The Push for Racial Equity in Child Welfare: Can Blind Removals Reduce Disproportionality?*

E. Jason Baron†

Ezra G. Goldstein‡

Joseph Ryan§

September 2022

Abstract

We conduct the first quantitative analysis of "blind removals," an increasingly popular reform that seeks to reduce the over-representation of Black children in foster care by eliminating biases in the removal decisions of investigators. We first show that over-representation in most foster care systems is driven by Black children being substantially more likely than White children to be investigated for maltreatment to begin with. Conditional on initial rates of investigation, investigators remove White and Black children similarly. Second, we find no evidence that blind removals impacted the already small racial disparities in the removal decision, but they substantially increased time to removal.

*We received valuable feedback from Anthony Bald, Patrick Bayer, Joseph Doyle, Brett Drake, Sarah Font, Marie-Pascale Grimon, Max Gross, Brian Jacob, Bridgette Lery, Nicholas Lovett, Chris Mills, Peter Nencka, Jessica Pryce, Emily Putnam-Hornstein, Luke Rodgers, Cullen Wallace, David Welsch, Cathy Xue, and participants at the University of Michigan's Causal Inference in Education Research Seminar and the University of Southern California Children's Data Network. We also thank the Child Welfare Director and Departmental Specialist in Kent County, and the many other child welfare employees across Michigan and Kent County for their help in understanding how the blind removals program works in practice and for bringing humanity to the data. We appreciate Brian Jacob, the Child and Adolescent Data Lab, the Education Policy Initiative, and the Youth Policy Lab for their generosity in sharing data. The project received approval from the University of Michigan's Institutional Review Board: HUM00104615. This research was funded with help from training grant R305B170015 from the U.S. Department of Education's Institute of Education Sciences. Some of the data used in this study 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 dataset collection personnel, the funding source, NDACAN, and Cornell University bear no responsibility for the analyses or interpretations presented here. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not reflect the views of any other entity.

†Duke University and NBER. Email: jason.baron@duke.edu.

‡Pennsylvania State University. Email: eggoldstein@psu.edu.

§University of Michigan. Email: joryan@umich.edu.

> "We are in the midst of a time in this country where a lot of attention is being paid to systemic racism. Guess what? The foster care system is not excused. After we acknowledge that racism and bias have roots deep in the child welfare system, we must take action to address the inequity."
>
> —Aysha Schomburg
> Associate Commissioner
> U.S. Children's Bureau
> May 2021

I Introduction

Child welfare and foster care involvement is surprisingly common in the U.S. By age 18, 37 percent of children—including up to 53 percent of Black children—will have a child welfare investigation for alleged abuse or neglect (Kim et al., 2017). Similarly, by age 18, up to 5 percent of children—including up to 9 percent of Black children—will have entered foster care at some point (Yi, Edwards and Wildeman, 2020). Child welfare leaders and scholars have long expressed concerns regarding racial disproportionality in the U.S. foster care system—the fact that Black children are represented in foster care systems at levels much higher than their numbers in the overall population. Specifically, while Black children make up under 14 percent of the U.S. child population, they account for nearly a quarter of all children in foster care systems in the U.S.

As the opening quote highlights, calls for reforms focused on racial equity have grown louder in recent months, as the nationwide push to examine structural racism in institutions ranging from higher education to criminal justice has reached the child welfare system and specifically foster care. While a number of initiatives such as diversity and anti-racism training for investigators have been piloted over the last few years (Reddy, Williams-Isom and Putnam-Hornstein, 2022), a program known as "blind removals" has recently seen a dramatic surge in popularity.

Blind removals propose to weed out implicit biases when child welfare investigators are weighing whether or not to remove a child from their home due to an allegation of abuse or neglect. The idea behind the program is straightforward, and is borrowed from experiments that mask demographic characteristics of musicians auditioning for orchestras (Goldin and Rouse, 2000). Specifically, blind removals work off of the following premise: if demographic information is not known to child welfare professionals deciding whether or not to remove a child, then implicit biases will not impact foster care placement decisions.

