Government Reactions to Tragedy: How Maltreatment Deaths Impact Child Protection*

Alice Heath†

This Draft: January 1, 2024
[First Draft: October 5, 2023]

Please click here for the latest version.

Abstract

Governments often change policy or practice following a tragedy. In theory, governments should use information from tragedies to optimally update. In practice, their reactions may not be well-calibrated and may depend on whether the tragedy is spotlighted in the media. This paper examines how child protection agencies react to maltreatment deaths and assesses the consequences for welfare. I first analyze newspaper archives to construct a dataset of publicized maltreatment deaths between 1999 and 2019. I then employ a staggered adoption event study to identify the impact of a death on child protection systems and child outcomes. Agencies react sharply to highly-publicized deaths, increasing removals by 19 percent. There is no detectable reaction to less-publicized deaths, suggesting agencies respond primarily to scrutiny rather than information. Highly-publicized deaths induce an increase in removals among children with the highest predicted risk of maltreatment and hospitalizations for maltreatment-related diagnoses among Medicaid recipients decline. But Black children's removal rates rise more than White children's even conditional on risk, increasing the Black-White removal rate gap. Agency reactions to tragedies therefore do not appear to be optimal, though parts of their reactions may be welfare-enhancing.

*Thank you to Jeffrey Liebman, Marcella Alsan, Mark Shepard and David Cutler for ongoing advice and support. Thank you also to Chris Avery, Jason Baron, Marianne Bitler, Tridevi Chakma, Will Dobbie, Alex Hoagland, Jennifer Hochschild, Kristen McCormick, Danny Schnieder, Richard Zeckhauser, and many workshop and seminar participants for advice and feedback. Thank you to Adrian Hackney for superb research assistance. This research has been supported by a James M. and Cathleen D. Stone PhD Scholar fellowship from the Multidisciplinary Program in Inequality and Social Policy at Harvard University. The data used in this publication, AFCARS Foster Care Files (2000–2019) and NCANDS Child and Agency Files (2000-2019), were obtained from the National Data Archive on Child Abuse and Neglect and have been used in accordance with its Terms of Use Agreement license. The Administration on Children, Youth, and Families, the Children's Bureau, the original data set collection personnel or funding source, NDACAN, Cornell University and their agents or employees bear no responsibility for the analyses or interpretations presented here.

†Harvard Kennedy School. Email: alice_heath@hks.harvard.edu

1 Introduction

Governments often change policy or practice following a tragedy. After the 9/11 attacks, governments dramatically increased airport security requirements and surveillance of Muslim communities. Police departments have reduced arrests following highly-publicized killings of Black men and women at the hands of police (Devi and Fryer 2020; Premkumar 2019). In 2021, Massachusetts banned open-water swimming in Walden Pond after a spate of drownings was reported in the Boston Globe.

In theory, governments should use all information, including what they learn from tragedies, to continuously update beliefs and adjust policy and policy implementation accordingly. In practice, governments may fail to update appropriately due to budget constraints, distorted incentives, or behavioral biases. In these circumstances, the scrutiny that accompanies a *highly-publicized* tragedy might provide a political opportunity or disciplining effect that improves decision-making. Conversely, scrutiny might activate biases or distort incentives, leading governments to make worse policy and practice decisions.

This paper examines how governments react to tragedies in the context of child protection. I investigate how local child protection agencies across the US react to maltreatment deaths of children in their jurisdiction and assess the consequences of their reactions for resident welfare. Do agencies increase removal rates following maltreatment deaths, and if yes are they reacting to information or scrutiny? Moreover, are agencies' reactions to maltreatment deaths well-calibrated: do they increase removals among the highest-risk children, preventing maltreatment and improving child outcomes? Or are agencies' reactions haphazard: do they also increase removals among children at low risk of maltreatment, which is less likely to benefit children and may harm them? I answer these questions by identifying how highly- and less-publicized maltreatment deaths impact child protection removal rates, the risk of children removed, and child health outcomes.

The US child protection system interacts with many children and involves high-stakes decisions, so it is a setting well-suited to studying how tragedies influence government decision-making. Charged with protecting children from maltreatment, agencies' primary tool is removing children they assess to be at high risk of maltreatment from their parents. Child maltreatment causes immediate and long-term harm (Currie and Spatz Widom 2010; Currie and Tekin 2012; Doyle and Aizer 2018), and removal can protect a child from those harms (Bald et al. 2022; Gross and Baron 2022). But removal - especially of lower-risk children - can also harm children, causing trauma and poor long-term outcomes (Doyle 2007; 2008). In the US, 3.5 million children are the subject of a child protection investigation each year, one in three will experience a child protection investigation before age 18, and one in twenty will enter foster care (KidsCount 2020; Wildeman and Emanuel 2014). These children are disproportionately Black, Native American and low-income (Bullinger, Raissian, and Schneider 2022; Drake and Pandey 1996; Yi, Edwards, and Wildeman 2020).[1] There is

[1]The cause of these disparities - which might reflect appropriate responses to risk or biased decision-making - is

a widely-held view among child protection practitioners that highly-publicized tragedies drive large swings in agency practices (see e.g. Lepore 2016), but limited work has examined this quantitatively.

To formalize hypotheses, I develop a stylized conceptual framework that outlines possible agency responses to tragedies. I focus on investigators, who assess children's maltreatment risk and remove them when the expected cost of removal is less than that of leaving them in their home (Doyle 2007). Socially optimal removal decisions can be distorted by error and subject to resource constraints. Child deaths can provide information that reduces error or a political opportunity that relaxes resource constraints, improving removal decisions. This response might be conditional on the death being accompanied by media scrutiny. Alternatively, child deaths - particularly those accompanied by media scrutiny - might activate biases, making removal decisions worse.

The framework yields two principal questions. First, do investigators use information from all deaths to continuously update their assessments of child risk and adjust decision-making, or do they react primarily to the scrutiny that accompanies a highly-publicized death? Second, when investigators do react to deaths, is their response well-calibrated and targeted at children with highest maltreatment risk, or is their response haphazard, also increasing removals among children with low risk of maltreatment?

To answer these questions, I first construct a new dataset of publicized child protection tragedies. I develop a text-analysis algorithm that searches full-text archives of 108 local newspapers across the US from 1999 to 2019. The algorithm identifies 60 distinct events where coverage of stories related to a child death and child protection was unusually high.

I then use several individual-level administrative datasets containing child protection and health outcomes to examine the impact of the tragedies. Individual- and state-level data from the National Data Archive on Child Abuse and Neglect provide information on agency removal decisions, reporting from the public, investigations, and substantiated maltreatment. Vital statistics data from the CDC provides information on child mortality. State Inpatient Data from the Healthcare Cost & Utilization Project provides information on inpatient stays for maltreatment-related diagnoses, including injuries and psychiatric diagnoses. An advantage of these outcome data is that mortality and hospitalizations are observable for the whole population, not only children reported to the child protection system.

My difference-in-differences identification strategy adapts recent staggered adoption event study methods (e.g. Callaway and Sant'Anna 2020; Cengiz et al. 2019; Sun and Abraham 2020) to my context, in which some jurisdictions experience multiple highly-publicized tragedies. The approach compares the time path of outcomes in each jurisdiction that experienced a highly-publicized tragedy to outcomes in event-specific sets of control jurisdictions that did not experience a highly-publicized tragedy. Control states include both never-treated and not-recently-treated jurisdictions; I exclude

the subject of exploration in economics and other fields (Baron et al. 2023; Jason Baron, Goldstein, and Ryan 2023; Putnam-Hornstein et al. 2013; Roberts 2022; Wulczyn et al. 2013) jurisdictions that recently experienced a highly-publicized tragedy from the set of controls to prevent dynamic treatment effects from contaminating the estimates (Callaway and Sant'Anna 2020).[2]

The analysis yields several results. First, agencies have large responses to highly-publicized tragedies, sharply increasing removal rates following a month with high news coverage. The rate that children age 0-9 are removed from their home for the first time increases by 19 percent in the 24 months after a tragedy. This increase is primarily driven by decisions to substantiate maltreatment and remove children, conditional on a report being made from the public and screened in.

A 19 percent increase in removal rates is large and unusual. On average, removal rates within jurisdictions *decrease* by 1.8 percent in years without a highly-publicized death. Jurisdictions also do not appear to have significant reactions to to less-publicized tragedies. On average, three children die every month in each US state, but jurisdictions only increase their year-on-year removal rates by more than 19 percent in 3.9 percent of control years. Moreover, the magnitude of the reaction increases with the intensity of news coverage; when there are fewer newspaper stories or when the story is not published in the first five pages of the newspaper, agencies increase removals by a much smaller amount or not at all. Agencies therefore adjust decision-making much more in response to highly-publicized deaths than to less-publicized deaths, suggesting that they react primarily to scrutiny rather than information.

Next, I assess whether agencies' reactions to highly-publicized deaths are well-calibrated. Agencies should target removals at children at high risk of maltreatment, for whom the expected social cost leaving them in their home is higher than the social cost of removal. Removing these children would prevent maltreatment and reduce adverse health outcomes. Alternatively, agencies might have a haphazard response, in which they also increase removals among children at lower risk of maltreatment, for whom the benefits of removal do not outweigh the costs. I assess which of these scenarios is more likely by estimating how removals change by predicted risk decile, and by estimating the impact on children's mortality and hospitalization rates for injury, psychiatric and maltreatment-related diagnoses.

Several pieces of evidence suggest that agencies remove more of the highest-risk children following a highly-publicized tragedy. First, direct analysis of future maltreatment risk reveals that agencies target increased removals at the top of the risk distribution. Applying machine learning techniques, I use jurisdiction-specific logistic lasso models to generate a risk score for each child screened in for an investigation. The score represents the predicted likelihood that the child would be re-investigated by the child protection agency within 6 months if they were left in their home (Baron et al. 2023). After a tragedy, removals increase the most among children in the highest risk decile, and do not increase among children in the lowest four risk deciles.

[2] In addition to the standard no-anticipation and parallel-trends assumptions, my adapted approach where jurisdictions experience more than one event requires one additional assumption: that dynamic treatment effects stabilize after 3 years. The results are robust to specifications that rely on different assumptions and sets of control jurisdictions.

Changes in child health outcomes also provide evidence that more of the highest-risk children are removed following highly-publicized deaths. Among Medicaid recipients, a subset of children particularly likely to interact with the child protection system, there is a reduction in child hospitalizations for injuries, for psychiatric diagnoses, and for a more specific set of diagnoses that scholars have identified as highly likely to be due to maltreatment. There is no detectable change in these hospitalizations among children with private insurance, who are less likely to interact with the child protection system. These results suggest that following a highly-publicized death, agencies remove high-risk children who would otherwise have been hospitalized. Six additional removals are associated with one fewer inpatient visit with an injury diagnosis.

Increased removals among the highest-risk children and the decrease in hospitalizations among Medicaid recipients suggest that elements of agencies' reactions are welfare-enhancing. There is limited evidence of a haphazard response: removals do not increase at all among children with lowest predicted re-maltreatment risk. Taken together, these findings are consistent with agencies having a well-calibrated response. A definitive welfare analysis requires knowing where the optimal removal threshold lies: this would determine whether some of the increased removals have negative social benefit, and whether those costs outweigh the benefits for children who are helped. This is beyond the scope of this paper, but the findings suggest agencies respond purposefully to the scrutiny they experience when a maltreatment death is highly-publicized.

These promising results are complicated by my finding that a highly-publicized death increases the already large racial gap in removal rates, and this does not appear to be driven by differences in underlying risk. Black children's removal rates are much higher at baseline - the baseline removal rate for Black and White children is 62.7 and 28.3 per 10,000 respectively. Following a highly-publicized death the Black-White gap increases by 8.67 children per 10,000, or 31 percent of the baseline gap. The change persists for at least 24 months. The increased gap could reflect agencies appropriately adjusting removal rates if marginal Black children were at higher risk before a tragedy. But in control periods, Black children who are left at home have lower risk; White children left at home are more likely to be re-investigated.[3] Moreover, following a highly-publicized tragedy Black children's removal rates increase more than White children's even after conditioning on risk level. Importantly, this result does not shed light on whether Black or White children should optimally be removed at higher or lower rates.[4] It does suggest that highly-publicized tragedies may induce updating error that varies for Black and White children - perhaps driven by racial stereotypes (Bordalo et al. 2016) - and results in an increase in the Black-White removal rate gap.

My findings demonstrate that local media plays an important role in driving state and local child protection practices. Agencies remove more children following highly-publicized tragedies and

[3] This is consistent with recent findings from Baron et al. (2023).

[4] It is possible that the removal rates of both Black and White children should be lower or higher; the optimal threshold depends on the social costs of removal and maltreatment. Recent work by Baron et al. (2023) and suggest that White children may be under-removed relative to Black children, given their respective risk of maltreatment.

respond much less, if at all, to tragedies that are less publicized. This is evidence that agencies do not use all information to continuously update policy and practice, but instead are primarily responsive to scrutiny. Agencies remove more high-risk children following highly-publicized tragedies, do not increase removals among the lowest risk children, and hospitalizations for maltreatment-related diagnoses, including injury and psychiatric diagnoses, decline among the Medicaid population. This is suggestive evidence that media scrutiny improves welfare for some children, perhaps by increasing resources available to agencies or focusing agency decision-making. However, racial disparities in removal rates increase even after conditioning on risk, which suggests the scrutiny induces some mis-optimization. The paper sheds light on how local media can impact decision-making and citizen welfare, and underscores the importance of judicious responses to tragic events by both government and the media.

Related literature. My paper contributes to a growing economic literature focused on child protection. Recognizing the large potential welfare impacts of child removal, several papers use quasi-randomly assigned investigators to study the effect of entry to foster care for children on the removal margin. Doyle (2007), 2008 and Roberts (2019) found that removal caused worse outcomes for marginal older children; papers studying other jurisdictions have found that removal improves the outcomes of marginal younger children (Bald et al. 2022; Gross and Baron 2022; Roberts 2019). These divergent results suggest that removal thresholds are different across jurisdictions and demographic groups, which raises the question of how agencies set those thresholds. This paper identifies a highly-publicized child death as an event that shifts a jurisdiction's removal threshold downwards and examines the consequences for welfare.

My paper also contributes to a literature that examines how media attention impacts a range of outcomes, including voting (Snyder, Jr., and Strömberg 2008), consumer behavior (DellaVigna and Gentzkow 2010), and stock prices (Kreitmeir, Lane, and Raschky 2020). Scholars have also shown that government and other public officials react to media attention focused on natural disasters with increased natural disaster relief (Eisensee and Strömberg 2007), on police killings with reduced police activity (Devi and Fryer 2020), and on crime with more punitive criminal sentences (Lim, Snyder Jr., and Strömberg 2015; Philippe and Ouss 2018). I show media attention focused on maltreatment deaths increases child protection activity and assess the quality of the response.

The theory that publicized maltreatment deaths drive child protection practices is often cited by practitioners and has been studied by some scholars. Many papers are qualitative (e.g. Chenot 2011; Thomlison and Blome 2012; Warner 2014); some have taken a quantitative approach. Gainsborough (2009), 2010 found scandals between 2002 and 2004 were associated with new legislation but no change in spending. Jagannathan and Camasso (2017) and Camasso and Jagannathan (2019) find that annual maltreatment fatalities are associated with increased child protection activity, but their empirical approach does not rule out confounding factors or reverse causality. Warburton et al.

(2014) use an interrupted time series to show that one highly-publicized death in Ontario increased removal rates of older children, decreased high school graduation rates and increased welfare receipt. My paper is the first to provide a conceptual framework that grounds agency responses to tragedy theoretically. I also develop a novel approach to identify multiple highly-publicized deaths across the US over two decades and study their impact using a robust identification strategy.

Finally, my paper contributes to work on inequalities that arise in high-stakes decisions. Child protection is setting with large racial disparities (Yi, Edwards, and Wildeman 2020).[5] Several recent papers show that unwarranted disparities by race exist in child protection (Baron et al. 2023) and criminal justice (Arnold, Dobbie, and Yang 2018; Kleinberg et al. 2018), and other papers show that information can cause updating that increases race and gender disparities (Angelova, Dobbie, and Yang 2022; Eren and Mocan 2018; Sarsons 2017). I find that media stories lead to disparate outcomes by race in the child protection context.

I begin the paper with background and a motivating example of a high-profile maltreatment death in New York City in 2005. I then outline my conceptual framework, which models investigator decisions and presents hypothetical government responses to tragedies. The methodology section describes how I construct a new dataset of newspaper stories about child deaths, the administrative data I use to measure the child welfare system responses and child outcomes, and the adapted staggered adoption approach that identifies the impact of the deaths. Finally I present the results and discuss the implications for resident welfare, the media, and government.

2 Background and Conceptual Framework

Child protection agencies in the US are charged with preventing child abuse and neglect in their jurisdictions; agencies are administered at either the state or county level. In all jurisdictions, children follow a similar pathway through the system. First, members of the public report suspected maltreatment to a hotline. Hotline screeners then decide whether to screen the report in for investigation. When reports are screened in an investigator will gather evidence, determine whether the report is "substantiated" - meaning the child was a victim of maltreatment - and decide whether the child should be removed from their family and placed in foster care.

Removing a child from their home is one of the most far-reaching interventions made by governments. Once a child is removed from their parents, the average length of stay in foster care is 21 months. Children in foster care live with other family members, with professional foster parents, or in institutional settings including group homes, hospitals and residential care facilities. While in foster care, child protection case workers continue to work with the child with the aim of finding a

[5]These disparities are just beginning to be explored by economists. Baron et al. (2023) find that racial disparities are driven by child protection workers screening in and removing high-risk Black children at higher rates, and conclude that White children may be under-removed. Jason Baron, Goldstein, and Ryan (2023) find that blind removals do not change racial disparities in decision-making.

permanent, safe family. The initial goal is usually to reunify the child with their parents. When this is not possible some children are adopted, and others remain in foster care until they are eighteen.

In the US, more than 3.5 million children are the subject of a report of maltreatment each year, around 50 percent of those reports are screened in for an investigation, and 5 percent of investigated reports result in a child being removed (Yi, Edwards, and Wildeman 2020). More than 200,000 children are removed from their homes each year, and one in twenty US children will spend some time in foster care before they are eighteen. These rates vary widely across states: as a portion of all children in the state, removal rates varied from fewer than 20 per 10,000 in Virginia, to 130 per 10,000 in West Virginia in 2018.

2.1 Motivating Example

On January 11 2006, 7-year-old Nixzmary Brown was found beaten to death in her family's Brooklyn apartment. Nixzmary's mother and stepfather were indicted for murder, manslaughter, assault, unlawful imprisonment, reckless endangerment, and sexual abuse. New York City's Administration for Children's Services (ACS) had received two hotline reports in May and December 2005 alleging physical abuse and educational neglect, and additional calls from school officials who had observed that Nixzmary was absent from school and suffering injuries of increasing severity. CPS investigated the family following the hotline calls but did not remove Nixzmary from her home.

The New York Post ran front page stories about Nixzmary's death for nine consecutive days and continued coverage through May. Headlines - some illustrated in Figure 1 - filled the front page and included "Shame on Them All", "Will She Get Justice?" and "*Now* City Acts!".

On January 18, ACS disciplined six workers, including three who were suspended without pay. On January 24, Mayor Bloomberg announced the creation of an interagency task force, a Department of Investigation examination into deaths of children known to ACS, and $16 million of additional funding for staff and training. In early February, City Council assembly members called on the ACS Director to resign. Over the following months, several city committees and task forces released reports and ACS reviewed all their open cases, hired staff, reorganized leadership, changed policy around investigations, instituted a new performance management system, and partnered with the New York Police Department and Department of Education to increase responses to school absences.

At the same time, removal rates in New York City increased sharply. Figure 1 illustrates the monthly number of removals in the City. After five years of steady decline - a common trend as agencies try to keep children safe with their families - removals doubled from 1,944 in 2005 to 3,876 in 2006. The increase was immediate and removals remained almost twice as high as before Nixzmary's death for the next five years.

Following this tragedy, the City introduced a requirement that ACS must report the annual number of deaths of children known to the agency. Since 2011, more than 50 children known to the agency have died each year. If agencies were using deaths to update their assessment of the risk of death, we would expect them to increase removal rates after every death. But the agency only increased removals following Nixzmary's death, which received intensive newspaper coverage. In this paper I investigate whether agencies across the US have larger reactions to highly-publicized tragedies than to less-publicized tragedies, examine whether agencies' reactions to deaths are well-calibrated, and explore the impact on child well-being.

2.2 Conceptual Framework

To develop hypotheses, I develop a stylized conceptual framework that builds on Doyle (2007)'s and Kleinberg et al. (2018)'s models of decision-making. For simplification, I focus on the last stage in the child protection pipeline: the investigator's removal decision. Child welfare investigator j observes child i and decides whether to remove the child or leave them in their home. I denote $Remove_{ij} = 1$ if the investigator chooses to remove the child, and $Remove_{ij} = 0$ if they do not remove the child. Families are drawn from some distribution of risk for future maltreatment Y, where $Y_i = 1$ if maltreatment is inflicted by the child's parents and $Y_i = 0$ otherwise. An investigator assesses the likelihood of maltreatment if the child is left at home, $p_{ij}(x_i, \pi) = Pr(Y_i = 1|Remove_i = 0)$, which is a function of the family's characteristics x_i and the likelihood of maltreatment in the population π.

In this framework, removing the child guarantees there is no future maltreatment from the child's parents, $Y_i = 0$.[6] But removing the child comes with some cost. This includes government costs, such as foster care payments and social worker time, and the short and long-term harm that both the family and child suffer when a child is removed from their home.[7] Investigator j therefore has the following payoffs for each family:

$$C_{ij}(Remove_{ij}, p_i) = \underbrace{a_j Remove_{ij}}_{\text{Harm due to removal}} + \underbrace{p_{ij}(x_i, \pi) b_j (1 - Remove_{ij})}_{\text{Harm due to maltreatment}}, \tag{1}$$

where a_j is the cost the investigator assigns to harm from maltreatment, and b_j is the cost the investigator assigns to harm from removal.

Optimal removal decisions: A fully informed, social-welfare maximizing investigator in an unconstrained agency chooses $Remove_{ij}$ for each family to minimize cost C_{ij}. A social welfare-optimizing investigator will use social costs, $a_j = a^S$ and $b_j = b^S$ and perfectly assess child risk, $p_{ij} = p_i$. If they remove the child, $C_{ij} = a^S$. If they do not remove the child, $C_{ij} = p_i b^S$. They

[6] In practice, children can still be maltreated by their parents after a removal, for example during visitation. In the model the cost associated with this can be incorporated into the cost of removal.

[7] These harms include trauma that results from removal and the risk of maltreatment while in foster care.

therefore remove the child if:

$$p_i(x_i, \pi) > \frac{a^S}{b^S} = T_{OPT}. \tag{2}$$

An investigator will remove families with risk of maltreatment, p_i, greater than the optimal threshold $T_{OPT} = a^S/b^S$, which is determined by the ratio of removal and maltreatment costs.

