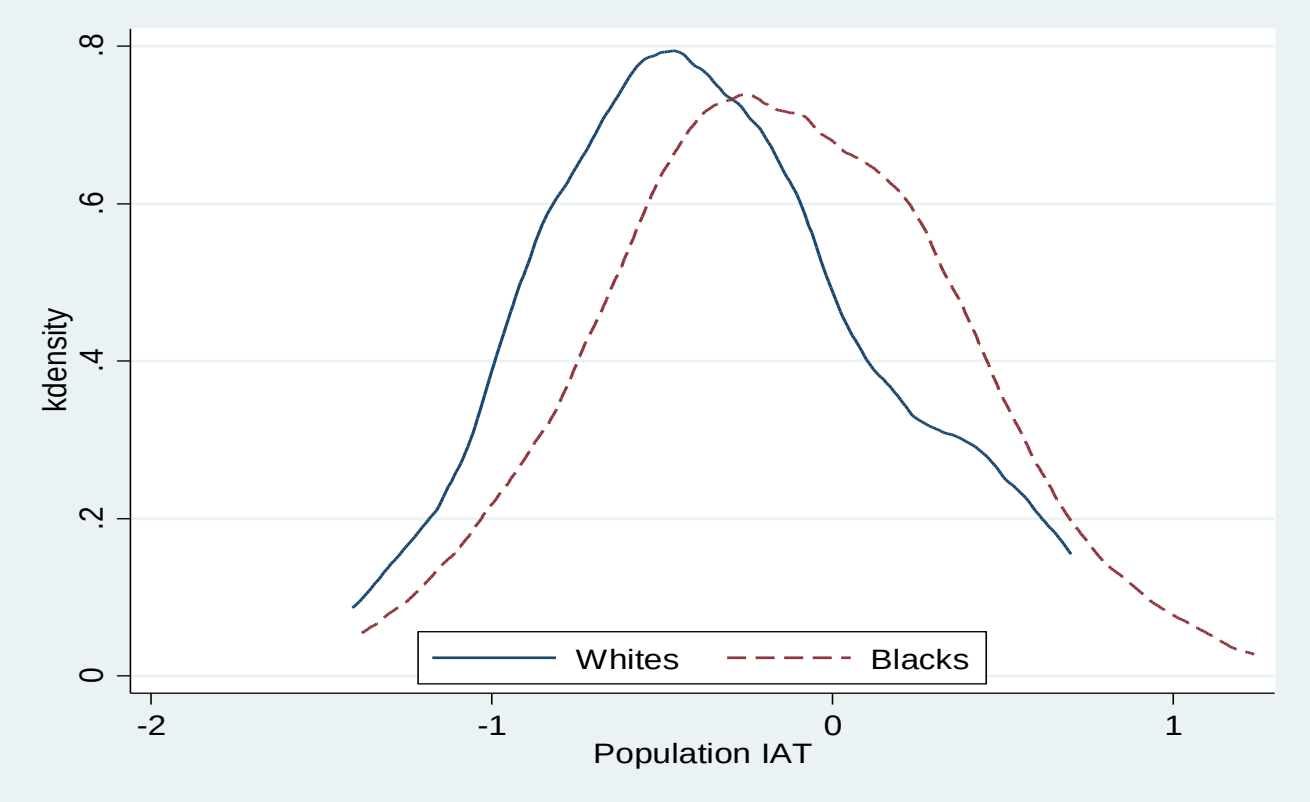
- Reveal cognitive processes of which individuals may not be aware (e.g., perception, stereotyping)
- Or may be uncomfortable disclosing (e.g., prejudice)

Limitations of IAT

- Weakly predicts discriminatory behavior (Oswald et al. 2013)
- Unstable, changes over time (Dasgupta and Greenwald 2001) – attenuation bias

Population IAT at baseline

Whites vs Blacks



Negative values: negative stereotypes on blacks vs whites

Empirical strategy

1. Effect of exposure to different race:

$$Y_{ijk1} = \alpha Y_{i0} + \beta MixRoom_{i0} + \gamma Race_i + \lambda X_{i0} + \mu X_{j0} + \delta_{k0} + \varepsilon_{ijk1}$$

$\beta > 0$: reduced prejudice against blacks

2. Heterogeneous effects by race of respondent:
   if improved attitude towards *other* group, sign of β should be *opposite* for whites & blacks
   → estimate separately

3. When dealing w/ multiple outcomes, adjust p-values for family-wise error rate (FWER), following Westfall & Young (1993)

Results I: Stereotypes

Effect of MixRoom on racial stereotypes

Dependent variable:	*Population IAT*		
Sample:	*Full Sample*	*Whites*	*Blacks*
	(1)	(2)	(3)
Mixed Room	-0.034	0.282**	-0.079
	(0.055)	(0.135)	(0.073)
Controls[(a)]	X	X	X
Roommate controls[(b)]	X	X	X
Mean of dep.var. in same race room	-0.166	-0.414	-0.084
R-squared	0.122	0.209	0.095
No. Obs.	495	116	329

Treatment closes the gap in Popul IAT b/w whites and blacks

What drives stereotypes reduction?

