Systemic Discrimination among Large US Employers

Pat Kline, Evan Rose, and Chris Walters

Presented by Myles Wagner and Anthony Yu

October 7, 2021

Outline

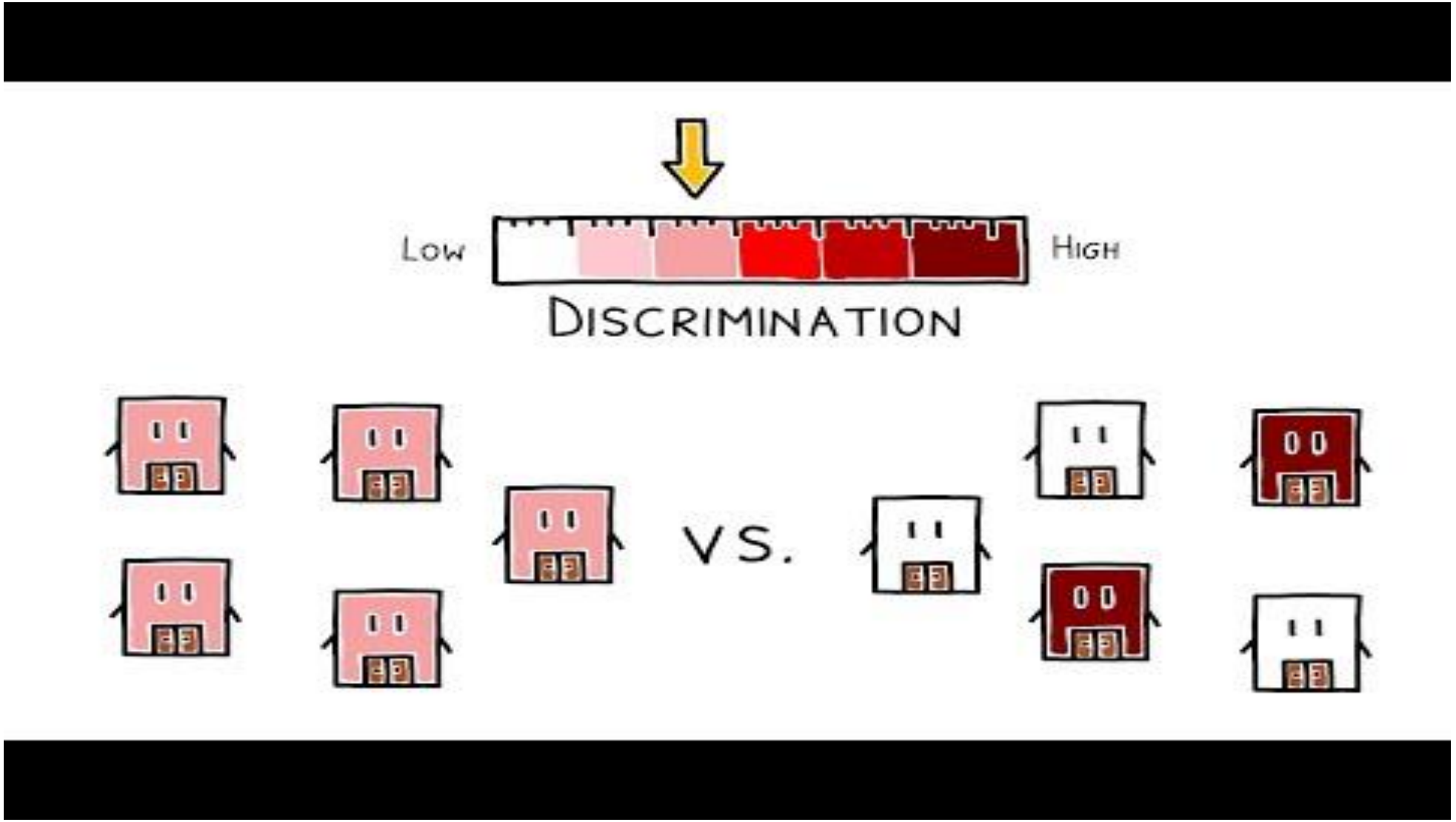
- Paper summary
  - Audit study
  - Main results

- Discussion
  - Paper-specific
  - Big picture

Discussion Points

- Questions to keep in mind throughout
  - What type of discrimination does an audit study measure?
  - What is the policy relevance of the findings in KRW (2021)?
  - How is health equity similar to and different from employment discrimination?