As investigators observe children's outcomes, they should update their prior belief about the underlying risk of maltreatment in the population π_0 to π_t.[8] They use this updated belief to inform their assessments of individual child risk. Using a simple Bayesian learning model:

$$E(\pi|Y_n, n) = \frac{Y_n + \alpha}{n + \alpha + \beta}, \tag{3}$$

where n is the number of children observed, Y_n is the number of those children who were maltreated, and α and β are fixed parameters that determine the prior belief $\pi_0 = \frac{\alpha}{\alpha+\beta}$. Equation (3) suggests that as investigators observe more children, the value of each additional observation becomes small. Continuously optimizing investigators should therefore only make small changes to their decision threshold following a single observation.

Investigator error: Error enters investigators' decision-making through estimates of maltreatment risk and social costs. Investigators estimate maltreatment risk $\hat{p}_i = p_i + \varepsilon_{ij}$, where ε_{ij} is a draw from some distribution that could be a function of both investigator and family characteristics. Investigator estimates of social costs may be distorted by behavioral factors including salience and over-weighting of personal costs.[9] An investigator's assessment of the ratio of harm from removal and harm from maltreatment therefore becomes $\frac{\gamma_j a^S}{b^S}$, where $\gamma_j > 0$ and reflects the distortion. For socially optimizing investigators, $\gamma_j = 1$ and there is no distortion. If $\gamma_j > 1$, the investigator over-weights the harm of removal and removes too few children, and if $\gamma_j < 1$ the investigator over-weights the harm from maltreatment and removes too many. Investigators with error will remove children for whom:

$$p_i > \frac{\gamma_j a^S}{b^S} - \varepsilon_{ij} = T_{SUBOPT_{ij}}. \tag{4}$$

[8] I assume a tragedy in which a child dies provides a signal of the risk of maltreatment in the population; it does not change the social costs associated with maltreatment or removal. The social cost associated with a child's removal includes staff costs, trauma to child, and the cost of foster parents; the social cost associated with a child's maltreatment are short- and long-term health, impacts on child development and costs of medical care. I assume the social planner's assessments of these costs do not change because a death has occurred or been publicized.

[9] The personal costs associated with leaving the child in their home, a^p, might include feelings of guilt and the risk of losing their job if the child then dies. Personal costs associated with removal, b^p, could be the time, paperwork, and difficult conversations with families.

Resource constraints: If an agency is resource-constrained investigators may be unable to remove all children with assessed risk above the removal threshold. For example, agencies might have too few placements available for children whom they would otherwise remove from their homes. Investigators would then remove the children they assess to have the highest risk, up to the agency threshold, and remove children for whom:

$$p_i > \max\{\frac{\gamma_j a^S}{b^S} - \varepsilon_{ij}, T_{AGENCY}\} = T_{SUBOPT_{ij}}. \tag{5}$$

A binding resource constraint would cause agencies to remove too few children.

These two factors can lead investigators to have thresholds that differ from the socially optimal threshold and that differ across children. Investigators using a threshold that is too low will remove low-risk children who would be better left at home. Investigators using a threshold that is too high will fail to remove children who should be removed. Moreover, if the error in risk assessment has a non-monotonic relationship with true risk, the risk threshold could become "fuzzy", where investigators remove too many low-risk children and fail to remove enough high-risk children.

2.3 Responses to Tragedy: Information or Scrutiny?

I first consider two possible scenarios: (A) investigators continuously use information they learn from maltreatment deaths to update beliefs and adjust decision-making, or (B) investigators primarily respond to scrutiny and react much more to highly-publicized deaths.

In the absence of sharp changes in underlying risk or social costs, continuously updating investigators should not sharply change removal decisions. A child death should cause investigators to update their assessments of population risk using equation 3. The death would need to be unusual to cause an unusually large change in removal rates. Moreover, if investigators are responding to all information, updating should not be dependent on the extent to which a death is publicized in local newspapers. Therefore, if investigators are continuously updating:

- *Prediction A1:* Removal rates will increase following a death only if the death is unexpected. If deaths occur often, the reaction to each death will be small.
- *Prediction A2:* Agency reactions will not be conditional on the death being publicized.

Conversely, if agencies have large responses to highly-publicized deaths only, this is evidence that they react primarily to scrutiny rather than information.

2.4 Responses to Tragedy: Targeted or Haphazard?

I will also assess whether agencies' reactions to tragedies are (A) targeted among children at the highest-risk of maltreatment, or (B) haphazard, increasing removals among low-risk children or inducing removal rate changes unrelated to underlying risk.

A targeted reaction to a tragedy would consist of investigators removing more children at high risk of maltreatment. This could take two forms. First, the tragedy might decrease investigator error such that more of the very highest-risk children children are removed, whom investigators had previously left in their homes despite high risk. This would be an unambiguously positive change and should reduce maltreatment. Second, investigators might shift their removal threshold downwards and remove more children who lie just below the previous threshold. The effects of a targeted threshold shift would be ambiguous: it could happen because error is reduced, activated, or resource constraints are relaxed, and the welfare effects would depend on the social costs of removal and maltreatment. I will therefore characterize the response but will not be able to make strong claims about where it is optimal, even if it is targeted. If the reaction is targeted:

- *Prediction B1:* Increased removals will be concentrated among high-risk children.
- *Prediction B2:* Removal rate increases will be associated with improved child outcomes.
- *Prediction B3:* Removal rate increases that are higher for specific demographic groups will be associated with higher underlying maltreatment risk for marginal children in those groups.

In contrast, tragedies might *activate* error and induce a haphazard response, in which investigators remove more low-risk children. Removing low-risk children would be less likely to cause a reduction in maltreatment. A haphazard response could also consist of changing removal rates for different demographic groups in a way that is unrelated to underlying risk. Targeted and haphazard responses are illustrated graphically in Figure 2.

3 Data

Each observation in the datasets I construct for my primary analyses is a jurisdiction-month. Child protection systems in the US can be administered by the state, county or through a hybrid system. I treat the state as the jurisdiction in the 41 states that are state or hybrid-administered, and in three additional states where counties are small and not observed in the data. I treat the county as the jurisdiction in the six states that are administered at the county level and where I observe large counties.[10] I restrict my analyses to the 55 jurisdictions - 44 states and 11 counties - where both the entire population of children in foster care and jurisdiction-specific news coverage are observable every year between 2000 and 2019. The child welfare datasets I use mask the county name in counties where fewer than 1,000 children are in foster care in a year. I exclude any county jurisdiction that is masked in any year because I am unable to observe the full time-path of child

[10]States with county-administered systems are California, Colorado, Minnesota, New York, North Carolina, North Dakota, Ohio, Pennsylvania and Virginia. I observe large counties in California, Colorado, Minnesota, New York, Ohio and Pennsylvania and treat these counties as jurisdictions. In North Carolina, North Dakota, and Virginia counties are unobservable in my data so I treat each state as a jurisdiction. Nevada and Wisconsin operate with hybrid systems and I again treat the states as jurisdictions.

protection outcomes. I also exclude 19 additional counties and one state, Idaho, where I do not have access to full-text archives for a newspaper over the whole time period; in these jurisdictions I am unable to observe whether a highly-publicized tragedy occurred. These criteria leave me with a sample of 55 jurisdictions that contains 81.7 percent of the population of children age 0-9 in the US.[11]

3.1 Local Newspaper Archives

To identify public tragedies in the child protection system I use two databases that contain full text articles and metadata from US local newspaper archives: the ProQuest US Newsstream and the Access World News Database. I restrict my analysis to newspapers in the databases where the full text is available for the full time period between 1999 and 2019. The ProQuest database allows more flexibility in search and analysis, but provides coverage of only a subset jurisdictions; I therefore supplement the newspapers available in ProQuest with an additional set from Access World News. When both databases contain full text for the same newspaper in all years, I analyze it in the ProQuest database and drop it from the set of Access World News newspapers. I assign newspapers to jurisdictions using the location of the newspaper contained in the metadata. The set of newspapers I use to identify public tragedies consists of 108 sources that cover 57 jurisdictions.[12]

For both news databases I first use search terms to identify a subset of stories in the archives that meet a set of high-level criteria. ProQuest provides access to an analysis platform where I can work with a dataset of stories containing each story's title, the newspaper, jurisdiction, date, page number, and full article text. I make my initial dataset using search terms including multiple variants of "child death" and "maltreatment". In the Access World News Database I am unable to create a dataset with the full article text for analysis. Instead I search for a set of criteria by jurisdiction - including the same "child death", "maltreatment" terms, and the jurisdiction-specific child protection agency name - and save the search results as html documents. I transform the html documents into a dataset with each story's title, the newspaper, jurisdiction, date, page number, and a preview of the article text that prioritizes the search terms used. These initial searches yield two datasets: 809,958 articles from Proquest and 11,544 stories from AWN. Section 4.1 outlines how I use these datasets of stories to identify high-profile tragedies in the child protection system.

3.2 Child Protection System Outcomes

To calculate child welfare system outcomes, I use administrative data from the Adoption and Foster Care Analysis and Reporting System (AFCARS) and the National Child Abuse and Neglect Data

[11] Appendix Table A5 lists the 55 jurisdictions that are included and Appendix Table A6 lists the jurisdictions that are excluded.

[12] The coverage of these newspapers across the US is illustrated in Figures A and B, and the full set of newspapers is listed in Appendix C.

System (NCANDS). Both datasets are administered by the Children's Bureau in the federal government's Administration for Children and Families. AFCARS Foster Care files contain child-level information that state and tribal agencies are federally mandated to submit to the federal government. The annual files contain the entire population of children in foster care in the US at any point in a given fiscal year. NCANDS Child Files contain case-level data that states and tribal agencies submit voluntarily. These files contain information on every report of alleged child abuse or neglect that was investigated or received an alternative response in the child protection agencies that submit data in a given fiscal year. NCANDS Agency Files contain agency-level information on all allegations of child abuse or neglect for a given year, and are also voluntarily submitted.

Calculating Rates: I focus on outcomes of children age 0-9 because they are the most likely to interact with the child protection system: from 2000 to 2018 67.2 percent of children removed were under age 10. I use population data from the CDC to convert monthly counts of child outcomes to annualized rates per 10,000 children age 0-9:

$$MonthRate_{jt} = \frac{MonthCount_{jt} * 12 * 10,000}{AnnualPopulation_{jt}} \tag{6}$$

The interpretation of this rate is: if a jurisdiction has a removal rate of 7 per 10,000 in a particular month, the jurisdiction could expect 7 per 10,000 children to be removed if the removal rate that month were sustained for a year. For my primary analyses of child welfare and health outcomes I divide by the population of children age 0-9. The CDC reports population annually; because my dataset is at the month level I assume the population changes linearly each month to avoid artificial jumps in rates that are driven by the annual reporting.

I also normalize the rates to obtain my estimates as percentage changes. Many of the outcomes I am examining are rare events, and at the month-jurisdiction level there are several zero counts. This prevents me from using logs to estimate percentage changes. Instead, within each sub-dataset I normalize rates by dividing them by the mean rate in $EventTime$ -24 to -4 for the particular jurisdiction j and tragedy d. Every rate is then expressed as a proportion of the mean rate in the $before$ period for that jurisdiction and tragedy:

$$NormalizedRemovalRate_{jtd} = \frac{MonthRate_{jtd}}{\sum_{t=-24}^{-4} MonthRate_{jtd}/21} \tag{7}$$

Removal Rates: My primary measure of a child protection agency's response is the rate that children aged 0-9 in the jurisdiction experience a first removal from their home. I calculate this rate using AFCARS data that contain the universe of children in foster care each year. Children in foster care appear once in that year's AFCARS file. The data includes the date of each child's first removal in their lifetime and the date of the most recent removal they have experienced. I cannot observe any additional removals that a child may have experienced. Because I observe every child's first removal, I use a count of the number of first removals in each jurisdiction and month as the primary measure of child protection activity. I also count the total number of removals that I observe as a secondary measure, but this will be an underestimate.

Termination of Parental Rights: Termination of parental rights (TPR) is a legal procedure where the parental rights of a child's existing parents are permanently removed. Judicial interpretations of the US constitution require the jurisdiction applying for termination of parental rights to prove the parents are unfit: "until the State proves parental unfitness, the child and his parents share a vital interest in preventing [the] erroneous termination of their natural relationship" *Stanley v. Illinois* 1972. Termination of parental rights must be approved by a judge, and termination is necessary before the child can be adopted by new parents. The dates of any TPR from both parents are included in the AFCARS datasets. Among children removed each month, I count the number who experience any TPR within two years of their removal date.

Maltreatment allegations and investigations: In the NCANDS Agency Files I observe the total number of maltreatment allegations reported by members of the public at the state-year level. I use this to calculate the rate that child maltreatment reports are made at the state-year level. In the NCANDS Child Files I observe children's interactions with the child protection system before removal. These include screened-in reports, where a hotline worker screens in a new maltreatment allegation from the public and refers it to the agency for an investigation or alternative response. I count the number of allegations that are screened in each month, treating multiple allegations for the same child on the same day as a single allegation. Because the NCANDS files include data from only a subset of states each year, my sample ("Investigation Sample") is smaller for outcomes in this dataset; it includes just the 40 jurisdictions listed in Table A5.[13]

Substantiated Child Maltreatment: Each screened-in allegation of maltreatment results in a finding of whether the allegation was substantiated: did the child protection system determine that maltreatment occurred or not? This information is included in the NCANDS files; I use it to calculate both the rate that children in the population are found to have experienced maltreatment each month, and the portion of screened-in allegations that are substantiated.

[13]State coverage for the NCANDS data has increased over time: in 2000 just 34 states submitted child-level data, whereas by 2019 all 50 states submitted data. I include observations from a jurisdiction in the "Investigation Sample" only when I can observe outcomes in the full 24 months before and after the date of a particular public tragedy.

3.3 Child Health Outcomes

Fatalities: Detailed mortality data from the CDC's National Vital Statistics System (NVSS) contains every fatality in the US and includes individual dates of death, cause of death, age at death, place of residence, and race. I use this data to compute the rate of total fatalities of children age 0-9 in each jurisdiction and month. I also compute the rate of accident and homicide fatalities among children age 0-9, because these are more likely to be caused by child abuse or neglect and are the type of fatality that child protection systems are charged with preventing.

Inpatient Hospital Stays for Injury, Psychiatric, and Maltreatment-Related Diagnoses: I use files from the Hospital Cost Utilization Project (HCUP) State Inpatient Database (SID) to calculate the rate of child inpatient stays due to injury and psychiatric care in jurisdictions for which I have data. The SID includes inpatient discharge records for all individuals in community hospitals in states that participate, regardless of payer. I restrict the sample to years in which hospitals used ICD-9 codes, due to inconsistency in diagnoses after the transition to ICD-10. The sample is smaller for these outcomes due to data availability and includes the 15 jurisdictions listed in Table A5.

I look at three outcomes. First, I count children with diagnoses, many of which are age-specific, that have been identified in other reviews (Leeb 2008; Schnitzer et al. 2011) as likely to be due to maltreatment. Second, I count children with a primary injury diagnosis and any injury diagnosis (the datasets list up to 10 diagnoses), because research has found that maltreatment-specific diagnosis codes can be unreliable for identifying maltreatment (Scott et al. 2009). Moreover, children who suffer neglect may also be more likely to experience accidental injuries that are not directly inflicted by a parent. Third, I count children hospitalized with a primary psychiatric diagnosis and with any psychiatric diagnosis, because abuse and neglect are associated with mental health conditions (Keyes et al. 2012).

3.4 Controls

I use NVSS data to count the number of working-age adults who die due to substance use in each month and jurisdiction; I use this as a proxy for parental opioid use in a jurisdiction or month, which has been shown to be correlated with child abuse and neglect (Chapman 2022). I calculate the rate as described in Section 3.2, but dividing by the adult population instead of the child population. I use data from the Bureau of Labor Statistics' Local Area Unemployment Statistics to calculate the jurisdiction-level unemployment rate in each jurisdiction and month, because unemployment is associated with rates of neglect (Brown and De Cao 2020). I use the log mean unemployment rate for the month and the previous two months as a control.[14]

4 Empirical Strategy

4.1 Identifying Highly-Publicized Maltreatment Deaths

I create a new dataset of highly-publicized tragedies in the child protection system by developing a text-analysis algorithm that identifies months in which a jurisdiction experienced an unusually high number of stories about children who died due to maltreatment.

The goal of the algorithm is to identify public tragedies that meet a set of criteria:

1. A child died;
2. Due to maltreatment;
3. In the local jurisdiction.

To find these tragedies, I develop an algorithm that searches for keywords in the full text of local newspaper archives between 1999 and 2019, selects stories that meet a set of criteria, and identifies months where the normalized number of stories exceeds a threshold. Table A3 summarizes the search terms the algorithm uses to identify stories; this process yields 20,103 stories from newspapers in the ProQuest database and 9,601 from Access World News newspapers. I then count the total number of these stories in each jurisdiction and month in my time frame. I can observe the page number for the stories identified so I also count the number of stories that are in the first five pages of the newspaper or a subsection of the newspaper.

Some jurisdictions have more newspapers or more stories about child deaths than others. To identify months with unusually high coverage I therefore normalize the counts within jurisdictions and obtain Z-scores: the number of standard deviations above the jurisdiction-specific mean.

$$ZScore_{jm} = \frac{StoryCount_{jm} - Mean(StoryCount_{jm})}{Std.Dev(StoryCount_{jm})} \quad (8)$$

I apply a cutoff rule and include only jurisdiction-months that are at least 3 standard deviations above the mean. Between 1999 and 2019, the algorithm identifies 246 jurisdiction-months where coverage exceeds the threshold of 3 standard deviations above the mean across the 60 jurisdictions in my dataset. My results are robust to cutoffs that are higher than this; at lower cutoffs the agency response is smaller, which I discuss in Section 5.

Many of these high-coverage months are clustered: jurisdictions experience several months within a short time frame where the number of stories that meet the criteria exceed the threshold (see Figure A1). In some cases a story about one child remains in the news for several months; in

[14]The lagged relationship between unemployment and maltreatment is not clear theoretically; Brown and De Cao (2020) find a contemporaneous imapct at the annual level.

others, the newspaper reports a similar story shortly after the first. It is unlikely that the timing of subsequent events in a cluster is exogenous, and jurisdictions' responses to earlier tragedies would generate trends that would confound estimates that rely on exogenous timing or parallel trends (Sun and Abraham 2020). I therefore treat the first month in which a jurisdiction meets the criteria as the date of the highly-publicized tragedy, and drop the jurisdiction from both treated and control jurisdictions for three years after any date that meets the criteria; this mitigates the identification challenge. My approach identifies the impact of a single month with high coverage, or a cluster of months if several have high coverage within a short time period. This process reduces the total number of tragedies to 60.

To assess the validity of the algorithm a research assistant reviewed a set of articles in high-coverage months identified by the algorithm and hand-coded whether they met the criteria. 90 percent of the stories are about maltreatment deaths. 58 percent of the stories are about children who died after a visit from the local child protection agency, and 5 percent were about children who died while in foster care. I include all events identified by the algorithm in the main analysis; the results are larger if I assess the impact of events confirmed to be about maltreatment deaths.

For each story, the research assistant also manually coded the name of the child, the child's date of death, and their age at death. This allows me to assess the timing of the newspaper coverage relative to the date of death. I use an algorithm that uses the name of the child to predict their race and gender. This allows me to assess whether jurisdiction responses vary by child race or gender.

4.2 Identifying the Impact of a Highly-Publicized Tragedy on Child Protection Activity and Child Outcomes

Identification Strategy and Assumptions: I identify the average causal impact of a highly-publicized tragedy on child protection activity and child outcomes using a difference-in-differences design that draws on the recent work on staggered adoption contexts. The approach compares the time path of outcomes in each jurisdiction that experienced a highly-publicized death (treated jurisdictions) to the time path of outcomes in an event-specific set of control jurisdictions that did not experience a highly-publicized death in that time period. My approach is equivalent to first conducting a dynamic difference-in-differences analysis for every highly-publicized death, and then aggregating those estimates to obtain a weighted average for each month relative to four months before the spike in news. The primary identifying assumption is that in the absence of a highly-publicized death, the treated jurisdictions would have experienced the same trends in outcomes as the control jurisdictions, after controlling for time-varying factors in the jurisdiction.

This approach is not a traditional two-way fixed effects design, where the dataset consists of the entire panel of jurisdiction-month observations and the regression controls for jurisdiction and month fixed effects, estimating treatment effects using dummies for whether the jurisdiction has experienced an event. I do not use this alternative approach for three reasons. First, several recent papers demonstrate that treatment dynamics can contaminate treatment effect estimates if already-treated jurisdictions are used as controls (Athey and Imbens 2021; Borusyak and Jaravel 2017; Callaway and Sant'Anna 2020; De Chaisemartin and d'Haultfoeuille 2020; Goodman-Bacon 2021; Sun and Abraham 2020). I address this concern by excluding from the set of controls jurisdictions that have experienced a highly-publicized death in the three years before the date of the event. This is similar to one of Callaway and Sant'Anna (2020)'s proposed approaches to use only not-yet-treated jurisdictions as controls. In my setting - where newspapers have existed for decades - it is unlikely that any jurisdiction is never-treated. I therefore also use long-ago-treated jurisdictions as controls. This requires an additional assumption that treatment effects dissipate within 3 years of a highly-publicized death.

Second, the same literature demonstrates that estimates from two-way fixed effects designs are weighted averages of the set of two-by-two comparisons within the data, where the weights depend on the number of observations before and after the date of the event. These weights are often non-intuitive, and estimates can be difficult to interpret. My design addresses this concern by balancing the dataset in event time: every treatment and control jurisdiction contains exactly 24 monthly observations before and after the date of the event. The estimates are therefore weighted only by the weights specified in the regression: the number of individuals in each jurisdiction.

Finally, some jurisdictions experienced multiple highly-publicized deaths between 2000 and 2019. Adapting the standard staggered adoption design allows me to exploit variation from more than one tragedy in each jurisdiction. My approach is similar to that taken by Lafortune, Rothstein, and Schanzenbach (2018) and Cengiz et al. (2019): I include multiple tragedies from the same jurisdiction if the tragedies occurred with at least three years between them. This requires the same assumption above: that treatment effects dissipate within three years. Specifically, I assume that in time periods at least three years after a highly-publicized death, the impact of a new event will be the same in expectation as a jurisdiction that has never experienced a highly-publicized death or that experienced one many years earlier. Appendix Tables A7 and A8 show that my results are robust to both different sets of controls and different assumptions about the length of time in which dynamic treatment effects dissipate.

Event Study Dataset Construction and Regression Specifications: To operationalize my identification strategy, I first construct a set of 60 sub-datasets: one for each of the events, $d \in [1, 60]$, I identified in the newspaper analysis. Each sub-dataset consists of jurisdiction-month observations in the month of the highly-publicized death, the 24 months before it, and the 24 months after it. The variable $EventTime \in [-24, 24]$ denotes the month relative to the month of highly-publicized death d, which is set at $EventTime = 0$.

