Unwarranted Disparity in High-Stakes Decisions: Race Measurement and Policy Responses

Jason Baron
Duke

Joseph Doyle
MIT

Natalia Emanuel
FRB:NY

Peter Hull
Brown

Joseph Ryan
Michigan

CRIW Race, Ethnicity, and Economic Statistics for the 21st Century
March 2024

Expressed views do not represent those of Michigan CPS, the Federal Reserve Bank of New York, or the Federal Reserve System.
Introduction

- There are widespread concerns about racial discrimination in many high-stakes settings, e.g. criminal justice, child protection...
  - To quantify discrimination in such settings, researchers must hold fixed specific factors that might lead to omitted variable bias (OVB)
  - E.g. racial disparities in foster care placement rates may reflect discrimination or underlying differences in the need for intervention
Introduction

There are widespread concerns about racial discrimination in many high-stakes settings, e.g. criminal justice, child protection...

To quantify discrimination in such settings, researchers must hold fixed specific factors that might lead to omitted variable bias (OVB)

E.g. racial disparities in foster care placement rates may reflect discrimination or underlying differences in the need for intervention

A recent quasi-experimental literature addresses OVB by conditioning on relevant potential outcomes (Arnold et al. ’22; Baron et al. ’24)

In contrast to conventional experimental studies of direct discrimination which condition on all observable “non-race” characteristics

This approach yields broader measures of discrimination, which account for indirect drivers that conventional studies typically hold fixed
Introduction

- There are widespread concerns about racial discrimination in many high-stakes settings, e.g. criminal justice, child protection...
  - To quantify discrimination in such settings, researchers must hold fixed specific factors that might lead to omitted variable bias (OVB)
  - E.g. racial disparities in foster care placement rates may reflect discrimination or underlying differences in the need for intervention

- A recent quasi-experimental literature addresses OVB by conditioning on relevant potential outcomes (Arnold et al. ’22; Baron et al. ’24)
  - In contrast to conventional experimental studies of direct discrimination which condition on all observable “non-race” characteristics
  - This approach yields broader measures of discrimination, which account for indirect drivers that conventional studies typically hold fixed

- This paper develops further tools to quantify indirect drivers of discrimination that conventional studies usually condition on
  - Race (mis)coding and crafting policy responses in multi-phase systems
Setting: Michigan Child Protective Services (CPS)

- 37% of U.S. children have contact with CPS by the age of 18
Setting: Michigan Child Protective Services (CPS)

- 37% of U.S. children have contact with CPS by the age of 18

- Broad mandate of CPS: place children in foster care when they are at imminent risk of maltreatment in the home; otherwise leave at home
Setting: Michigan Child Protective Services (CPS)

- 37% of U.S. children have contact with CPS by the age of 18

Broad mandate of CPS: place children in foster care when they are at imminent risk of maltreatment in the home; otherwise leave at home

- Natural *unwarranted disparity* (UD) measure: racial gaps among children with the same potential for future maltreatment in the home
Setting: Michigan Child Protective Services (CPS)

- 37% of U.S. children have contact with CPS by the age of 18

- Broad mandate of CPS: place children in foster care when they are at imminent risk of maltreatment in the home; otherwise leave at home
  - Natural *unwarranted disparity* (UD) measure: racial gaps among children with the same potential for future maltreatment in the home

- Baron et al. (2024) use quasi-random screener and investigator assignment to estimate UD at both stages and how they interact
Baron et al. (2024) Main Approach and Findings

- Idea: use quasi-random assignment to overcome selective observability of future at-home maltreatment potential (“identification at infinity”)
  - Works because placement rates are low: some screeners/investigators remove almost no kids from home, revealing unselected distributions
Baron et al. (2024) Main Approach and Findings

- Idea: use quasi-random assignment to overcome selective observability of future at-home maltreatment potential ("identification at infinity")
  - Works because placement rates are low: some screeners/investigators remove almost no kids from home, revealing unselected distributions

- Three main findings:
  1. Significant UD in the decisions of both screeners and investigators
  2. Screeners account for 13-19% of overall UD in foster care placement
  3. Placement UD is concentrated among children with subsequent maltreatment potential (i.e. high-risk cases)
New Tool 1: Exploring the Role of Racial Misclassification