Testing information story

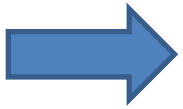
- Effect should be larger for people whose roommate is a bigger "surprise" compared to ex ante beliefs

Roommate characteristics

1. Academic ability – proxied by UCT entry score
2. Altruism/niceness – proxied by answers to survey questions on helping others, developing opportunitistic friendships,…

Respondent's priors

- Measured through IATs on positive/negative characteristics and academic performance of different groups



Build indicators of positive/negative surprise based on combinations of priors & roommate realizations

Stereotypes reduction & «surprise»

Sample:	*Full sample* (1)	*Whites* (2)	*Blacks* (3)	*Full sample* (4)	*Whites* (5)	*Blacks* (6)
Panel A: Dependent variable: Population IAT						
	Surprise defined on the basis of					
	Academic Ability			*Altruistic Attitudes*		
Mixed Room	0.001	0.259**	-0.022	-0.015	0.293**	-0.077
	(0.061)	(0.130)	(0.085)	(0.062)	(0.150)	(0.084)
Mixed Room*Positive Surprise	-0.038	0.133	-0.109	-0.114	-0.015	-0.074
	(0.160)	(0.432)	(0.201)	(0.115)	(0.299)	(0.163)
Mixed Room*Negative Surprise	-0.144	-0.067	-0.226*	0.056	-0.085	0.354***
	(0.093)	(0.306)	(0.134)	(0.092)	(0.202)	(0.127)
Mean of dep. var. in same race room	-0.166	-0.414	-0.083	-0.166	-0.414	-0.083
R-squared	0.126	0.214	0.099	0.126	0.212	0.103
No. Obs.	489	115	324	489	115	324

Omitted category: people in same race rooms whose beliefs were confirmed

Results II: Academic performance

Impact on GPA

Dependent variable:	*GPA*		
	Full sample	*Whites*	*Blacks*
	(1)	(2)	(3)
Mixed Room	0.170*	-0.119	0.258**
	(0.096)	(0.244)	(0.122)
	[0.177]	[0.921]	[0.083]*
UCT entry score	7.826***	10.001***	5.616***
	(1.085)	(2.566)	(1.321)
Roommate's UCT entry score	-0.149	1.882	-0.088
	(0.679)	(1.621)	(0.875)
Controls	X	X	X
Roommate controls	X	X	X
Academic program FE	X	X	X
Mean of dep. var. in same race room	-0.018	0.673	-0.242
R-squared	0.421	0.562	0.388
No. Obs.	487	116	325

Magnitude:
.26 std dev.

Closes 1/3 of the gap b/w blacks & whites

Impact on other academic indicators

Dependent variable:	*Number of exams passed*		
	Full sample	*Whites*	*Blacks*
	(4)	(5)	(6)
Mixed Room	0.463**	-0.364	0.715***
	(0.198)	(0.561)	(0.252)
	[0.036]**	[0.730]	[0.011]**
UCT entry score	13.239***	9.943	11.595***
	(2.270)	(6.763)	(2.940)
Roommate's UCT entry score	-0.434	0.938	-0.784
	(1.297)	(3.605)	(1.759)
Controls	X	X	X
Roommate controls	X	X	X
Academic program FE	X	X	X
Mean of dep. var. in same race room	4.944	6.481	4.444
R-squared	0.704	0.702	0.714
No. Obs.	487	116	325

Impact on other academic indicators

Dependent variable:	*Number of exams passed*			*Eligible to continue*		
	Full sample	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(4)	(5)	(6)	(7)	(8)	(9)
Mixed Room	0.463**	-0.364	0.715***	0.126***	0.033	0.169***
	(0.198)	(0.561)	(0.252)	(0.031)	(0.067)	(0.042)
	[0.036]**	[0.730]	[0.011]**	[0.000]***	[0.730]	[0.000]***
UCT entry score	13.239***	9.943	11.595***	0.611*	-0.187	0.710
	(2.270)	(6.763)	(2.940)	(0.362)	(0.773)	(0.483)
Roommate's UCT entry score	-0.434	0.938	-0.784	0.226	0.672	0.167
	(1.297)	(3.605)	(1.759)	(0.184)	(0.443)	(0.250)
Controls	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X
Academic program FE	X	X	X	X	X	X
Mean of dep. var. in same race room	4.944	6.481	4.444	0.869	0.922	0.849
R-squared	0.704	0.702	0.714	0.319	0.408	0.416
No. Obs.	487	116	325	487	116	325

Index of academic performance

Dependent variable:	*Index of Performance*		
	Full sample	*Whites*	*Blacks*
	(10)	(11)	(12)
Mixed Room	0.333***	-0.099	0.481***
	(0.110)	(0.270)	(0.145)
UCT entry score	7.846***	7.653***	6.321***
	(1.218)	(2.425)	(1.595)
Roommate's UCT entry score	0.094	2.094	-0.001
	(0.712)	(1.719)	(0.976)
Controls	X	X	X
Roommate controls	X	X	X
Academic program FE	X	X	X
Mean of dep. var. in same race room	-0.031	0.727	-0.280
R-squared	0.442	0.405	0.452
No. Obs.	487	116	325

Is the effect mediated by roommate's prejudice?

