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?
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 \( \Delta_{fj} = \beta_f + \beta_1 X_j + e_{fj} \)
Local prejudice matters for black-white gap

Figure 5: Relationships between contact gaps and establishment characteristics

a) Race

- Local demographics:
  - % county Black (11026)
  - % block Black (6958)
  - % block female (6958)

- Local sentiment:
  - County race IAT (10943)
  - County gender IAT (11000)
  - DMA animus (10943)

- Establishment characteristics:
  - % managers non-white (10387)
  - % managers female (10387)
  - Log emp (10645)

- Region:
  - Midwest (11026)
  - South (11026)
  - West (11026)

b) Gender

- Local demographics:
  - % county Black (10720)
  - % block Black (6791)
  - % block female (6791)

- Local sentiment:
  - County race IAT (10639)
  - County gender IAT (10694)
  - DMA animus (10639)

- Establishment characteristics:
  - % managers non-white (10094)
  - % managers female (10094)
  - Log emp (10343)

- Region:
  - Midwest (10720)
  - South (10720)
  - West (10720)

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

P-value for joint sig w/o firm FE: 0.35, w/ firm fe: 0.97
Smaller gaps at profitable firms, fed contractors, and centralized firms

Figure 6: Relationships between contact gaps and firm characteristics

(a) Race

- Firm performance
  - Log employment
  - Sales / emp
  - Profit / emp
  - GD score

- Legal compliance
  - DOL viol / emp
  - Empl-discr cases / emp
  - Federal contractor

- Firm diversity
  - % board Black
  - % board female
  - % managers non-white
  - % managers female
  - Has chief diversity officer
  - GD diversity score

- Callback patterns
  - Callback centralization

(b) Gender

- Firm performance
  - Log employment
  - Sales / emp
  - Profit / emp
  - GD score

- Legal compliance
  - DOL viol / emp
  - Empl-discr cases / emp
  - Federal contractor

- Firm diversity
  - % board Black
  - % board female
  - % managers non-white
  - % managers female
  - Has chief diversity officer
  - GD diversity score

- Callback patterns
  - Callback centralization

Effect on job white-black contact gap

Bivariate ▪ Multivariate

P-value for joint significance: 0.000

Effect on job male-female contact gap

Bivariate ▪ Multivariate

P-value for joint significance: 0.000
Estimating distribution of firm fixed effects

- Want to know population distribution of $\Delta_f$

- Let $z_f = \frac{\Delta_f}{s_f}$ 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 $\hat{\mu} = d\hat{G}_\mu$

- Deconvolution estimator of density

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

Figure 7: Deconvolution estimates of firm-level discrimination distributions

a) Race

Implied firm mean gap: 0.0212
Implied between firm SD: 0.0183

b) Gender

Implied firm mean gap: -0.00121
Implied between firm SD: 0.0264
Lorenz curve
Estimating firm-specific estimates

- Observe $\hat{\Delta}_f$, want to know population $\Delta_f$

- 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

a) Race

Race characteristics
- White - Black adj wage
- % ind Black
- % mgmt - % ind Black
- White - Black col share

Gender characteristics
- Male - female adj wage
- % ind female
- % mgmt - % ind female
- Male - female col share

Concentration
- Top 4 sales share

Regression coefficient on white-Black gap
- Posterior mean
- Linear shrinkage

b) Gender

Race characteristics
- White - Black adj wage
- % ind Black
- % mgmt - % ind Black
- White - Black col share

Gender characteristics
- Male - female adj wage
- % ind female
- % mgmt - % ind female
- Male - female col share

Concentration
- Top 4 sales share

Regression coefficient on male-female gap
- Posterior mean
- Linear shrinkage
Gaps by industry

Figure 9: Posterior means by industry

- **Race**
  - Engineering services
  - Health services
  - Auto/repair services
  - Business services
  - Accommodation
  - Insurance/real estate
  - Securities brokers
  - Banks/credit
  - Other retail
  - Eating/drinking
  - Furnishing stores
  - Apparel stores
  - Auto dealers/services
  - Food stores
  - General merchandise
  - Building materials
  - Wholesale nondurable
  - Wholesale durable
  - Electric/gas
  - Communications
  - Freight/transport
  - Other manufacturing
  - Apparel manufacturing
  - Food products

- **Gender**
  - Engineering services
  - Health services
  - Auto/repair services
  - Business services
  - Accommodation
  - Insurance/real estate
  - Securities brokers
  - Banks/credit
  - Other retail
  - Eating/drinking
  - Furnishing stores
  - Apparel stores
  - Auto dealers/services
  - Food stores
  - General merchandise
  - Building materials
  - Wholesale nondurable
  - Wholesale durable
  - Electric/gas
  - Communications
  - Freight/transport
  - Other manufacturing
  - Apparel manufacturing
  - Food products

Posterior mean white/Black contact rate gap

Posterior mean male/female contact rate gap
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

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

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 $\hat{q}_p$ as $\min_{j \geq p} FDR(j)$, 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