Systemic Discrimination Among Large U.S. Employers

Patrick Kline, Evan K. Rose, and Christopher Walters

May, 2022

Who discriminates?

- Title VII of Civil Rights Act of 1964: illegal to discriminate on the basis of race, sex, color, religion, and national origin
- Large literature uses correspondence studies to measure market-average discrimination against these protected characteristics (Bertrand and Duflo, 2017; Baert, 2018; Quillian e.a., 2017)
- Limited empirical evidence on whether disparate treatment is concentrated in particular companies (e.g., Bertrand and Mullainathan, 2004; Agan and Starr, 2018, 2020)
- To what extent is discrimination "endemic" to particular firms?

Systemic discrimination

The Equal Employment Opportunity Commission (EEOC) is also interested in "systemic" discrimination in particular firms, which they define as:

> *A pattern or practice, policy and/or class cases where the discrimination has a broad impact on an industry, profession, company or geographic location.*

Obama admin EEOC chair Jenny Yang (2016):

> *Tackling systemic discrimination---where a discriminatory pattern or practice or policy has a broad impact on an industry, company or geographic area---is central to the mission of EEOC.*

FY2020: 538 "systemic" investigations, mostly focused on firms. Nearly 1,000 additional compliance evaluations of fed contractors by Office of Federal Contract Compliance (OFCCP).

Today

New correspondence experiment designed to measure bias by large U.S. employers

- Targeted design: sample entry-level jobs from 100+ Fortune 500 firms
- Apply to as many as 125 *geographically distinct* jobs from each firm
- 8 applications to each job
- Sample size: 84,000 applications (20x Bertrand and Mullainathan, 2004)
- Experiment organized in 5 waves spanning the COVID pandemic

Design allows us to test whether firms exhibit nationwide patterns of discrimination

Goals: Measurement and detection

Characterize firm component of discrimination

- Variance decompositions quantifying heterogeneity across firms
- Contrast with industry, state, and job title
- Correlates of discrimination
- Distributional estimates

Assess prospects for detecting discrimination by particular employers

- Empirical Bayes posterior estimates for individual firms
- Control over false discoveries

Related literature

- Audit and correspondence experiments for measuring racial discrimination (Daniel, 1968; Wienk et al., 1979; Heckman and Siegelman, 1993; Heckman, 1998; Bertrand and Mullainathan, 2004; Pager et al., 2009; Nunley et al., 2015; Bertrand and Duflo, 2017; Quillian et al, 2017; Baert, 2018; Gaddis, 2018; Neumark, 2018)
- Other characteristics: effects of sex, age, LGBTQ, national origin, criminal record, unemployment, education (Pager, 2003; Oreopolous, 2011; Tilcsik, 2011; Kroft et al., 2013; Arceo-Gomez and Campos-Vasquez, 2014; Deming et al., 2016; Farber et al., 2016; Agan and Starr, 2018, 2020; Neumark et al., 2019; Pedulla, 2020)
- Differences across firms / industries / geography (Bertrand and Mullainathan, 2004; Rooth, 2007; Charles and Guryan, 2008; Pager, 2016; Banerjee et al., 2018; Agan and Starr, 2018, 2020; Christensen et al., 2020)
- Detection of unit-level biases (Glover et al., 2017; Chan et al., 2019; Kline and Walters, 2021; Avivi et al., 2021; Goncalves and Mello, 2021)
- Empirical Bayes / false discovery rates (Benjamini and Hochberg, 1995; Efron et al., 2001; Storey, 2002; Armstrong, 2015; Efron, 2016; Gu and Koenker, 2020; Basu et al., 2021)

Experimental design

Sampling frame (I/II)

Holding companies split into brands with separate hiring portals (e.g., Berkshire Hathaway into Geico, McLane, Fruit of the Loom, etc.)	Fortune 500
InfoGroup and Burning Glass data merged to measure geographic distribution of establishments and vacancies	123 firms with sufficient expected geographic scope
Hiring platforms investigated to test for feasibility of submitting fictitious applications	108 feasible to audit

Sampling frame (I/II)

Holding companies split into brands with separate hiring portals (e.g., Berkshire Hathaway into Geico, McLane, Fruit of the Loom, etc.)

Fortune 500

InfoGroup and Burning Glass data merged to measure geographic distribution of establishments and vacancies

123 firms with sufficient expected geographic scope

Hiring platforms investigated to test for feasibility of submitting fictitious applications

108 feasible to audit

Compustat: U.S. employment at 108 sampled firms totaled ∼**15M** in 2020

Sampling frame (II/II)

4 not sampled in wave 1 due to COVID interruption; 9 firms dropped before completion due to technological constraints; 19 added in wave 2 or later; 4 posted insufficient jobs to sample in all waves

72 sampled in all waves

36 sampled in subset of waves

Job sampled from universe of entry-level vacancies posted on each firm's hiring portal; most recently posted job prioritized

25 vacancies in distinct counties sampled each wave

One pair of applications (1 black and 1 white name) sent every 1-2 days; gender (50% male), age (uniform age 20-60), gender identity (5% gender-neutral, 5% same-gender pronouns), and sexual orientation (10% LGBTQ student club, 10% other club) unconditionally randomly assigned

8 applications sent to each vacancy

Resume characteristics

Job applications manipulate employer perceptions of several protected characteristics:

- Race & gender: distinctive first names obtained from Bertrand and Mullainathan (2004) + NC data on speeding tickets. Last names from Census
- Age: year of high school graduation

Stratify on race (4B/4W), unconditional random assignment of gender, age, as well as LGBTQ affiliation and gender identity

Random assignment of job-appropriate experience, high school, associate degree, resume design, answers to personality tests, etc.

Fully automated sampling of vacancies and submission of apps

Example resumes

Joshua Erickson

J|E

Preferred Pronouns: They | Them | Theirs

(224)-478-1806
joshuaerickson9@gmail.com
124 Carol Louise Dr
Caseyville, IL

Education History

Young Magnet High School | Chicago, IL
1990 to 1994

Previous Employment

Retail Associate | O Fallon, IL
11/2019 to Present
Good Feet Store

1. *Performed visual merchandising in sales areas.*
2. *Wrote up inventory logs daily.*

Host | Marine, IL
10/2018 to 11/2019
Phyl's Chet Rose's Tavern

1. *Communicated efficiently with all restaurant staff.*
2. *Monitored guests needs and workflow of the restaurant seating customers accordingly.*
3. *Recognized for hard work, dedication, dependability, prompt and reliable attendance, and willingness to work overtime as needed.*

Professional References

Juliet Romero: Previous supervisor at Good Feet Store
Cassandra Edwards: Previous supervisor at Phyl's Chet Rose's Tavern

---

MR

Maurice Randle

(781)-790-4717
3620 232Nd St
Bothell, WA
mrandle667@verizonmail.me

Previous Employment

Retail Associate | Seattle, WA 9/2018 to Present
Oiselle Running
Reference | Salvador Porter | (206) 160-2193
I. Received, unpacked, tagged, and issued sales floor merchandise.
II. Participated in year-end inventory and cycle counts.
III. Served as a consultant to help customer make the right selection.

Cashier | Bellevue, WA 1/2017 to 9/2018
Crossroads Farmers Market
Reference | Ezequiel Stephens | (425) 885-1919
I. Operated registers, scanners, scales and credit card/debit card terminals.
II. Served customers with a friendly demeanor and positive attitude.
III. Maintained clean and orderly checkout areas and completed other general cleaning duties, such as mopping floors and emptying trash cans.

Cashier | Redmond, WA 7/2015 to 1/2017
Redmond Marriott Town Center
Reference | Kayley Gonzalez | (206) 538-2874
I. Used coupons effectively & discounts.
II. Other responsibilities included scanning items, processing payments, applying coupons, providing change.
III. Operated scanners, scales, cash registers, and other electronics on a daily basis.