Dependent variable:	*GPA*		
Sample:	*Full sample*	*Whites*	*Blacks*
	(1)	(2)	(3)
Mixed Room	0.306**	-0.124	0.507***
	(0.127)	(0.214)	(0.191)
	[0.017]**	[0.769]	[0.009]***
Mixed Room * Roommate pop IAT baseline	0.466**	0.635	0.660**
	(0.201)	(0.442)	(0.319)
	[0.053]*	[0.254]	[0.097]*
Roommate pop IAT baseline	-0.262**	-0.234	-0.329**
	(0.116)	(0.274)	(0.162)
Controls	X	X	X
Roommate controls	X	X	X
Academic program FE	X	X	X
Mean of dep. var. in same race room	-0.008	0.701	-0.233
R-squared	0.441	0.634	0.394
No. Obs.	364	85	248

Move white roommate's IAT from -.36 to 0 → +.26 std dev GPA black

Impact on academic performance – year 2

Dependent variable:	*GPA*			*Number of exams passed*		
Sample:	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(1)	(2)	(3)	(4)	(5)	(6)
Mixed Room	-0.003	-0.091	0.041	0.771***	0.557	0.875**
	(0.052)	(0.094)	(0.076)	(0.283)	(0.823)	(0.401)
	[0.999]	[0.682]	[0.902]	[0.016]	[0.785]	[0.059]
UCT entry score	3.491***	7.348***	1.630**	3.675	-6.199	6.199
	(0.601)	(1.137)	(0.747)	(3.102)	(9.477)	(4.128)
Roommate's UCT entry score	0.456	0.663	0.756**	1.734	2.311	1.699
	(0.312)	(0.634)	(0.318)	(1.800)	(4.396)	(2.430)
Controls	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X
Academic program fixed effects	X	X	X	X	X	X
Mean of dep. var. in same race room	0.382	0.685	0.241	5.450	6.100	5.112
R-squared	0.505	0.788	0.448	0.534	0.658	0.590
No. Obs.	343	104	200	343	104	200

Impact on academic performance – year 2

Dependent variable:	*Eligible to continue*			*Index of Performance*		
Sample:	*Full sample*	*Whites*	*Blacks*	*Full sample*	*Whites*	*Blacks*
	(7)	(8)	(9)	(10)	(11)	(12)
Mixed Room	0.099**	-0.025	0.106*	0.278**	-0.045	0.365*
	(0.041)	(0.076)	(0.062)	(0.137)	(0.308)	(0.213)
	[0.027]	[0.918]	[0.137]			
UCT entry score	0.775*	0.636	0.762	6.395***	9.415***	4.392**
	(0.461)	(0.795)	(0.742)	(1.434)	(2.582)	(2.084)
Roommate's UCT entry score	0.623*	0.678	0.658	1.762*	2.223	2.201**
	(0.352)	(0.460)	(0.486)	(0.984)	(1.613)	(1.104)
Controls	X	X	X	X	X	X
Roommate controls	X	X	X	X	X	X
Academic program fixed effects	X	X	X	X	X	X
Mean of dep. var. in same race room	0.890	0.957	0.874	-0.073	0.545	-0.349
R-squared	0.264	0.519	0.378	0.448	0.657	0.475
No. Obs.	342	104	199	342	104	199

Why does black students' performance improve?

- Not because (white) roommate is more skilled: roommate's UCT score does not predict own performance
- Not (entirely) driven by people studying together (same faculty)

Other possible channels

- Role modeling: learning how to navigate the system
- Fewer opportunities for joint distractions
- Anxiety reduction: increased inter-racial friendships and dating (not conscious dancing w/ or dating a person of another race)

Results on explicit attitudes & prosocial behavior

Improvements in the following indexes:

1. Friendships
   - %friends and study mates of diff. race (actual & ideal), how often hang out w/ diff. race
2. Attitudes
   - Talk about race, affirmative action, dancing/dating other group
3. Pro-social behavior
   - Volunteer, money to charity, cooperate in prisoner's dilemma

Residential choices in year 2

- Students in mixed rooms no more likely to exit residence system or change roommate