YouTube video



Background

- Illegal to discriminate in hiring on basis of race, sex, color, religion, and national origin
- Large literature that uses audit studies to measure market-level averages of differences in contacts by race
- Less literature documenting whether there is variation across firms
  - **Paper:** Do all firms discriminate a similar amount or are some firms really bad while most are okay?

Large audit study

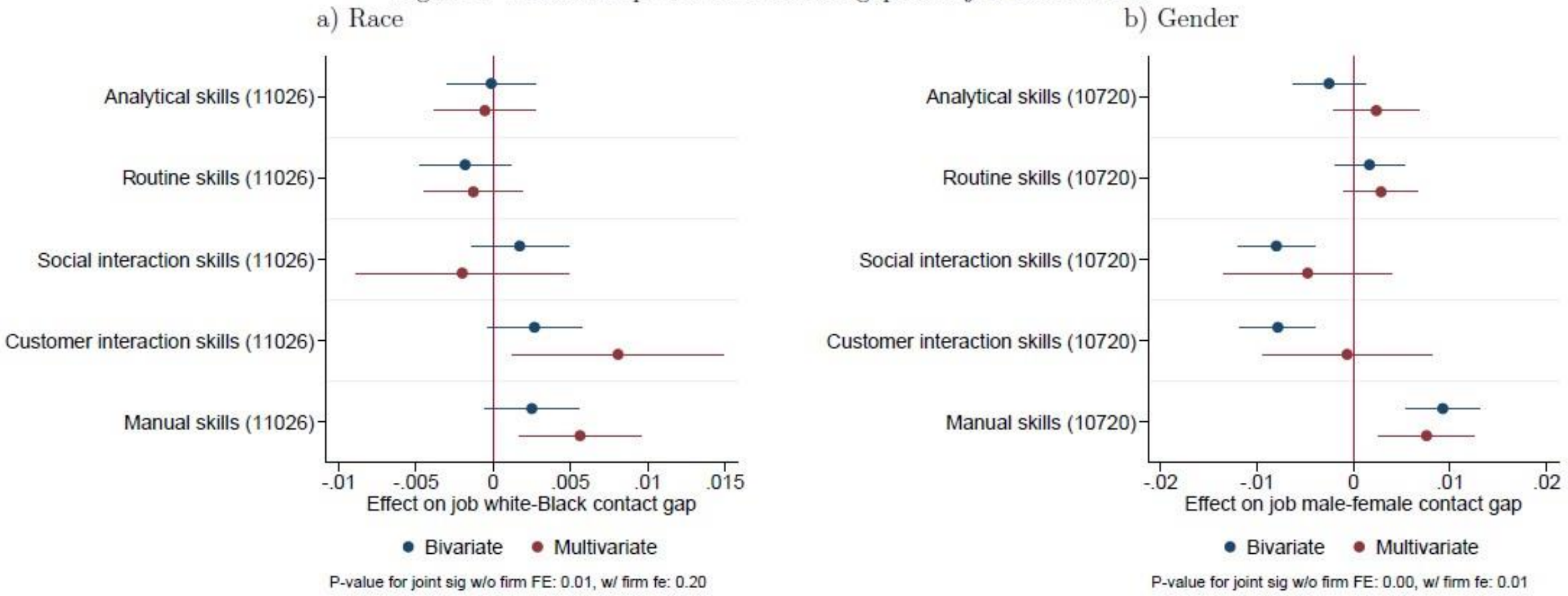
- Sample entry-level jobs from 100+ Fortune 50 firms
- Apply to 125 geographically distinct jobs from each firm
- 8 applications to each job
- Sample size: 84,000 application (20x Bertrand and Mullainathan 2004, 84x Button et al. 2021)
  - ***Massive***
- Experiment organized in 5 waves
  - Randomized: names, age, race
- Main outcome variable is whether application was contacted or contact gap
  - Full data: application by job by firm level

Average effects (mean 30-day callback rate)

- Black-white gap -0.0205 (0.0017)
- Male-female gap 0.00064 (0.003)
- Over 40-under 40 gap -0.0059 (0.003)
- Average differences by race and age are statistically significantly different