For each sub-dataset I begin with 49 monthly observations (one for each $EventTime \in [-24, 24]$)

for the "treated" jurisdiction that experienced the tragedy, and 49 observations for each of the other 54 jurisdictions. I then drop all jurisdictions in the sub-dataset that experienced a highly-publicized death in the three years before the date of event d, or in the two years after event d. This prevents treatment effects from tragedies in control jurisdictions that occurred before public tragedy d or that occur after public tragedy d from contaminating the estimates. All other jurisdictions are included in the set of control jurisdictions: both those that have no previous tragedy, and those where the previous tragedy was more than three years before the date of public tragedy d. This process results in 60 sets of control jurisdictions; one for each public tragedy. The mean number of control jurisdictions in each sub-dataset is 25.05.

I combine this set of 60 sub-datasets to form one large dataset and run the following specification to estimate the impact of a public tragedy on child protection activity and child outcomes:

$$Y_{jtd} = \alpha_j + \beta Z_{jt} + \sum_{\tau=-24}^{24} (EventTime_{dt}^{\tau} * \gamma_d^{\tau}) + \sum_{\tau=-24}^{24} \delta_\tau (EventTime_{dt}^{\tau} * Treated_{jd}) + \varepsilon_{jtd} \tag{9}$$

Y_{jtd} is the outcome of interest in jurisdiction j, month t, and sub-dataset for highly-publicized death d. α_j are a set of jurisdiction fixed effects, which control for factors that generate non-time-varying differences across jurisdictions. These might include impacts that are driven by the size of the jurisdiction, demographic differences including child poverty, and differences in child welfare systems such as average removal thresholds or the availability of preventative services. Z_{jt} are a set of time-varying controls for each jurisdiction-month, which account for factors that could confound the estimates if they vary with both jurisdiction and time. I include the number of adult substance use deaths in the jurisdiction-month, to control for the increases in opioid use that occurred at different times across jurisdictions during my time period, and which might cause both a highly-publicized death and a change in removal rates or other outcomes. I also include lagged log unemployment rates. The terms $\sum_{\tau=-24}^{24} (EventTime_{dt}^{\tau} * \gamma_d^{\tau})$ are a set of month fixed effects for each sub-dataset that flexibly control for time trends that occur within the control jurisdictions for each event.

The final terms $\sum_{\tau=-24}^{24} \delta_\tau (EventTime_{dt}^{\tau} * Treated_{jd})$ identify the average difference between the treated and control jurisdictions in each time period $\tau \in [-24, 24]$ relative to a tragedy. I cluster at the jurisdiction level and use population weights. The coefficients illustrated in the event study figures are δ_τ, and the 24 coefficients $\delta_{\tau \geq 1}$ identify the dynamic path of the average treatment effect of a death in the 24 months following the month the death had high news coverage.

This specification identifies the impact of a highly-publicized death on the jurisdiction's rate of child removal, investigation, substantiation, and TPR within 2 years of removal. To identify race-specific removal rates I use specification (1) with the removal rate of Black children and White children as the dependent variable, and weight by the population of Black and White children respectively. I conduct the same regression to identify the effect of a highly-publicized death on mortality and hospitalization rates.

For table estimates, which present aggregated estimates of the impact in the 2 years after an event relative to the 2 years before it, I use an analogous long-regression specification:

$$
\begin{aligned}
Y_{jtd} =& \alpha_j + \beta Z_{jt} + \sum_{\tau=-24}^{24} (EventTime_{dt}^{\tau} * \gamma_d^{\tau}) + \\
& \delta_0 Treated_{jd} + \delta(After_{dt} * Treated_{jd}) + \kappa(Zero_{dt} * Treated_{jd}) + \varepsilon_{jtd}. \qquad (10)
\end{aligned}
$$

In this specification, δ identifies the average change in the outcome in months 2-24 after an event. The $Zero$ term follows the approach taken by Deshpande and Li (2019) and dummies out time period where I cannot identify whether outcomes occurred before or after the month of a highly-publicized death. I dummy out the change one month before and after the event because the dates in my child welfare datasets can be incorrect by up to 16 days due to masking for privacy concerns.[15] I dummy out months between $EventTime = -4$ and $EventTime = -1$ to exclude anticipation effects: by construction, the month of the highly-publicized tragedy is the month when news coverage exceeded a threshold, but news coverage and the child protection response could have started before that month. I code $EventTime \in [-4, 1]$ as $Zero$ in all analyses of all outcomes to maintain consistency; for transparency I show the time path in the $Zero$ periods in all event study graphs and report the coefficients on $Zero$ in tables.

I also run a triple difference regression to identify the impact of a highly-publicized death on the gap in Black-White removal rates. For this specification the dataset contains observations for both Black and White removal rates in each jurisdiction-month subdataset, and each term in Equation 10 is interacted with a dummy for whether the observation for the dependent variable is the Black removal rate. The full specification is in Equation 12 in the Appendix. In this triple difference specification, the term δ_4 on $(After_{dt} * Treated_{jd} * Black)$ identifies the change in the gap in Black-White removal rates.

These regressions identify the impact of everything that occurs as a result of a highly-publicized death on children's outcomes. I examine how child protection activity responds along several margins: reports from the public, investigations, substantiations, and removals. But it is likely that

[15] In the AFCARS and NCANDS data, dates are masked to protect children's privacy. In the AFCARS data the child's date of birth is re-coded to day 15 in the month in which they were born. Every other date is then re-coded so the lengths of time between their date of birth and interactions with the child protection system are correct. Dates of removal, investigation, TPR, etc. may therefore be incorrect by up to 16 days. In the NCANDS data, the report date is re-coded to the 8th or 23rd of the month, and subsequent dates are adjusted again to preserve the length of time between events. This recording is most consequential for the observability of child welfare responses in the months immediately before and after a highly-publicized death. For example, if the highly-publicized death occurred on June 2, a child whose removal date is listed as May 29 might in fact have been removed after the highly-publicized death. Similarly, a child whose removal date is listed as June 7 might have been removed before the date of the event. I therefore dummy out the month of the event and the months before and after the tragedy in my analyses.

agencies and possibly the general public respond along additional margins. For example, funding for agencies might increase, children might remain with foster parents longer before being reunified, and families might expose their children to fewer risks. For this reason I do not treat public tragedies as an instrument for the impact of a removal. Instead my approach identifies the impact of a jurisdiction's whole response on children's outcomes.

4.3 Identifying the Impact if a Highly-Publicized Tragedy on the Risk of Children Removed

Risk model. To assess how removal rates change for children with different levels of risk, I build a predictive model that generates a risk score for children screened in by child protection agencies. The model uses information about allegations of maltreatment to predict the likelihood a child left in their home will receive an additional screened-in report within 6 months.

I build the model using the NCANDS Child Files, which consist of all maltreatment allegations that were screened in for investigation or alternative response. I first restrict the sample to jurisdictions and years where children are linkable across years, which allows me to see all allegations for each child across the entire time frame. In the NCANDS data, each observation is a child-allegation pair. Since multiple allegations can be investigated at the same time and new reports made during an ongoing investigation are often assigned to the same investigator, I transform the dataset into child-investigation observations. I do this by merging observations in cases when a child receives a new allegation before the disposition date of an existing one.[16]

I follow Baron et al. (2023) and use whether an investigated child experiences a new screened-in allegation within 6 months of the original investigation as my primary outcome.[17] I run jurisdiction-specific Lasso-regularized logistic models on a training dataset that consists of 20 percent of the data in "untreated" time periods; I exclude children investigated in the three years following a highly-publicized tragedy. I stratify the 20 percent sample on month and whether the child was removed.

Lasso regularization improves predictive power by dropping or shrinking predictors to prevent over fitting. The set of possible predictors X includes child and family characteristics, maltreatment allegation types, whether the child had previously been a victim of maltreatment, the number of allegations of maltreatment, and the source of the allegations. I estimate the likelihood that an investigated child is re-investigated within 6 months by choosing a vector of coefficients β that

[16] I treat the first date of the investigation as the date of the first report, and the last date as the date the child is removed or, if the child is not removed, the date that the last allegation is disposed.

[17] Baron et al. (2023) argue new allegations are least likely to be endogenous since new allegations are often assigned to the same investigator who then makes decisions about substantiation and removal.

minimize the following expression:

$$\sum_{i=1}^{n}[Re - Report_i x_i\beta - log(1 + e^{x_i\beta})] + \lambda\sum_{j=1}^{m}|\beta_j|, \tag{11}$$

where $i \in (1, n)$ indexes investigations, and $j \in (1, m)$ indexes predictors. $Re - Report_i$ indicates whether the child experienced a screened-in re-report within 6 months, X is the set of predictors, and β the set of coefficients. $\lambda \geq 0$ is the tuning parameter that determines how by much the coefficients are reduced and is selected using 10-fold cross validation within the training dataset. I run the estimation on a separate training dataset for each jurisdiction; this allows each jurisdiction to have its own risk function and maximizes use of the data since jurisdictions provide slightly different sets of predictors. The full predictive model therefore consists of a set of jurisdiction-specific models; Figure A10 depicts the range of variable coefficients across jurisdictions.

I test the predictive power of the full model using a test dataset containing the remaining 80 percent of the sample. I apply the model to all children in the test dataset and estimate a predicted risk score for each investigation.

Validity of the risk model. After generating a risk score for every child in the test dataset I assign each investigation to jurisdiction-specific risk deciles; a tenth of all children investigated in each jurisdiction fall into each decile. I test the model by regressing the outcomes - re-report and removal - on the risk deciles. For re-reports I restrict the sample to children left at home, which is the same restricton used for the training sample. For removals I use the whole test sample, including children who were removed. To obtain the mean outcomes of children in each decile I add the risk decile coefficients to the intercept, which is the likelihood of each outcome for children in the first decile. These means are presented graphically in Figures A11 and A12 and show that both outcomes increase monotonically with risk decile. I did not train the model to predict removals, so the relationship is a reassuring test that suggests that investigators use a risk calculus is that similar to the score produced by the risk model.

Figure A13 presents the same regressions for Black and White children and shows that conditional on predicted risk decile, the re-report rate of Black and White children is very similar. Moreover, conditional on predicted risk decile, Black and White children have a similar likelihood of being removed. The large Black-White removal rate gap is driven by a larger portion of Black children having higher predicted risk.

Estimating the impact by risk decile. I count the number of allegations, substantiations and removals by month and risk decile and multiply the counts by $\frac{1}{0.8}$ because the test dataset is an 80 percent sample. I then assess the impact of a highly-publicized tragedy on the removal rates of children in different risk groups using the difference-in-differences strategy described in Section

4.1, using the long difference Equation 10 separately for each risk decile. I use the rate of children removed in each decile as a portion of the whole population of children age 0-9 as the outcomes since the investigation rate for each decile may change.

5 Estimates of the Impact of Tragedies

5.1 First Stage: Newspaper Stories About Tragedies

The analysis of newspaper archives yields 60 events in which newspapers in a jurisdiction published an unusually high number of stories about child maltreatment deaths. The timing of these events appears to follow a crowded out, crowded in dynamic. High coverage of child deaths is crowded in by other similar stories: a month with high coverage in a jurisdiction is several times more likely to occur if another month with high coverage has occurred recently. But high coverage of deaths is crowded out by other local and national news that is happening at the same time: no high-coverage month occurs in the same month that a jurisdiction that has a team in a national football, basketball, baseball, or hockey final, nor in the same month that a jurisdiction experiences a mass shooting or a short-term weather emergency (see Figure 8). Stories about maltreatment deaths are also less likely to be covered in months in which there is a local or national election.[18]

The dynamic event study plot in Figure 3 shows the first stage: there is a large increase in newspaper coverage in month 0, which is the date of the spike in newspaper coverage, in jurisdictions in which the 60 included public tragedies occurred. The normalized count is close to zero in months -24 to -1, and in month 0 it increases to over 4 standard deviations above the mean. This occurs by construction: an event occurs when a jurisdiction-month has a normalized count that exceeds 3 standard deviations above the mean. There is no observable pre-trend. News coverage remains high for around 8 months following month 0 and is especially high in months 1-3 after the event; this reflects the "crowding in" dynamic described above.

5.2 Impact of Highly-Publicized Tragedies on Removal Rates

Removal rates increase sharply following a highly-publicized child protection tragedy, and the rise is sustained for at least 24 months. The two panels in Figure 4 show the time path of removal rates around a highly-publicized tragedy. Panel (a) plots the raw weighted average removal rates for jurisdictions that experienced the tragedy and the control jurisdictions that did not experience a tragedy. Before a tragedy, the average removal rates in treated jurisdictions was below the removal rate in control jurisdictions. Following a tragedy, the removal rate in the treated jurisdictions jumps to above the rate in control jurisdictions and remains high in the following 24 months.

[18]Data on dates of mass shootings come from The Violence Project (Peterson and Densley 2019), dates of weather emergencies from NOAA National Centers for Environmental Information (NCEI). US Billion-Dollar Weather and Climate Disasters https://www.ncei.noaa.gov/access/billions/, dates of sports finals from google searches.

Panel (b) plots the event time coefficients δ_τ from Equation 9. The coefficients represent the dynamic causal impact of tragedy on removal rates in the 24 months following a news spike relative to month -4. Before month -4 the coefficients are statistically indistinguishable from 0 and do not appear to be trending systematically. By month 2, removal rates have increased sharply and in all subsequent months the difference in removal rates is positive and highly significant.

The *Zero* period, from month -3 to 1, is grayed out, since it includes both months in which news could be anticipated and where the date of removal cannot be cleanly observed as before or after the month of the news spike. There is an increase in removal rates during this period that begins in month -3. Removals reported in month -3 are not impacted by the data masking; they certainly occurred before the date of the news spike. This pre-trend before the news spike suggests jurisdictions have information about the deaths before the month when coverage is intense enough to meet my threshold. Across the 60 events, the median time between the date of death and the news spike is 2 months, so it is plausible that agencies anticipate coverage before it occurs.

Estimates from the long regression are reported in Table 3: in the 24 months after a highly publicized death, the rate at which children age 0-9 are removed from their homes for the first time increases by a highly significant 6.16 per 10,000 children annually from the baseline of 38.41 per 10,000 in months [-24 to -4] before the news coverage. The normalized estimate reports this is a 19 percent increase relative to jurisdictions' pre-event mean. This is a large increase: between 2000 and 2018 the average year-on-year change in removal rates was -1.8 percent. In years without a highly-publicized death, only 3.9 percent of year-to-year removal rate changes within a jurisdiction were greater than 19 percent.

The primary analyses report the change in removal rates following all events identified by the text analysis algorithm. In Table 4 I show the impact of different types of events. Among events confirmed to be about child deaths, the estimate of the impact on removal rates rises to 19.7 percent of the baseline. Among events involving children known to CPS who died within 2 months of the news spike, the impact is even larger – a 24.2 percent increase.

My primary measure of removal rates captures the rate that children are removed from their home for the first time, because I can observe every child's first removal. I also measure the total removal rate in a jurisdiction and the rate that children are re-removed; both measures are under-estimates because I see a maximum of two removals per child each year. Nevertheless, the results are consistent with the primary measure: total removals increase by 13.9 percent (se = 3.6 percent), which is highly significant, and repeat removal rates increase by a marginally significant 6.7 percent (se = 3.5 percent). Jurisdictions' primary response to a tragedy is therefore to increase removals among children who have never been removed before, rather than those whom they know about.

5.3 Mechanisms: Response By Intensity of News Coverage

To assess how the agency response varies with news coverage intensity I run my analysis with different thresholds for including events. The primary analysis tests months where the number of stories exceeds 3 standard deviations above the mean and at least one story is in the first 5 pages of the newspaper. In the analyses below I adjust this threshold. To ensure that months identified with the lower thresholds do not have a higher rate of false positives I include only high-coverage months that my RA confirmed included a story about a child maltreatment death.

I first include all months where stories are at least 1 standard deviation above the mean and at least one story is in the first 5 pages. As in the original analysis, I exclude jurisdictions from both control and new treatment groups for three years after a month that exceeds the threshold. This process yields a new set of events. I then run dynamic and long difference event studies for sets of stories in this new dataset with different intensity of news coverage: highest (>4 sd above the mean), high ($3-4$ sd above the mean), medium ($2-3$ sd above the mean) and low ($1-2$ sd above the mean). Figures 6 show the dynamic event studies for each of these groups. Events with both the highest coverage (>4 sd) and high coverage (between 3 and 4 sd) see large and significant increases in removal rates: Table 7 presents the long difference estimates as 20.7 percent and 14.1 percent respectively. Events with the lowest coverage are followed by a small initial rise in removal rates that then declines within a year of the event, and a non-significant 6.6% increase in removal rates over the next 2 years.

I then examine whether the size of the response varies depending on whether the stories were published in the front of the newspaper. For this analysis I keep the threshold for stories at 3 sd above the mean, but include both jurisdictions that had stories in the first five pages of the newspaper and jurisdictions that did not have stories in the first five pages. I estimate the child protection removal response to each of these groups. Figure 5 and Table 7 report the results. Deaths with coverage at the front of the paper have a strong and highly significant 23.1 percent increase in removal rates. Deaths with high levels of coverage but no stories in the front of the paper have a much smaller 5.7 percent response that is statistically indistinguishable from zero.

Finally, increases in removal rates of the magnitude I find are unusual. A randomization inference test, illustrated in A6, shows that a reaction of the same size or larger occurs in only 5.4 percent of randomly assigned sets of month-jurisdiction pairs, or once every 20 months. Similarly, across all jurisdictions and years, only 3.9 percent of year-to-year changes are larger than 19 percent in control periods. In contrast, three maltreatment deaths occur each month in the average state.

These findings demonstrate that jurisdictions are responsive to the size of the news coverage. This could be because deaths that generate larger news stories are different and provide special information that makes a large reaction appropriate. But coverage appears to be driven by local and national news, not just features of the particular death: as discussed in Section 4.1, highly-publicized deaths are less likely to occur when there is local sports team in a final, mass shooting, extreme weather event, or a local or national election.

5.4 Mechanisms: Public Reporting or Agency Decision-Making?

The decision to remove a child comes at the end of a long series of decisions made by both members of the public and child protection agency actors. The increase in removals could be driven by the public making more reports about children at risk of maltreatment, by hotline workers screening in more cases for investigation, or by investigators making different removal decisions for the families they assess. To test what is driving the increase in removals I examine how these processes change following a public tragedy. Data on these mechanisms is not available for all jurisdictions, so I use sub-samples with a smaller set of jurisdictions to assess them. Figures A9 and A8 show that the removal results reported above remain in these samples.

Figure 10 shows the estimates of the impact of tragedies on the number of reports from the public, the number of screened-in reports, and the rate of removal conditional on a report being screened-in. Using normalized rates as the outcome, I find a small and insignificant impact of the tragedy on the number of reports from the public: public reports increase by 2 percent (se = 4 percent). I estimate a small and marginally significant 4.2 percent increase in the rate of screened in reports (se = 2.5 percent). The ratio of removals to newly screened-in reports increases by a highly significant 10.3 percent (se = 3.8 percent). The increase in removal rates is therefore driven by decisions made within the child protection agency, rather than increased reporting from the general public.

5.5 Longer-Term Outcomes: TPR and Child Health

The large increase in removals is accompanied by a rise in children who are removed and subsequently permanently separated from their parents within 2 years: the long estimate in Table 5 shows 0.823 additional children per 10,000 experience a Termination of Parental Rights (TPR) in the 24 months following a tragedy, which is an 8.3 percent rise. Some of the additional children removed as a result of the death were therefore at sufficiently high risk of maltreatment that a court agreed the children should be legally separated from their parents. But the increase in TPRs is smaller than the increase in removals, and the rate of children experiencing a TPR *conditional* on experiencing a removal decreases by 8 percent. Children who are removed following a tragedy therefore appear to be lower-risk on average, though this result could also be driven by judges changing the threshold at which they grant a TPR.

Panel (a) in Figure 11 presents the impact of a public tragedy on accident and homicide fatalities of children age 0-9. In the 24 months following a tragedy, the fatality rate remains flat and statistically indistinguishable from the rate in month -4. The long difference estimate is also indistinguishable from 0, and the 95 percent confidence intervals rule out a change of more than approximately 5 percent from a pre-event mean of 1.06 per 10,000[19]. Table 5 shows there is also no detectable change in the all-cause fatality rate for children age 0-9.

Panel (b) in Figure 11 shows the time path of child inpatient visits due to injury following a tragedy and shows no detectable change. The long estimate, presented in Table 5, is 0.02 per 10,000 children age 0-9. This is very small relative to the pre-treatment mean of 30.58 per 10,000 and not statistically significant. However, the rate at which Medicaid recipients are hospitalized for with a primary injury diagnosis as a portion of the whole child population age 0-9 declines by 0.923 per 10,000 from a baseline of 12.50 (7.4 percent), and by 1.566 per 10,000 for hospitalizations with any injury diagnosis from a baseline of 16.40 (9.5 percent); these estimates are significant at the 5 percent level.

The decrease in injuries is striking because children in foster care are automatically eligible for Medicaid, so the Medicaid population is likely to increase with removal rates, which would drive an increase in the Medicaid injury rate. Probing the result further, I find that half the effect occurs among children under 3, who are most at risk of maltreatment. Moreover, I find a significant 0.452 per 10,000 decrease in diagnoses that have been found to be highly associated with maltreatment among Medicaid recipients, from a baseline of 4.49 (10.1 percent). I also examine children's hospitalizations for psychiatric care and a significant decline in hospitalizations of children with any psychiatric diagnosis and a decline, though insignificant, in Medicaid recipient children hospitalized with a primary psychiatric diagnosis. Across all the hospitalization outcomes I find no detectable impact on children with private insurance, a subset of children less likely to interact with the foster care system and who therefore represent a plausible additional control group.

The hospitalization estimates are calculated using a smaller sample than the primary analyses; the estimate for the change in removal rates following a highly-publicized tragedy in this sample is 8.96 per 10,000 children.[20] A highly-publicized tragedy should not be viewed as an instrument for removal since many other factors in the child protection agency change. But the magnitudes of the removal and Medicaid injury rate changes suggest that removing 6 additional children following a public tragedy is associated with 1 fewer injury hospitalization among Medicaid recipients.

5.6 Impact on Removal Rates Conditional on Maltreatment Risk

Analysis of the impact on removal rates by predicted risk shown in Figure 14 reveals that the largest increase in removal rates occurs at the highest risk decile. As a portion of all children in the population, removals of children in the tenth decile increase by 2.5 per 10,000, which is a 30 percent increase. In contrast, removals of children in the 7th, 8th and 9th deciles increase by less than 0.5 per 10,000 children. There is no detectable impact on children in the lowest 4 deciles; the

[19]The upper bound of the CI is 0.663 per 100,000 children and the lower bound -0.546 per 100,000
[20]This is larger than in the full sample (6.16).

point estimates are very small and negative. This evidence suggests that the increase removals is concentrated among the highest-risk children following a highly-publicized tragedy.