Conclusions

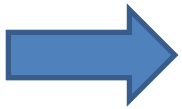
Our work brings together two disciplines

Social psychology

- Diversity, identity and stereotype formation
- *Integration policies* to change individual attitudes, reduce prejudice & inter-group conflict

Economics

- Ethnic diversity negatively correlated w/ growth, public good provision, trust, quality of institutions
- If peer effects, segregation widens disparities among groups
- *Integration policies* to reduce gaps in outcomes



Need to jointly consider attitudinal change & performance gains

II. Revealing Stereotypes. Evidence from Immigrants in Schools

A. Alesina
Harvard

M. Carlana
HKS

E. La Ferrara
Bocconi

P. Pinotti
Bocconi

Motivation

Organizations are increasingly promoting interventions to increase awareness on implicit stereotypes of their employees

Motivation

HARVARD UNIVERSITY | Office of the Senior Vice Provost
Faculty Development & Diversity

Best Practices for Conducting Faculty Searches

Harvard's faculty recruitment strategies are founded on the principle of **inclusive excellence**—maximizing excellence and diversity simultaneously. Embracing inclusive excellence is essential for keeping Harvard productive, creative, competitive, and successful in training the next generation of leaders in every field.

Each faculty search is a chance to attract the broadest talent pool. And to shape the future faculty. We can remain strong where we already are strong, and strike out in new intellectual directions as well.

We offer here the high points for practical strategies for productive and equitable searches:

- **Planning a search**
- **Recruiting a broad and deep group of candidates**
- **Evaluating candidates fairly**
- **Managing informative campus visits**

Yet we have no evidence of what would happen if (biased) people took IAT and were told their score

We strongly encourage every search committee member to take at least one IAT. (Registration is free, and the first test takes no more than 15 minutes.) The tests are not meant to challenge your conscious attitudes, but to reveal the extent to which you may nevertheless associate groups like "female" with "family" and "male" with "career." As members of this intellectual community, you will surely find it of some interest to discov that your mind contains associations of which you are unaware. You can find the IATs at https://implicit.harvard.edu.

Research questions

1. Do stereotypes lead to discrimination against immigrants?
   - We compare blindly graded tests & tests (non-blindly) graded by own teachers, during the same week
   - We find that teachers penalize immigrants relative to natives, & penalty is stronger for teachers w/ more negative stereotypes

2. If individuals become aware of their stereotypes, do they change their behavior?
   - Randomized experiment: reveal own stereotypes to teachers

