Two Papers About Child Protection

Alice Heath
March 2024
Child Protection

- Child protection systems are charged with preventing abuse and neglect
  - Primary tool: investigating and placing at-risk children in foster care

- Systems have broad reach...

- 37% of children investigated by age 18 (Kim et al. 2017)
- 5% placed in foster care (10% of Black children) (Yi, Edwards, and Wildeman 2020)

- ... important consequences ...

- Maltreatment harms children; foster care can help (Bald et al. 2022; Currie and Spatz Widom 2010; Gross and Baron 2022)
- Unnecessary removal harms children (Doyle 2007a; Doyle 2008; Roberts 2019)

- ... and large racial disparities

- 53% of Black children investigated, 29% of White
- 10% of Black children removed, 5% of White

Alice Heath 1/29
Child Protection

- Child protection systems are charged with preventing abuse and neglect
  - Primary tool: investigating and placing at-risk children in foster care

- Systems have broad reach...
  - 37% of children investigated by age 18 (Kim et al. 2017)
  - 5% placed in foster care (10% of Black children) (Yi, Edwards, and Wildeman 2020)

Maltreatment harms children; foster care can help (Bald et al. 2022; Currie and Spatz Widom 2010; Gross and Baron 2022)

Unnecessary removal harms children (Doyle 2007a; Doyle 2008; Roberts 2019)

... and large racial disparities
- 53% of Black children investigated, 29% of White
- 10% of Black children removed, 5% of White
Child Protection

- Child protection systems are charged with preventing abuse and neglect
  - Primary tool: investigating and placing at-risk children in foster care

- Systems have broad reach...
  - 37% of children investigated by age 18 (*Kim et al. 2017*)
  - 5% placed in foster care (10% of Black children) (*Yi, Edwards, and Wildeman 2020*)

- ... important consequences ...
  - Maltreatment harms children; foster care can help (*Bald et al. 2022; Currie and Spatz Widom 2010; Gross and Baron 2022*)
  - Unnecessary removal harms children (*Doyle 2007a; Doyle 2008; Roberts 2019*)
Child Protection

- Child protection systems are charged with preventing abuse and neglect
  - Primary tool: investigating and placing at-risk children in foster care

- Systems have broad reach...
  - 37% of children investigated by age 18 (Kim et al. 2017)
  - 5% placed in foster care (10% of Black children) (Yi, Edwards, and Wildeman 2020)

- ... important consequences ...
  - Maltreatment harms children; foster care can help (Bald et al. 2022; Currie and Spatz Widom 2010; Gross and Baron 2022)
  - Unnecessary removal harms children (Doyle 2007a; Doyle 2008; Roberts 2019)

- ... and large racial disparities
  - 53% of Black children investigated, 29% of White
  - 10% of Black children removed, 5% of White

Kim et al. 2017
Yi, Edwards, and Wildeman 2020
Bald et al. 2022
Currie and Spatz Widom 2010
Gross and Baron 2022
Doyle 2007a
Doyle 2008
Roberts 2019
   • Are racial disparities due to differences in underlying risk or discrimination?
   • Where in system does discrimination play a role?

   • What is the role of tragedies in driving child protection decisions?
   • How do reactions to tragedies impact racial disparities?
Baron et al: Research Question

- Puzzle: Large racial disparities in child protection
- Multi-phase system: (1) hotline call screeners, (2) investigators
- Are disparities at each stage due to **underlying risk** (OVB) or to **discrimination**?

**Figure I: Child Protection in Michigan**

Notes: The figure describes the child protection process in Michigan. Both screeners and investigators are quasi-randomly assigned, as described in the text. The percentages on screening-in and out refer to all calls received; percentages thereafter refer to investigated cases.
Unwarranted Disparities: Racial differences in screener and investigator decision rates, conditional on maltreatment potential (Arnold, Dobbie, and Yang 2018)

- Maltreatment potential:
  - Among children who will be maltreated, what is the difference in screen-in/removal rates?
    \[ \Delta_{j1} = E[D_{ij} | R_i = b, Y_i^* = 1] - E[D_{ij} | R_i = w, Y_i^* = 1] \]
  - Among children who will not be maltreated*, what is the difference in screen-in/removal rates?
    \[ \Delta_{j0} = E[D_{ij} | R_i = b, Y_i^* = 0] - E[D_{ij} | R_i = w, Y_i^* = 0] \]
  - Overall UW: \[ \Delta_j = \Delta_{j0}(1 - \mu) + \Delta_{j1}\mu \]