The same figure shows that in deciles 3 to 10, the removal rate of Black children increase more than the removal rate of White children conditional on risk decile. For example, in the 5th decile the removal rate of Black children increases by ten times the rate of White Children: the Black removal increases by 1.024 per 10,000 Black children in the population, whereas the rate for White children increases by a non-significant 0.099 per 10,000. This is striking given the finding in Section 4 that Black and White children are equally likely to experience a re-report conditional on risk decile. Without knowledge of the social costs of removal and maltreatment I am unable to say whether Black and White children should optimally be removed at higher rates or lower rates, but the disparity in removal rates after conditioning on risk is suggestive evidence of mis-optimization.

5.7 Impact on Black-White Removal Rate Gap

The disparities in removals by race conditional on risk produce an increase in the Black-White gap in removal rates. Figure 13 presents the weighted raw mean removal rates in the 60 jurisdictions that experienced tragedies around the date of the event for all children, for Black children and for White children. Both Black and White children experience an increase in removal rates immediately after a public tragedy, but Black children experience a much larger increase despite having higher removal rates before an event. This is consistent with the two graphs in Figure 13, which present the regression coefficients from Equation 10 with the removal rate of Black children and the removal rate of White children as the outcome.

The coefficients in Table 2 report results from the long difference regressions: the Black removal rate increases by 12.67 children per 10,000 whereas the White removal rate increases by 4.00 children per 10,000 annually. Because Black children are removed at twice the rate of White children at baseline (60.44 per 10,000 for Black children vs. 32.58 per 10,000 for White children), the percentage change in removal rates of Black and White children is similar, at 17.2 percent and 15.4 percent respectively. However, the triple difference in Table 2 shows that this represents a large increase in the Black-White removal rate *gap* in the 24 months following a public tragedy.

This large increase in the Black-White removal gap might be optimal if the removal rate was much lower than optimal for Black families before the tragedy. But the risk analysis in section 5.6 found that Black removal rates increased more than White even conditional on risk. Moreover, the average predicted re-maltreatment risk of Black children removed was *lower* than the average predicted re-maltreatment risk of White children in treated jurisdictions in the baseline periods, suggesting that marginal Black children may have been less at-risk than marginal White children.[21] The disparity also cannot be explained by the race of the child who died: agencies have larger

[21] This is consistent with recent work from Baron et al. (2023) who find that high-risk White children are removed at lower rates than high-risk Black children.

responses to the deaths of Black children (see Table 4), but Figure A16 demonstrates that removal rates also increase more for Black children following White children's deaths. The increase in the gap following a highly-publicized tragedy therefore suggests some agency mis-optimization by race.

6 Discussion

6.1 Do Governments React to Information or Scrutiny?

The conceptual framework outlined two predictions in the case that governments use information they learn from all deaths to update continuously:

- *Prediction A1:* Removal rates will increase following a death only if the death is unexpected. If deaths occur often, the reaction to each death will be small.
- *Prediction A2:* Agency reactions will not be conditional on the death being publicized.

My findings show that agencies primarily react to scrutiny rather than information. In the average state, three children die due to maltreatment each month, so a death should not be unexpected for investigators. Yet the agency response to a highly-publicized death is unusually large. The randomization inference illustrated in Figure A6 demonstrates that agencies increase their removal rates with a similar magnitude in only 5 percent of randomly assigned months, or once every 20 months; across all jurisdictions in years without a highly-publicized death, only 3.9 percent of year-to-year removal rate changes were larger than 19 percent. Moreover, agencies respond more when a death is more highly-publicized. Agencies therefore do not use information to continuously optimize their removal decisions, but instead are primarily responsive to scrutiny.

6.2 Are Government Reactions Targeted or Haphazard?

The conceptual framework outlined three predictions in the case that governments have a targeted reaction to maltreatment deaths:

- *Prediction B1:* Increased removals will be concentrated among high-risk children.
- *Prediction B2:* Removal rate increases will be associated with improved child outcomes.
- *Prediction B3:* Removal rate increases that are higher for specific demographic groups will be associated with higher underlying maltreatment risk for marginal children in those groups.

Several findings are consistent with predictions B1 and B2, supporting the hypothesis that agencies have a targeted response to highly-publicized tragedies. First, removal rates are concentrated among children at high predicted risk of re-maltreatment: there are large increases in removal rates among children in the highest risk decile and no detectable change in removal rates in the bottom four deciles. Second, following a highly-publicized tragedy, more children are permanently separated

(TPR) from their parents, which suggests either that more of the highest risk children are removed, or that judges are lowering the threshold for a TPR. There is also a decrease in hospitalizations for injuries, and for psychiatric- and maltreatment-related diagnoses among Medicaid recipients: removing additional high risk children improves health outcomes.

However, analysis of the agency response by child race reveals some evidence of a less well-calibrated response. Following a highly-publicized tragedy, the removal rates of Black children increase more than the removal rates of White children for children in risk deciles 3 to 10. This produces a 31 percent increase in the already large Black-White gap in removal rates. An increased gap in removal rates could be optimal if Black children are under-removed at baseline. But Black removal rates increase more than White rates even after conditioning on risk, which does not support this explanation. The increase in Black removal rates is consistent with work in other settings that finds information shocks can increase racial gaps, sometimes in ways that are suggestive of bias and generate more punitive outcomes for Black residents (e.g. Angelova, Dobbie, and Yang 2022; Eren and Mocan 2018). I cannot make claims about whether Black and White children should optimally be removed at higher or lower rates,[22] but the difference in impact by race suggests highly-publicized tragedies induce mis-optimization that may be driven by stereotyping or other biases.

6.3 Welfare and Policy Implications

Agencies react to highly-publicized tragedies by removing more children, the removals are concentrated among children at high risk of maltreatment, and the reaction is associated with a decrease in hospitalizations of Medicaid recipients. The social cost of maltreatment is high: a recent estimate of per-victim nonfatal maltreatment costs is \$1,063,677 (2023 USD) (Peterson, Florence, and Klevens 2018).[23] Preventing maltreatment is therefore valuable. But unnecessary removals are also costly: for example, removals among marginal children in Illinois increased crime, benefit receipt, and teen pregnancy (Doyle 2007; 2008).[24]

Without observing the full set of social costs and benefits it is difficult to assess whether agencies' reactions to highly-publicized child maltreatment deaths improve net welfare. I find that six additional removals are associated with one fewer hospitalization for injury. Under the assumption that the only benefit of the six removals is preventing one incident of maltreatment and using Peterson, Florence, and Klevens (2018)'s estimate of the cost of maltreatment, agencies' reactions will improve net welfare as long as the social cost of each of the five additional removals is less

[22] Baron et al. (2023) recently argue that high-risk White children are under-removed, whereas sociologist Roberts (2022) argues that Black children are over-removed.

[23] This value incorporates monetized QALYs, which estimate the value of pain, suffering, and grief reported in surveys of maltreatment victims, in addition to lost productivity. The figure reported in Peterson, Florence, and Klevens (2018) is \$830,928 (2015 USD); I update to 2023 USD using the CPI.

[24] Doyles' papers assessed the impact of removal among marginal children in Illinois, at a time when removal rates in the state were higher than in most jurisdictions now.

than \$212,735. If additional maltreatment is prevented, that threshold would be higher. On the other hand, some of the evidence in the paper suggests that children removed are less high-risk than average, so maltreatment prevented may be less severe and the threshold would be lower. Determining the precise welfare effect is beyond the scope of this paper; the discussion above underscores the importance of child protection agencies responding judiciously to highly-publicized deaths and focusing removals among children most likely to benefit.

As agencies make efforts to reduce maltreatment in their jurisdictions, they might consider whether alternative policies or practices could have the same maltreatment-prevention benefits of removal with fewer costs. Work increasingly demonstrates the effectiveness of cash transfers in reducing maltreatment (e.g. Bullinger, Packham, and Raissian 2023; Cancian, Yang, and Slack 2013; Rittenhouse 2023); these programs also generate a range of other benefits for low-income families, including those are not at risk of abuse or neglect, and could be promising alternatives.

Finally, many policy advocates suggest agencies should focus more on prevention programs that keep children safely in their homes, and argue that progress made in these efforts is undone every time there is a child death. My paper does not make claims about the efficacy of prevention programs, but my results help assess the quantitative relevance of the advocates' argument. Between 2002 and 2018, jurisdictions reduced removal rates by an average of 1.8% from year-to-year. A highly-publicized death increased removal rates by 19%, offsetting 9 years of reductions in removal rates that might otherwise have occurred.

7 Conclusion

This paper examines how child maltreatment deaths drives child protection policy and practice and the consequences for citizen welfare. I find that media scrutiny influences decision-making in child protection agencies: following a highly-publicized maltreatment death, agencies respond by substantially increasing the rate at which they remove children from their homes. Agencies do not respond to less-publicized tragedies, which suggests the large reactions are primarily driven by scrutiny. There is evidence that the scrutiny induces a well-calibrated response: removal rates increase most among children at highest-risk of re-maltreatment, there is no change in removals among the lowest-risk children, and hospitalizations for injuries, psychiatric care and maltreatment-related diagnoses decrease among child Medicaid recipients. Agencies' reactions are therefore likely to be welfare-enhancing for some children. But racial disparities grow following highly-publicized tragedies, including disparities conditional on risk, which suggests part of the response is less well-calibrated and may be driven by bias. This finding is in line with other papers that have found information shocks can induce racial bias and exacerbate racial disparities. The paper highlights the ambiguous potential impacts of media attention on policy and citizen welfare, and underscores the importance of well-informed and judicious media and policy responses to tragic events.

References

Angelova, V., Dobbie, W., and Yang, C. S. (2022). "Algorithmic Recommendations and Human Discretion".

Arnold, D., Dobbie, W., and Yang, C. S. (2018). "Racial Bias in Bail Decisions". *The Quarterly Journal of Economics* 133(4). Publisher: Oxford Academic, 1885–1932.

Athey, S. and Imbens, G. W. (2021). "Design-based analysis in difference-in-differences settings with staggered adoption". *Journal of Econometrics.* Publisher: Elsevier.

Bald, A. et al. (2022). "The Causal Impact of Removing Children from Abusive and Neglectful Homes". *The Journal of political economy* 130(7). Place: Chicago Publisher: The University of Chicago Press, 1919–1962.

Baron, E. J. et al. (2023). *Racial Discrimination in Child Protection.*

Bordalo, P. et al. (2016). "Stereotypes". *The Quarterly Journal of Economics* 131(4). Publisher: MIT Press, 1753–1794.

Borusyak, K. and Jaravel, X. (2017). "Revisiting event study designs". *Available at SSRN 2826228.*

Brown, D. and De Cao, E. (2020). *Child Maltreatment, Unemployment, and Safety Nets.* Rochester, NY.

Bullinger, L. R., Packham, A., and Raissian, K. M. (2023). *Effects of Universal and Unconditional Cash Transfers on Child Abuse and Neglect.* National Bureau of Economic Research.

Bullinger, L. R., Raissian, K. M., and Schneider, W. (2022). "The power of the future: Intergenerational income mobility and child maltreatment in the United States". *Child Abuse & Neglect.* The rise of economic inequality and its consequences for child protection 130, 105175.

Callaway, B. and Sant'Anna, P. H. (2020). "Difference-in-differences with multiple time periods". *Journal of Econometrics.* Publisher: Elsevier.

Camasso, M. J. and Jagannathan, R. (2019). "Conceptualizing and testing the vicious cycle in child protective services: The critical role played by child maltreatment fatalities". *Children and Youth Services Review* 103, 178–189.

Cancian, M., Yang, M.-Y., and Slack, K. S. (2013). "The Effect of Additional Child Support Income on the Risk of Child Maltreatment". *Social Service Review* 87(3), 417–437.

Cengiz, D. et al. (2019). "The Effect of Minimum Wages on Low-Wage Jobs*". *The Quarterly Journal of Economics* 134(3), 1405–1454.

Chapman, A. (2022). “The opioid crisis and child maltreatment across counties and time in the United States, 2007–2017”. *The ANNALS of the American Academy of Political and Social Science* 703(1). Publisher: SAGE Publications Sage CA: Los Angeles, CA, 139–161.

Chenot, D. (2011). “The Vicious Cycle: Recurrent Interactions Among the Media, Politicians, the Public, and Child Welfare Services Organizations”. *Journal of Public Child Welfare* 5(2). Publisher: Routledge _eprint: https://doi.org/10.1080/15548732.2011.566752, 167–184.

Currie, J. and Spatz Widom, C. (2010). “Long-Term Consequences of Child Abuse and Neglect on Adult Economic Well-Being”. *Child Maltreatment* 15(2). Publisher: SAGE Publications Inc, 111–120.

Currie, J. and Tekin, E. (2012). “Understanding the cycle: Childhood maltreatment and future crime”. *Journal of Human Resources* 47(2). Publisher: University of Wisconsin Press, 509–549.

De Chaisemartin, C. and d’Haultfoeuille, X. (2020). “Two-way fixed effects estimators with heterogeneous treatment effects”. *American Economic Review* 110(9), 2964–96.

DellaVigna, S. and Gentzkow, M. (2010). “Persuasion: Empirical Evidence”. *Annual Review of Economics* 2(1), 643–669.

Deshpande, M. and Li, Y. (2019). “Who is screened out? Application costs and the targeting of disability programs”. *American Economic Journal: Economic Policy* 11(4), 213–48.

Devi, T. and Fryer Roland G., J. (2020). *Policing the Police: The Impact of “Pattern-or-Practice” Investigations on Crime*. Working Paper 27324. Series: Working Paper Series. National Bureau of Economic Research.

Doyle, J. J. and Aizer, A. (2018). “Economics of Child Protection: Maltreatment, Foster Care, and Intimate Partner Violence”. *Annual Review of Economics* 10(1), 87–108.

Doyle, J. J. (2007). “Child Protection and Child Outcomes: Measuring the Effects of Foster Care”. *American Economic Review* 97(5), 1583–1610.

— (2008). “Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care”. *Journal of Political Economy* 116(4). Publisher: The University of Chicago Press, 746–770.

Drake, B. and Pandey, S. (1996). “Understanding the relationship between neighborhood poverty and specific types of child maltreatment”. *Child Abuse & Neglect* 20(11), 1003–1018.

Eisensee, T. and Strömberg, D. (2007). “News droughts, news floods, and US disaster relief”. *The Quarterly Journal of Economics* 122(2). Publisher: MIT Press, 693–728.

Eren, O. and Mocan, N. (2018). “Emotional Judges and Unlucky Juveniles”. *American Economic Journal: Applied Economics* 10(3), 171–205.

Gainsborough, J. F. (2009). "Scandals, Lawsuits, and Politics: Child Welfare Policy in the U. S. States". *State Politics & Policy Quarterly* 9(3). Publisher: Sage Publications, Inc., 325–355.

— (2010). *Scandalous Politics: Child Welfare Policy in the States*. Google-Books-ID: SE4QliDME4cC. Georgetown University Press. 223 pp.

Goodman-Bacon, A. (2021). "Difference-in-differences with variation in treatment timing". *Journal of Econometrics*. Publisher: Elsevier.

Gross, M. and Baron, E. J. (2022). "Temporary Stays and Persistent Gains: The Causal Effects of Foster Care". *American economic journal. Applied economics* 14(2), 170–199.

Jagannathan, R. and Camasso, M. J. (2017). "Social outrage and organizational behavior: A national study of child protective service decisions". *Children and Youth Services Review* 77, 153–163.

Jason Baron, E., Goldstein, E. G., and Ryan, J. (2023). "The Push for Racial Equity in Child Welfare: Can Blind Removals Reduce Disproportionality?" *Journal of Policy Analysis and Management* 42(2). Publisher: Wiley Online Library, 456–487.

Keyes, K. M. et al. (2012). "Childhood maltreatment and the structure of common psychiatric disorders". *The British Journal of Psychiatry* 200(2). Publisher: Cambridge University Press, 107–115.

KidsCount (2020). *Child Wellbeing Indicators & Data — KIDS COUNT Data Center*. Library Catalog: datacenter.kidscount.org. URL: https://datacenter.kidscount.org/data/share/USA~1~35~36,37,38,41,40~char~0#USA/1/35/36,37,38,41,40/char/0 (visited on 04/20/2020).

Kleinberg, J. et al. (2018). "Human Decisions and Machine Predictions*". *The Quarterly Journal of Economics* 133(1), 237–293.

Kreitmeir, D., Lane, N., and Raschky, P. A. (2020). *The Value of Names - Civil Society, Information, and Governing Multinationals*.

Lafortune, J., Rothstein, J., and Schanzenbach, D. W. (2018). "School Finance Reform and the Distribution of Student Achievement". *American Economic Journal: Applied Economics* 10(2), 1–26.

Leeb, R. T. (2008). *Child maltreatment surveillance: Uniform definitions for public health and recommended data elements*. Centers for Disease Control and Prevention, National Center for Injury ...

Lepore, J. (2016). "Baby Doe". *The New Yorker* 1.

Lim, C. S. H., Snyder Jr., J. M., and Strömberg, D. (2015). "The Judge, the Politician, and the Press: Newspaper Coverage and Criminal Sentencing across Electoral Systems". *American Economic Journal: Applied Economics* 7(4), 103–135.

Peterson, C., Florence, C., and Klevens, J. (2018). "The economic burden of child maltreatment in the United States, 2015". *Child Abuse & Neglect* 86, 178–183.

Peterson, J. and Densley, J. (2019). *The Violence Project database of mass shootings in the United States.*

Philippe, A. and Ouss, A. (2018). ""No Hatred or Malice, Fear or Affection": Media and Sentencing". *Journal of Political Economy* 126(5). Publisher: The University of Chicago Press, 2134–2178.

Premkumar, D. (2019). *Public Scrutiny and Police Effort: Evidence from Arrests and Crime After High-Profile Police Killings.* SSRN Scholarly Paper ID 3715223. Rochester, NY: Social Science Research Network.

Putnam-Hornstein, E. et al. (2013). "Racial and ethnic disparities: A population-based examination of risk factors for involvement with child protective services". *Child abuse & neglect* 37(1). Publisher: Elsevier, 33–46.

Rittenhouse, K. (2023). "Income and Child Maltreatment: Evidence from a Discontinuity in Tax Benefits". *Available at SSRN 4349231.*

Roberts, D. (2022). *Torn Apart: How the Child Welfare System Destroys Black Families–and How Abolition Can Build a Safer World.* Basic Books.

Roberts, K. V. (2019). "Foster care and child welfare".

Sarsons, H. (2017). "Interpreting signals in the labor market: evidence from medical referrals". *Job Market Paper.*

Schnitzer, P. G. et al. (2011). "Identification of ICD codes suggestive of child maltreatment". *Child abuse & neglect* 35(1). Publisher: Elsevier, 3–17.

Scott, D. et al. (2009). "The utility and challenges of using ICD codes in child maltreatment research: a review of existing literature". *Child abuse & neglect* 33(11). Publisher: Elsevier, 791–808.

Snyder, J. M., Jr., and Strömberg, D. (2008). *Press Coverage and Political Accountability.*

Stanley v. Illinois (1972). Num Pages: 645 Publisher: Supreme Court Series Number: No. 70-5014.

Sun, L. and Abraham, S. (2020). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects". *Journal of Econometrics.*

Thomlison, B. and Blome, W. W. (2012). "Hold the presses: A commentary on the effects of media coverage of fatalities on the child welfare system". *Journal of Public Child Welfare* 6(3). Publisher: Taylor & Francis, 243–254.

Warburton, W. P. et al. (2014). "The impact of placing adolescent males into foster care on education, income assistance, and convictions". *Canadian Journal of Economics/Revue canadienne d'économique* 47(1). Publisher: Wiley Online Library, 35–69.

Warner, J. (2014). "'Heads must roll'? Emotional politics, the press and the death of Baby P". *British Journal of Social Work* 44(6). Publisher: Oxford University Press, 1637–1653.

Wildeman, C. and Emanuel, N. (2014). "Cumulative Risks of Foster Care Placement by Age 18 for U.S. Children, 2000–2011". *PLOS ONE* 9(3). Publisher: Public Library of Science, e92785.

Wulczyn, F. et al. (2013). "Poverty, social disadvantage, and the black/white placement gap". *Children and Youth Services Review* 35(1). Publisher: Elsevier, 65–74.

Yi, Y., Edwards, F. R., and Wildeman, C. (2020). "Cumulative Prevalence of Confirmed Maltreatment and Foster Care Placement for US Children by Race/Ethnicity, 2011–2016". *American journal of public health (1971)* 110(5). Place: Washington Publisher: American Public Health Association, 704–709.