Background and Data

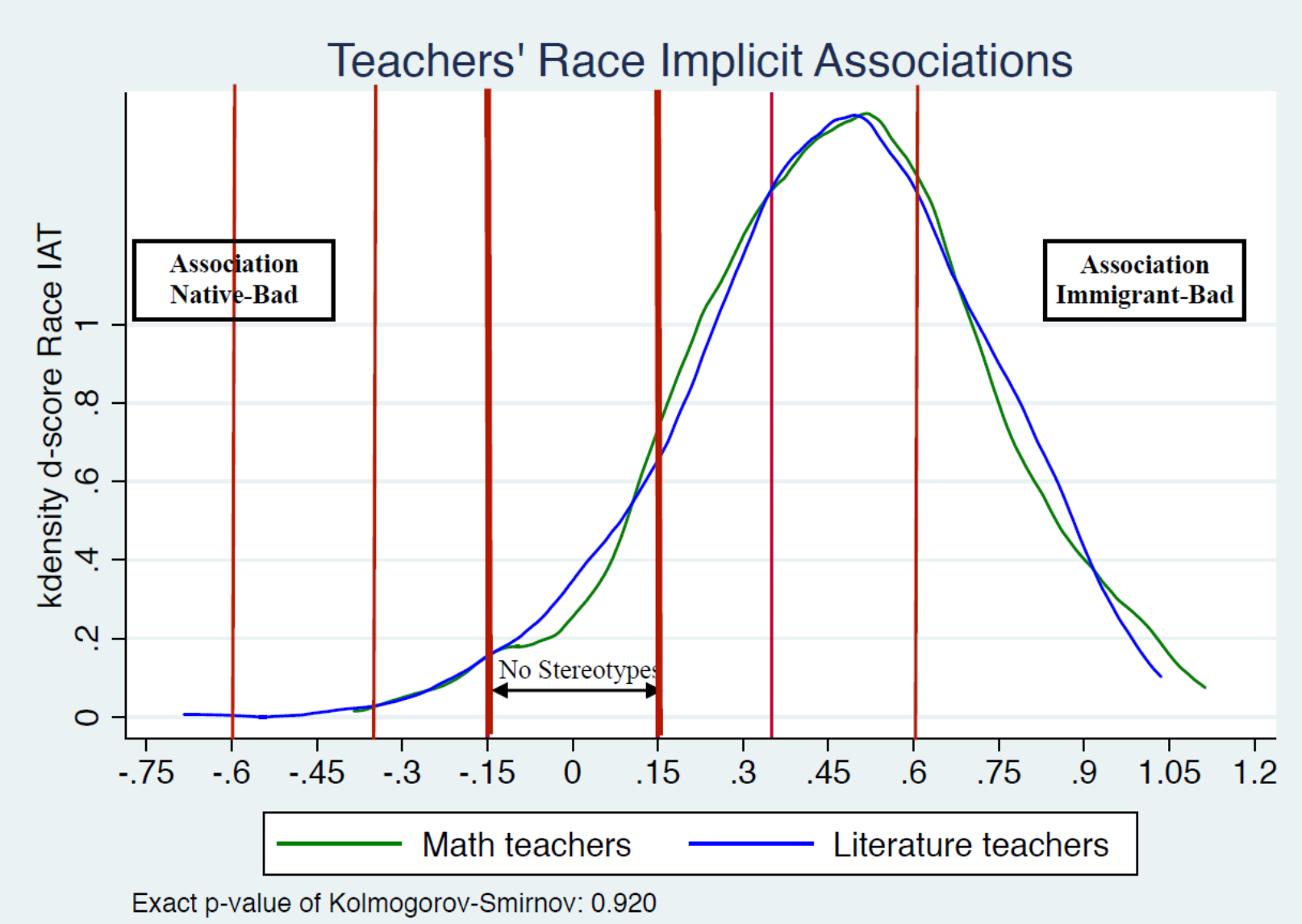
Sample & data sources

- 102 middle schools in Northern Italy
- 1.384 math and literature teachers
- Middle School: grade 6 to 8, same teachers & classmates for 3 years

- **Teacher survey**: IAT, demographic information, explicit beliefs & attitudes towards immigrants

- **Student data** from 2 administrative sources:
  - Italian Ministry of Education (MIUR): teacher-assigned grades in math and literature
  - National Evaluation Agency (INVALSI): standardized test scores in grade 8, family background

Distribution of teachers' IAT



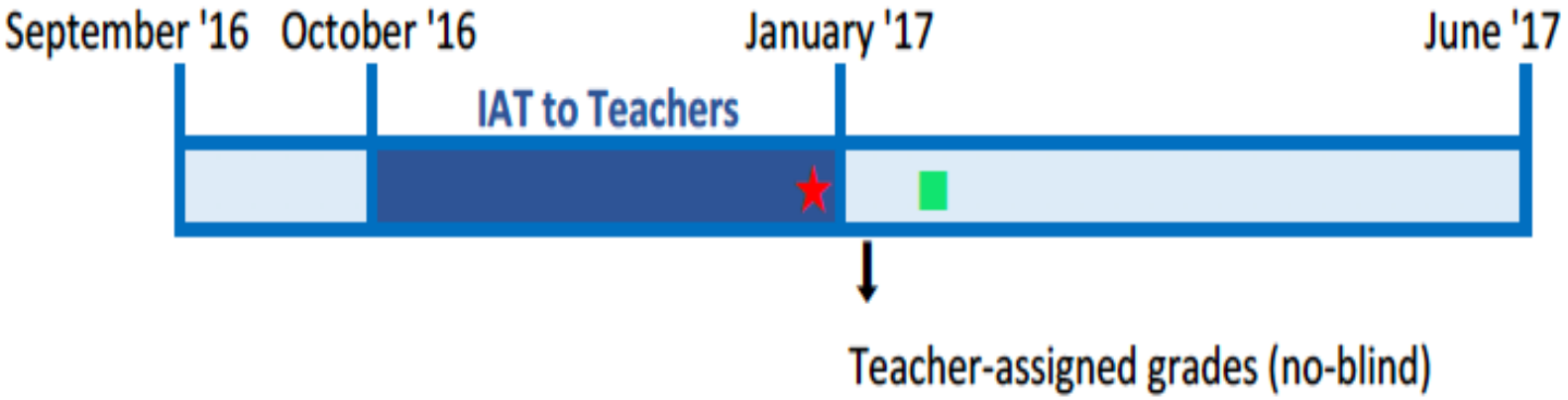
Experimental design

- We offered all teachers the possibility of receiving an email w/ feeedback on their own IAT
  - > 80% of teachers chose to get it → ITT & LATE
  - Choice to received feedback uncorrelated w/ teacher characteristics

- Text of email
  - Brief description of what IAT does
  - Placement into "slight", "moderate" or "strong", based on Greenwald et al. (2009)

Timing of experiment

- We randomized the timing of the feedback *at school level*: 2 weeks before vs. 2 weeks after end-of-semester grading