Education History

Everett Community College 1995 to 1997
Everett, WA
Associates | Marketing

Naches Valley High School 1991 to 1995
Naches, WA
General Studies

Skills

Communication
Prioritizing tasks
Highly detail oriented

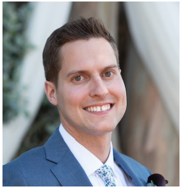
Team *RandRes*

Hadar Avivi

Ross Chu

Ben Scuderi

Jake Anderson



Brian Collica

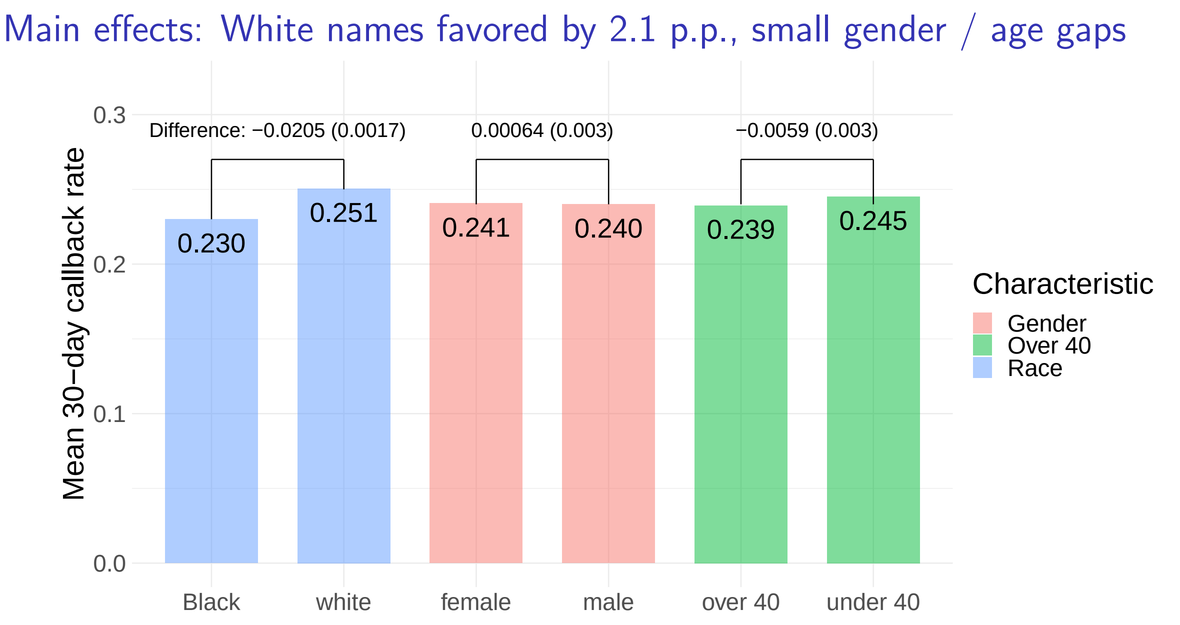
Nicole Gandre

The Randres Corps: May Adberg, Jason Chen, Stephanie Cong, Simon Duabis, Daniel Dychala, Samuel Gao, Alexandra Groscost, Victoria Haworth, Camille Hillion, Ben Keltner, Mary Kruberg, Jiaxin Lei, Carol Lee, Collin Lu, Oliver McNeil, Riley Odom, Sarah Phung, Eric Phillips, Stephanie Ross, Marcus Sander, Pat Tagari, Quinghuai Tan, Lydia Wen, Zijun Xu, Xilin Ying, Andy Zhong, Leila Zhua, Yingjia Zhang, and Yiran Zi

A first look at the data

Summary stats

	A. All firms			B. Balanced sample		
	White	Black	Difference	White	Black	Difference
Resume characteristics						
Female	0.499	0.499	-0.001	0.500	0.498	0.003
Over 40	0.535	0.535	0.000	0.534	0.533	0.002
LGBTQ club member	0.081	0.082	-0.001	0.079	0.080	-0.001
Academic club	0.040	0.042	-0.002	0.039	0.042	-0.003*
Political club	0.042	0.042	0.001	0.042	0.041	0.001
Gender-neutral pronouns	0.041	0.041	-0.001	0.040	0.040	0.000
Same-gender pronouns	0.043	0.042	0.001	0.042	0.041	0.001
Associate degree	0.476	0.485	-0.009**	0.478	0.485	-0.006*
N applications	41837	41806	83643	32703	32665	65368
N jobs			11114			8667
N firms			108			72
1/2/3/4/5 waves			3/4/15/16/72			


Main effects: White names favored by 2.1 p.p., small gender / age gaps
Mean 30−day callback rate
0.3
0.2
0.1
0.0
Difference: −0.0205 (0.0017)
0.00064 (0.003)
−0.0059 (0.003)
0.230
0.251
0.241
0.240
0.239
0.245
Black
white
female
male
over 40
under 40
Characteristic
Gender
Over 40
Race

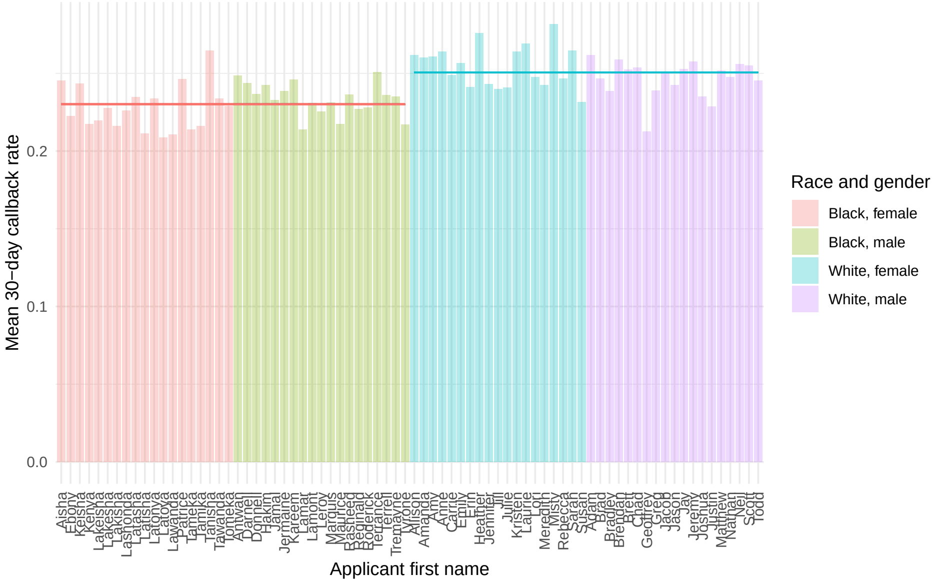
Interactions

	OLS			Logit		
	(1)	(2)	(3)	(4)	(5)	(6)
	White	Black	Difference	White	Black	Difference
Female	0.00716*	-0.00694*	0.0141**	0.0388*	-0.0398*	0.0786**
	(0.00423)	(0.00412)	(0.00579)	(0.0229)	(0.0236)	(0.0322)
Over 40	-0.0104**	-0.00125	-0.00915	-0.0562**	-0.00711	-0.0491
	(0.00428)	(0.00413)	(0.00590)	(0.0231)	(0.0236)	(0.0328)
Political club	-0.00207	-0.00229	0.000220	-0.0109	-0.0126	0.00171
	(0.0107)	(0.0105)	(0.0150)	(0.0562)	(0.0587)	(0.0815)
Academic club	0.00341	0.0147	-0.0113	0.0173	0.0806	-0.0633
	(0.0111)	(0.0107)	(0.0155)	(0.0576)	(0.0574)	(0.0817)
LGBTQ club	-0.0165**	0.00631	-0.0228**	-0.0889**	0.0349	-0.124**
	(0.00787)	(0.00763)	(0.0110)	(0.0431)	(0.0419)	(0.0601)
Same-gender pronouns	-0.00971	-0.0165	0.00681	-0.0515	-0.0934	0.0420
	(0.0106)	(0.0101)	(0.0146)	(0.0571)	(0.0587)	(0.0816)
Gender-neutral pronouns	-0.0106	-0.0103	-0.000279	-0.0564	-0.0578	0.00138
	(0.0108)	(0.0105)	(0.0150)	(0.0581)	(0.0598)	(0.0830)
Associate degree	0.00573	-0.00152	0.00724	0.0309	-0.00869	0.0396
	(0.00431)	(0.00412)	(0.00584)	(0.0233)	(0.0236)	(0.0325)
Constant	0.201***	0.185***	0.0160***	-1.377***	-1.485***	0.108***
	(0.00848)	(0.00820)	(0.00621)	(0.0514)	(0.0538)	(0.0366)
N	41837	41806	83643	41837	41806	83643
χ^2 stat for joint significance			14.71			14.54
p-value			0.0650			0.0687

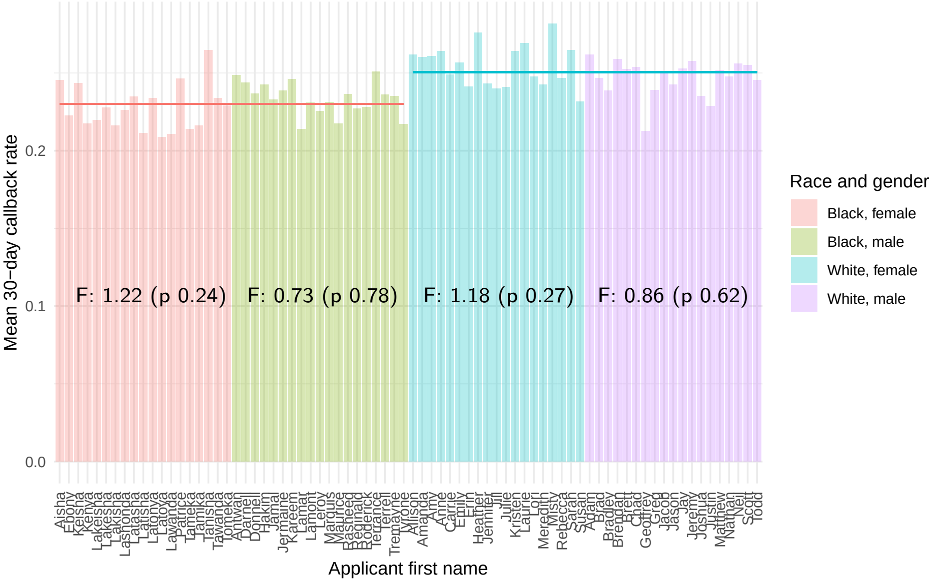
Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Insignificant name effects within race / gender cell



Insignificant name effects within race / gender cell



Firm, state, and industry variation

Defining terms

Contact gap at job j of firm f is Δ_{fj}

- e.g., for race, Δ_{fj} is white contact rate - Black contact rate

Firm mean contact gap is $\mathbb{E}[\Delta_{fj}] = \Delta_f$

- Measures expected contact gap at randomly sampled job from firm f

Defining terms

Contact gap at job j of firm f is Δ_{fj}

- e.g., for race, Δ_{fj} is white contact rate - Black contact rate

Firm mean contact gap is $\mathbb{E}[\Delta_{fj}] = \Delta_f$

- Measures expected contact gap at randomly sampled job from firm f

Random sampling of jobs + random assignment + SUTVA imply:

$$\mathbb{E}\left[\hat{\Delta}_{fj}|\Delta_{fj}\right] = \Delta_{fj}, \quad \mathbb{E}\left[\hat{\Delta}_{fj}\right] = \Delta_f .$$

Does Δ_f differ between firms?