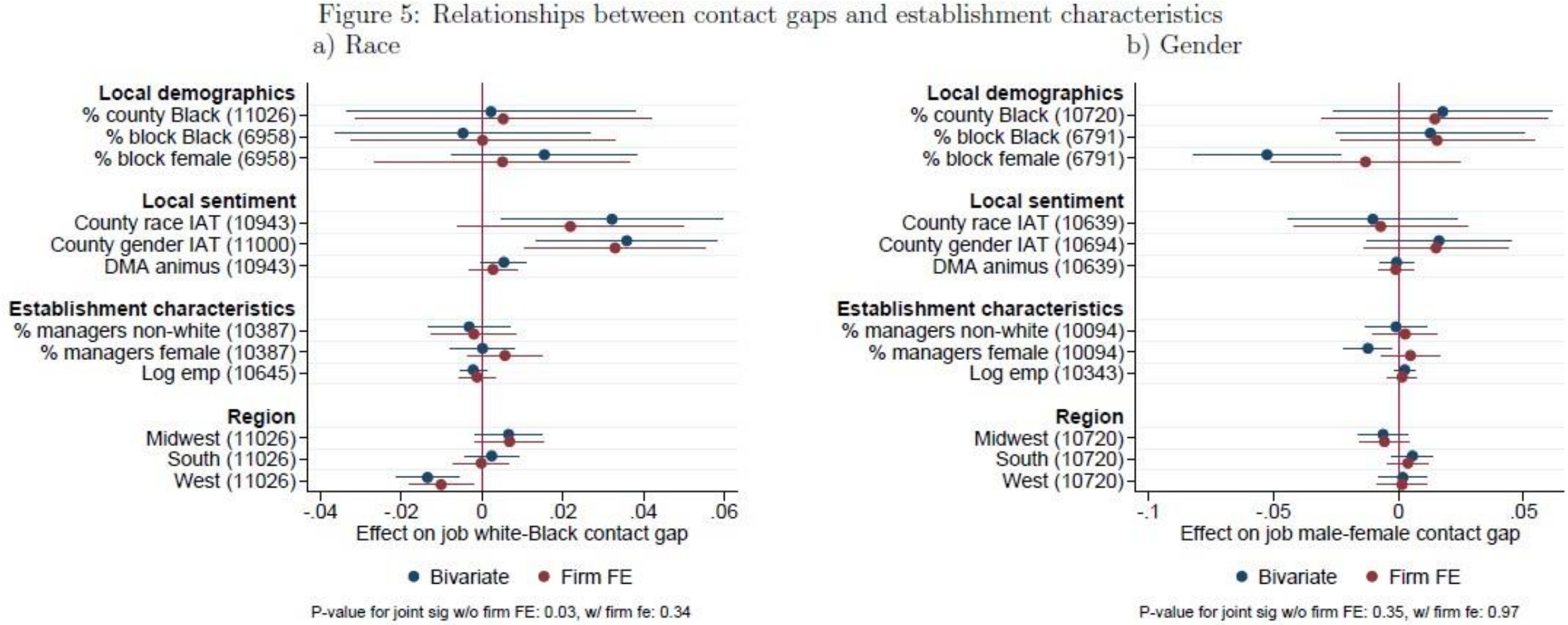
Gaps by job task content ($\widehat{\Delta}_{fj} = \beta_f + \beta_1 X_j + e_{fj}$)