Results

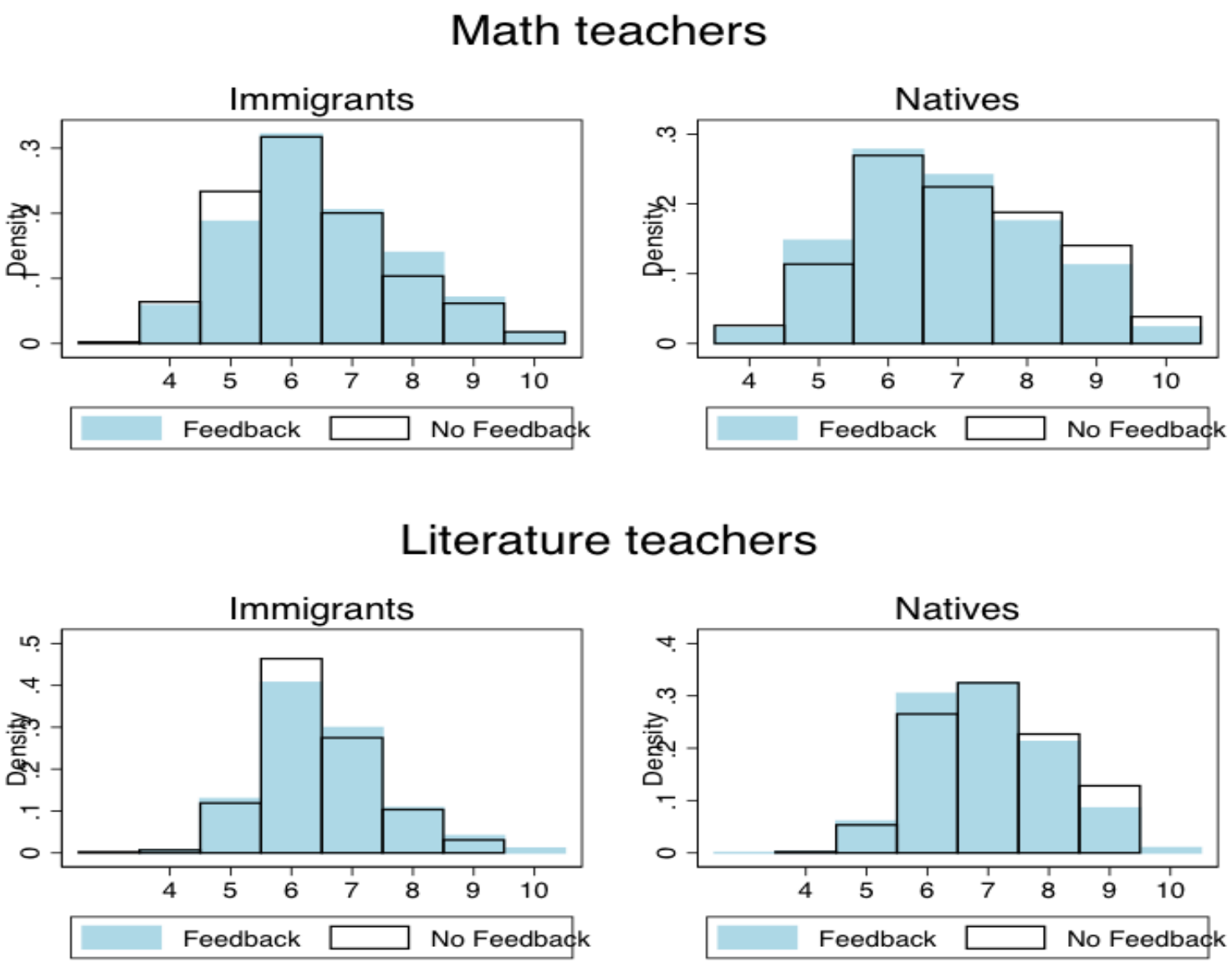
ITT effect on teacher-assigned grades

	(1)	(2)	(3)	(4)	(5)	(6)
Dep Var:	**Grades given by Math Teacher**			**Grades given by Literature Teacher**		
Panel A: Intention to Treat						
Early Feedback*Imm	0.392***	0.437***	0.439***	0.312***	0.302***	0.288***
	(0.142)	(0.129)	(0.126)	(0.103)	(0.083)	(0.087)
Early Feedback	-0.153	-0.176*	-0.155	-0.150*	-0.160**	-0.147*
	(0.100)	(0.094)	(0.095)	(0.084)	(0.072)	(0.075)
Immigrant	-0.713***	-0.687***	0.915	-0.697***	-0.685***	-0.107
	(0.091)	(0.222)	(1.274)	(0.055)	(0.131)	(1.361)
Obs.	5141	5141	5141	5138	5138	5138
R^2	0.023	0.108	0.118	0.037	0.167	0.174
Mean dep. var.	6.83	6.83	6.83	6.95	6.95	6.95
Student Controls	No	Yes	Yes	No	Yes	Yes
Student Controls*Imm	No	Yes	Yes	No	Yes	Yes
Teacher Controls	No	No	Yes	No	No	Yes
Teacher Controls*Imm	No	No	Yes	No	No	Yes