	(1) χ^2 test of heterogeneity	(2) p-value for no discrim against:	Contact gap SD (3) Bias-corrected	(4) Cross-wave	(5) Cross-state
Race	276.5	W: 1.00	0.0185	0.0168	0.0178
	[0.000]	B: 0.00	(0.0031)	(0.0032)	(0.0031)
Gender	205.2	M: 0.00	0.0267	0.0287	0.0269
	[0.000]	F: 0.05	(0.0038)	(0.0035)	(0.0038)
Over 40	144.6	Y: 0.22	0.0103	0.0044	0.0086
	[0.011]	O: 0.02	(0.0069)	(0.0158)	(0.0082)

Classic χ^2 test for whether all Δ_f are equal

Do all firms discriminate in same direction?

	(1) χ^2 test of heterogeneity	(2) p-value for no discrim against:	Contact gap SD (3) Bias-corrected	(4) Cross-wave	(5) Cross-state
Race	276.5	W: 1.00	0.0185	0.0168	0.0178
	[0.000]	B: 0.00	(0.0031)	(0.0032)	(0.0031)
Gender	205.2	M: 0.00	0.0267	0.0287	0.0269
	[0.000]	F: 0.05	(0.0038)	(0.0035)	(0.0038)
Over 40	144.6	Y: 0.22	0.0103	0.0044	0.0086
	[0.011]	O: 0.02	(0.0069)	(0.0158)	(0.0082)

Test if all Δ_f have same sign, implying common direction of discrimination (Bai, Santos, and Shaikh, 2021)

Substantial variation in discrimination across firms

	(1) χ^2 test of heterogeneity	(2) p-value for no discrim against:	Contact gap SD (3) Bias-corrected	(4) Cross-wave	(5) Cross-state
Race	276.5	W: 1.00	0.0185	0.0168	0.0178
	[0.000]	B: 0.00	(0.0031)	(0.0032)	(0.0031)
Gender	205.2	M: 0.00	0.0267	0.0287	0.0269
	[0.000]	F: 0.05	(0.0038)	(0.0035)	(0.0038)
Over 40	144.6	Y: 0.22	0.0103	0.0044	0.0086
	[0.011]	O: 0.02	(0.0069)	(0.0158)	(0.0082)

Estimate of standard deviation of Δ_f, correcting for sampling variance with standard errors (e.g., Krueger and Summers, 1998; Aaronson et al., 2007, Kline, Saggio, Sølvsten, 2020)

Substantial variation in discrimination across firms

			Contact gap SD		
	(1) χ^2 test of heterogeneity	(2) *p*-value for no discrim against:	(3) Bias-corrected	(4) Cross-wave	(5) Cross-state
Race	276.5	W: 1.00	0.0185	0.0168	0.0178
	[0.000]	B: 0.00	(0.0031)	(0.0032)	(0.0031)
Gender	205.2	M: 0.00	0.0267	0.0287	0.0269
	[0.000]	F: 0.05	(0.0038)	(0.0035)	(0.0038)
Over 40	144.6	Y: 0.22	0.0103	0.0044	0.0086
	[0.011]	O: 0.02	(0.0069)	(0.0158)	(0.0082)

Generalize using covariance between wave- and state-specific gaps within firm

Geography less important than firm

	(1) Race	(2) Gender	(3) Over 40
State	0.0076	-	-
	(0.0034)		
	[0.038]	[0.668]	[0.583]
Industry	0.0141	0.0190	0.0048
	(0.0021)	(0.0029)	(0.0053)
	[0.000]	[0.000]	[0.112]
Job title SOC3 code	0.0136	0.0111	0.0034
	(0.0025)	(0.0043)	(0.0105)
	[0.000]	[0.007]	[0.527]
Hiring platform intermediary	0.0059	0.0024	0.0024
	(0.0025)	(0.0088)	(0.0071)
	[0.008]	[0.049]	[0.212]

Cross-state variability in race effects ≈25% of that across firms

Gender and age insignificant

At least half of each firm component explained by industry

	(1) Race	(2) Gender	(3) Over 40
State	0.0076	-	-
	(0.0034)		
	[0.038]	[0.668]	[0.583]
Industry	0.0141	0.0190	0.0048
	(0.0021)	(0.0029)	(0.0053)
	[0.000]	[0.000]	[0.112]
Job title SOC3 code	0.0136	0.0111	0.0034
	(0.0025)	(0.0043)	(0.0105)
	[0.000]	[0.007]	[0.527]
Hiring platform intermediary	0.0059	0.0024	0.0024
	(0.0025)	(0.0088)	(0.0071)
	[0.008]	[0.049]	[0.212]

Industry explains **58%** of firm race gaps and 51% of firm gender gaps

"Bi-directional" discrimination against both men and women across industries

Job titles important, but not conditional on firm

	(1) Race	(2) Gender	(3) Over 40
State	0.0076	-	-
	(0.0034)		
	[0.038]	[0.668]	[0.583]
Industry	0.0141	0.0190	0.0048
	(0.0021)	(0.0029)	(0.0053)
	[0.000]	[0.000]	[0.112]
Job title SOC3 code	0.0136	0.0111	0.0034
	(0.0025)	(0.0043)	(0.0105)
	[0.000]	[0.007]	[0.527]
Hiring platform intermediary	0.0059	0.0024	0.0024
	(0.0025)	(0.0088)	(0.0071)
	[0.008]	[0.049]	[0.212]

Large job title variance for race, but jointly insignificant in two-way model controlling for firm dummies

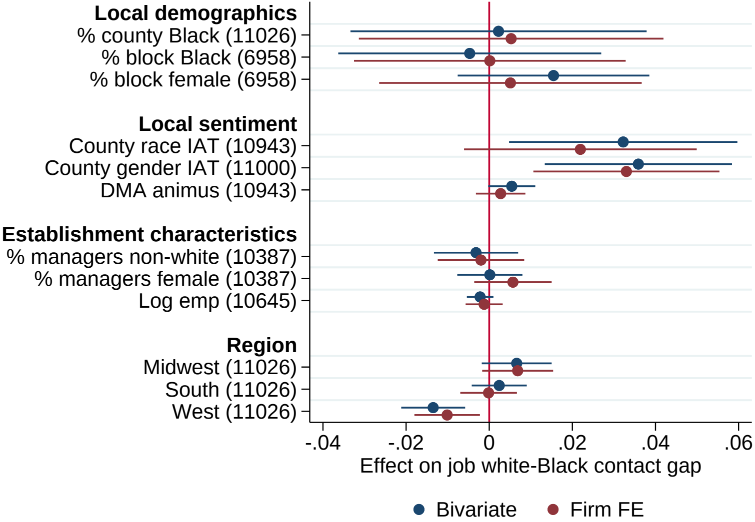
Gender job title variation smaller and also explained by firm

Job title and state insignificant conditional on firm FE

	Race		Gender		Over 40	
	State	Job title	State	Job title	State	Job title
SD firm effects	0.0176	0.0150	0.0253	0.0255	0.0096	0.0088
SD job title / state effects	0.0003	-	-	0.0080	0.0004	-
Covariance	0.0000	0.0001	0.0000	0.0002	0.0000	0.0002
N jobs	11026	11026	10720	10720	10652	10652
N firms	108	108	108	108	108	108
N job titles / states	51	47	51	47	51	47
N job titles / states $>$ 1 firm	51	43	51	43	51	43
Mean gap	0.0196	0.0196	0.0023	0.0023	0.0037	0.0037
p-value firm effects	0.000	0.0008	0.000	0.000	0.071	0.040
p-value job title / state effects	0.186	0.327	0.482	0.237	0.86	0.459