* Maltreatment here = re-investigated within 6 months
Identification at Infinity

Problem: Selective Observability

- Only observe potential maltreatment among children who are left in home
- For children removed, can’t see if they would have been maltreated
- Have all information to estimate UD except rates of Black and White maltreatment in population

Intuition for Solution:

- Imagine a randomly assigned screener who screens everyone out
- Among these children, we can see maltreatment potential
- Since screener was randomly assigned, this is a good estimate for maltreatment rates in the population
- In absence of this screener, can extrapolate from workers with low placement rates

Child protection is a good setting for this strategy because most children are not removed

- Implement using quasi-randomly assigned investigators in Michigan
Panel A: All Hotline Calls
Panel B: Screened-in Calls
Complement with non-parametric bounds

- Upper bound: assume all children placed would have been maltreated
- Lower bound: assume all children placed would not have been maltreated
- Bounds relatively tight because very few children are placed
Findings: Unwarranted Disparities

- **White children more likely to be maltreated**
  - Both among children reported to hotline (14.1% vs. 12.9%), and among screened-in calls (17.5% vs. 15.5%)

- **UD in call screens: Black children 5pp more likely to be screened in**
  - (12% disparity relative to average of 60%)

- **UD in investigations: Black children 1.7pp more likely to be investigated**
  - (50% disparity relative to average of 3.4%)

- → **UD in placements: Black children 1.1pp more likely to be removed**
  - 55% larger than average among all calls.

Alice Heath
Findings: Unwarranted Disparities

- **White children more likely to be maltreated**
  - Both among children reported to hotline (14.1% vs. 12.9%), and among screened-in calls (17.5% vs. 15.5%)

- **UD in call screens: Black children 5pp more likely to be screened in**
  - (12% disparity relative to average of 60%)

- **UD in investigations: Black children 1.7pp more likely to be investigated**
  - (50% disparity relative to average of 3.4%)

- → **UD in placements: Black children 1.1pp more likely to be removed**
  - 55% larger than average among all calls.

- Majority of UD in placements comes from investigations.
- Disparities larger than controlled observational disparities
Table II: Estimates of Mean Maltreatment Risk and Unwarranted Disparity

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Subsequent maltreatment risk</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>All calls</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black children</td>
<td>0.129</td>
<td>0.155</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>White children</td>
<td>0.141</td>
<td>0.175</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.003)</td>
</tr>
<tr>
<td><strong>Panel B: Unwarranted disparity (UD)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Screeners</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average across decision-makers</td>
<td>0.050</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.002)</td>
</tr>
<tr>
<td><strong>Investigators</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel C: Placement UD and decompositions</strong></td>
<td>Equation (8)</td>
<td>Equation (9)</td>
</tr>
<tr>
<td>Placement UD</td>
<td>0.011</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Screener share (%)</td>
<td>12.5</td>
<td>18.6</td>
</tr>
<tr>
<td></td>
<td>(2.5)</td>
<td>(2.0)</td>
</tr>
<tr>
<td>Investigator share (%)</td>
<td>87.5</td>
<td>81.4</td>
</tr>
<tr>
<td></td>
<td>(2.5)</td>
<td>(2.0)</td>
</tr>
<tr>
<td>Number of screeners</td>
<td>162</td>
<td>162</td>
</tr>
<tr>
<td>Number of investigators</td>
<td>814</td>
<td>814</td>
</tr>
</tbody>
</table>
Unwarranted disparities exist. Who drives them and who do they hurt?

- Investigators amplify UDs among high risk children: UD is 5.8pp
- Investigators mitigate UDs among low risk children: UD is 0.8pp
- Racial concordance plus more white investigators plays a large role: investigators more likely to give benefit of the doubt to concordant families?
Unwarranted disparities exist. Who drives them and who do they hurt?

- Investigators amplify UDs among high risk children: UD is 5.8pp
- Investigators mitigate UDs among low risk children: UD is 0.8pp
- Racial concordance plus more white investigators plays a large role: investigators more likely to give benefit of the doubt to concordant families?

Welfare Implications: under-removing White children?

- Future maltreatment potential decreases with investigator placement rates
- In Michigan, positive effects of foster care for marginal children (Gross and Baron, Baron and Gross)
- These effects are larger for White children → marginal White child is higher-risk / has more to gain from removal
Panel B: Maltreatment potential ($Y^* = 1$)
Lessons / Discussion

- UD/Discrimination exists—driven by under-removal of White children?
  - Counters received wisdom in field that Black children are over-removed?
  - (Peter Hull characterized the findings as evidence that CPS discriminates against Black children)
  - What are channels for these findings to reach practitioners?