LATE on teacher-assigned grades

	(1)	(2)	(3)	(4)	(5)	(6)
Dep Var:	**Grades given by Math Teacher**			**Grades given by Literature Teacher**		
Panel B: Local Average Treatment Effect						
Email*Imm	0.501***	0.552***	0.554***	0.403***	0.392***	0.366***
	(0.171)	(0.161)	(0.156)	(0.132)	(0.112)	(0.114)
Email	-0.206	-0.234*	-0.208	-0.202*	-0.214**	-0.194*
	(0.131)	(0.127)	(0.128)	(0.111)	(0.099)	(0.101)
Immigrant	-0.713***	-0.659***	0.998	-0.697***	-0.624***	-0.306
	(0.090)	(0.226)	(1.252)	(0.054)	(0.142)	(1.408)
Obs.	5141	5141	5141	5138	5138	5138
R^2	0.022	0.105	0.116	0.035	0.161	0.169
F- stat	84.5	94.2	101.1	106.8	125.9	138.8
Mean dep. var.	6.83	6.83	6.83	6.95	6.95	6.95
Student Controls	No	Yes	Yes	No	Yes	Yes
Student Controls*Imm	No	Yes	Yes	No	Yes	Yes
Teacher Controls	No	No	Yes	No	No	Yes
Teacher Controls*Imm	No	No	Yes	No	No	Yes

ITT – where in the distribution?



Testing information channel

Do teachers respond b/c they have learnt something new?

2 strategies:

a. Awareness / explicit bias: our treatment should convey new info, unless teacher is already aware of his/her bias

b. Precision: updating in response to our treatment should depend on precision of the signal

a. Impact by explicit bias

We use 2 proxies to measure if teacher is aware of his/her bias

1. WVS question on whether "Immigrants should have same right to jobs as natives"
   - Those who answer "No" are explicitly biased → our message should come as less of a surprise
   - We expect smaller (or no) effect on this sub-sample

2. "Prejudice": dummy=1 if teacher said that it is "Likely" or "Extremely likely" that immigrants disproportionately attend vocational track b/c of prejudice in school or workplaces

Heterogeneous effects by explicit bias

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	**Grades given by Math Teacher**			**Grades given by Literature Teacher**		
Early Feedback*Imm	0.439***	0.051		0.288***	-0.200	
	(0.126)	(0.236)		(0.087)	(0.185)	
Immigrant	0.915	0.727		-0.107	-0.037	
	(1.274)	(1.238)		(1.361)	(1.234)	
Early Feedback	-0.155	-0.131		-0.147 *	0.080	
	(0.095)	(0.180)		(0.075)	(0.159)	
Early Feedback*Imm*WVS		0.545 *			0.663 ***	
		(0.293)			(0.187)	
Early Feedback*WVS		-0.138			-0.288*	
		(0.225)			(0.168)	
Early Feedback*Imm*Prejudice						
Early Feedback*Prejudice						
Obs.	5141	5141		5138	5138	
R^2	0.118	0.120		0.174	0.179	
Student Controls	Yes	Yes		Yes	Yes	
Student Controls*Imm	Yes	Yes		Yes	Yes	
Teacher Controls	Yes	Yes		Yes	Yes	
Teacher Controls*Imm	Yes	Yes		Yes	Yes	

Heterogeneous effects by explicit bias

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	**Grades given by Math Teacher**			**Grades given by Literature Teacher**		
Early Feedback*Imm	0.439***	0.051	0.296***	0.288***	-0.200	0.229 **
	(0.126)	(0.236)	(0.126)	(0.087)	(0.185)	(0.104)
Immigrant	0.915	0.727	1.050	-0.107	-0.037	-0.310
	(1.274)	(1.238)	(1.412)	(1.361)	(1.234)	(1.372)
Early Feedback	-0.155	-0.131	-0.144	-0.147 *	0.080	-0.068
	(0.095)	(0.180)	(0.112)	(0.075)	(0.159)	(0.080)
Early Feedback*Imm*WVS		0.545*			0.663 ***	
		(0.293)			(0.187)	
Early Feedback*WVS		-0.138			-0.288*	
		(0.225)			(0.168)	
Early Feedback*Imm*Prejudice			0.491**			0.249
			(0.259)			(0.217)
Early Feedback*Prejudice			-0.091			-0.195
			(0.166)			(0.127)
Obs.	5141	5141	5141	5138	5138	5138
R^2	0.118	0.120	0.121	0.174	0.179	0.179
Student Controls	Yes	Yes	Yes	Yes	Yes	Yes
Student Controls*Imm	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Controls	Yes	Yes	Yes	Yes	Yes	Yes
Teacher Controls*Imm	Yes	Yes	Yes	Yes	Yes	Yes