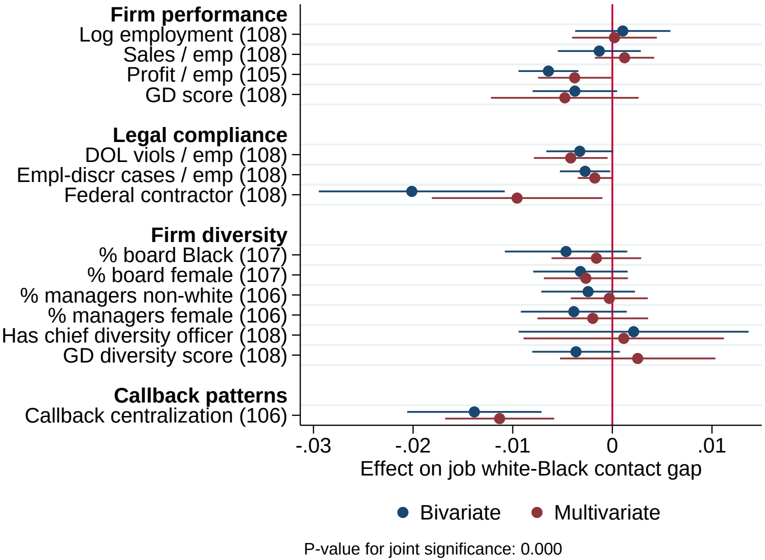
Correlates of discrimination

Best establishment level predictors are local sentiment but signal is weak

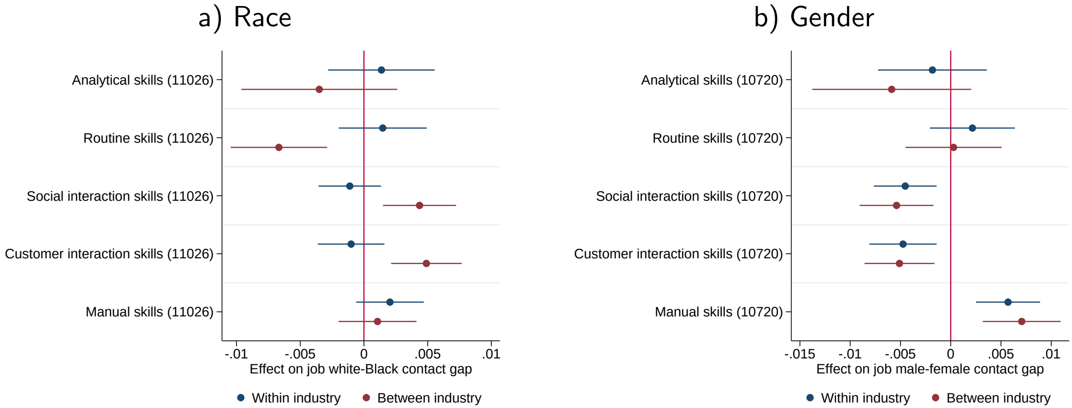


P-value for joint sig w/o firm FE: 0.03, w/ firm fe: 0.34

Smaller gaps at profitable firms, fed contractors, and centralized firms



Contact gaps concentrated in customer facing sectors



The distribution of discrimination

Beyond variances: the distribution of discrimination

Investigate other features of the distribution of Δ_f using hierarchical model:

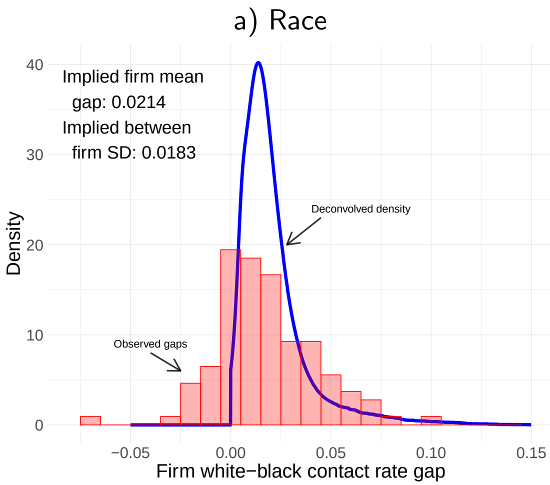
$$\hat{\Delta}_f | \Delta_f, s_f \sim N\left(\Delta_f, s_f^2\right)$$

$$\Delta_f | s_f \sim G$$

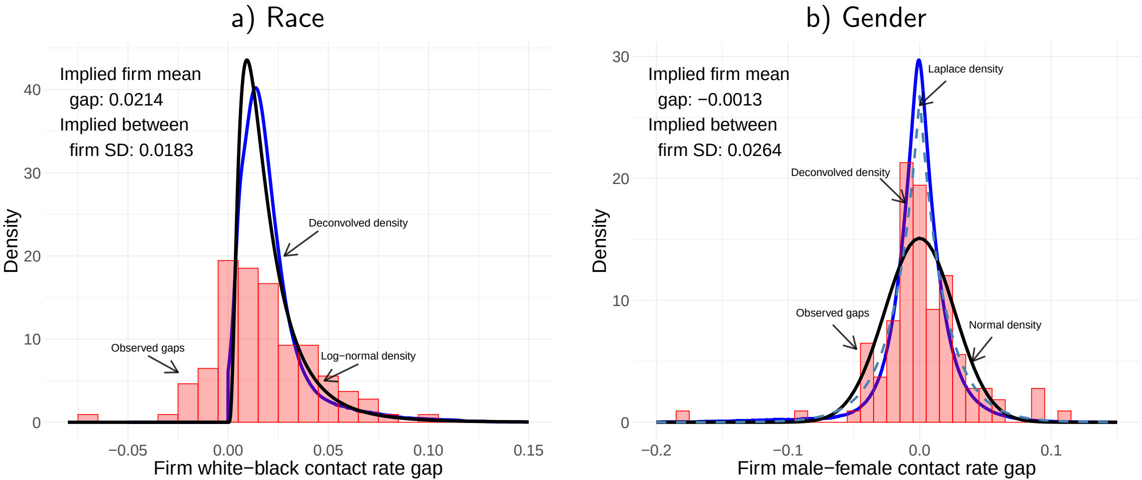
Apply Efron (2016) Empirical Bayes (EB) deconvolution estimator to extract underlying distribution G from noisy estimates $\hat{\Delta}_f$

- Shape constraint: impose no discrimination against whites
- Choose regularization to match bias-corrected variance estimate

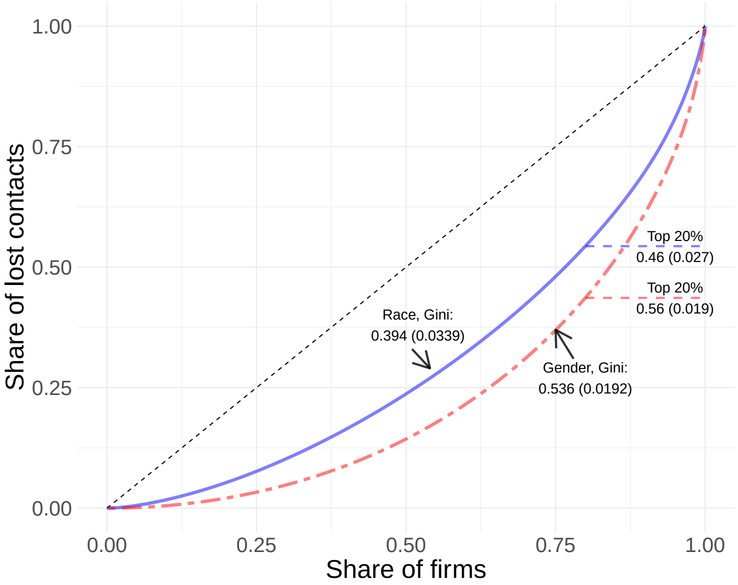
Discrimination deconvolved



Comparison to parametric families

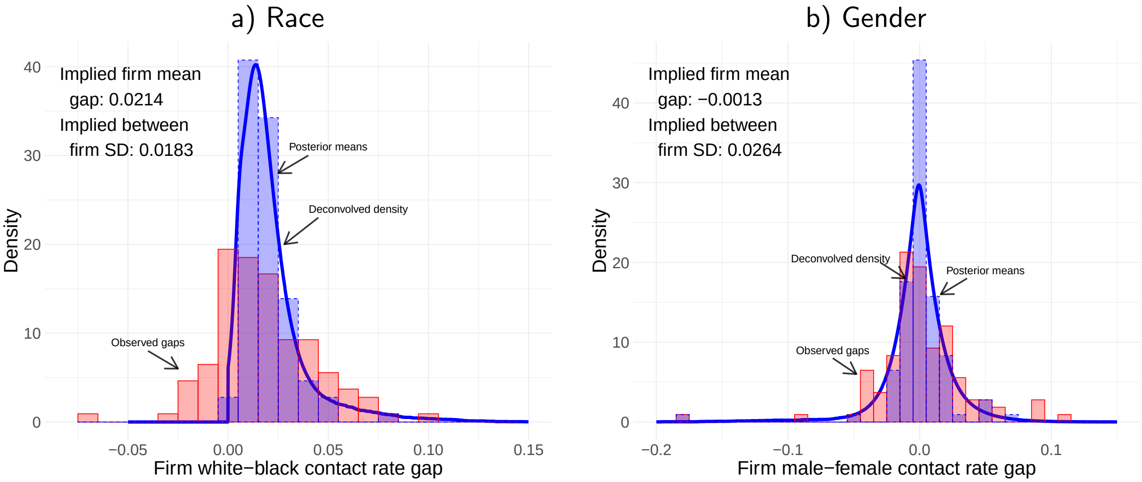


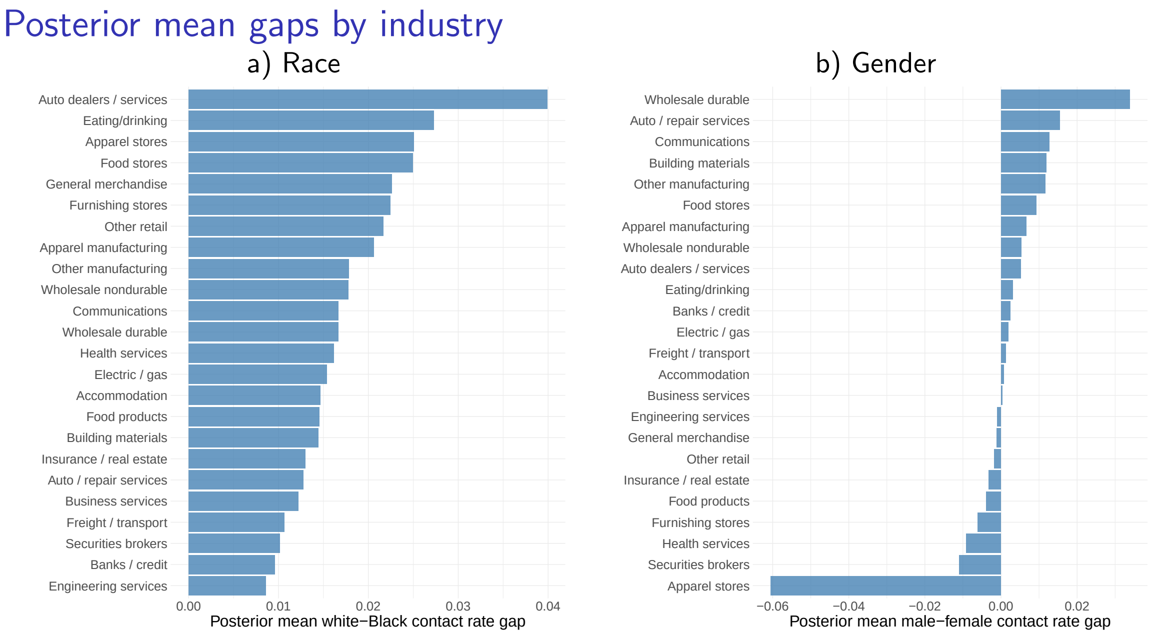
Lorenz curves: Top 20% of firms explain ∼50-60% of lost contacts