- Application of “identification at infinity” and bounds
  - Are they convincing?
  - Other settings where these techniques could be used to examine discrimination?

- Multi-phase systems
  - Landed on this focus late.
  - Where else are there multi-phase systems?
Two Papers Today

   - Are racial disparities due to differences in underlying risk or discrimination?
   - Where in system does discrimination play a role?

   - Role of media attention in driving child protection decisions.
   - Examine impact on racial disparities.
1,750 maltreatment deaths occur each year in the US.
1,750 maltreatment deaths occur each year in the US.
1,750 maltreatment deaths occur each year in the US.
This Paper: Child Protection Responses to Maltreatment Deaths

Her smile hides a life of unimaginable pain. When seven-year-old Nixmary died, she weighed less than a four-year-old. Everyone — school, family, child workers — failed her. Her story shames all who could have saved her.

Experts fear child abuse deaths may spawn ‘foster-care panic’

By David Carr

NEW YORK — There is no argument that Nixmary Brown’s death was horrific: the 7-year-old Brooklyn girl was allegedly tortured and beaten by her stepfather. There is, however, concern and contention over the next phase for New York City’s child protection agency and the families it monitors.

Some child-welfare advocates worry that the tragedy is being compounded by what they call “foster-care panic.” The phenomenon, which has occurred in several states over the years, happens in high-profile cases after an abused child in a family already under scrutiny is killed and authorities then sharply increase the number of children removed from their parents.

Often, some experts argue, the result is a harmful overreaction to public pressure. Too many children have been worse as a more positive presence in the community — that’s huge progress,” Arsham said. “We sit now right on the cusp of losing that if we’re not careful.”

Arsham’s staff includes Violet Rittenhour, who said her two children were taken into foster care for a year in 2001 because caseworkers felt, that as a single mother juggling work and college, she couldn’t properly care for them.

Despite measured statements from Mayor Michael Bloomberg and other officials, Rittenhour believes most frontline child protection workers have been affected by intense news coverage of Nixmary’s death and are so wary of being blamed for a similar tragedy that they unnecessarily recommend removal of children.

“They’re still acting on the premise of, ‘When in doubt, remove,’” Rittenhour said.

Austin insisted that New York
Research Questions

1. Do child protection agencies react to maltreatment deaths?

2. If they do react, is the reaction due to information or scrutiny?

3. Are reactions to maltreatment deaths:
   A. Well-calibrated and likely to benefit children, or;
   B. Poorly calibrated and perhaps harming children?
Research Questions

1. Do child protection agencies react to maltreatment deaths?

2. If they do react, is the reaction due to information or scrutiny?

3. Are reactions to maltreatment deaths:
   A. Well-calibrated and likely to benefit children, or;
   B. Poorly calibrated and perhaps harming children?

Use full-text newspaper archives to identify publicized maltreatment deaths across the US between 1999 and 2019, then employ staggered adoption event study to assess impact.
Investigator Decision

Investigator $j$ faces cost $C_{ij}$ for each child $i$:

$$C_{ij}(\text{Remove}_{ij}, p_i) = a_j \text{Remove}_{ij} + p_{ij}(x_i, \pi_j)b_j(1 - \text{Remove}_{ij}),$$

where $a_j$ and $b_j$ are the social costs due to removal and maltreatment, respectively.

(Doyle 2007b; Kleinberg et al. 2018)

Social-welfare maximizing investigator uses social costs, $a^S$ and $b^S$, perfectly assesses child risk $p_i$, and removes child when $p_i$ is greater than threshold:

$$p_i(x_i, \pi) > \frac{a^S}{b^S} = T_{OPT}.$$
Targeted or Haphazard Response

(a) Targeted: Reduced Error

(b) Targeted: Threshold Shift

(c) Haphazard: Increased Error
Highly-Publicized Maltreatment Deaths Included in Analysis

60 events between 2003 and 2018 across 55 jurisdictions

- 52 about maltreatment deaths. Use all 60 for main analysis.