Figure 4: Relationships between contact gaps and job task content



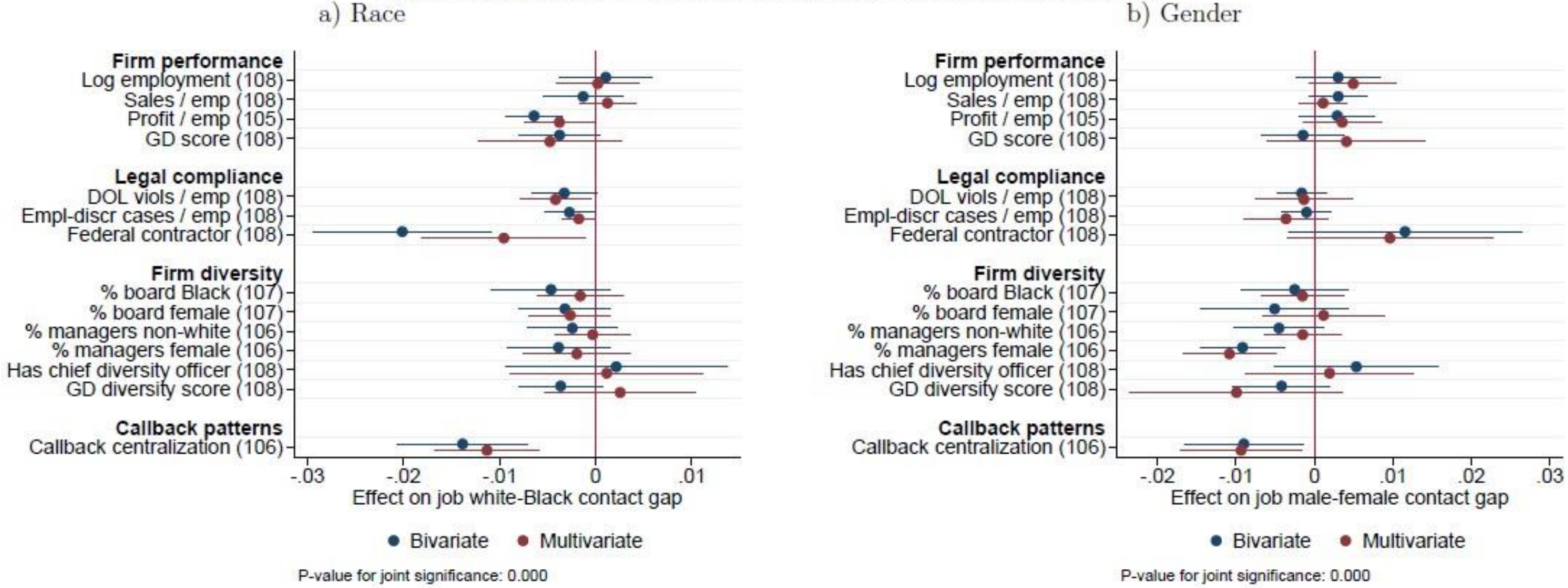
Local prejudice matters for black-white gap

Figure 5: Relationships between contact gaps and establishment characteristics



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

Figure 6: Relationships between contact gaps and firm characteristics



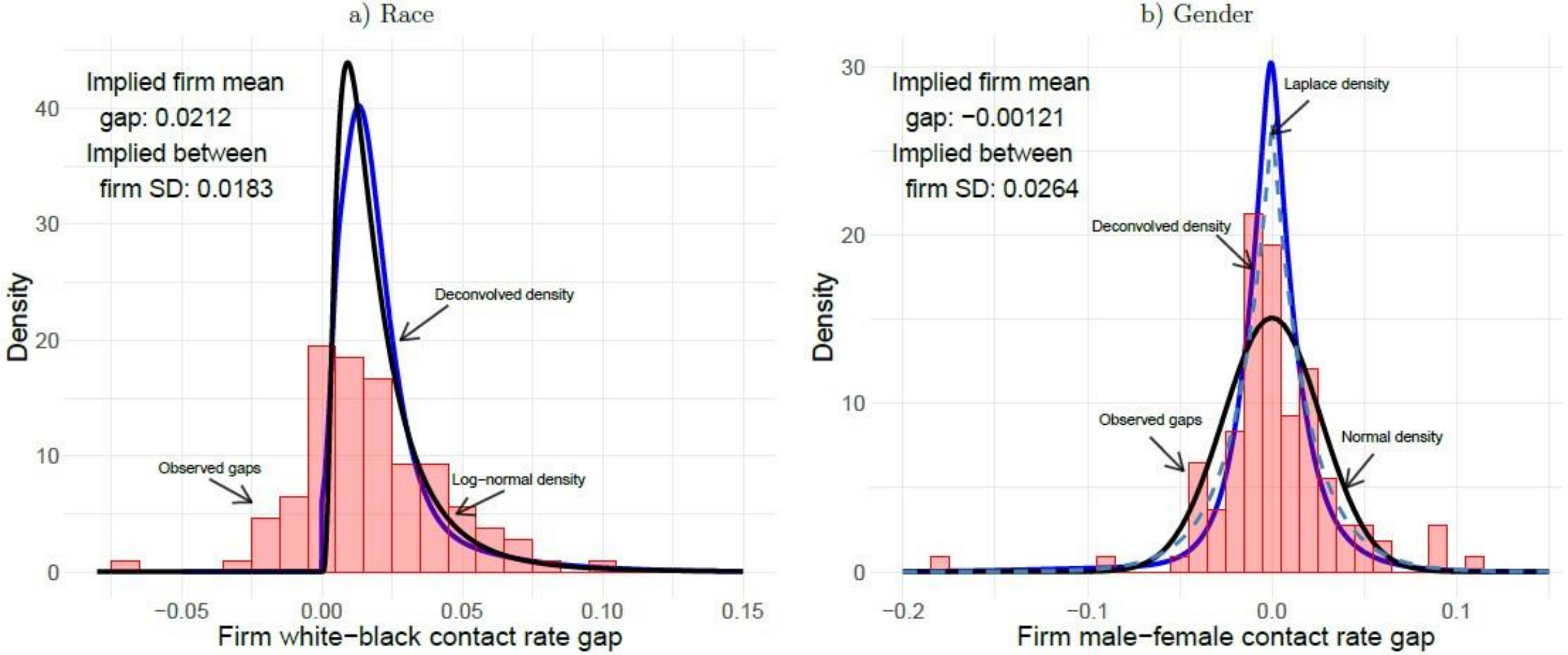
Estimating distribution of firm fixed effects