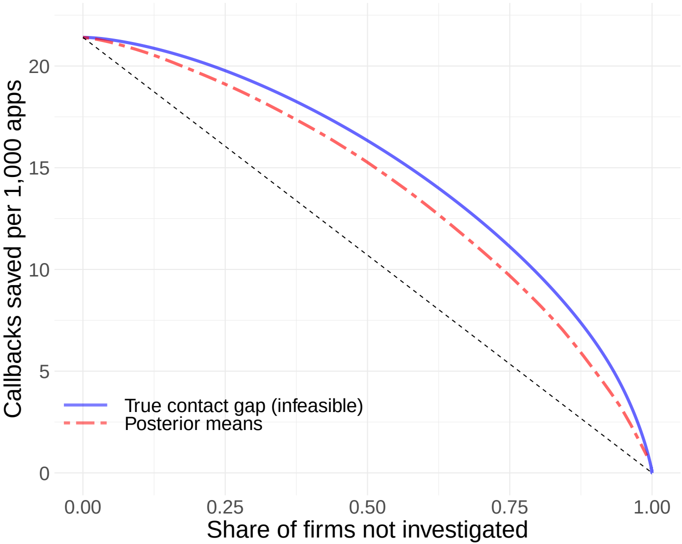
Detecting discriminators

EB approach: Treat deconvolved density as prior to form posterior means




Posterior mean gaps by industry
a) Race
Auto dealers / services
Eating/drinking
Apparel stores
Food stores
General merchandise
Furnishing stores
Other retail
Apparel manufacturing
Other manufacturing
Wholesale nondurable
Communications
Wholesale durable
Health services
Electric / gas
Accommodation
Food products
Building materials
Insurance / real estate
Auto / repair services
Business services
Freight / transport
Securities brokers
Banks / credit
Engineering services
0.00
0.01
0.02
0.03
0.04
Posterior mean white−Black contact rate gap
b) Gender
Wholesale durable
Auto / repair services
Communications
Building materials
Other manufacturing
Food stores
Apparel manufacturing
Wholesale nondurable
Auto dealers / services
Eating/drinking
Banks / credit
Electric / gas
Freight / transport
Accommodation
Business services
Engineering services
General merchandise
Other retail
Insurance / real estate
Food products
Furnishing stores
Health services
Securities brokers
Apparel stores
−0.06
−0.04
−0.02
0.00
0.02
Posterior mean male−female contact rate gap

Detection possibilities: posterior means are highly informative



Detecting *any* discrimination

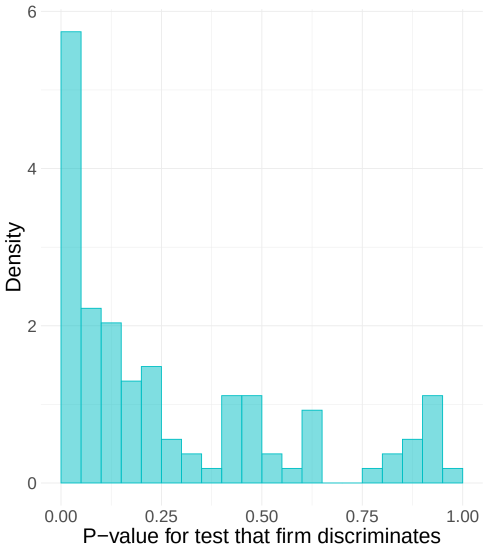
EB posterior means are best guess of each firm's contact gap, but possible that some firms with large $\bar{\Delta}_f$ have true contact gaps of exactly zero

Focusing on whether $\Delta_f > 0$ may lead to different prioritization of firms (Gu and Koenker, 2020)

"Extensive margin" of discrimination has direct legal relevance, since direct Title VII prohibits *any* discrimination based upon protected characteristics

Next: Use multiple-testing methods to examine impact of controlling False Discovery Rates vs. focusing on expected gaps

Multiple testing: P-values reflect mix of false and true nulls



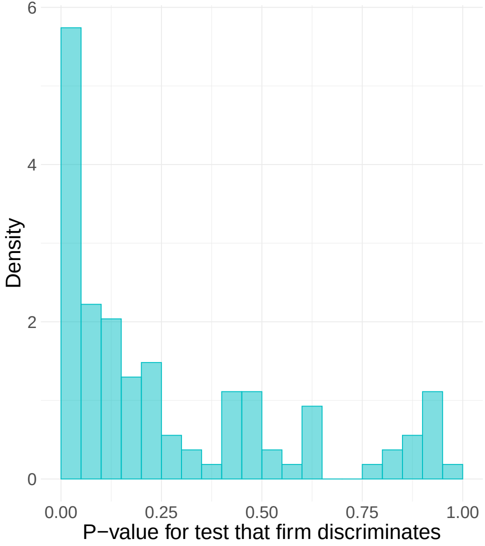
Compute p-values for test of $\Delta_f = 0$

P-values generated by a two-type mixture:

- $\Pr(\hat{p}_f < p | \Delta_f = 0) = p$ (True nulls)
- $\Pr(\hat{p}_f < p | \Delta_f \neq 0) \geq p$ (False nulls)

Fraction of firms with $\Delta_f = 0$ is $\pi_0 \in [0, 1]$

Multiple testing: Goal is to control *False Discovery Rate*

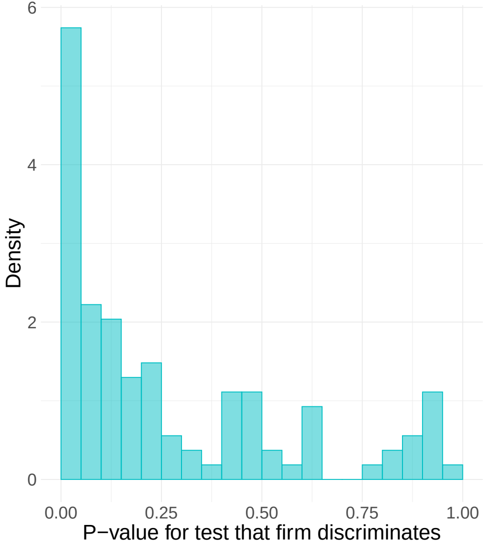


False Discovery Rate of rule rejecting nulls with $\hat{p}_f$ below p is:

$$FDR\left(p\right) = \Pr\left(\Delta_f = 0 | \hat{p}_f < p\right) = \frac{p\pi_0}{F_{\hat{p}}\left(p\right)}$$

Base decisions on $\hat{q}_f = \widehat{FDR}(\hat{p}_f)$

Multiple testing: Goal is to control *False Discovery Rate*



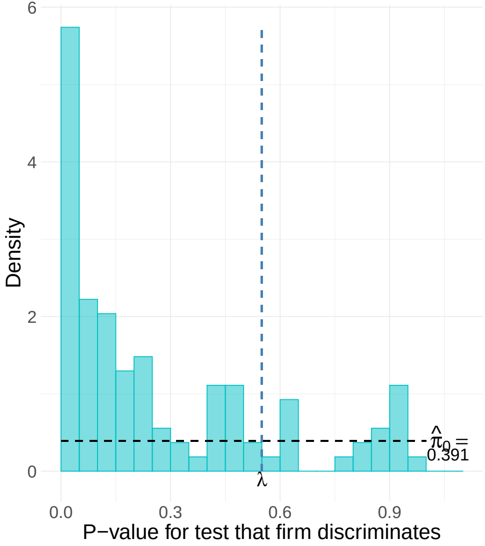
False Discovery Rate of rule rejecting nulls with $\hat{p}_f$ below p is:

$$FDR\left(p\right) = \Pr\left(\Delta_f = 0 | \hat{p}_f < p\right) = \frac{p\pi_0}{F_{\hat{p}}\left(p\right)}$$

Base decisions on $\hat{q}_f = \widehat{FDR}(\hat{p}_f)$

e.g., if $\hat{q}_f = 0.05$ then we expect *at least* 19 out of every 20 firms with p-values below $\hat{p}_f$ to have $\Delta_f \neq 0$.