8 Tables and Figures

Figure 1
Motivating Example:
A Highly Publicized Tragedy in New York City

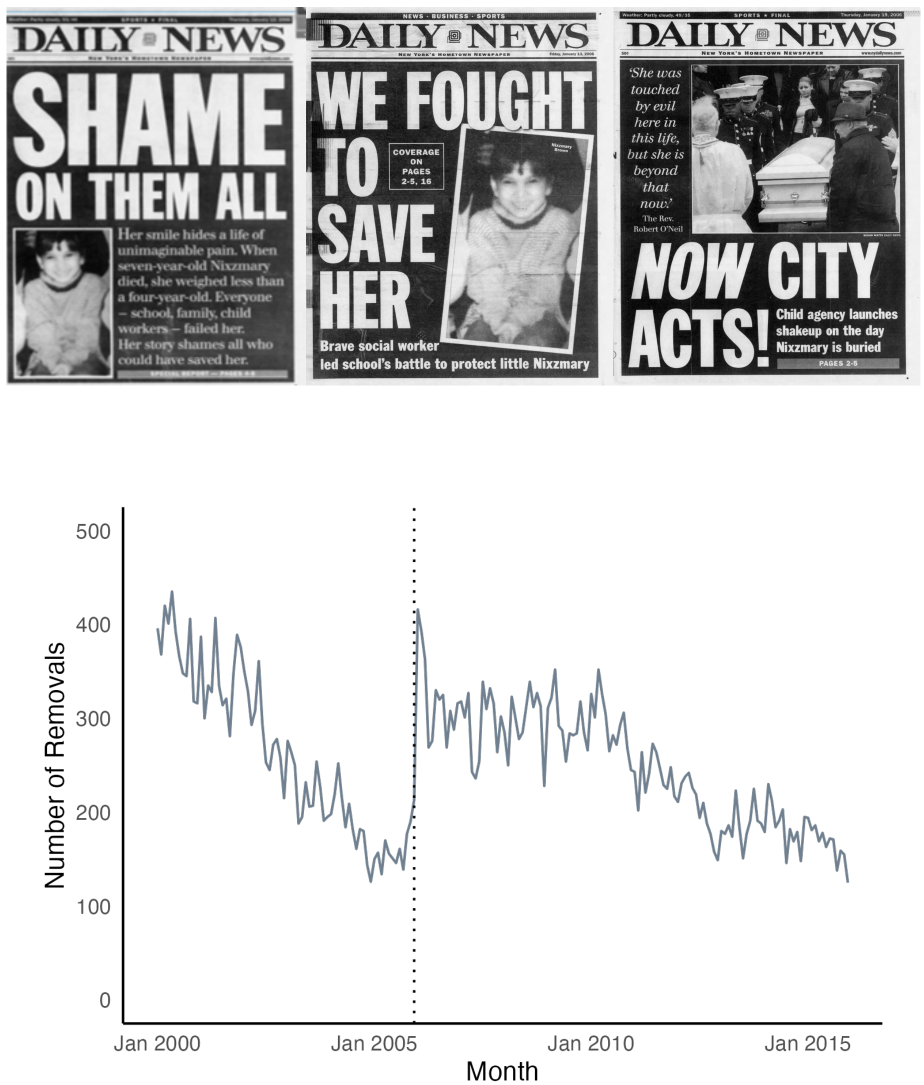
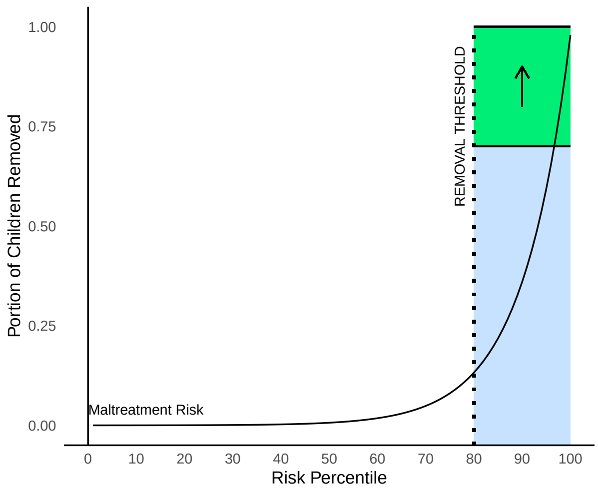
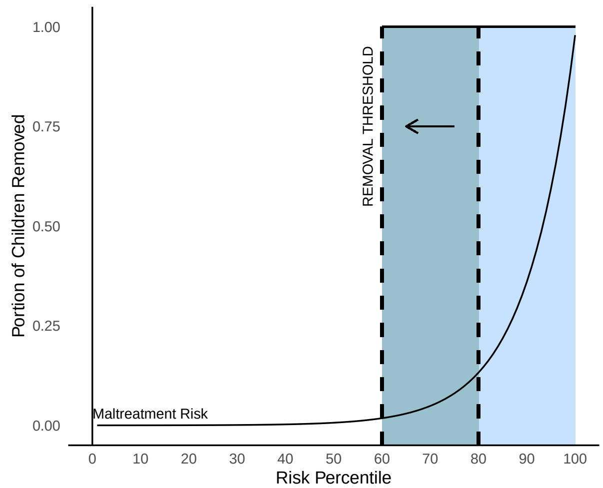


Figure 1: Front page newspaper stories in the New York Post covering Nixzmary Brown's death and monthly number of first removals in New York City. The black line is December 2005, the month before Nixzmary Brown's death in January 2006.

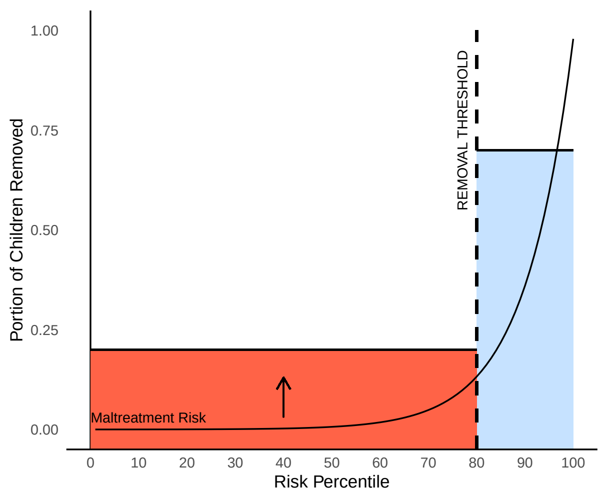
Figure 2
Conceptual Framework: Targeted vs. Haphazard Responses to a Tragedy



(a) Targeted Response: Reduced Error



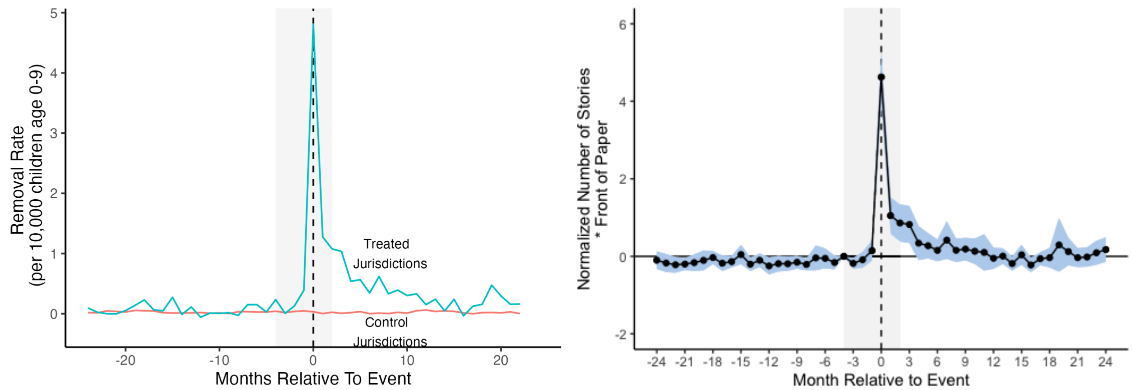
(b) Targeted Response: Threshold Shift



(c) Haphazard Response: Increased Error

Figure 2: Fully informed, social welfare-optimizing investigators will remove children whose risk of maltreatment falls above the threshold a^S/b^S, where a^S is the social cost of removal, and b^S is the social cost of maltreatment. In practice, the threshold may differ from the socially optimal threshold due to error or agency resource constraints. Following a tragedy, agencies could react with (a) a targeted response in which they remove more high-risk children above the original removal threshold, (b) a targeted response in which they shift the threshold down, removing more children just below the original threshold, or (c) a haphazard response, where error increases and more low-risk children far down the risk distribution are removed. The reaction may also be a mix of these scenarios.

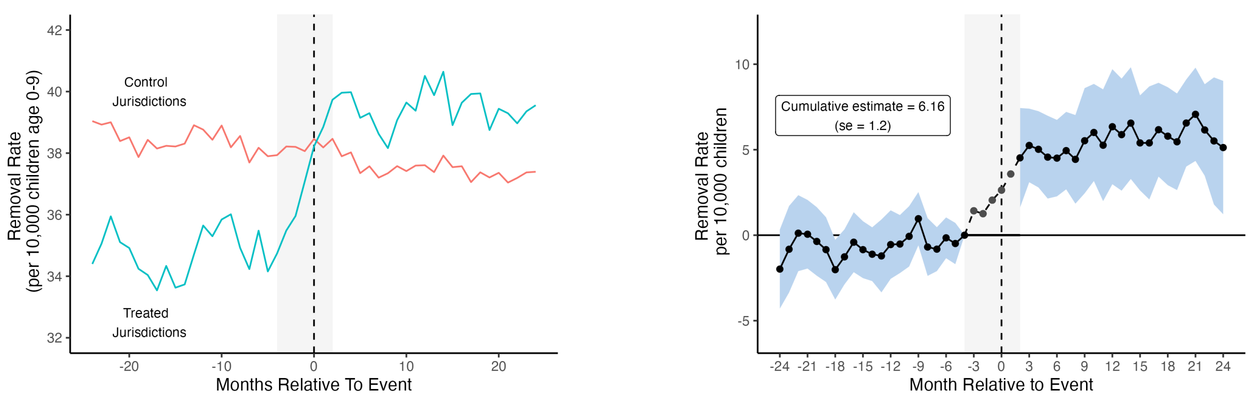
Figure 3
First Stage: Intensity of News Coverage



(a) Raw Normalized Story Count (b) Story Count Event Study Estimates

Figure 3: Panel (a) shows mean normalized counts of stories identified by the text analysis algorithm in control jurisdictions and treatment jurisdictions in the 24 months before and after a highly-publicized tragedy. Panel (b) shows equivalent dynamic difference-in-differences coefficients: mean normalized counts of stories in a treated jurisdiction around the time of a high profile abuse death, relative to month -4. All estimates weighted by number of children age 0-9.

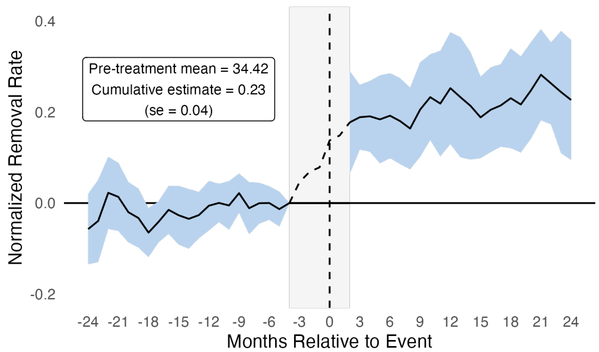
Figure 4
Impact of Highly-Publicized Tragedy on Removal Rates



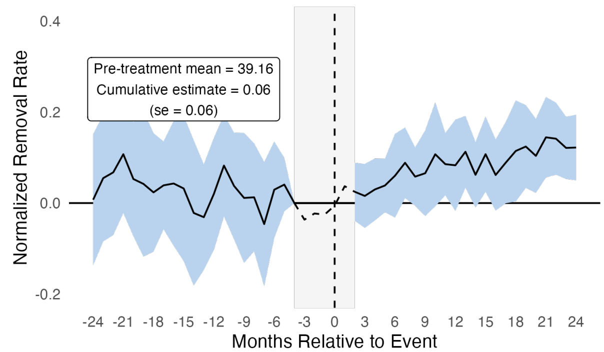
(a) Stories on Front Page (b) Stories not on Front Page

Figure 4: Panel (a) shows weighted mean removal rates per 10,000 children age 0-9 in control jurisdictions and treatment jurisdictions in the 24 months before and after a highly-publicized tragedy. Panel (b) shows equivalent dynamic difference-in-differences coefficients: average change in removal rate in a treated jurisdiction around the time of a high profile abuse death, relative to month -4. All estimates are weighted by number of children age 0-9.

Figure 5
Impact of Tragedy: Variation By Front Page Coverage



(a) Raw Removal Rates (Full Sample)



(b) Raw Removal Rates (Full Sample)

Figure 5: Time path of removal rates for public tragedies by whether a story was on the front page. All tragedies in these samples had number of stories at least 3 sd above the mean.

Figure 6
Impact of Tragedy: Variation By Intensity of News Coverage

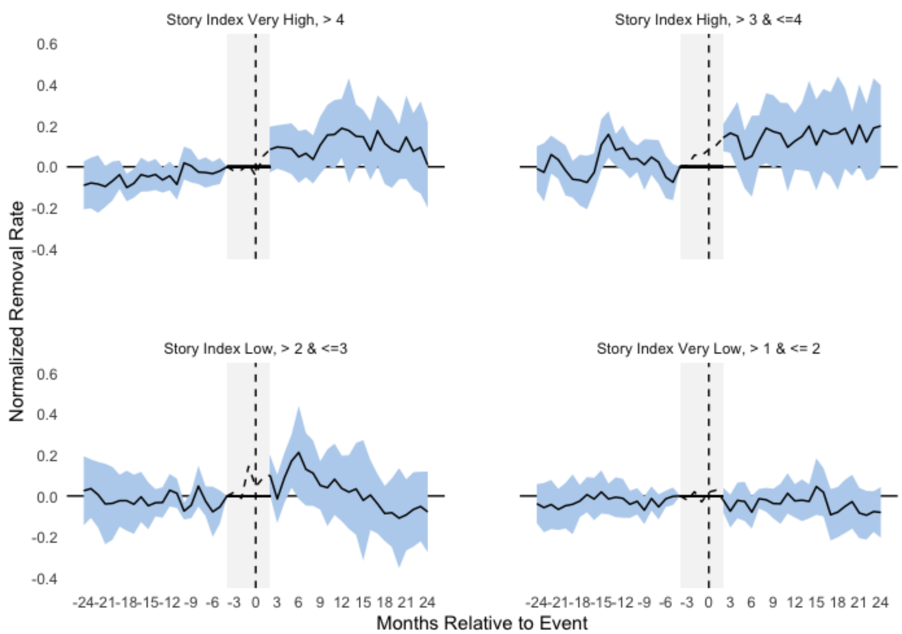


Figure 6: Time path of removal rates for tragedies with varying intensity of newspaper coverage. In these samples all events had front page coverage. When the number of stories in a month is 3 or more standard deviations above the mean, the removal rate increase is high and sustained. When the number of stories is between 1 and 2 standard deviations above the mean, there is no response.

Figure 7
Number of Deaths Known to CPS
and Number of Highly-Publicized Events

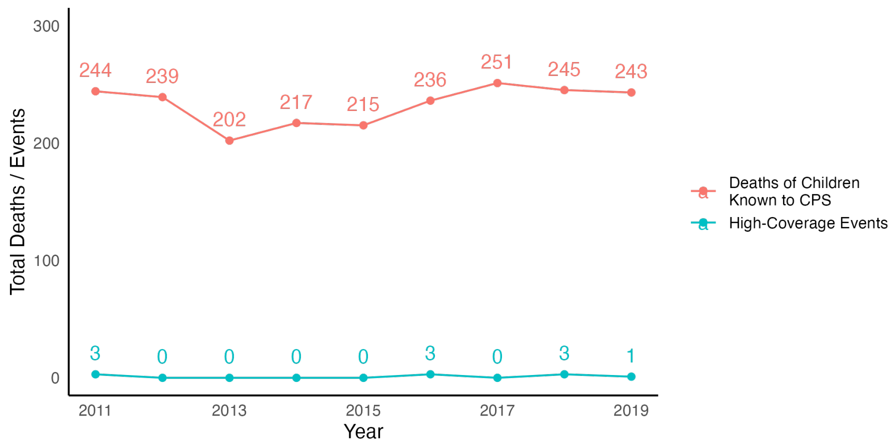


Figure 7: Annual number of deaths of children already known to the child protection agency in Arizona, Florida, Maine, New York City and Texas. Deaths in Texas, Florida and Arizona are maltreatment deaths. Deaths in New York City and Maine are all deaths of children previously known to the child protection agency.

Figure 8
Coverage of Tragedies
and Other News

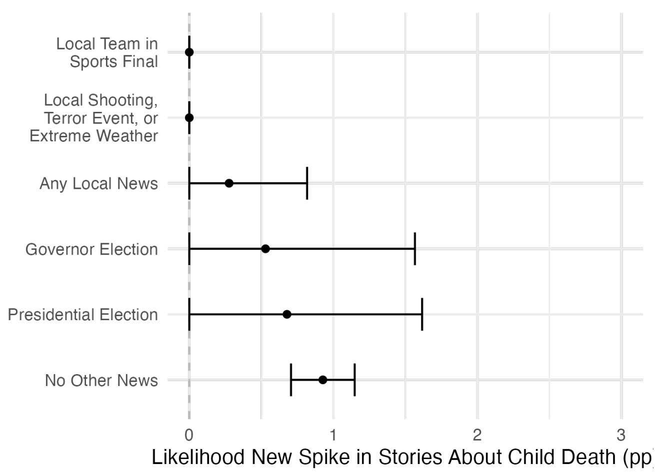
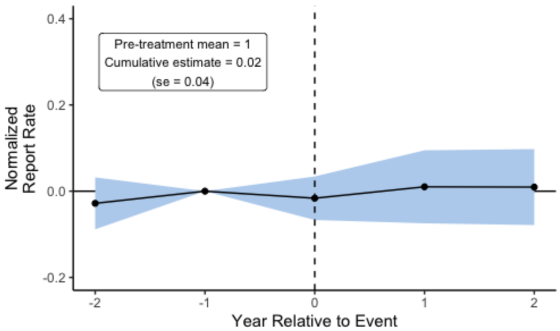
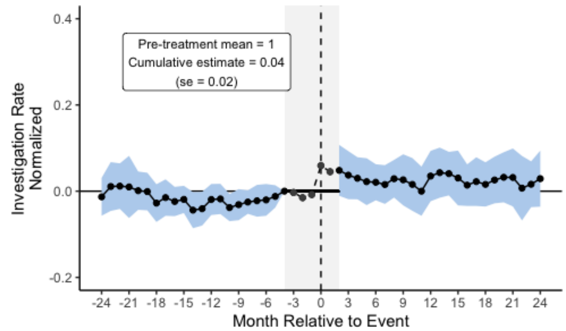


Figure 8: What predicts months with high coverage within jurisdictions.

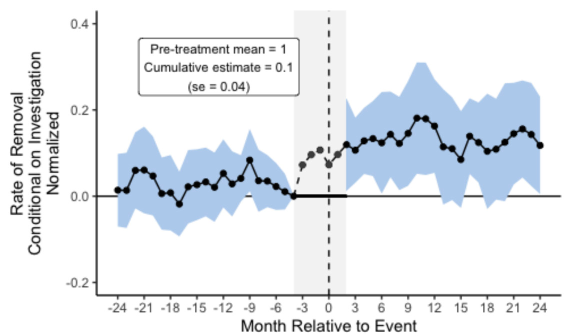
Figure 9
Mechanisms: What Drives Increased Removal Rates?



(a) Normalized Report Rate
(Reports Sample)



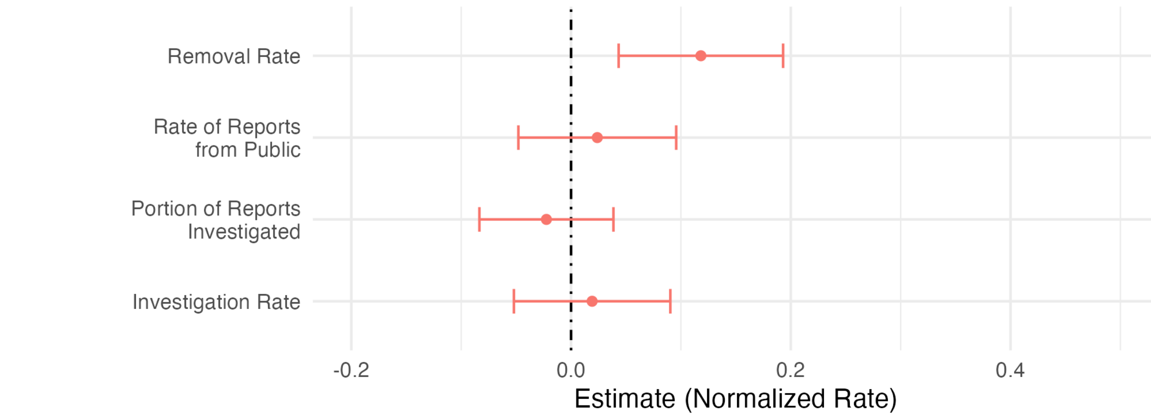
(b) Normalized Investigation Rate
(Investigation Sample)



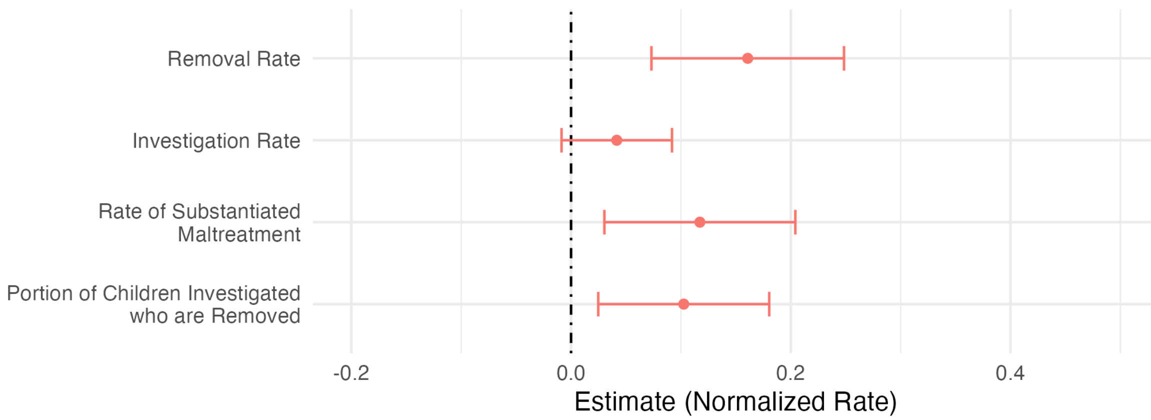
(c) Normalized Ratio of Removals
to New Investigations
(Investigation Sample)

Figure 9: The 16% increase in removal rates observed in the investigation sample is driven primarily by child protection response: reports from the public (observed only at the state-year level) increase by a non-significant 1% in the two years following an event; investigations screened in increase by a non-significant 4%; ratio of removals to new investigations increases by a significant 10%.

Figure 10
Mechanisms: What Drives Increased Removal Rates? Long Estimates.



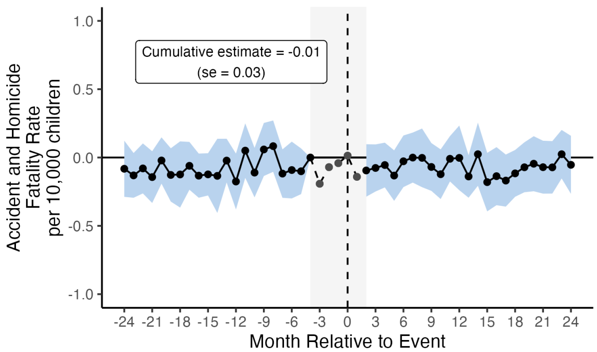
(a) Mechanisms: Reports from the Public, Investigations and Removals in Annual Data



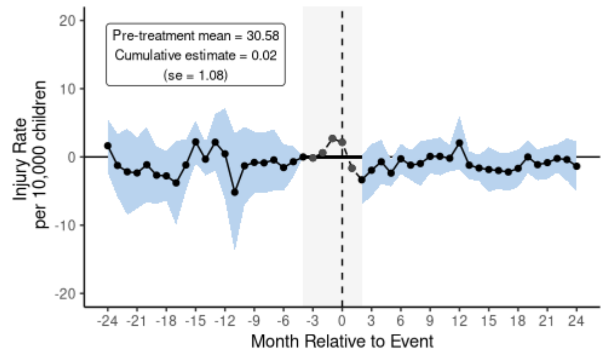
(b) Mechanisms: Investigations, Substantiations and Removals in Investigation Sample

Figure 10: The 16% increase in removal rates observed in the investigation sample is driven primarily by child protection response: reports from the public (observed only at the state-year level) increase by a non-significant 1% in the two years following an event; investigations screened in increase by a non-significant 4%; ratio of removals to new investigations increases by a significant 10%.