- We study how racial (mis)classification by CPS investigators affects estimates of unwarranted disparity
  - Screeners access a state-wide administrative database (MIBridges) with self-reported race; investigators can re-code upon meeting the child
  - Baron et al. (2024) only have potentially re-coded CPS race data and limit samples to cases involving children coded as either white or Black
  - Concern that re-coding obscures discrimination among self-reported white vs. Black children (Luh, 2022; Finlay et al. 2024)
New Tool 1: Exploring the Role of Racial Misclassification

- We study how racial (mis)classification by CPS investigators affects estimates of unwarranted disparity
  - Screeners access a state-wide administrative database (MIBridges) with self-reported race; investigators can re-code upon meeting the child
  - Baron et al. (2024) only have potentially re-coded CPS race data and limit samples to cases involving children coded as either white or Black
  - Concern that re-coding obscures discrimination among self-reported white vs. Black children (Luh, 2022; Finlay et al. 2024)

- We link CPS data to self-reported race from educational records and extend the Baron et al. (2024) quasi-experimental approach
  - Develop a decomposition of unwarranted disparity into components that are revealed by accurate coding and obscured by miscoding
New Tool 1: Exploring the Role of Racial Misclassification

- We study how racial (mis)classification by CPS investigators affects estimates of unwarranted disparity
  - Screeners access a state-wide administrative database (MI Bridges) with self-reported race; investigators can re-code upon meeting the child
  - Baron et al. (2024) only have potentially re-coded CPS race data and limit samples to cases involving children coded as either white or Black
  - Concern that re-coding obscures discrimination among self-reported white vs. Black children (Luh, 2022; Finlay et al. 2024)

- We link CPS data to self-reported race from educational records and extend the Baron et al. (2024) quasi-experimental approach
  - Develop a decomposition of unwarranted disparity into components that are revealed by accurate coding and obscured by miscoding

- Note: we don’t assume educational records have the “ground truth” or even that race is a static/objective concept (Agadjanian 2022)
Racial Misclassification in CPS Data

- Overall white/Black shares are similar across the two datasets
- Most misclassification occurs from white/Black → multiracial/other
### Characteristics of Self-Reported White/Black Children

Well- and misclassified similar on observable characteristics

Misclassified children tend to have lower re-investigation rates

<table>
<thead>
<tr>
<th></th>
<th>Well-Classified by CPS</th>
<th>Misclassified by CPS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>White (1)</td>
<td>Black (2)</td>
</tr>
<tr>
<td>Female</td>
<td>0.479</td>
<td>0.480</td>
</tr>
<tr>
<td>Age at investigation</td>
<td>7.371</td>
<td>6.953</td>
</tr>
<tr>
<td>Had a previous investigation</td>
<td>0.494</td>
<td>0.481</td>
</tr>
<tr>
<td>No. of previous investigations</td>
<td>1.171</td>
<td>1.030</td>
</tr>
<tr>
<td>Physical abuse allegation</td>
<td>0.280</td>
<td>0.302</td>
</tr>
<tr>
<td>Alleged parent perpetrator</td>
<td>0.913</td>
<td>0.908</td>
</tr>
<tr>
<td>Alleged other relative perpetrator</td>
<td>0.045</td>
<td>0.061</td>
</tr>
<tr>
<td>Foster care placement</td>
<td>0.025</td>
<td>0.037</td>
</tr>
<tr>
<td>Re-investigated within six months</td>
<td>0.175</td>
<td>0.143</td>
</tr>
<tr>
<td>Observations</td>
<td>118,583</td>
<td>46,382</td>
</tr>
</tbody>
</table>
Decomposing Unwarranted Disparity

\[
\begin{align*}
\text{Total UD} &= E[D \mid R^* = w, Y^* = y] - E[D \mid R^* = b, Y^* = y] \\
\text{Revealed UD} &\approx E[D \mid R^* = R = w, Y^* = y] - E[D \mid R^* = R = b, Y^* = y] \omega_R \\
\text{Obscured UD} &= E[D \mid R^* = w, R \neq w, Y^* = y] - E[D \mid R^*, R \neq b, Y^* = y] \omega_O
\end{align*}
\]