Multiple testing: At least 60% of firms discriminate against Black names



Efron et al. (2001) upper bound:

$$\pi_0 \leq \min_{p \in [0,1]} f_{\hat{p}}(p)$$

Storey (2002) estimator: for $\lambda \in [0, 1)$

$$\hat{\pi}_0(\lambda) = \frac{\sum_{f=1}^{108} 1\{\hat{p}_f > \lambda\}}{(1-\lambda)\,108}$$

Because true nulls over-represented close to 1, tighter bound, more variance as $\lambda \to 1$

Many firms detected to be discriminating with low $\hat{q}_f$

	Race	Gender
	One-tailed	Two-tailed
	Bootstrapped λ	
$\hat{\pi}_0$	0.391	0.833
$\#$ q-values ≤ 0.05	23	1
$\#$ q-values ≤ 0.1	45	5
λ	0.550	0.300

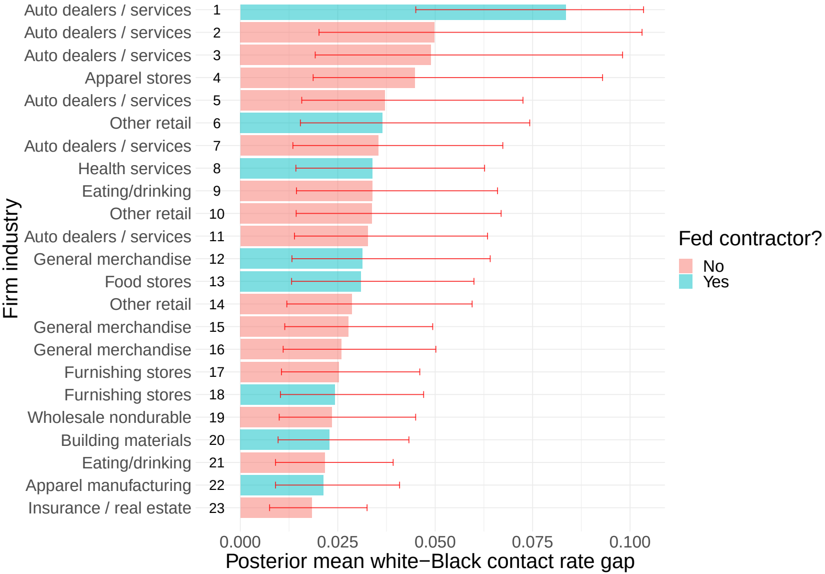
23 firms have q-values below 0.05, implying about 1 expected to have $\Delta_f = 0$

Many firms detected to be discriminating with low $\hat{q}_f$

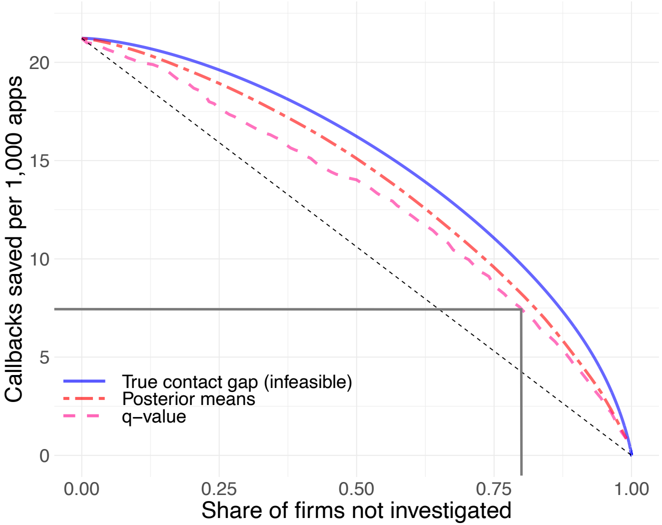
	Race	Gender
	One-tailed	Two-tailed
	Bootstrapped λ	
$\hat{\pi}_0$	0.391	0.833
# q-values ≤ 0.05	23	1
# q-values ≤ 0.1	45	5
λ	0.550	0.300

Higher π_0 estimate for gender produces higher q-values

Firms with $\hat{q}_f < .05$, sorted by posterior mean (brackets are 95% EBCIs)



Firms with q-values less than 0.05 responsible for $\approx$ 40% of lost callbacks



Conclusion

Actionable intelligence?

Many large companies exhibit nationwide patterns of disparate treatment; others don't

Callback centralization / fed contractor results suggest policies and structure of some firms may leave them more susceptible to bias

Identities of > 20 firms demonstrably discriminating against Black names constitute potentially actionable intelligence for enforcers of anti-discrimination laws

- Can investigations by EEOC / OFCCP, etc. identify and eliminate problematic firm-wide practices underlying these patterns?
- Hierarchical detection of job-level discrimination? (Avivi et al., 2021)

Or is sunlight the best disinfectant?

Potential to release “discrimination report card” for public consumption

Becker (1957): workers can (partially) evade bias via sorting

- Sectors and identities of egregious discriminators not obvious, especially for race

Firms themselves may also be unaware of bias in their organizations

- Public scrutiny may lead to positive reforms, at risk of patronizing equilibria (Coate and Loury, 1993)

Appendix material

A sampling of recent systemic cases (racial discrimination)

- Dillard Department Store, E.D. Ark., No. 4:30-cv-01152, filed September 29, 2020 - Alleging that Defendant did not promote African American employees into managerial positions because of their race and did not recruit African American college students into its Executive Development Program.

- Personnel Staffing Group, N.D. Ill., No. 1:20-cv-02683, filed June 24, 2020 - Alleging that Defendant failed to assign or refer employees and applicants and subjected employees/applicants to unequal terms and conditions based on race and sex (black, female).

- Helados La Tapatia, Inc., E.D. Cal., No. 1:20-cv-00722, filed May 22, 2020 - Alleging that Defendant discriminated in recruitment and hiring for unskilled entry-level positions based on national origin (non-Hispanic) and race (white, black, and Asian), and discharged charging party because of his race and/or national origin, non-Hispanic white.

A sampling of recent systemic cases (gender / age discrimination)

- Sactacular Holdings, LLC d/b/a Adam and Eve, E.D.N.C., No. 5:19-cv-00402, filed Sept. 12, 2019 - Alleging that Defendant refused to hire men into sales associate positions.
- American Freight, N.D. Ala., No. 2:19-cv-00273, filed Feb. 14, 2019 - Alleging that Defendant engaged in a pattern or practice of failing to hire female employees into warehouse positions because of sex.
- LTI Services, N.D. Ind., No. 3:20-cv-00304, filed April 9, 2020 - Alleging that Defendant failed to hire qualified females for a client seeking to staff warehouse receiving positions.
- Jet Propulsion Laboratory, C.D. Cal., No. 2:20-cv-03131, filed April 3, 2020 - Alleging that Defendant discriminated against employees age 40 and older in layoffs and in rehire decisions.

EEOC vs. Target Corp

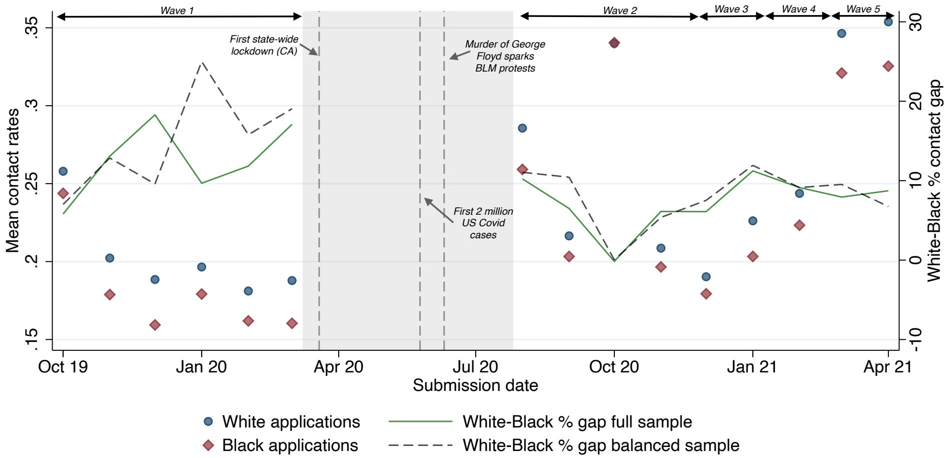
Complaint involved a group of individuals who claimed they were not hired at Target due to race.

One individual, Kalisha White, applied and was told the manager was "too busy" for an interview. She reapplied under the name "Sarah Brucker" and was granted an interview.

EEOC eventually prevailed and won a settlement + consent decree against Target; M. Bertrand was an expert witness.

Claim that manager was "too busy" viewed as a pretext for racial discrimination.

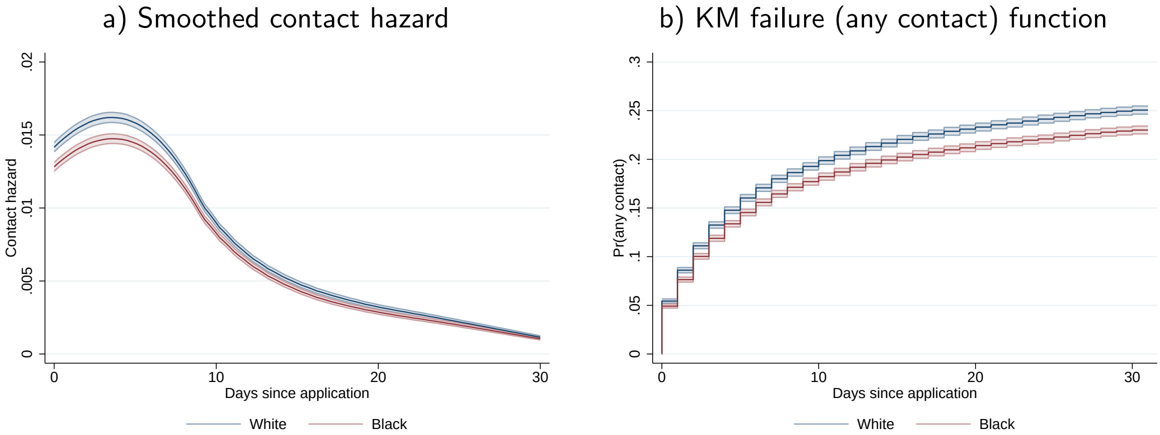
No explicit intent need here...instead courts ruled "they could admit into evidence expert testimony to the effect that the employer may have racially identified the applicants as African American on the basis of their names."