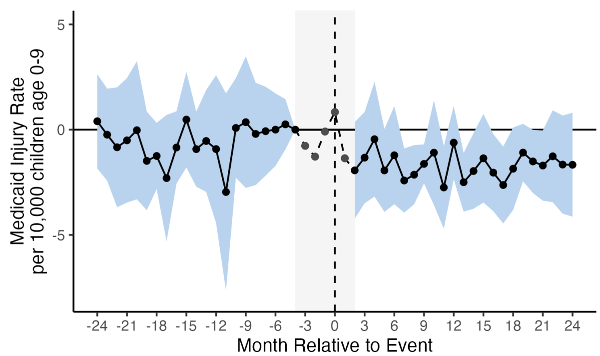
Figure 11
Impact of Highly-Publicized Tragedy on Child Health Outcomes



(a) Accident and Homicide
Fatality Rate
(Full Sample)



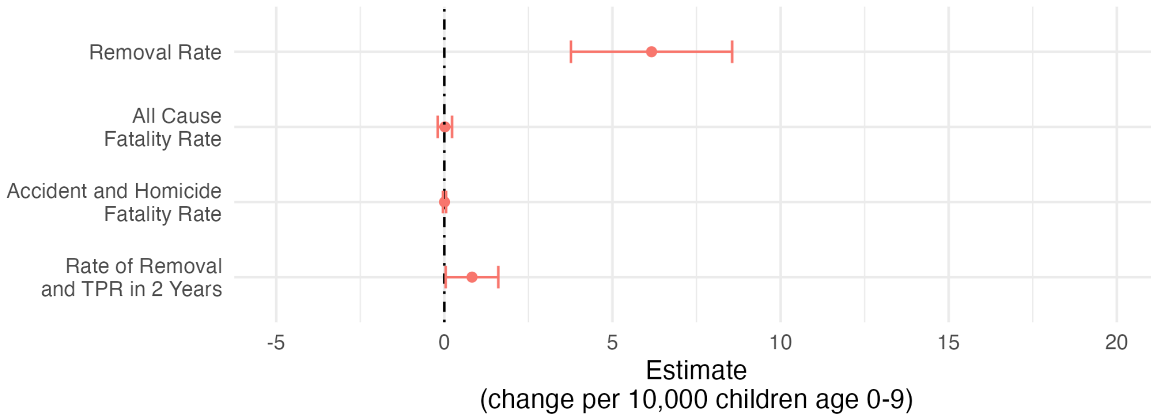
(b) Injury Rate
(Injury Sample)



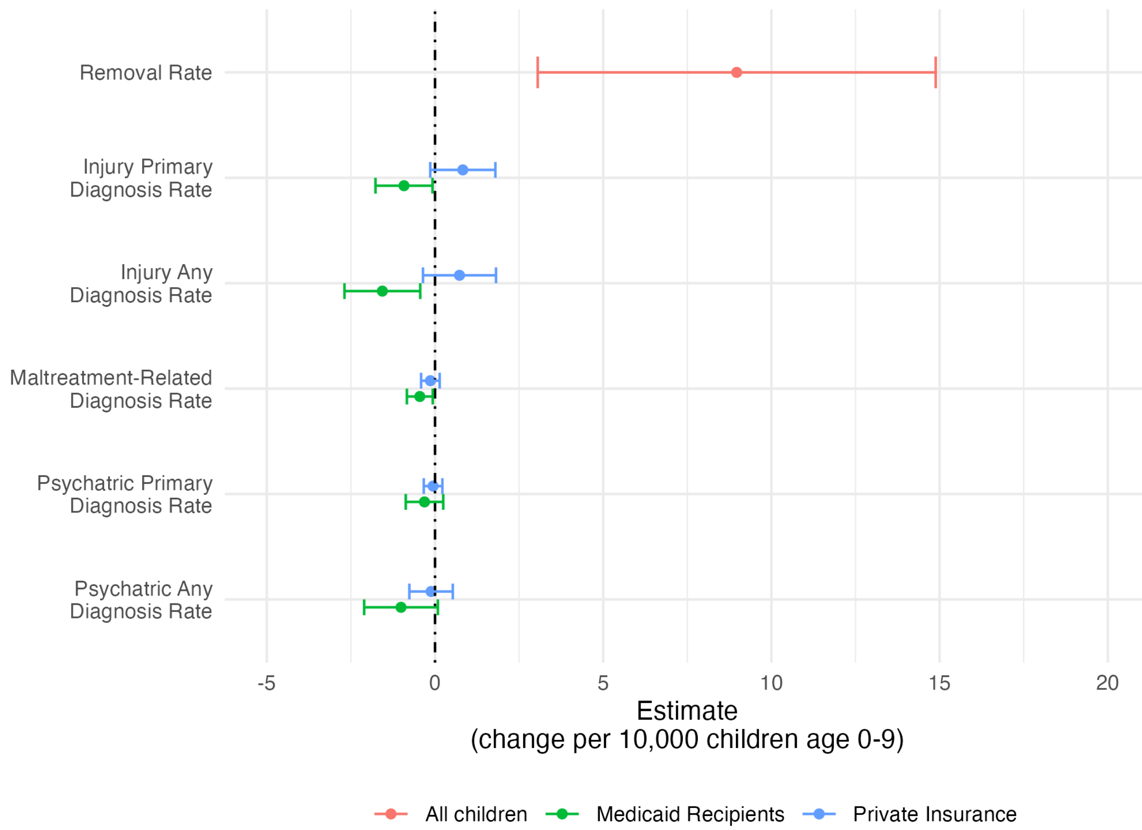
(c) Medicaid Injury Rate
(Injury Sample)

Figure 11: Impact of highly-publicized maltreatment deaths on health outcomes.

Figure 12
Impact of Highly-Publicized Tragedy on Child Outcomes: Long Estimates



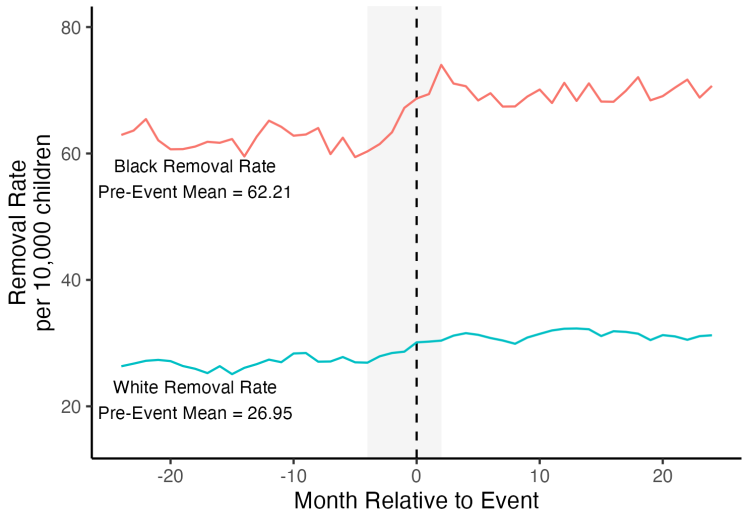
(a) Outcomes: Full Sample



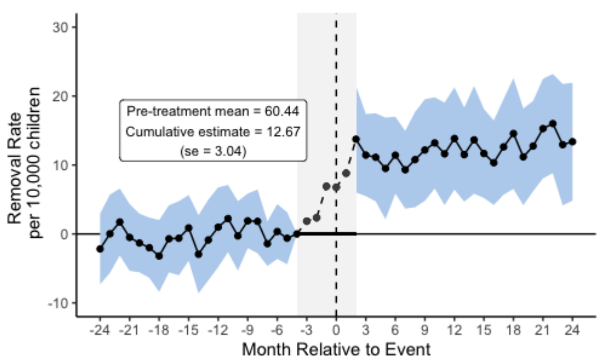
(b) Outcomes: Injury Sample

Figure 12: There is no detectable impact of a highly-publicized tragedy on child mortality, mortality due to accident or homicide, or hospitalizations due to injury among the whole population. Among Medicaid recipients, there is a decrease in hospitalizations due to injuries at a rate that suggests six additional removals are associated with one fewer injury. There is also a decrease in hospitalizations of children with psychiatric diagnoses, and maltreatment-related diagnoses.

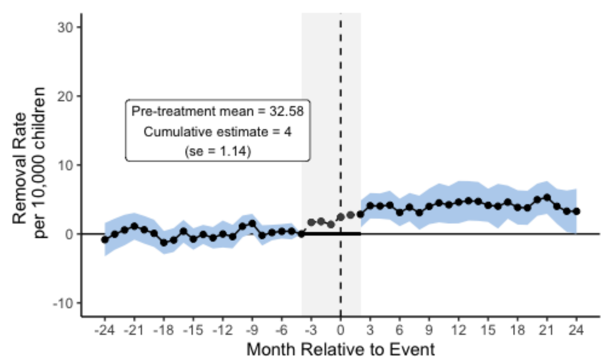
Figure 13
Impact of Highly Publicized Tragedy on
Removal Rates of Black and White Children



(a) Black and White Raw Removal Rates
Treated Jurisdictions



(b) Black Removal Rate
Event Study Coefficients



(c) White Removal Rate
Event Study Coefficients

Figure 13: Following a highly-publicized maltreatment death, the Black removal rate increases by over two times the increase in the White removal rate, despite the higher Black removal rate at baseline.

Figure 14
Removal Rate Increase
By Risk Decile and Race

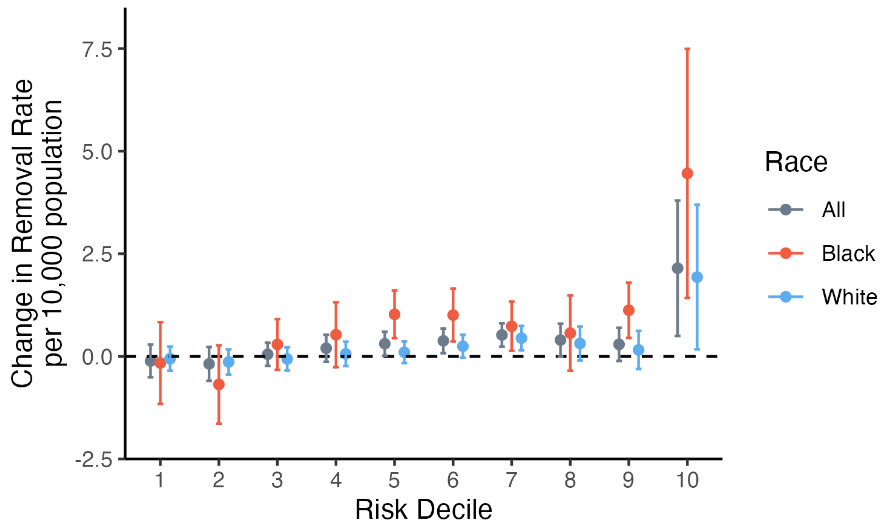


Figure 14: Estimates of impact of a highly-publicized tragedy for children in each predicted risk decile, by race. Impact is change in removal rate per 10,000 children age 0-9. Regressions are weighted by total population of children, Black children, and White children respectively.

Table 1
Summary Statistics for
Control and Treatment Jurisdictions in Each Sample

	Full Sample		Hospital'n Sample		Invest. Sample	
	Control	Treated	Control	Treated	Control	Treated
Removal Rate	38.41	34.76	35.82	35.73	38.03	35.16
Rate of Removal and TPR in 2 Years	7.65	7.46	6.55	6.00	8.19	7.75
Portion of Removals with TPR in 2 Years	20.42	21.57	19.18	15.54	22.06	22.26
Accident and Homicide Fatality Rate	1.11	1.06	0.70	0.85	1.06	1.02
All-Cause Fatality Rate	8.21	8.19	6.71	8.17	7.94	7.95
Investigation Rate					362.14	340.90
Substantiated Maltreatment Rate					94.98	83.78
Ratio of New Investigations to Removals					11.35	14.02
Number of Events	60		13		44	
Number of Jurisdictions	55		15		40	
Average jurisdictions in control	25.05		4.42		17.41	

Table 1: Pre-treatment Summary Statistics: mean for Treated and Control jurisdictions in months -24 to -4 relative to the date of the highly-publicized maltreatment death, for the three samples used to identify the impact of events on outcomes. Statistics are weighted by the total population of children age 0-9 each month, and by the population of White or Black children for the race-specific removal rates. All statistics are expressed in rate per 10,000 children in the population unless otherwise noted.

Table 2
Impact of a Highly-Publicized Maltreatment Death

	Removal Rate	Removal Rate (Black)	Removal Rate (White)	Removal Rate (Triple Diff)
Treated*After	6.163***	12.667***	3.998***	3.998***
	(1.196)	(3.044)	(1.136)	(1.136)
Treated*After*Black				8.669**
				(2.743)
Treated	−2.242	−2.397	−1.722	−1.722
	(1.577)	(3.011)	(1.111)	(1.111)
Black				11.306
				(15.637)
Num.Obs.	76 525	76 525	76 525	153 050
Num.Clusters.	55	55	55	55

Table 2: Long difference estimates for impact of a highly-publicized death by race. The fourth model is a triple difference specification that identifies the change in the Black-White gap in removal rates. All regressions include jurisdiction fixed effects and time-varying controls for adult substance use and unemployment. Standard errors are clustered at the jurisdiction level.

Table 3
Impact of a Highly-Publicized Maltreatment Death on Child Welfare System Outcomes

	Est.	S.E.	Clusters	N
Removal Rate				
Absolute	6.163***	1.196	55	76,525
Normalized	0.189***	0.039	55	76,525
Black Removal Rate				
Absolute	12.667***	3.044	55	76,525
Normalized	0.172***	0.041	55	76,525
White Removal Rate				
Absolute	3.998***	1.136	55	76,525
Normalized	0.154**	0.052	55	76,525
Total Removal Rate				
Absolute	4.579***	1.157	55	76,525
Normalized	0.139***	0.036	55	76,525
Repeat Removal Rate				
Absolute	0.363+	0.189	55	76,525
Normalized	0.067+	0.035	55	76,525
Rate of Removal and TPR in 2 Years				
Absolute	0.823*	0.390	55	76,525
Normalized	0.083+	0.050	55	76,525
Portion of Removals with TPR in 2 Years				
Absolute	−0.815+	0.452	55	76,525
Normalized	−0.091**	0.034	55	76,525

Table 3: Long difference estimates for impact of a public tragedy on child protection system outcomes. All regressions include jurisdiction fixed effects and time-varying controls for adult substance use and unemployment. Standard errors are clustered at the jurisdiction level.

Table 4
Impact of a Highly-Publicized Maltreatment Death on Removal Rates by Type of Event

	All Events	Child Death	Known to CPS	Known and Recent	Black Deaths	White Deaths
Long Diff. Estimate	0.189***	0.197***	0.230***	0.242***	0.190***	0.108***
	(0.039)	(0.041)	(0.043)	(0.063)	(0.033)	(0.031)
Num.Obs.	76 525	66 529	44 590	24 108	26 950	21 364
Clusters	55	53	52	51	52	50
Num Events	60	52	35	19	21	17

Table 4: Long difference estimates for sub-samples of events, categorized using data hand-coded by a research assistant. All regressions include jurisdiction fixed effects and time-varying controls for adult substance use and unemployment. Standard errors are clustered at the jurisdiction level.

Table 5
Impact of a Highly-Publicized Maltreatment Death on Child Hospitalizations

	Est.	S.E.	Events	Clusters	N
Accident and Homicide Fatality Rate					
All Children Age 0-9	0.006	0.023	60	55	76,525
All Cause Fatality Rate					
All Children Age 0-9	0.017	0.106	60	55	76,525
Injury Hospitalization Rate: Primary Injury Diagnosis					
All Children Age 0-9	−0.098	0.750	13	15	3,170
Medicaid Recipients Age 0-9	−0.923*	0.395	13	15	3,170
Private Payers Age 0-9	0.826+	0.451	13	15	3,170
Injury Hospitalization Rate: Any Injury Diagnosis					
All Children Age 0-9	−0.807	0.974	13	15	3,170
Medicaid Recipients Age 0-9	−1.566**	0.525	13	15	3,170
Private Payers Age 0-9	0.728	0.506	13	15	3,170
Medicaid Recipients Payers Age 0-3	−0.872**	0.301	13	15	3,170
Medicaid Recipients Payers Age 0-5	−1.220**	0.399	13	15	3,170
Hospitalization Rate: Maltreatment-Related Diagnosis					
All Children Age 0-9	−0.589	0.370	13	15	3,170
Medicaid Recipients Age 0-9	−0.452*	0.179	13	15	3,170
Private Payers Age 0-9	−0.138	0.129	13	15	3,170
Psychiatric Hospitalization Rate: Primary Psychiatric Diagnosis					
All Children Age 0-9	−0.374	0.368	13	15	3,170
Medicaid Recipients Age 0-9	−0.313	0.260	13	15	3,170
Private Payers Age 0-9	−0.058	0.129	13	15	3,170
Psychiatric Hospitalization Rate: Any Psychiatric Diagnosis					
All Children Age 0-9	−1.054	0.655	13	15	3,170
Medicaid Recipients Age 0-9	−1.011*	0.511	13	15	3,170
Private Payers Age 0-9	−0.118	0.301	13	15	3,170

Table 5: Long difference estimates of child health outcomes following a highly-publicized tragedy. All outcomes are expressed per 10,000 children age 0-9. The fatality results come from the full sample. The hospitalization results from a smaller sample due to data availability. In the hospitalization sample, the estimate of the impact on removal rates is 8.96 per 10,000 children age 0-9. This suggests 10 additional removals are associated with 1 fewer hospitalization of a Medicaid recipient with a primary injury diagnosis, and 6 additional removals are associated with 1 fewer hospitalization of a Medicaid recipient with any injury diagnosis.

Table 6
Mechanisms: Impact of a Highly-Publicized Maltreatment Death on Child Welfare System Removal Pathway

	Est.	S.E.	Clusters	N
Removal Rate				
Absolute	5.453***	1.385	40	32,928
Normalized	0.161***	0.043	40	32,928
Investigation Rate				
Absolute	14.541*	7.152	40	32,928
Normalized	0.042+	0.025	40	32,928
Substantiated Maltreatment Rate				
Absolute	14.470**	4.890	40	32,928
Normalized	0.117**	0.043	40	32,928
Rate of Removal Conditional on Investigation				
Absolute	1.021	1.371	40	32,928
Normalized	0.103**	0.038	40	32,928

Table 6: Long difference estimates from sample use to estimate investigation statistics. This sample consists of a smaller subset of 40 jurisdictions that have data on investigations. Absolute estimates are per 10,000 children age 0-9. Normalized estimates are expressed as a ratio of the mean absolute rate in months -24 to -4.

Table 7
Impact of Maltreatment Deaths
with Varying Intensity of Media Coverage

	Removal Rate (Norm)	Removal Rate (Norm)	Removal Rate (Norm)	Removal Rate (Norm)	Removal Rate (Norm)
Long Diff Estimate	0.231***	0.057	0.207*	0.141*	0.066
	(0.044)	(0.055)	(0.100)	(0.059)	(0.101)
Num.Obs.	33 337	15 141	10 339	5537	5488
Clusters	34	32	32	46	49
Sample	Front Page, $> 3sd$	Not Front Page, $> 3sd$	Front Page, $> 4sd$	Front Page, $3 - 4sd$	Front Page, $2 - 3sd$

Table 7: Long difference estimates of events on removal rate, by intensity of news coverage. Coverage intensity is measured both by whether a story that met the criteria was in the first five pages of the newspaper, and number of stories measured by number of standard deviations relative to the mean.

Table 8
Impact of a Highly-Publicized Maltreatment Death by Predicted Risk Decile and Race

	Estimate					
	Est.	S.E.	N	Clusters	Pre-event Mean	Perc. Change
Decile 1						
Black	−0.162	0.495	37,730	47	2.487	-6.50
White	−0.057	0.147	37,730	47	0.751	-7.65
Decile 2						
Black	−0.685	0.474	37,730	47	3.182	-21.51
White	−0.137	0.152	37,730	47	1.084	-12.68
Decile 3						
Black	0.290	0.307	37,730	47	3.219	9.02
White	−0.064	0.140	37,730	47	1.177	-5.45
Decile 4						
Black	0.527	0.393	37,730	47	3.197	16.49
White	0.060	0.148	37,730	47	1.368	4.38
Decile 5						
Black	1.024***	0.288	37,730	47	3.404	30.07
White	0.099	0.131	37,730	47	1.594	6.21
Decile 6						
Black	1.006**	0.321	37,730	47	3.302	30.48
White	0.247+	0.138	37,730	47	1.669	14.78
Decile 7						
Black	0.733*	0.298	37,730	47	3.685	19.89
White	0.444**	0.148	37,730	47	1.837	24.15
Decile 8						
Black	0.566	0.456	37,730	47	5.236	10.80
White	0.313	0.206	37,730	47	2.391	13.11
Decile 9						
Black	1.123***	0.335	37,730	47	6.923	16.23
White	0.156	0.230	37,730	47	3.377	4.63
Decile 10						
Black	4.459**	1.508	37,730	47	13.780	32.36
White	1.931*	0.876	37,730	47	6.449	29.95

Table 8: Estimate of impact of a highly-publicized tragedy on removal rate per 10,000 children age 0-9, by risk decile and race of children removed.