- Want to know population distribution of Δ_f
  - Let $z_f = \frac{\widehat{\Delta}_f}{s_f}$ and $\mu_f = \frac{\Delta_f}{s_f}$ and normality $z_f \mid \mu_f \sim N(\mu_f, 1), \mu_f \sim G_\mu$
    - Spline and penalized MLE to recover $\hat{g}_\mu = d\hat{G}_\mu$
  - Deconvolution estimator of density

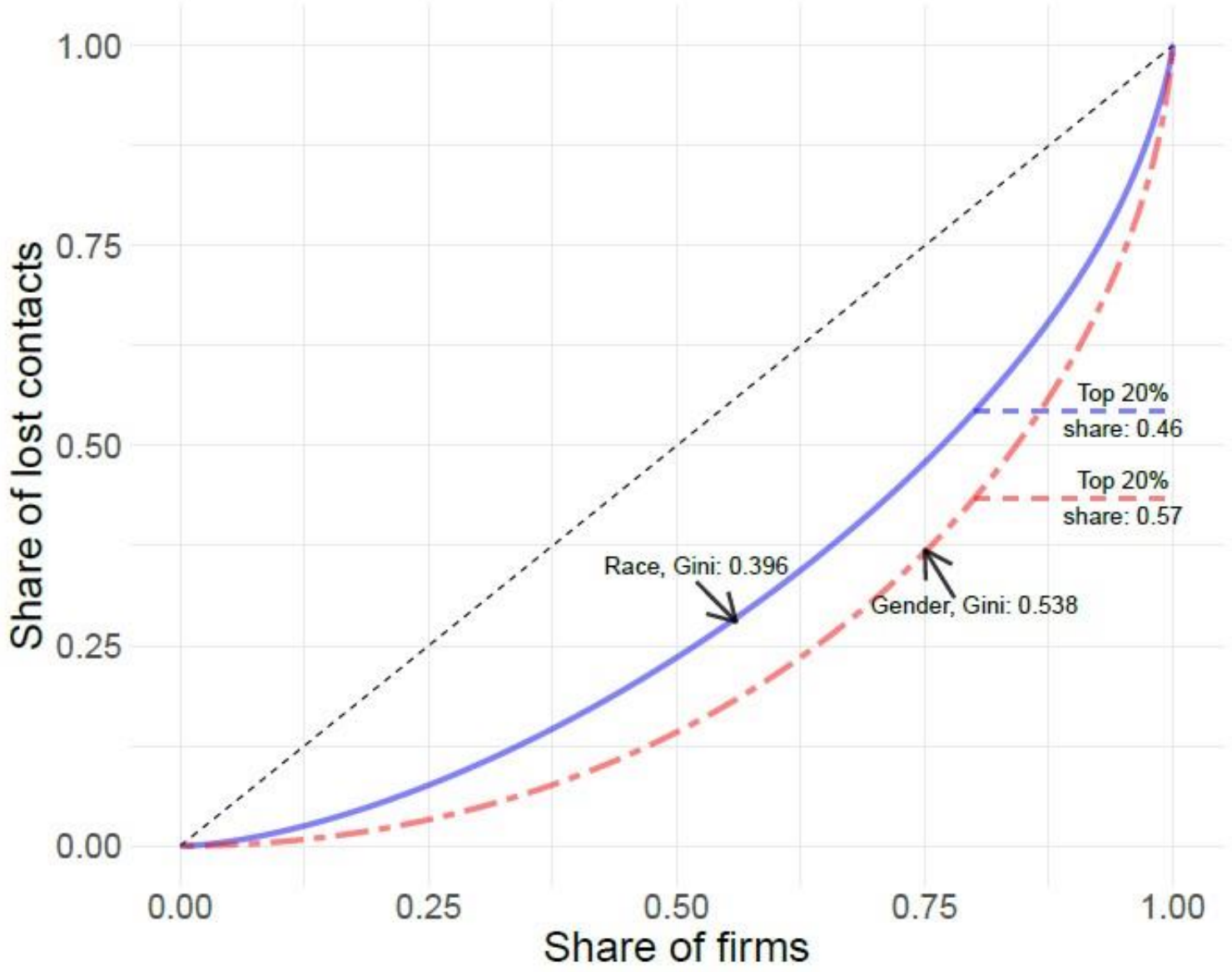
$$\hat{g}_\Delta(x) = \int e^{-t}\ \hat{g}_\mu(e^{-t}x)\ \hat{g}_{\ln s}\ (t)dt$$