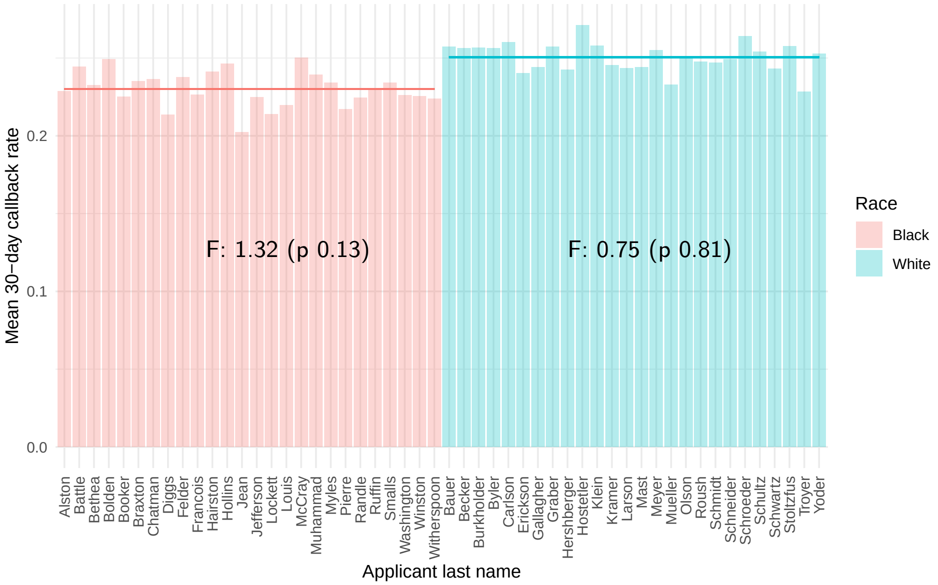
Average Black/white contact gap of 2.1pp, or 9%

- 36% avg. gap reported in meta-analysis of Quillian et al. (2017)
- Level diffs of 3pp in Bertrand and Mullainathan (2004) and 2.6pp in Nunley et al. (2015)
- Discrimination less severe among large firms? (Banerjee et al. 2018)

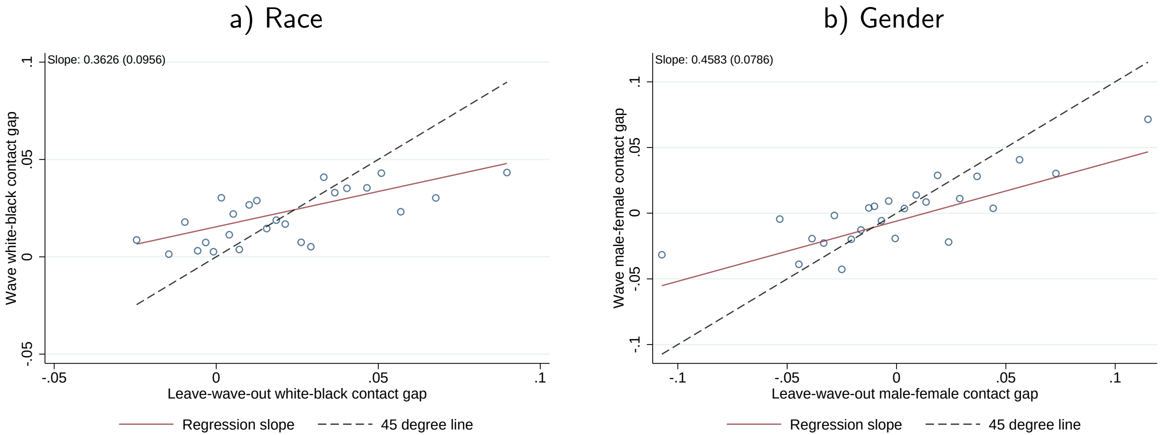
Contact gap stabilizes by 30 days



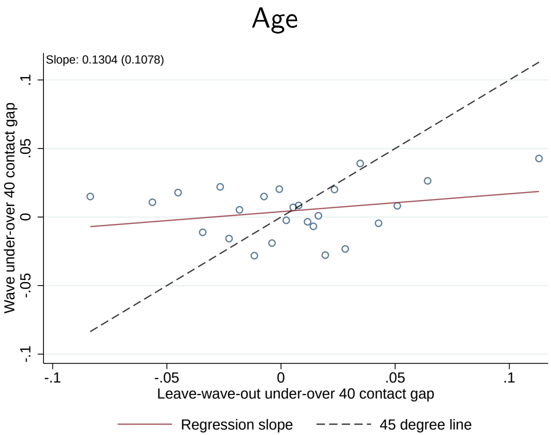
Callbacks by applicant last name



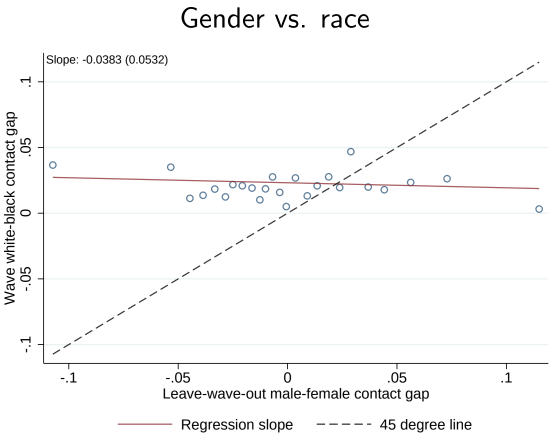
Cross-wave stability suggests sizable firm component



Weak cross-wave relationship for age



No relationship between firm race and gender gaps cross waves



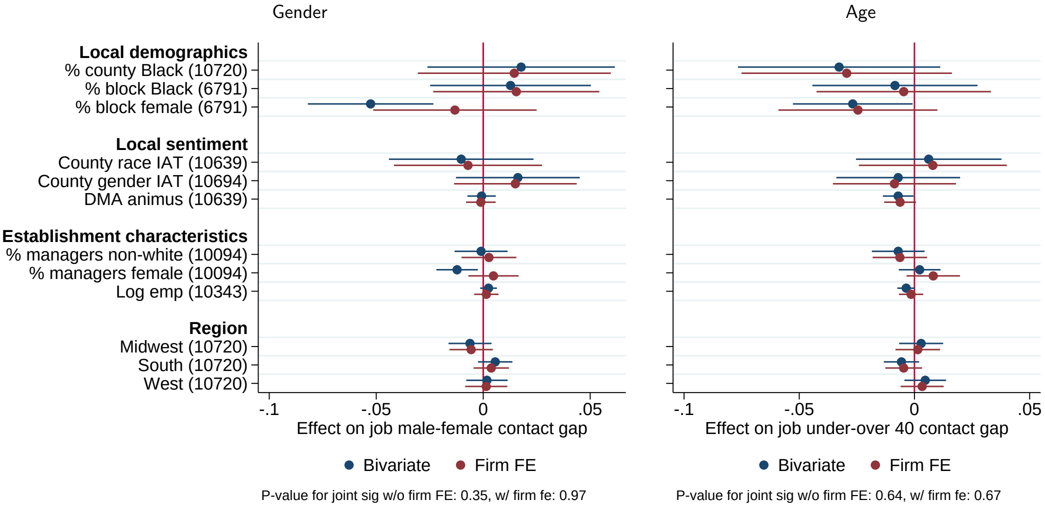
Std dev $\approx$ mean in levels, log odds, and log proportions

	LPM		Logit		Poisson	
	(1) Intercept	(2) Slope	(3) Intercept	(4) Slope	(5) Intercept	(6) Slope
Mean	0.2547	-0.0187	-1.2715	-0.1102	-1.6046	-0.0853
	(0.0036)	(0.0018)	(0.0276)	(0.0152)	(0.0238)	(0.0131)
Std. dev.	0.1607	0.0186	0.9755	0.1155	0.7047	0.0837
	(0.0035)	(0.0035)	(0.0385)	(0.0360)	(0.0382)	(0.0341)
Corr. w/own slope	-0.4010	1.000	0.0519	1.000	0.0685	1
	(0.1098)	-	(0.2074)	-	(0.3092)	-
Corr. w/LPM slope	-0.4010	1.000	-0.4274	0.8944	-0.5045	0.8075
	(0.1098)	-	(0.1068)	(0.2095)	(0.1149)	(0.3074)
Number of firms	103		103		103	