Prior to blind removals, investigators had immense discretion over the decision to remove and place a child in foster care. Once an investigator deemed a child was in imminent harm and recommended removal, all that was needed to file a court petition to remove was approval from the investigator's supervisor. Under blind removals, once an investigator deems the child is in imminent harm and recommends removal, all relevant paperwork is provided to clerks who redact any information that could possibly reveal the child's race and socio-economic status. The investigator who conducted the initial assessment of risk now presents the redacted case to a "blind removals committee" of 10–12 child welfare professionals at a "blind-removal meeting." At the conclusion of the meeting, those in attendance, including the initial investigator, come to a consensus and determine whether or not a court petition will be filed to remove. Therefore, while the initial investigator does know the race of the child, the remaining committee members do not.

Blind removals were pioneered in Nassau County, NY in 2010, but the popularity of the program has drastically increased in the last few years. Driven largely by its intuitive appeal and growing calls to reduce disproportionality, agencies across the country have expressed a growing interest in adopting the program. At the time of writing, a bill before the California Legislature would adopt the program in select counties and New York issued an administrative order in October 2020 calling for the program to be implemented statewide. Similarly, agencies in Los Angeles County, Austin, Chicago, and Baltimore have expressed interest in the program as well.[1]

Yet as enthusiasm around the program has rapidly increased, so has skepticism. Some critics of blind removals argue that the program does not go far enough and that small tweaks to the system cannot remedy the purported racial biases embedded in child welfare. Other critics argue that race *should* be considered at the removal decision and that higher standards should be put in place to remove Black children.[2] Finally, others worry that the program will be ineffective, place an unnecessary burden on a profession that already suffers from high rates of turnover, and increase the time that it takes to remove a child.

Despite the growing debate surrounding blind removals, there exists no causal evidence on their effectiveness. This paper fills this gap in the literature and provides the first quantitative examination of blind removal programs. There are only two counties in the country that have fully implemented the program: Nassau County in 2010 and Michigan's Kent County in August 2019. Examining the causal effects of Nassau County's program on foster care placement outcomes has been difficult due to data availability and quality concerns as well as concurrent reforms in the county that make it difficult to estimate the effect of blind removals alone. Due to these challenges with studying Nassau County's reform, we focus our analysis on the only other county to have implemented blind removals: Michigan's Kent County, home to Grand Rapids.

We begin broadly by asking whether blind removals are well-suited to achieve their intended goal in the first place: to reduce racial disproportionality—the over-representation of Black children in foster care systems. Disproportionality in foster care systems is the result of many decisions and

[1] See, for example Loudenback, J (2021), *Color-Blind Ambition*, The Imprint. Accessed at: `https://www.imprintnews.org/` (June 17, 2021) and Cosgrove, J (2021), *Why are Black children removed from homes at high rate? L.A. County plans 'blind removal' pilot*, Los Angeles Times. Accessed at: `https://www.latimes.com/` (July 14, 2021).

[2] For example, the Minnesota legislature is considering the Minnesota African American Family Preservation Act, which places higher standards for the removal of Black children.

factors, both external and internal to the child welfare system. There are at least two, non-mutually exclusive reasons why a foster care system could exhibit disproportionality. The first reason relates to factors external to the child welfare system. In 2019 the poverty rate in the U.S. was 10.5 percent. However, the poverty rate for Black families was nearly 20 percent. These income disparities coupled with the fact that child maltreatment is positively related to economic hardship could contribute to racial disproportionality in child welfare systems (Lindo, Schaller and Hansen, 2018). The second reason relates to factors internal to the child welfare system. Implicit biases in decisions made by child welfare personnel could contribute to disproportionality. For instance, investigators weighing allegations of abuse and neglect could judge parents of color more harshly and may be more likely to remove their children relative to White parents, even when the circumstances are similar.

Understanding the role of each of these explanations in shaping racial disproportionality in foster care systems is crucial to assess the potential effectiveness of programs such as blind removals. For instance, if disproportionality primarily stems from biases in decision-making within the system, and particularly at the decision to remove, then programs that seek to address implicit biases within the system such as blind removals may hold promise. However, if the probability of removal conditional on having an investigation is similar by race, and disproportionality is instead driven by differential initial rates of investigation, then programs such as blind removals will have limited effects on disproportionality.