Appendix

Appendix Figure A1
Highly-Publicized Tragedies

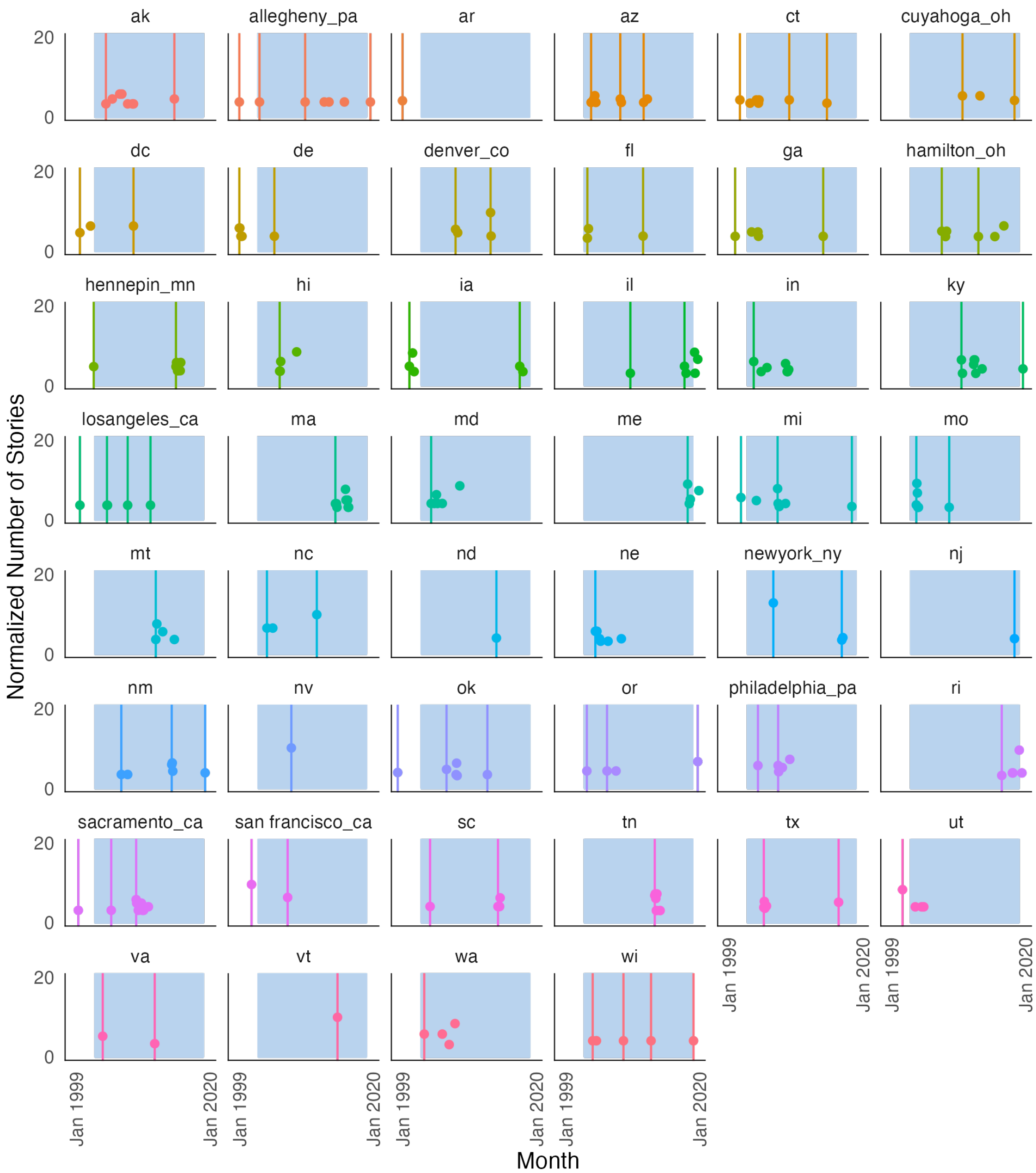


Figure A1: Points represent the timing of news spikes in each jurisdiction. News spikes often occur in clusters; I count discrete events as those that occur at least 3 years after a previous spike. New events are depicted with a vertical line in the this figure. Events that occurred between 2002 and 2018 are included in the analysis; this allows me to see whether events occurred in the three years before the first event, and allows a 24 month period to observe outcomes after the event.

Appendix Figure A2
Jurisdictions in the Analytic Dataset

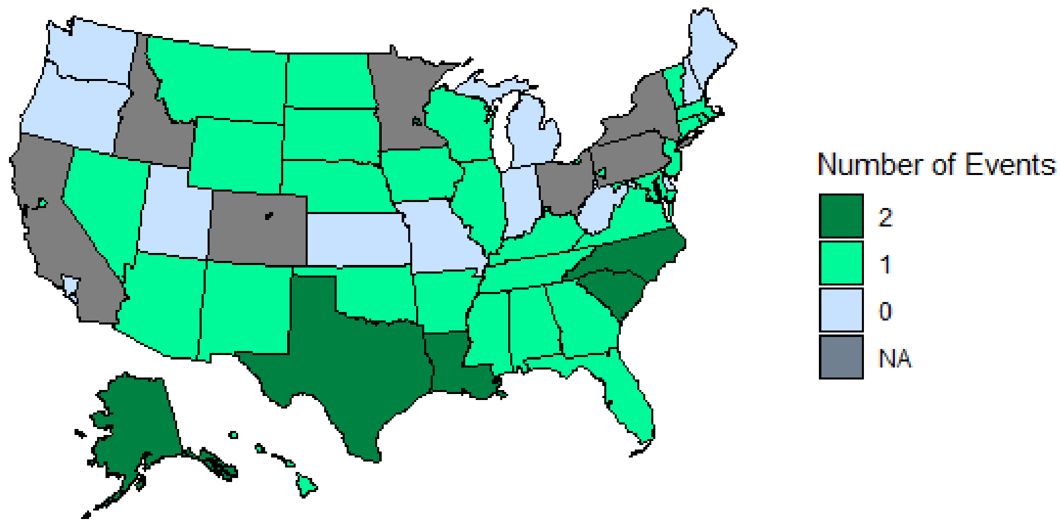


Figure A2: Jurisdictions that experienced a high-profile maltreatment death included in the analysis are green; jurisdictions that did not experience a high-profile maltreatment death are blue. Most child welfare jurisdictions are states, but in nine states child welfare is administered at the county level: California, Colorado, Minnesota, New York, North Carolina, North Dakota, Ohio, Pennsylvania, and Virginia. When counties are masked in the child welfare data due to small numbers I exclude them from the analysis; these counties are gray.

Appendix Figure A3
Newspaper Sources by Jurisdiction

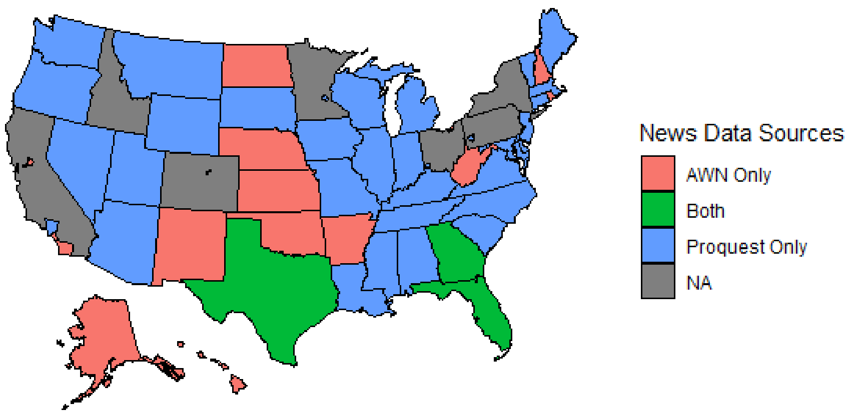


Figure A3: Newspaper source by jurisdiction.

Appendix Figure A4
Number of Newspapers by Jurisdiction

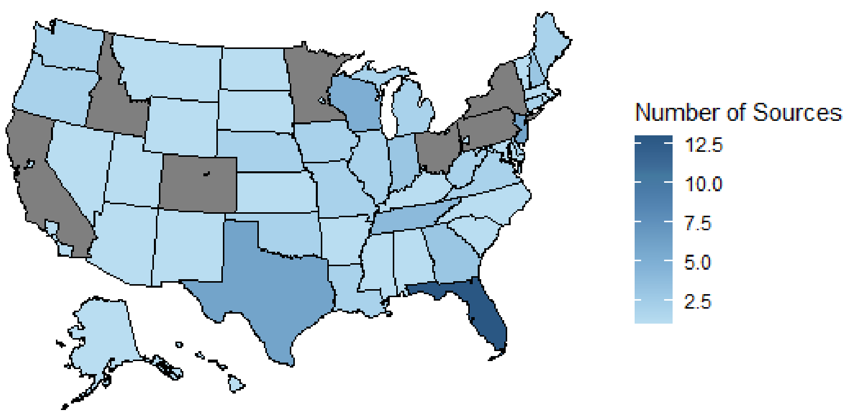
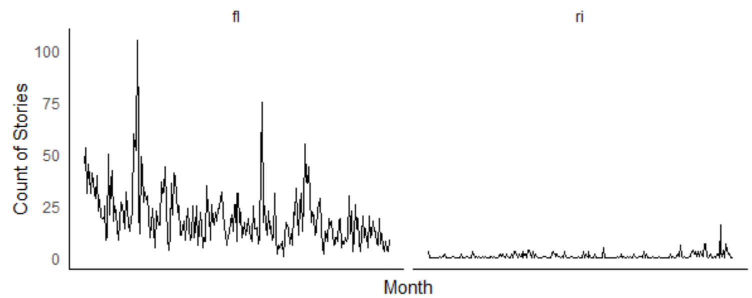
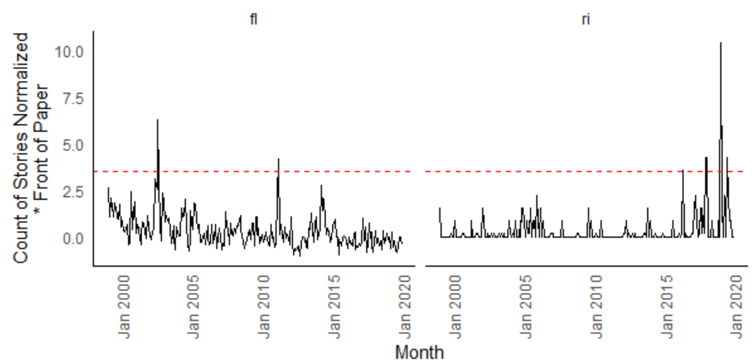


Figure A4: Number of newspapers by jurisdiction.

Appendix Figure A5
Illustration of Story Count Normalization



(a) Monthly Story Counts



(b) Monthly Story Normalized Counts

Figure A5: Normalization process: because newspaper markets in different jurisdictions vary in size, I normalize counts by calculating a Z-score and use the Z-score to assess how unusual the coverage is. These two graphs show the normalization process for two example jurisdictions: Florida and Rhode Island.

Appendix Figure A6
Randomization Inference

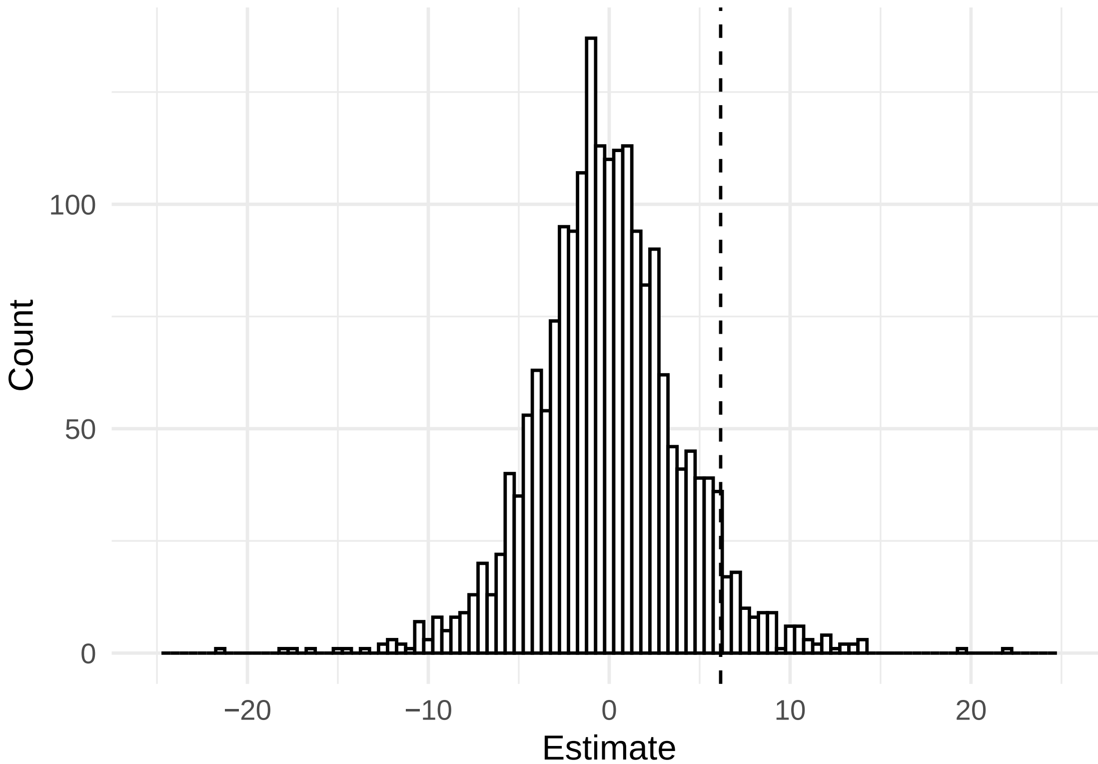


Figure A6: Histogram of long difference removal rate estimates for randomly assigned sets of jurisdiction-month pairs in jurisdictions in the full analysis sample. The dashed line falls at x = 6.16, which is the removal rate estimate produced in the primary analysis. 5.4% of randomly assigned sets of jurisdiction-months have removal rates higher than this estimate.

Appendix Figure A7
Year-to-Year Removal Rate Changes

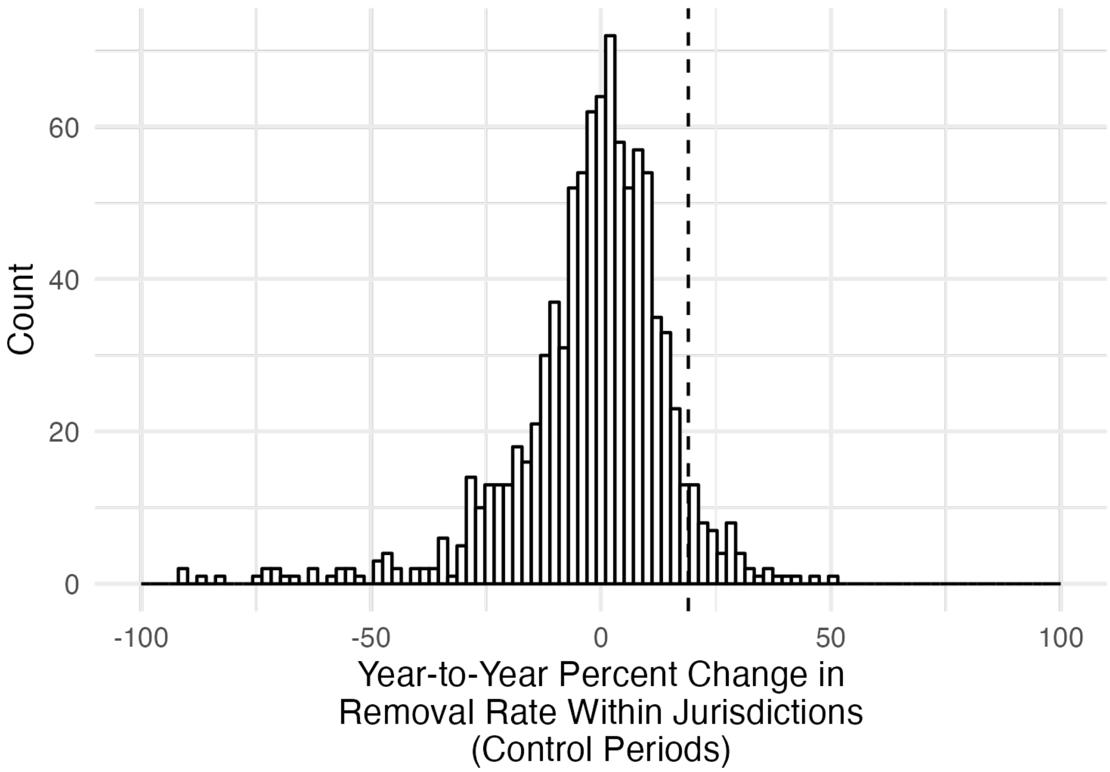
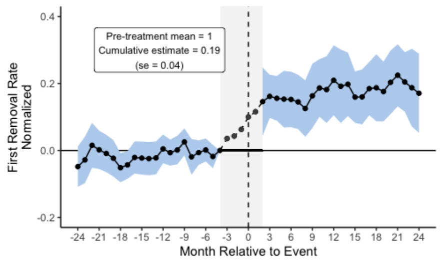
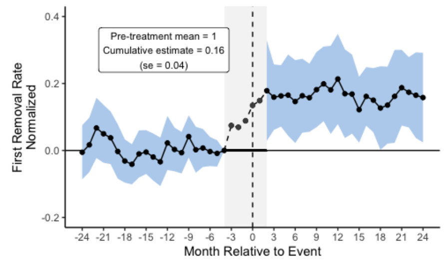


Figure A7: Histogram of control period year-to-year removal rate changes within jurisdictions in the full analysis sample. The dashed line falls at x = 19%, which is the percent removal rate estimate produced in the primary analysis. 3.9% of all year-to-year removal rate changes are larger than this in control periods.

Appendix Figure A8
Impact of a Highly Publicized Tragedy on Normalized Removal Rates in Each Sample



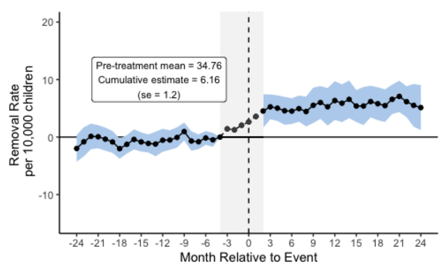
(a) Normalized Removal Rate
(Full Sample)



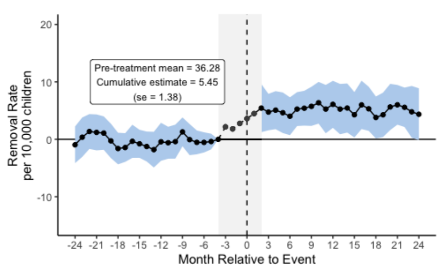
(b) Normalized Removal Rate
(Investigation Sample)

Figure A8: The percentage change in removal rates is very similar in both the full sample and the investigation sample, which is used to assess pre-removal decisions.

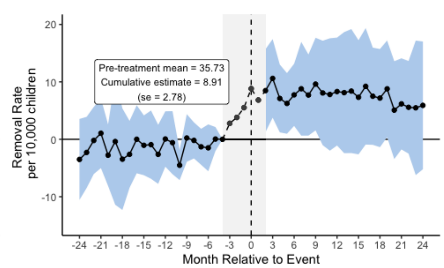
Appendix Figure A9
Impact of a Highly Publicized Tragedy on Removal Rates in Each Sample



(a) Removal Rate
(Full Sample)



(b) Removal Rate
(Investigation Sample)



(c) Removal Rate
(Injury Sample)

Figure A9: In all three samples there is a large and sustained increase in removal rates following a public tragedy.

Appendix Figure A10
Risk Model Coefficients

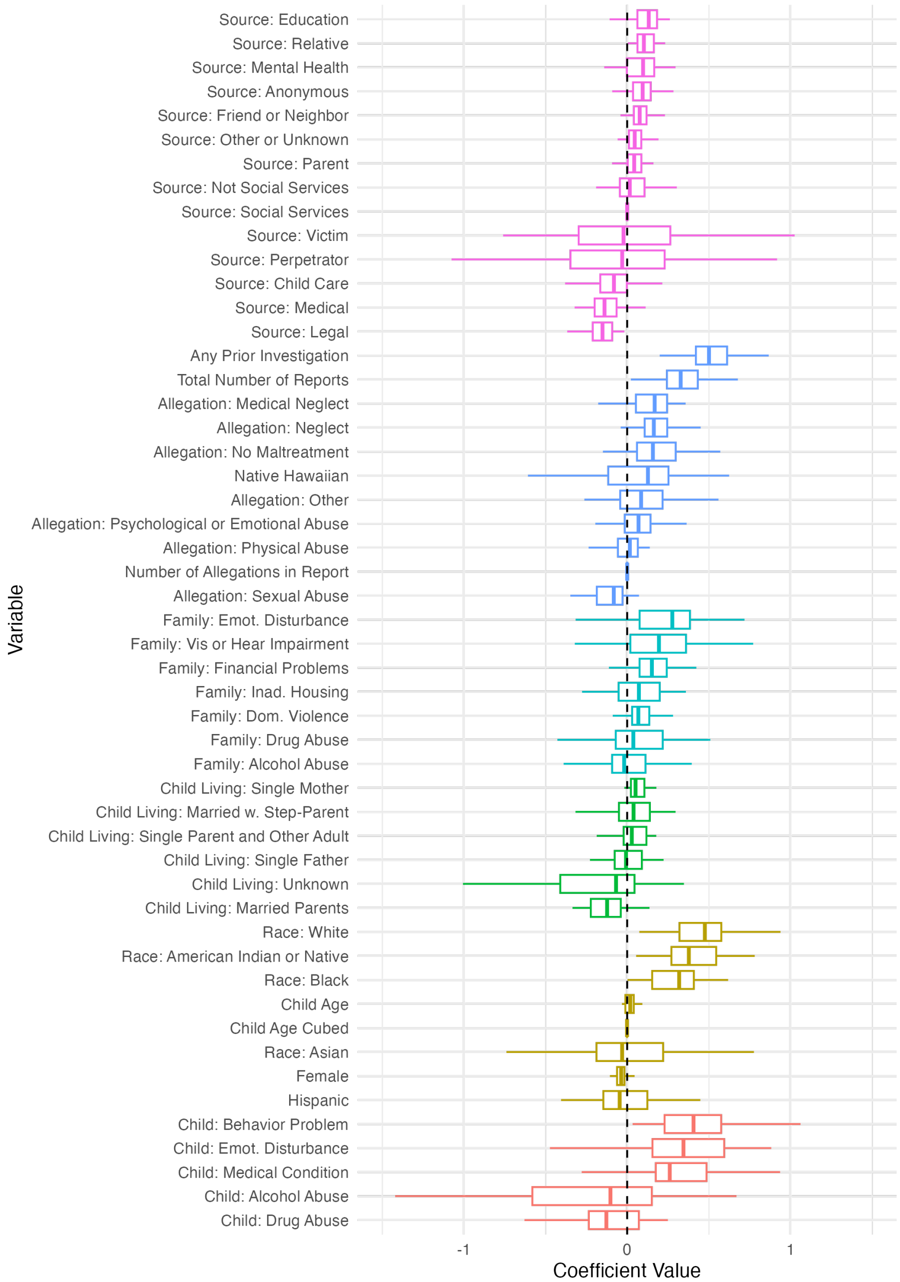


Figure A10: The risk model consists of multiple jurisdiction-specific lasso-logistic models built on 20% training datasets of children who were investigated but not removed from their homes. The model predicts the likelihood the investigated child is re-investigated within 6 months of the close of the investigation. This plot depicts the range of the magnitude of the coefficients on each variable included in the models across jurisdictions.

Appendix Figure A11
Model Predictive Power:
Children Re-Reported Within 6 Months by Predicted Risk Decile.

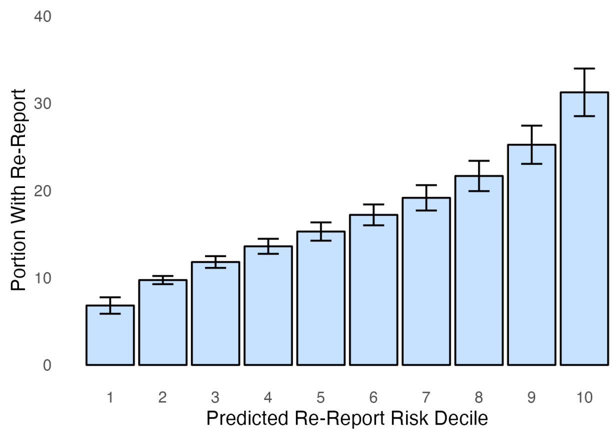


Figure A11: Among children who were investigated and not initially removed in the test dataset, portion of children in each predicted risk decile who were rereported within 6 months of the original investigation disposition date. Test dataset is an 80% sample of investigations in usual times, i.e. not after a highly-publicized tragedy. A separate train dataset was used to train the model.