Figure 3: Events used in analysis
Administrative Data to Construct Outcome Data

- **Child protection response**
  - **Removals** from AFCARS and NCANDS, state-submitted admin data
  - **Pre- and post- removal pipeline**: reports, screened-in reports, TPR
  - Characteristics of child, parents, and maltreatment allegation

- **Child health outcomes**
  - **Mortality** from National Vital Statistics System
  - **Hospitalizations** from State Inpatient Database

- **Time-varying controls**
  - Adult opioid deaths from NVSS
  - Unemployment from BLS
Empirical Approach: Adapted Staggered Adoption Event Study

- Key insights from recent staggered adoption literature:
  - Dynamic treatment effects contaminate estimates (Callaway and Sant’Anna 2020; De Chaisemartin and d'Haultfoeuille 2020; Sun and Abraham 2020)
  - Unbalanced panels generate non-intuitive weights (Goodman-Bacon 2021)
  - → Use clean controls and balanced panels

Alice Heath
Empirical Approach: Adapted Staggered Adoption Event Study

- Key insights from recent staggered adoption literature:
  - Dynamic treatment effects contaminate estimates \((\text{Callaway and Sant’Anna 2020; De Chaisemartin and d'Haultfoeuille 2020; Sun and Abraham 2020})\)
  - Unbalanced panels generate non-intuitive weights \((\text{Goodman-Bacon 2021})\)
  - \(\rightarrow\) Use clean controls and balanced panels

- In my adaptation, control states include both:
  - “Never treated” jurisdictions: 14 with no event in time period
  - “Not yet nor recently treated” jurisdictions: no event in previous 3 years \((\text{Lafortune, Rothstein, and Schanzenbach 2018})\)
Empirical Approach: Adapted Staggered Adoption Event Study

- Key insights from recent staggered adoption literature:
  - Dynamic treatment effects contaminate estimates ([Callaway and Sant’Anna 2020; De Chaisemartin and d'Haultfoeuille 2020; Sun and Abraham 2020])
  - Unbalanced panels generate non-intuitive weights ([Goodman-Bacon 2021])
  - Use clean controls and balanced panels

- In my adaptation, control states include both:
  - “Never treated” jurisdictions: 14 with no event in time period
  - “Not yet nor recently treated” jurisdictions: no event in previous 3 years ([Lafortune, Rothstein, and Schanzenbach 2018])

- Identifying assumptions:
  - Parallel trends, no anticipation, no dynamic treatment effects after 3 years
Removal Rates Increase by 19% Following a Highly-Publicized Tragedy

Figure 4: Impact of Highly-Publicized Tragedies on Removal Rate
Scrutiny or Information: Smaller Response to Less-Publicized Tragedies

Figure 5: Impact on Removal Rates (Percent Change): Highly-Publicized Tragedies in Front of Newspaper and Not in Front of Newspaper.

→ Agency reaction driven primarily by scrutiny, not information
Figure 6: Long Difference Estimates (Rate per 10,000 Children Age 0-9)
Estimate Removal Rate Change for Children in Each Decile

Long-difference approach used in main results for each risk decile.

**Figure 7:** Removal Probability by Predicted Risk Decile in Test Dataset
Long-difference approach used in main results for each risk decile.

Figure 8: Removal Probability by Predicted Risk Decile and Race in Test Dataset
1. Agencies increase removal rates by 19% following highly-publicized tragedies
   - Little reaction to less-publicized tragedies.
   - Reaction primarily driven by scrutiny.
1. Agencies increase removal rates by 19% following highly-publicized tragedies
   - Little reaction to less-publicized tragedies.
   - → Reaction primarily driven by scrutiny.

2. Reactions appear well-calibrated:
   - Removal rates increase most for children in top risk decile.
   - Hospitalizations decline among Medicaid population.
   - → Well-calibrated, but hard to rule out overreaction.
Summary of Findings

1. Agencies increase removal rates by 19% following highly-publicized tragedies
   - Little reaction to less-publicized tragedies.
   - → Reaction primarily driven by scrutiny.

2. Reactions appear well-calibrated:
   - Removal rates increase most for children in top risk decile.
   - Hospitalizations decline among Medicaid population.
   - → Well-calibrated, but hard to rule out overreaction.

3. Average effects mask heterogeneity by race:
   - Black removal rates increase more even conditional on predicted risk.
   - 25% increase in already large Black-White gap.
   - → Scrutiny induces some mis-calibration.
Both papers suggest removal decisions by race are miscalibrated
  - Discrimination paper makes welfare claim: white children under-removed
  - Hard to make strong welfare claims in tragedies paper beyond miscalibration

If UD exist at baseline, reactions to tragedies don’t appear to close them and may widen them?

If had worker data within a state, could examine how UD change following a tragedy

Responses to tragedies by worker race