Therefore, our paper begins by examining the sources of disproportionality in Kent County and Michigan foster care systems prior to the implementation of blind removals. To do so, we obtained a comprehensive administrative dataset from the Michigan Department of Health and Human Services (MDHHS) containing the universe of child maltreatment investigations in Michigan between January 2010 and March 2021. The data link investigations to their placement records, which allows us to determine whether a child was removed or not following a specific investigation.

Our analysis reveals that, prior to blind removals, virtually all disproportionality in Kent County's foster care system was introduced at the initial point of child welfare contact—the probability of being the subject of a child maltreatment investigation—and not at the removal decision. Specifically, we match the universe of Michigan public education records to the universe of child maltreatment investigations and show that the probability of ever being the alleged victim in a child maltreatment investigation in Kent County varies drastically by race: Black children are almost three times more likely than White children to ever be the alleged victim in a child maltreatment investigation. However, conditional on being the victim in a substantiated maltreatment investigation, removal rates by race are quite similar.

This pattern is also true of the average Michigan county more broadly. Black children in Michigan are twice as likely as White children to be the subject of a maltreatment investigation, but have nearly identical removal rates conditional on an investigation. In fact, we show with a simple counterfactual exercise that equalizing the removal probabilities of White and Black children conditional on an initial investigation in Michigan has a negligible effect on overall disproportionality, while equalizing the probability that a child is the subject of an investigation nearly eliminates it.

Notably, the lack of racial disparities—or differences in probabilities by race—at the removal decision in Kent County and Michigan are also common nationwide. Using the 2018 National Child Abuse and Neglect Data System (NCANDS), a nationwide dataset comprised of investigations for child abuse and neglect throughout the U.S., we show that the cases of Kent County and the state of Michigan do not appear to be outliers, but rather are broadly representative of most counties that report data to NCANDS. These results suggest that policies that target the removal decision such as blind removals may have a limited effect on overall disproportionality in the system.

Importantly, one could imagine a scenario in which blind removals may impact racial disproportionality even when disparities in the removal decision are already small to begin with. Specifically, if biases in child welfare reports are present and a differentially higher number of minor allegations are reported for Black children relative to White children, then equal removal rates conditional on initial investigations would be a sign of bias in the removal decision—in other words, optimal removal rates for Black children conditional on investigation should be lower than those of White children. As a result, if blind removals eliminate implicit biases on the part of the investigator, then in this scenario one might expect the removal rate of Black children (conditional on the initial investigation) to decrease to lower levels than that of White children—which could potentially impact disproportionality in the system.

We show that this scenario is unlikely as both case characteristics and substantiation rates for White and Black children tend to be quite similar in both Kent County and Michigan more broadly. Specifically, in Kent County, the share of investigations for White children that are for child abuse (versus neglect) is roughly 64 percent relative to 61 percent for Black children. Similarly, conditional on an investigation, the substantiation rate is roughly 25.5 percent for White children relative to 27.2 percent for Black children.

Still, in the final part of the article we examine whether blind removals had any impact on the already small racial disparities at the removal decision in Kent County. Our identification strategy is simple: for each investigator we compute a White-Black removal rate differential—the difference between the investigator's removal rate for White and Black children. We then compare the difference in Kent County's average removal rate differential before and after blind removals and relative to the change in control counties in a classic 2×2 difference-in-differences (DID) framework. Importantly, the results in this final section should be interpreted with caution as our setting is not ideal for a causal examination of the effects of blind removals. As mentioned above, the program was implemented in August 2019, soon before the COVID-19 pandemic. Moreover, inference is complicated by the fact that the treatment group consists of only one county.

We find suggestive evidence that blind removals led to a decline in the removal rate for both White and Black children. However, we find no evidence that blind removals had any effect on disparities at the removal decision; the decline in the removal rate for White children was similar to that of Black children. Therefore, while the program may have led to an overall decline in removals, it had no effects on its intended outcome: to reduce the over-representation of Black children in foster care.