Heterogeneity in callback rates

Figure 7: Deconvolution estimates of firm-level discrimination distributions



Lorenz curve

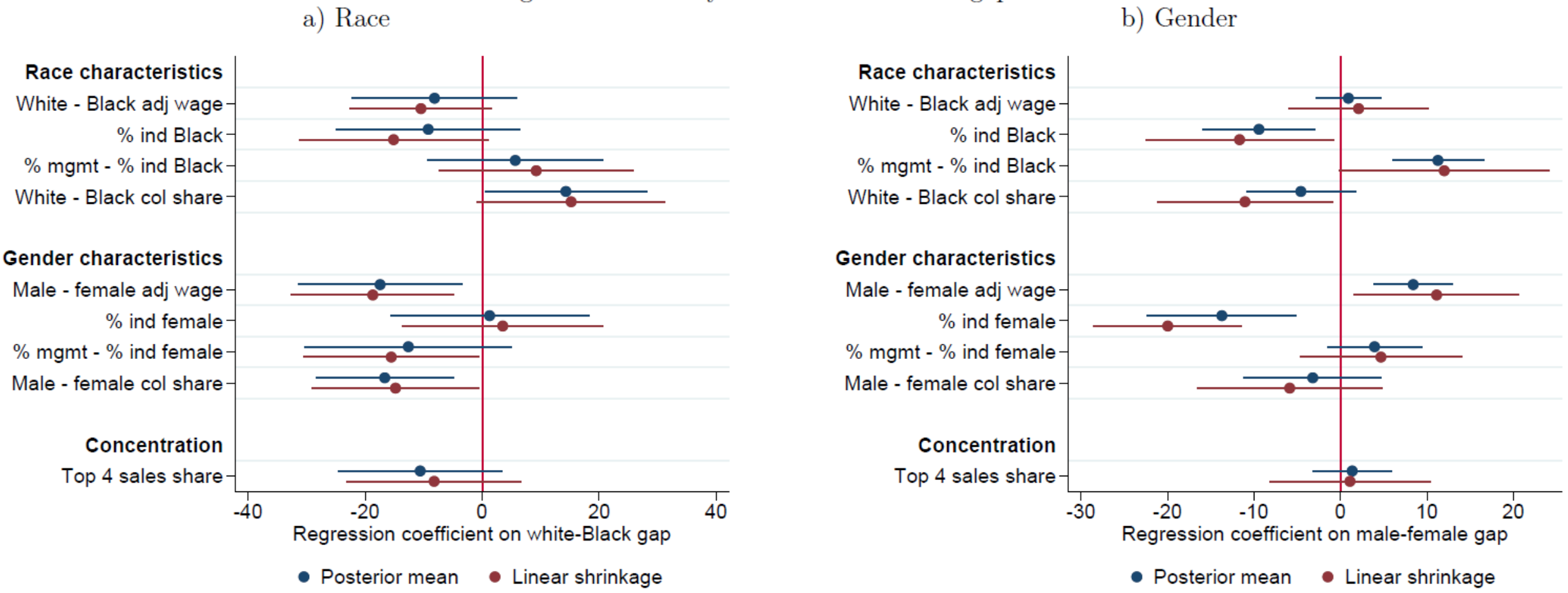


Estimating firm-specific estimates

- Observe $\widehat{\Delta}_f$, want to know population Δ_f
- Methods
  - Empirical Bayes
  - Linear Shrinkage
- Main idea: observed gaps unbiased, but imprecise; can get more precise measure with EB or linear shrinkage