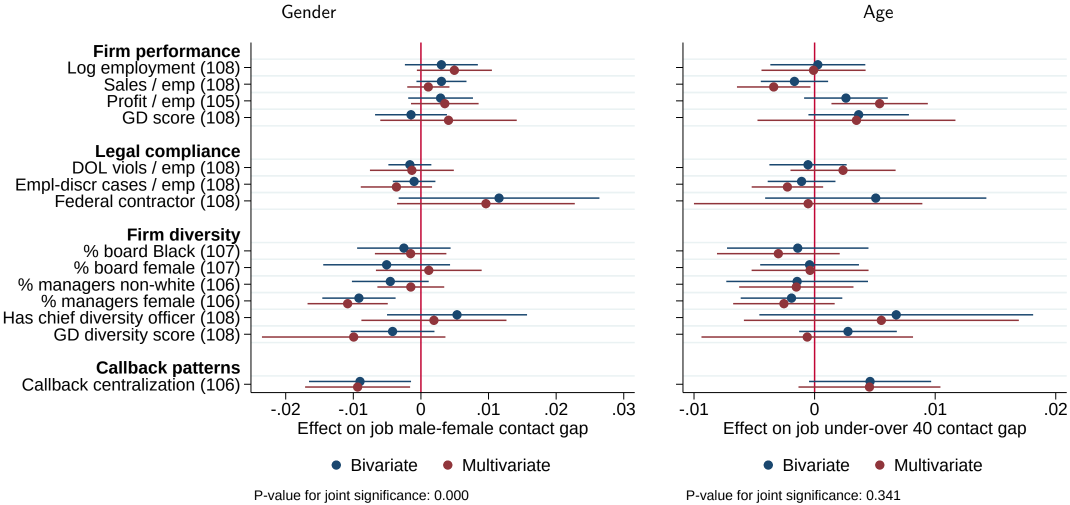
Insignificant variance components for LGBTQ clubs and pronouns

			Contact gap SD		
	(1) χ^2 test of heterogeneity	(2) *p*-value for no discrim against:	(3) Bias-corrected	(4) Cross-wave	(5) Cross-state
LGBTQ Club Member	88.0	W: 1.00	-	-	-
	[0.885]	B: 0.98			
Gender Neutral Pronouns	126.5	Y: 0.92	0.0198	0.0177	0.0147
	[0.076]	O: 0.65	(0.0156)	(0.0176)	(0.0208)

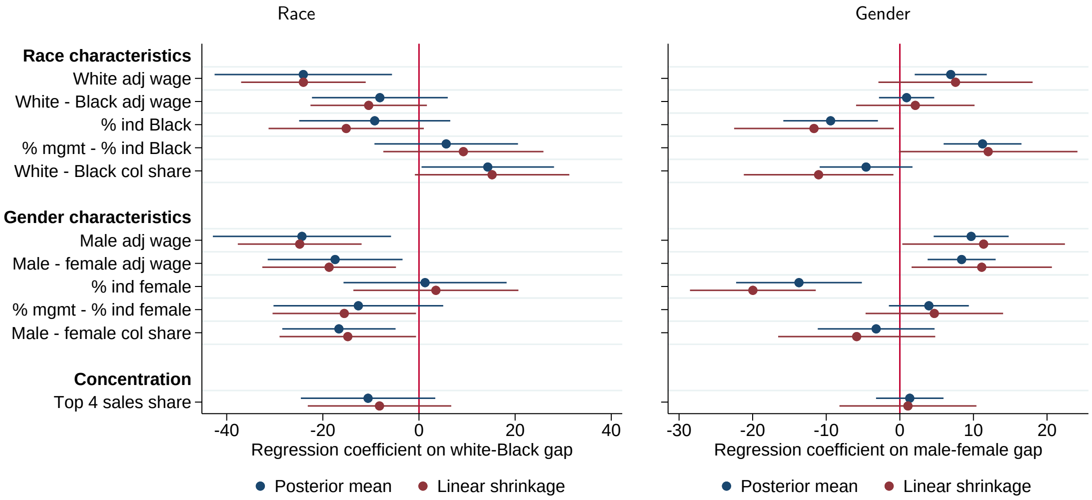
Limited relationship between establishment Xs and gender / age gaps



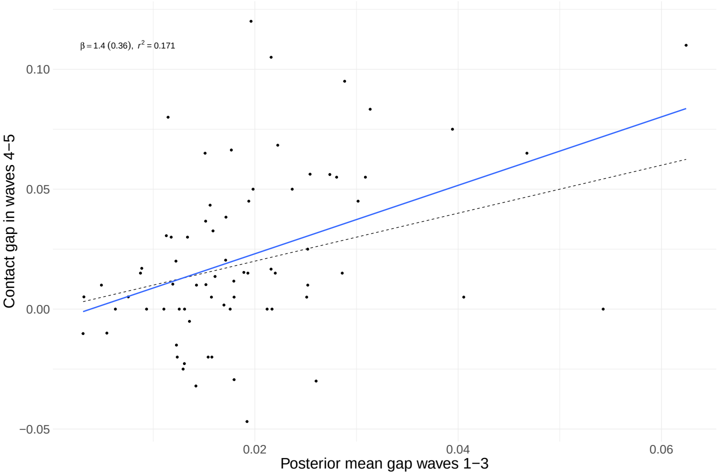
Same with firm covariates



Relationship between posterior mean contact gaps and industry characteristics

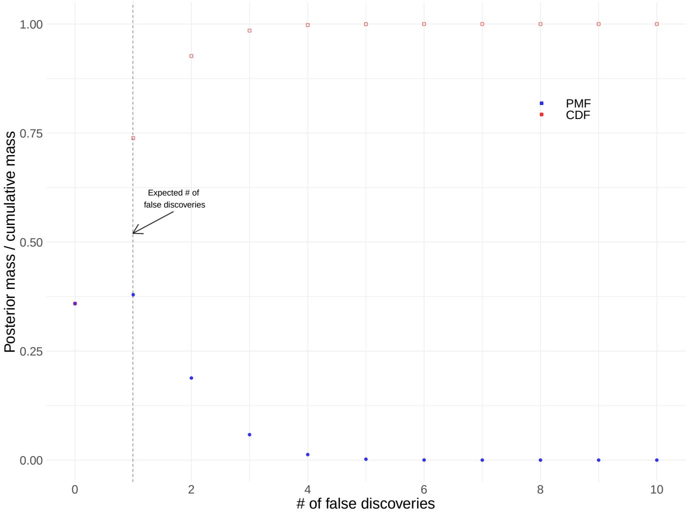


Out-of-sample predictive power of racial contact gap posteriors



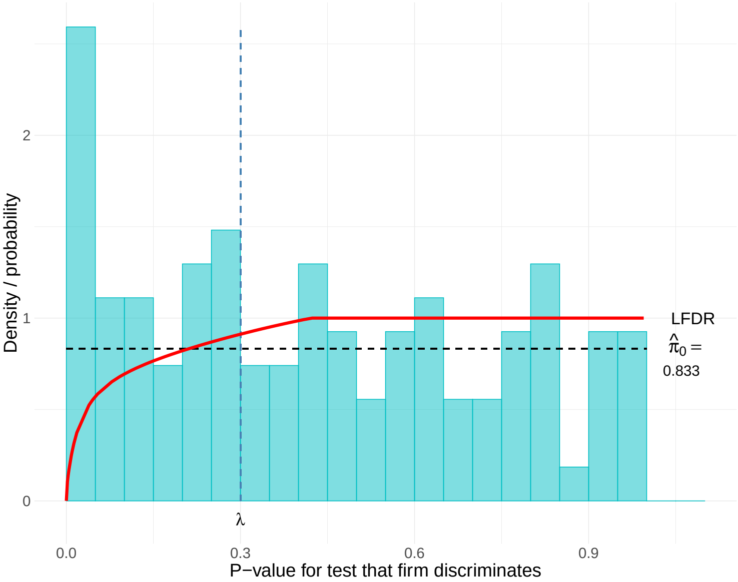
Notes: This figure plots posterior mean white-Black contact gaps using data from waves 1-3 against observed gaps in waves 4-5 for the sample of firms included in all five waves. The blue line is the least squares fit. Adjusting for the noise in the wave 4 and 5 estimates yields a bias corrected correlation between predictions in later waves and the latent true contact gaps of 0.71.

Posterior false discovery distribution among 23 firms with low q-values



Notes: This figure plots EB posterior estimates of the probability mass function and cumulative distribution of false discoveries among the 23 firms with q-values below 0.05 for race. Posterior was calculated using the Poisson binomial distribution implied by the 23 firms' LFDR estimates.

At least 17% of firms discriminate by gender



Robustness of selections to estimator of π_0 bound

	Race		Gender	Age
	One-tailed	Two-tailed	Two-tailed	Two-tailed
	Bootstrapped λ			
$\hat{\pi}_0$	0.391	0.541	0.833	0.833
# q-values $<= 0.05$	23	8	1	0
# q-values $<= 0.1$	45	21	5	1
λ	0.550	0.350	0.300	0.400
	Randomization inference p-values			
$\hat{\pi}_0$	0.370	0.455	0.808	0.802
# q-values $<= 0.05$	35	24	8	1
# q-values $<= 0.1$	55	36	10	1
λ	0.550	0.450	0.450	0.400
	Smoothed			
$\hat{\pi}_0$	0.451	0.882	0.854	0.832
# q-values $<= 0.05$	21	4	1	0
# q-values $<= 0.1$	40	18	5	1
	95% upper CI for π_0			
$\hat{\pi}_0$	0.602	0.696	1.000	1.000
# q-values $<= 0.05$	20	4	1	0
# q-values $<= 0.1$	31	18	5	1

Bounding job-level prevalence

Suppose $1-\phi_f$ of jobs at firm f have $\Delta_{fj}=0$, so that $\Delta_f=\phi_f\dot{\Delta}_f$

Variance of job-level gaps can be written:

$$\sigma_f^2=\phi_f\dot{\sigma}_f^2+\phi_f(1-\phi_f)\dot{\Delta}_f^2$$

where $\dot{\sigma}_f^2$ is the variance of gaps in subset of jobs that discriminate

Substituting yields:

$$\phi_f\geq\frac{\Delta_f^2}{\sigma_f^2+\Delta_f^2}$$

Can use to bound prevalence of job-level discrimination among firms with $q-$values below a given threshold

Bounds on job-level prevalence of discrimination