- \(D\): Foster care placement; \(Y^*\): Future maltreatment potential
- \(R^*\): Self-reported race; \(R\): CPS-recorded race
 Decomposing Unwarranted Disparity

\[
\begin{align*}
\text{Total UD} & = E[D \mid R^* = w, Y^* = y] - E[D \mid R^* = b, Y^* = y] \\
\approx & \ E[D \mid R^* = R = w, Y^* = y] - E[D \mid R^* = R = b, Y^* = y] \omega_R \\
& + E[D \mid R^* = w, R \neq w, Y^* = y] - E[D \mid R^*, R \neq b, Y^* = y] \omega_O \\
\end{align*}
\]

- \( D \): Foster care placement; \( Y^* \): Future maltreatment potential
- \( R^* \): Self-reported race; \( R \): CPS-recorded race

- Builds on Bohren et al. (2023) general discrimination decomposition
  - Quasi-experimental identification, building on Baron et al. (2024)
75% of UD among self-reported Black vs. white children with future maltreatment potential is revealed by the racial codes in CPS data

25% obscured by investigators re-coding children as multiracial/other
We study how policy responses to UD can be crafted and shaped by systemic biases in multi-phase systems like CPS

- Policymakers increasingly use predictive analytic tools to guide CPS involvement (Samat et al. 2021)
- If UD accumulates over multiple phases, intervention at later stages may need to “overcorrect” in order to undo UD from earlier stages
New Tool 2: Designing Appropriate Policy Responses

- We study how policy responses to UD can be crafted and shaped by systemic biases in multi-phase systems like CPS
  - Policymakers increasingly use predictive analytic tools to guide CPS involvement (Samat et al. 2021)
  - If UD accumulates over multiple phases, intervention at later stages may need to “overcorrect” in order to undo UD from earlier stages

- We predict future maltreatment potential with machine learning models (separately by race) then estimate algorithmic UD from using different race-specific thresholds (building on Arnold et al. 2021)
  - First consider a counterfactual which lowers the white risk threshold to eliminate UD among high-risk cases at the investigation phase
  - Then consider how much lower the risk threshold would need to be to eliminate UD among high-risk cases overall
Reducing the white threshold by 5.7pp from status quo eliminates UD at the investigation phase; further 2.3pp reduction offsets screener UD.
Summary

- Studies of racial discrimination can be meaningfully impacted by conditioning on “endogenous” racial codings or on earlier decisions
  - 25% of unwarranted disparity in CPS investigator decisions is obscured by investigators’ re-coding of children’s self-reported race
  - Adjustments to algorithmic risk thresholds at the investigation phase need to be 40% larger to offset earlier UDs by CPS call screeners

The empirical tools developed & applied here may prove useful in other high-stakes settings where unwarranted disparity is a concern. E.g. criminal justice, where decisions are often made in multiple phases and racial codings can be affected by the decision-makers.

Key ingredient: quasi-experimental variation that lets researchers condition on potential outcomes (e.g. future maltreatment potential)
Summary

- Studies of racial discrimination can be meaningfully impacted by conditioning on “endogenous” racial codings or on earlier decisions
  - 25% of unwarranted disparity in CPS investigator decisions is obscured by investigators’ re-coding of childrens’ self-reported race
  - Adjustments to algorithmic risk thresholds at the investigation phase need to be 40% larger to offset earlier UDs by CPS call screeners

- The empirical tools developed & applied here may prove useful in other high-stakes settings where unwarranted disparity is a concern
  - E.g. criminal justice, where decisions are often made in multiple phases and racial codings can be affected by the decision-makers
  - Key ingredient: quasi-experimental variation that lets researchers condition on potential outcomes (e.g. future maltreatment potential)
Summary

- Studies of racial discrimination can be meaningfully impacted by conditioning on “endogenous” racial codings or on earlier decisions
  - 25% of unwarranted disparity in CPS investigator decisions is obscured by investigators’ re-coding of children’s self-reported race
  - Adjustments to algorithmic risk thresholds at the investigation phase need to be 40% larger to offset earlier UDs by CPS call screeners

- The empirical tools developed & applied here may prove useful in other high-stakes settings where unwarranted disparity is a concern
  - E.g. criminal justice, where decisions are often made in multiple phases and racial codings can be affected by the decision-makers
  - Key ingredient: quasi-experimental variation that lets researchers condition on potential outcomes (e.g. future maltreatment potential)

Thank you!