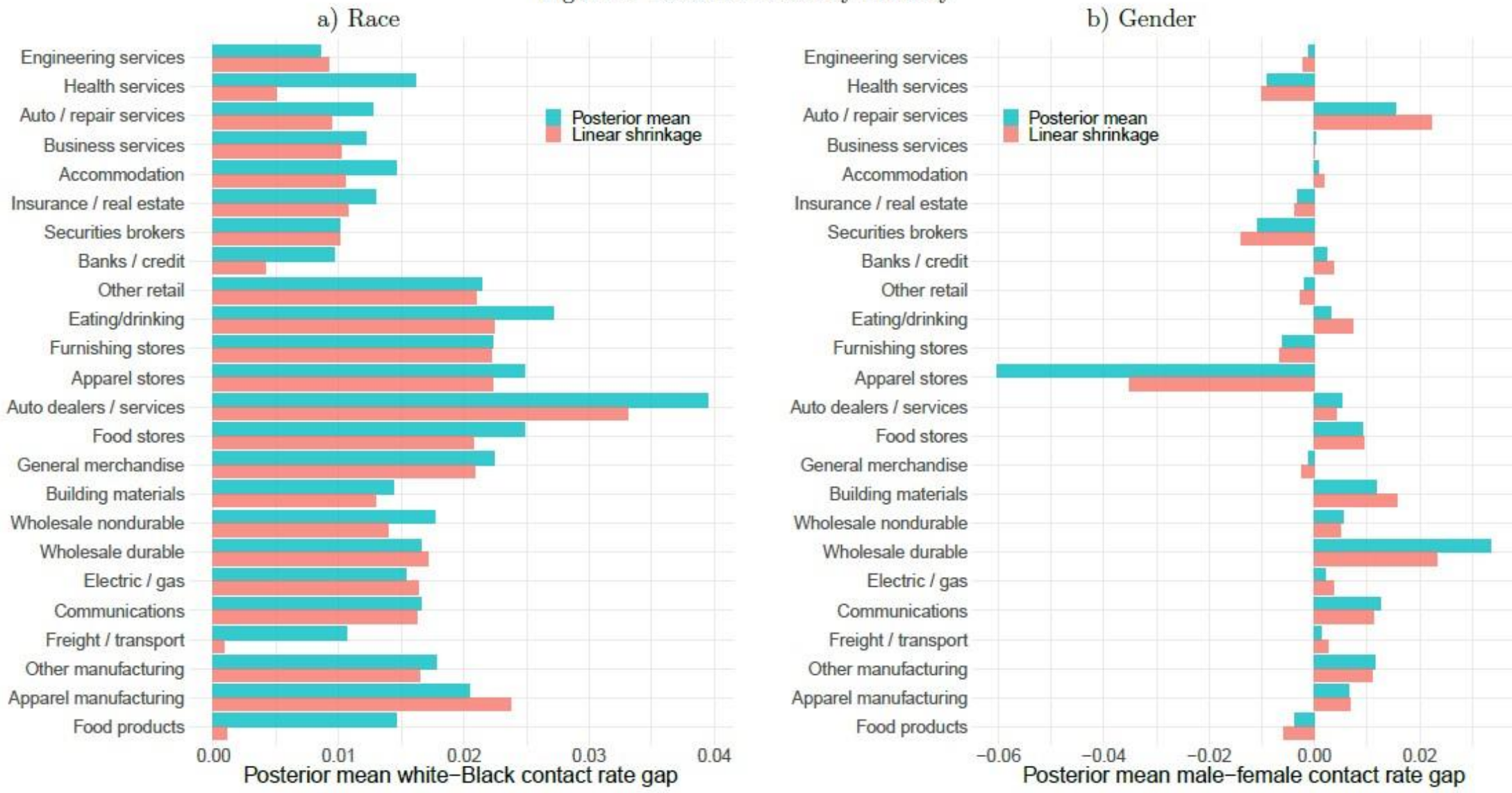
Contact gaps

Figure 10: Industry correlates of contact gaps



Gaps by industry

Figure 9: Posterior means by industry



Policy relevance

- Employment discrimination is illegal (Title VII of Civil Rights Act)
- How can a regulator use this method to find discriminatory firms?
- Outline regulator preferences that rely on experimental evidence of contact gaps to make decisions on which firms to investigate
  - Two types of preferences: care if gap is large, care if there is any gap
- Main takeaway:
  - 23 firms with posterior p-value < 0.05 that are likely to be discriminatory (see Table 9)
  - What is the best use of this information?
    - Investigation
    - Information