Appendix Figure A12
Investigator Removal Decisions:
Children Removed by Predicted Risk Decile.

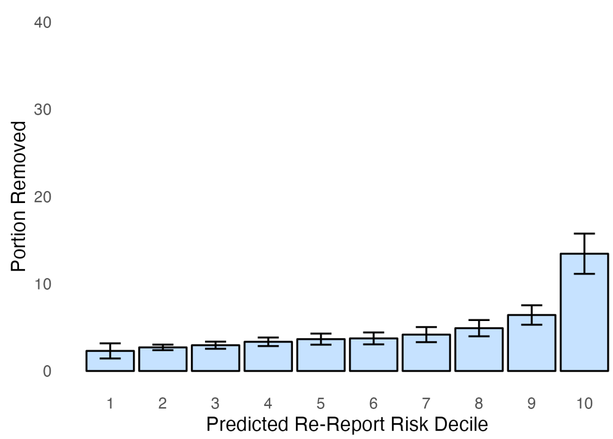


Figure A12: Among all children in the test dataset, portion of investigated children removed by predicted risk decile. Test dataset is an 80% sample of distinct investigations in usual times, i.e. not after a highly-publicized tragedy. A separate train dataset was used to train the model.

Appendix Figure A13
Predictive Model: Outcomes By Race

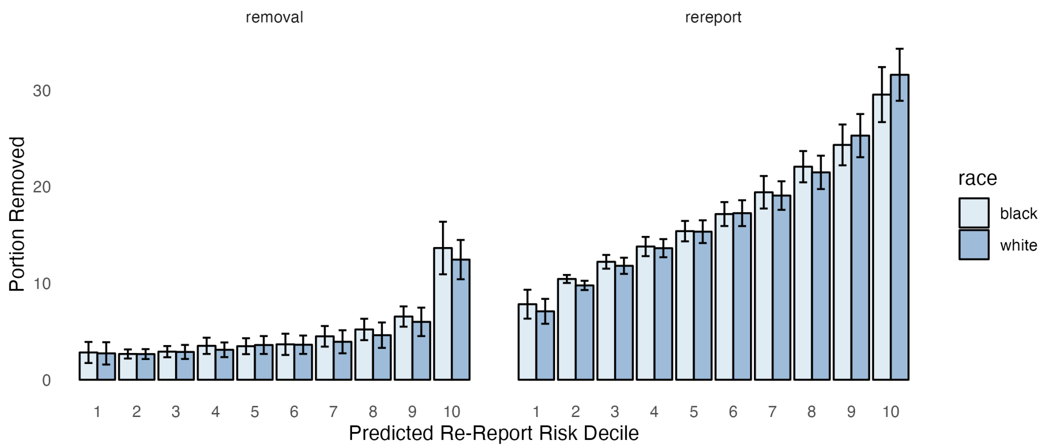
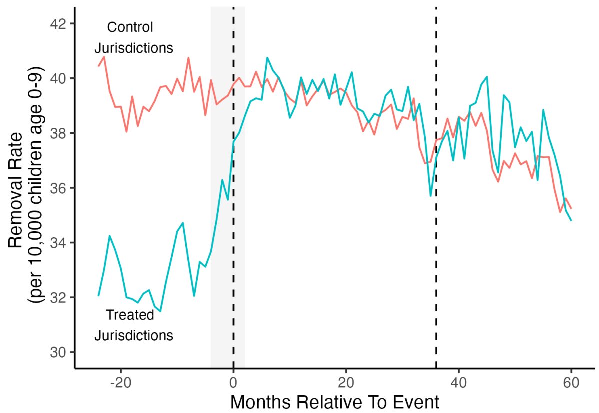
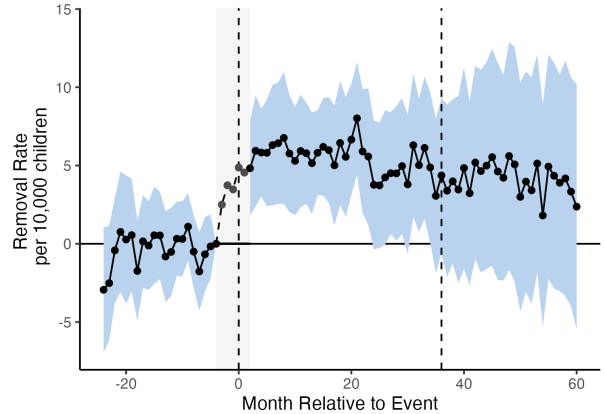


Figure A13: removal shows portion of all children investigated who were removed, by risk decile and race. rereport shows portion of all children who were not removed who were then rereported within 6 months, by risk decile and race.

Appendix Figure A14
Long-Term Impact of a Highly Publicized Tragedy on Removal Rates



(a) Raw Removal Rates (Full Sample)



(b) Removal Rate Event Study Estimates (Full Sample)

Figure A14: Panel (a) shows weighted mean removal rates per 10,000 children age 0-9 in control jurisdictions and treatment jurisdictions in the 24 months before and after a highly-publicized tragedy. Panel (b) shows equivalent dynamic difference-in-differences coefficients: average change in removal rate in a treated jurisdiction around the time of a high profile abuse death, relative to month -4. All estimates are weighted by number of children age 0-9.

Appendix Figure A15
Heterogeneity in Impact of a Highly-Publicized Tragedy on Removal Rates

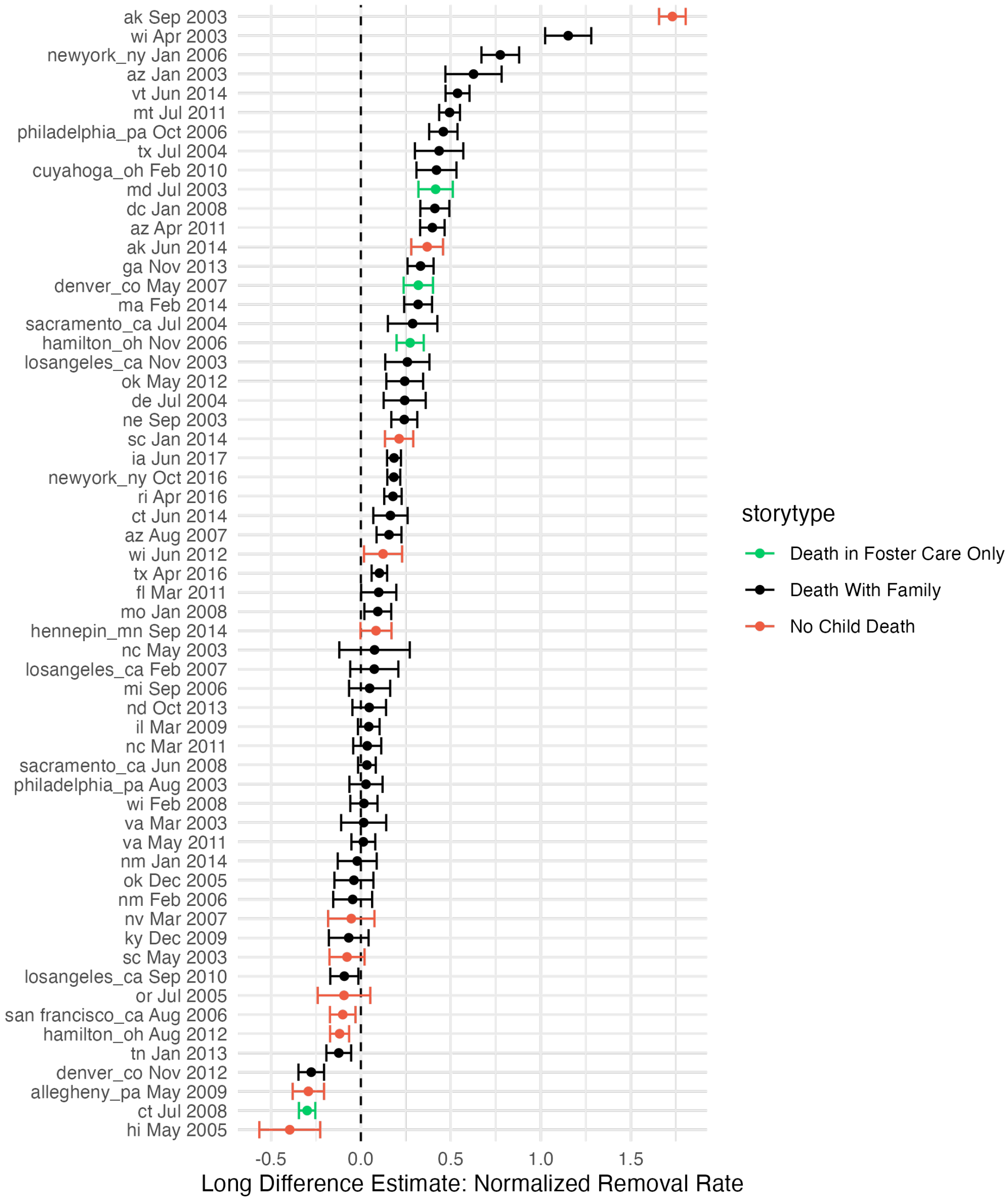
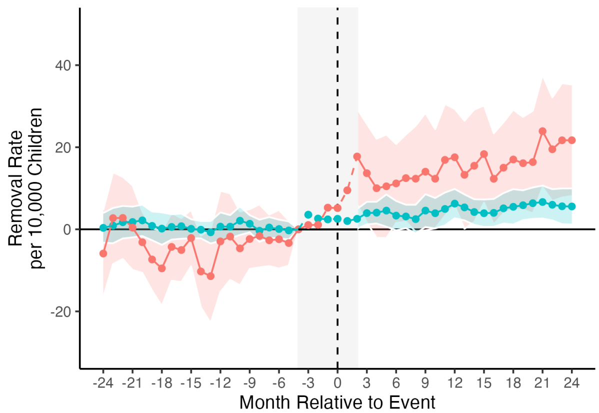
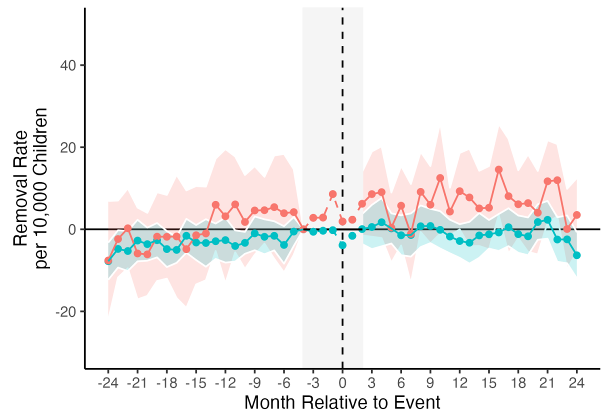


Figure A15: Event-specific impact of a highly-publicized maltreatment death on removal rates.

Appendix Figure A16
Impact of Highly Publicized Tragedy on Removal Rates of Black and White Children by Race of Child Who Died



(a) Removal Rate of Black and White Children After Death of a Black Child



(b) Removal Rates of Black and White Children After Death of a White Child

Figure A16: Red line is removal rate of Black children, green is removal rate of White children. Removal rates increase more for Black children regardless of the race of the child who died.

Appendix Table A1
Impact of Confirmed Highly-Publicized Maltreatment Death

	Estimate (Long Difference)			
	Est.	S.E.	Clusters	N
Removal Rate	6.526***	1.308	53	59,976
Removal Rate (Normalized)	0.192***	0.043	53	59,976
Black Removal Rate	13.672***	3.313	53	59,976
Black Removal Rate (Normalized)	0.187***	0.047	53	59,976
White Removal Rate	4.112***	1.237	53	59,976
White Removal Rate (Normalized)	0.151**	0.057	53	59,976
Total Removal Rate	4.735***	1.217	53	59,976
Total Removal Rate (Normalized)	0.138***	0.037	53	59,976
Repeat Removal Rate	0.420*	0.198	53	59,976
Repeat Removal Rate (Normalized)	0.078*	0.035	53	59,976
Rate of Removal and TPR in 2 Years	0.833*	0.413	53	59,976
Rate of Removal and TPR in 2 Years (Normalized)	0.089+	0.052	53	59,976
Portion of Removals with TPR in 2 Years	−0.869+	0.467	53	59,976
Portion of Removals with TPR in 2 Years (Normalized)	−0.103**	0.036	53	59,976

Table A1: Long difference estimates from subset of events confirmed to be about maltreatment deaths.

Appendix Table A2
Events Included in Primary Analyses

Month	Jurisdiction	Child Death	Name of child
Jan 2003	az	✓	Anndreah Robertson
Mar 2003	va	✓	Leatrice Chambers
Apr 2003	wi	✓	Cristian Cisneros
May 2003	nc	✓	Gabriel Duckett
May 2003	sc	✓	
Jul 2003	md	✓	Ciara Jobes
Aug 2003	philadelphia_pa	✓	Porchia Bennett
Sep 2003	ak		
Sep 2003	ne	✓	Diana Molina
Nov 2003	losangeles_ca	✓	Isaac Lopez
Jul 2004	de	✓	Madison Jade
Jul 2004	sacramento_ca	✓	Akira and Alexia Noel
Jul 2004	tx	✓	Davontae Williams
May 2005	hi	✓	
Jul 2005	or		
Dec 2005	ok	✓	Kelsey Smith-Briggs
Jan 2006	newyork_ny	✓	Nixzmary Brown
Feb 2006	nm	✓	
Aug 2006	san francisco_ca		
Sep 2006	mi	✓	Isaac Lethbridge
Oct 2006	philadelphia_pa	✓	Danieal Kelly
Nov 2006	hamilton_oh	✓	Marcus Fiesel
Feb 2007	losangeles_ca	✓	Darion Wheat
Mar 2007	nv		
May 2007	denver_co	✓	Chandler Grafner
Aug 2007	az	✓	Dustin Rhodes
Jan 2008	dc	✓	Brittany Jacks, Tatianna Jacks, N'Kiah Fogle, Aja Fogle
Jan 2008	mo	✓	Landon Eitel
Feb 2008	wi	✓	Anastasia Vang
Jun 2008	sacramento_ca	✓	Valeeya Brazile
Jul 2008	ct	✓	Michael Brown Jr.
Mar 2009	il	✓	Benjamin Kingan
May 2009	allegheny_pa		
Dec 2009	ky	✓	Robert Ross Jr.
Feb 2010	cuyahoga_oh	✓	Alexandria Hamilton
Sep 2010	losangeles_ca	✓	
Mar 2011	fl	✓	Noah Fake
Mar 2011	nc	✓	Aubrey Littlejohn
Apr 2011	az	✓	Schala Vera
May 2011	va	✓	Angeli Callender
Jul 2011	mt	✓	October Perez
May 2012	ok	✓	
Jun 2012	wi		
Aug 2012	hamilton_oh	✓	
Nov 2012	denver_co	✓	Jessica Ridgeway
Jan 2013	tn	✓	
Oct 2013	nd	✓	Hana Williams
Nov 2013	ga	✓	Emani Moss
Jan 2014	nm	✓	Omaree Varela
Jan 2014	sc	✓	
Feb 2014	ma	✓	No Name
Jun 2014	ak		
Jun 2014	ct	✓	
Jun 2014	vt	✓	Name Redacted
Sep 2014	hennepin_mn		
Apr 2016	ri	✓	Leah Harrington
Apr 2016	tx	✓	Leiliana Wright
Oct 2016	newyork_ny	✓	Myls Dobson
Jun 2017	ia	✓	Sabrina Ray
Sep 2017	il	✓	Steven Figueroa

Table A2: Names of children who suffered maltreatment deaths in each jurisdiction-month included in analysis.

Appendix Table A3
Criteria Used to Identify Stories about Maltreatment Deaths

	ProQuest	**Access World News**
Child death	(<child> <death>) OR (<death> of <child>)	<child> AND <death>
Due to Maltreatment	<maltreat*> OR <abus*> OR <neglect*> OR <invest*> OR <visit>	<maltreat*> OR <abus*> OR <neglect*> OR <invest*> OR <visit>
Agency	<Generic agency name> OR <Jurisdiction-specific name>	<Generic agency names> OR <Jurisdiction-specific name>
In jurisdiction	<County name> (if county)	<County name> (if county)
Highly-publicized	In first 5 pages of newspaper or section	In first 5 pages of newspaper or section

Table A3: Criteria used to identify newspaper stories about public tragedies from full text in each database.

Equation 12
Triple Difference Regression Specification

$$
\begin{aligned}
Y_{jtd} = &\alpha_j + \beta_0 Z_{jt} + \beta_1 Black + \mu_j(\alpha_j * Black) + \beta_2(Z_{jt} * Black) + \\
&\sum_{\tau=-24}^{24}(EventTime_{dt}^{\tau} * \gamma_d^{\tau}) + \sum_{\tau=-24}^{24}(EventTime_{dt}^{\tau} * \gamma_d^{\tau} * Black) + \\
&\delta_0 Treated_{jd} + \delta_1 Treated_{jd} * Black + \\
&\delta_3(After_{dt} * Treated_{jd}) + \delta_4(After_{dt} * Treated_{jd} * Black) + \\
&\kappa_0(Zero_{dt} * Treated_{jd}) + \kappa_1(Zero_{dt} * Treated_{jd} * Black) + \varepsilon_{jtd}. \qquad (12)
\end{aligned}
$$

Table A4: Triple Difference Specification used to estimate impact of highly-publicized deaths on Black-White removal rate gap.

Appendix Table A5
Jurisdictions Included in Each Sample

State	County	FullSample	InvsSample	InjurySample
Alabama		✓		
Alaska		✓		
Arizona		✓	✓	✓
Arkansas		✓	✓	
California	Losangeles	✓	✓	✓
California	Orange	✓	✓	
California	Sacramento	✓	✓	✓
California	San Francisco	✓	✓	
Colorado	Denver	✓	✓	✓
Connecticut		✓	✓	
Delaware		✓	✓	
District Of Columbia		✓	✓	
Florida		✓	✓	
Georgia		✓		
Hawaii		✓	✓	
Illinois		✓	✓	
Indiana		✓	✓	
Iowa		✓	✓	
Kansas		✓	✓	
Kentucky		✓	✓	✓
Louisiana		✓	✓	
Maine		✓	✓	
Maryland		✓		
Massachusetts		✓	✓	✓
Michigan		✓		✓
Minnesota	Hennepin	✓	✓	
Mississippi		✓	✓	
Missouri		✓	✓	
Montana		✓	✓	
Nebraska		✓	✓	
Nevada		✓		✓
New Hampshire		✓	✓	
New Jersey		✓	✓	✓
New Mexico		✓	✓	
New York	Newyork	✓	✓	✓
North Carolina		✓		✓
North Dakota		✓		
Ohio	Cuyahoga	✓	✓	
Ohio	Hamilton	✓	✓	
Oklahoma		✓	✓	
Oregon		✓		
Pennsylvania	Allegheny	✓	✓	
Pennsylvania	Philadelphia	✓	✓	
Rhode Island		✓		✓
South Carolina		✓	✓	
South Dakota		✓		
Tennessee		✓		
Texas		✓	✓	
Utah		✓	✓	✓
Vermont		✓		✓
Virginia		✓	✓	
Washington		✓	✓	✓
West Virginia		✓	✓	
Wisconsin		✓		
Wyoming		✓		

Table A5: Jurisdictions Included in Each Sample.

Appendix Table A6
Jurisdictions Not Included in Each Sample

State	County	FullSample	InvsSample	InjurySample
California	Alameda			
California	Contracosta			
California	Fresno			
California	Kern			
California	Other			
California	Riverside			
California	Sanbernardino			
California	Sandiego			
California	Sanjoaquin			
California	Sanmateo			
California	Santaclara			
California	Tulare			
Colorado	Adams			
Colorado	Elpaso			
Colorado	Jefferson			
Colorado	Other			
Idaho				
Minnesota	Other			
Minnesota	Ramsey			
New York	Erie			
New York	Other			
Ohio	Franklin			
Ohio	Montgomery			
Ohio	Other			
Ohio	Summit			
Pennsylvania	Other			
Puerto Rico				

Table A6: Jurisdictions Not Included in Any Sample. County listed as Other refers to the full set of counties in the state that are masked in the child protection data due to small population size. In Idaho and other named counties my datasets of newspaper stories did not contain a newspaper where I could see full-text stories for the full time series.

Appendix Table A7
Robustness Check: Long Difference Estimates with Different Sets of Controls

	Est.	S.E.	N. Events	Clusters	Obs.
Never- and Not-Recently Treated Controls (Primary Analysis)					
Absolute	6.163***	1.196	60	55	76,525
Normalized	0.189***	0.039	60	55	76,525
Not-Recently Treated Controls					
Absolute	7.316***	1.317	60	45	47,635
Normalized	0.214***	0.042	60	45	47,635
Never-Treated Controls					
Absolute	3.251*	1.370	60	50	31,828
Normalized	0.128***	0.038	60	50	31,828

Table A7: Robustness check: Long difference estimates of the impact of a highly-publicized death on absolute and normalized (% change) removal rates, for samples constructed with different sets of control jurisdictions. Estimate magnitudes are robust to control period lengths varying from 2 to 8 years. Precision declines alightly as the control length increases due to fewer events, clusters and observations, but all estimates remain significant at the 1% level. In the primary analysis I use a control period length of 3 years.

Appendix Table A8
Robustness Check: Long Difference Estimates with Different Control Periods

	Est.	S.E.	N. Events	Clusters	Obs.
Control Period: 2 Years					
Absolute	6.024***	1.216	70	55	97,495
Normalized	0.178***	0.038	70	55	97,495
Control Period: 3 Years (Primary Analysis)					
Absolute	6.163***	1.196	60	55	76,525
Normalized	0.189***	0.039	60	55	76,525
Control Period: 4 Years					
Absolute	6.098***	1.181	52	55	62,074
Normalized	0.202***	0.042	52	55	62,074
Control Period: 5 Years					
Absolute	6.392***	1.306	45	55	50,026
Normalized	0.218***	0.049	45	55	50,026
Control Period: 6 Years					
Absolute	6.539***	1.358	40	52	42,435
Normalized	0.223***	0.050	40	52	42,435
Control Period: 7 Years					
Absolute	6.562***	1.528	36	51	36,996
Normalized	0.232***	0.054	36	51	36,996
Control Period: 8 Years					
Absolute	6.546***	1.633	34	46	33,090
Normalized	0.243***	0.057	34	46	33,090

Table A8: Robustness check: Long difference estimates of the impact of a highly-publicized death on absolute and normalized (% change) removal rates, for samples constructed with different control periods after one month with unusually high coverage. Estimate magnitudes are robust to control period lengths varying from 2 to 8 years. Precision declines slightly as the control length increases due to fewer events, clusters and observations, but all estimates remain significant at the 1% level. In the primary analysis I use a control period length of 3 years.