This fact alone is of little concern if the program were costless to implement. However, there are a number of potential unintended consequences of the program that we discuss in Section V.C. We show that the program substantially increased the amount of time it takes to remove a child following a child welfare investigation: the median time to removal in Kent County increased by roughly 9 days (or 30 percent) relative to counties in the control group.[3] We find no evidence that this increased time to removal resulted in an increase in the proportion of foster care placements in family settings (as opposed to congregate care)—a measure of placement quality. Taken together, our results imply that blind removals may not be well-suited to significantly reduce racial disproportionality in most child welfare systems. We discuss policy implications of these findings in detail in Section VI.

This paper makes several contributions to the existing literature. First and foremost, it provides the first quantitative examination of blind removals, a popular reform currently considered for implementation in a number of state and local child welfare agencies. The only evidence to date on the effectiveness of blind removals is a qualitative study of Nassau County's program. The study used interviews and focus groups to examine the perceptions of blind removals among child welfare workers in Nassau, and concluded that the program was widely perceived to be successful at reducing disproportionality (Pryce et al., 2019).[4] Our study shows that, even though blind removals are widely perceived to be successful at reducing racial disproportionality, (1) they target a point in child welfare involvement with relatively smaller racial disparities, (2) do not impact racial disparities at the removal decision, and as a result (3) have limited effects on overall racial disproportionality in foster care.[5] While our analysis focuses on a particular county, most implementations of blind removals currently considered across the country closely mirror Kent County's program. Furthermore, as discussed above, our main result that most disproportionality is introduced at the investigation level—and not at the removal decision—holds across most foster care systems in the country.

This study also contributes to the literature focused on racial equity concerns in child welfare systems. There is a large literature documenting racial disproportionality in foster care systems across the U.S. (Edwards et al., 2021; Putnam-Hornstein et al., 2013, 2021; Wildeman and Emanuel, 2014). However, few studies consider disparities by race/ethnicity in the transition probabilities at each step of child welfare involvement (Bartholet et al., 2011). By showing that most disproportionality in foster care systems is introduced at the initial allegation level—as opposed to at the removal decision—our study sheds light on which policies and practices are likely to be most successful at reducing disproportionality.

Our paper also contributes to the small literature in economics examining foster care systems.

[3] It is important to note that during the time from the initial investigation to the time of removal from the home, child welfare personnel continue involvement and consultation with the family.

[4] The findings of this study have been a key driver of the surge in popularity of the program. In fact, the study's findings became widespread after a TED talk by the lead author, titled "To Transform Child Welfare, Take Race Out of the Equation" and published in 2018, garnered over 1.4 million views.

[5] Interestingly, numerous conversations with Kent County staff yield similar qualitative conclusions to those in Nassau County. Specifically, county staff members tend to (1) overestimate the role of the removal decision in shaping disproportionality in the system and (2) perceive the program to be successful at reducing disproportionality.

This literature is mostly focused on estimating the causal effects of foster care on children's outcomes as opposed to sources of racial disproportionality. For instance, Doyle (2013), Doyle (2008), Doyle (2007), Gross and Baron (2022), Baron and Gross (2022), Bald et al. (2022*a*), Roberts (2019), and Warburton et al. (2014) all examine the effects of foster care placements on children's outcomes. Grimon (2021) examines the effects of child welfare investigations on parents, and Lovett and Xue (2020) and Font and Mills (2020) examine the relative effectiveness of different placement types (kinship, unrelated caregivers, and congregate care).[6]

It is important to note that the findings of the literature on the causal effects of foster care are mixed. While papers such as Doyle (2008) and Doyle (2007) find that foster care has negative effects on children at the margin of placement, other studies suggest either much less harmful or even positive effects of foster care (Bald et al., 2022*a*; Baron and Gross, 2022; Gross and Baron, 2022; Roberts, 2019). This paper takes no stance on whether reducing or increasing removal rates is helpful or detrimental to children. Rather, we evaluate the effects of the blind removal program on its intended goal of reducing racial disproportionality without asking whether or not reducing disproportionality is welfare improving.