Table 9: Estimates of racial discrimination for firms with q-values below 0.05

q-value rank	Industry	Federal Contractor?	Contact gap	Std. err.	p-value	q-value	Posterior mean	Posterior 5th pctile	Posterior 95th pctile
1	Auto dealers / services	Yes	0.0952	0.0197	0.0000	0.0001	0.0833	0.0439	0.1034
2	Auto dealers / services	No	0.0507	0.0143	0.0003	0.0061	0.0348	0.0133	0.0670
3	Auto dealers / services	No	0.0738	0.0220	0.0005	0.0073	0.0481	0.0190	0.0974
4	Auto dealers / services	No	0.0787	0.0249	0.0010	0.0103	0.0489	0.0199	0.1021
5	Apparel stores	No	0.0733	0.0250	0.0022	0.0158	0.0440	0.0185	0.0917
6	Other retail	No	0.0469	0.0159	0.0020	0.0158	0.0282	0.0118	0.0587
7	Other retail	Yes	0.0605	0.0219	0.0033	0.0176	0.0359	0.0153	0.0731
8	General merchandise	Yes	0.0520	0.0187	0.0031	0.0176	0.0309	0.0131	0.0631
9	Auto dealers / services	No	0.0613	0.0240	0.0060	0.0194	0.0366	0.0157	0.0712
10	Eating/drinking	No	0.0560	0.0222	0.0064	0.0194	0.0334	0.0143	0.0648
11	Other retail	No	0.0560	0.0214	0.0050	0.0194	0.0333	0.0142	0.0658
12	Auto dealers / services	No	0.0540	0.0215	0.0068	0.0194	0.0323	0.0138	0.0623
13	Food stores	Yes	0.0511	0.0204	0.0069	0.0194	0.0305	0.0131	0.0589
14	General merchandise	No	0.0427	0.0170	0.0068	0.0194	0.0255	0.0109	0.0493
15	Furnishing stores	Yes	0.0400	0.0159	0.0066	0.0194	0.0239	0.0102	0.0462
16	Wholesale nondurable	No	0.0386	0.0158	0.0080	0.0199	0.0232	0.0099	0.0442
17	Apparel manufacturing	Yes	0.0350	0.0142	0.0078	0.0199	0.0210	0.0090	0.0401
18	Building materials	Yes	0.0373	0.0157	0.0093	0.0218	0.0226	0.0096	0.0425
19	Health services	Yes	0.0544	0.0240	0.0132	0.0292	0.0335	0.0142	0.0615
20	Furnishing stores	No	0.0400	0.0183	0.0152	0.0322	0.0250	0.0105	0.0452
21	Eating/drinking	No	0.0340	0.0159	0.0172	0.0346	0.0214	0.0090	0.0385
22	General merchandise	No	0.0423	0.0210	0.0229	0.0439	0.0275	0.0114	0.0486
23	Insurance / real estate	No	0.0278	0.0140	0.0257	0.0472	0.0182	0.0075	0.0320

Notes: This table reports estimates of white-Black contact gaps for the 23 individual firms with q-values less than 0.05. P-values and q-values come from one-sided tests of the null hypothesis that the firm does not discriminate against Black applicants. To ensure that q-values are non-decreasing for nested decision thresholds, we follow Storey (2002, 2003) in estimating $\hat{q}_f$ as $\min_{t \geq \hat{p}_f} \widehat{FDR}(t)$, which implies firms with different p-values may have the same q-value. Posterior means and percentiles are empirical Bayes posteriors constructed using the estimated distribution in Figure 7 as the prior.

Discussion

Regarding paper

- Randomization of names vs. randomization of race
- How important are differential contact rates matters for wage inequality or inequality on other measures?
  - Contact rates vs. actual hiring (discretion)
- Discrimination mechanisms?

Big picture

- What can audit studies pick up? What can they not pick up?
  - Audit studies pick up disparate treatment, not disparate impact

Relation to health equity

- What are the barriers to health equity?
  - Discrimination a la audit study (disparate treatment)
  - Disparate impact
    - Supply vs. demand → may or may not have underlying discriminatory intent