b. Impact by precision of the signal

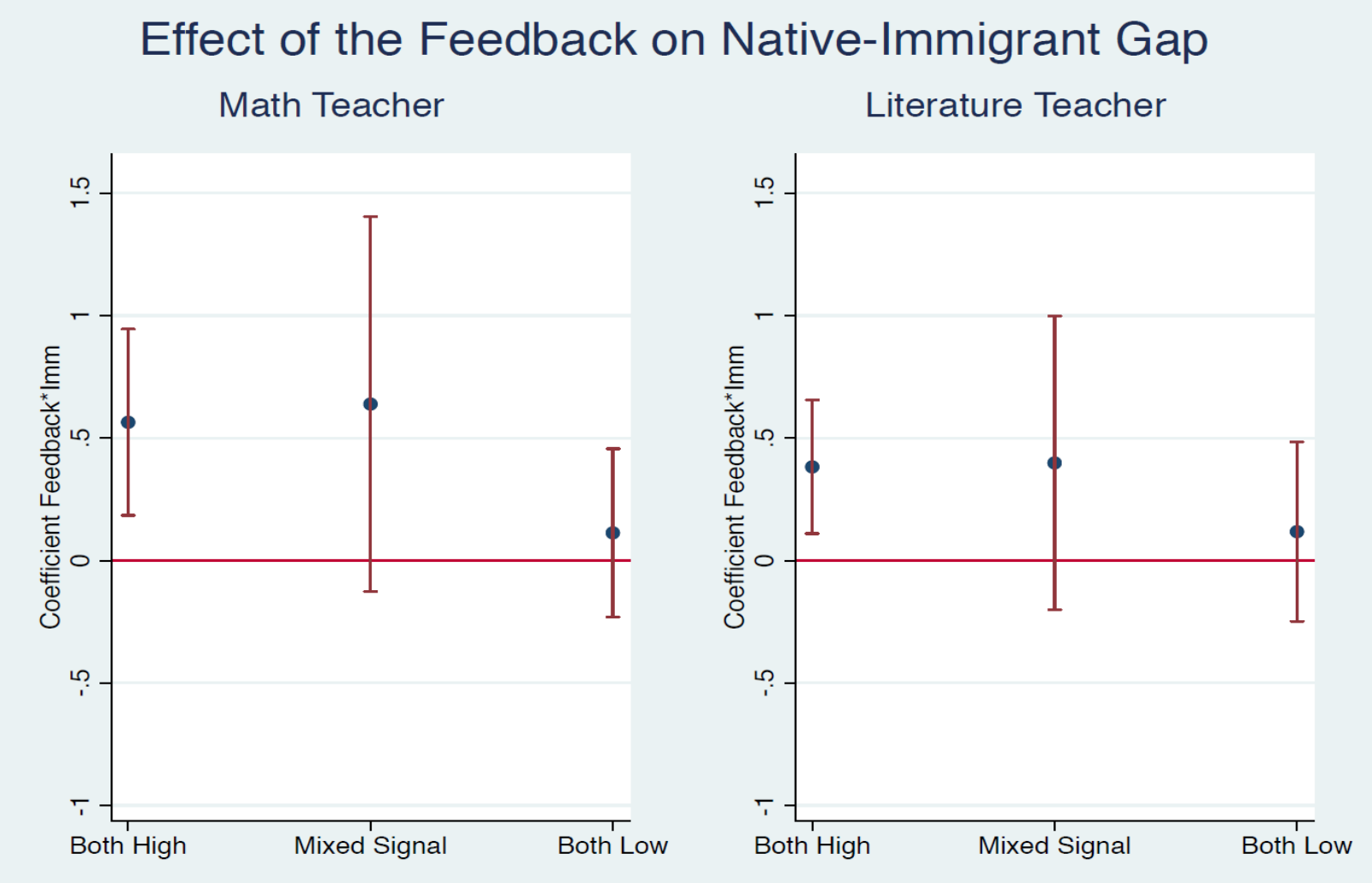
Each teacher received 2 feedbacks for native-immigrant IAT: **male** names, **female** names

- Positively correlated: 50% teachers have a moderate/severe bias in both, 20% slight/no bias in both
- But 30% teachers received one feedback as moderate/severe and one as slight/no bias

- Expect stronger impact on those who received high/high, weaker on those that received mixed signal, no impact on low/low

Impact by precision of the signal

Coeff. on “Early feedback*Immigrant”



Conclusions

- Interventions aimed at increasing awareness of implicit racial stereotypes can help counteract discrimination (e.g., committee members taking IAT)

- Caveat – possible over-reaction:

  In our sample literature teachers displayed less bias in grading (as measured by correlation b/w grade gap & teacher's IAT), yet they also adjust grading upwards in response to experiment

- More work needed to quantify extent of response & extend to different context (e.g., hiring or judicial decisions)

Thank you!