II Background

Roughly one in five Michigan public school students were the alleged victim in a maltreatment investigation by third grade and one in 60 were removed from their homes and placed in foster care (Gross and Baron, 2022; Ryan et al., 2018). This section describes the maltreatment investigation and removal processes in Michigan and Kent County before and after the implementation of blind removals.

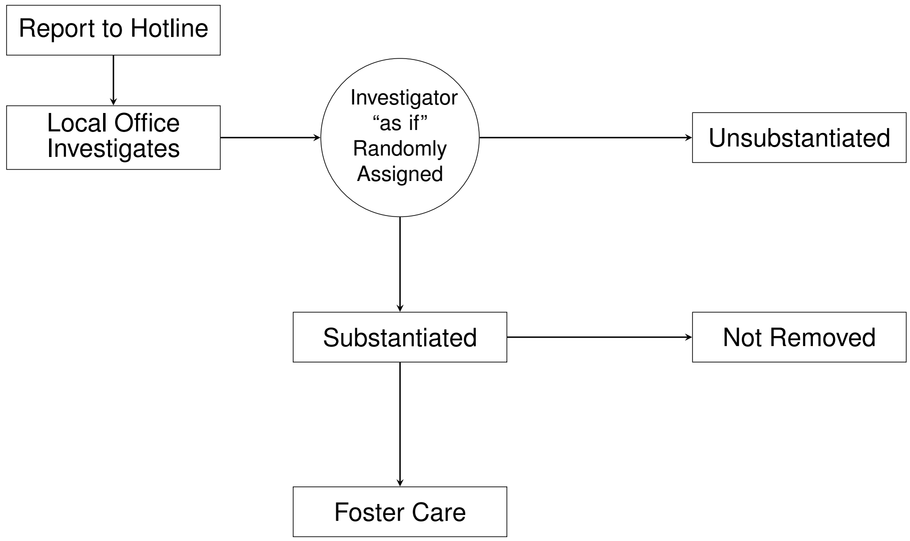
II.A Child Welfare System Pre-Blind Removals

Panel A of Figure 1 describes the child maltreatment investigation process in Michigan and Kent County prior to the implementation of blind removals.[7] The process begins when a call is made to an intake hotline to make an allegation of child abuse (e.g., bruises or burns) or neglect (e.g., lack of supervision or food deprivation). Anyone can make a call to report suspected maltreatment, but the most common reporters are education and law enforcement personnel (Baron, Goldstein and Wallace, 2020).

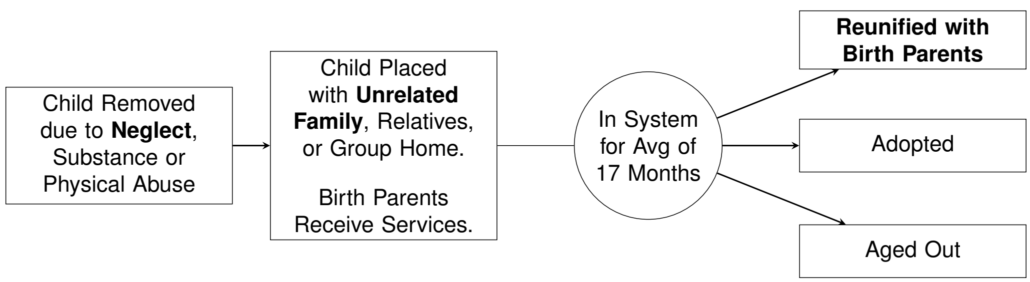
[6] See Bald et al. (2022*b*) for a succinct summary of this literature.

[7] Kent County is the fourth-most populous county in Michigan. It is located in Western Michigan and is home to Grand Rapids, the state's second-largest city.

Figure 1: Child Welfare and Foster Care in Michigan



(a) Child Welfare Process



(b) Foster Care Process

Notes: The figure describes the child maltreatment investigation and foster care processes in Michigan. "Substantiated" means that the investigator found enough evidence to support the allegation of abuse or neglect. Conditional on substantiation, in cases with the most intensive risk the child is also removed from the home and placed in foster care. Panel B highlights the most common experiences in foster care in bold font, according to The Adoption and Foster Care Analysis and Reporting System (AFCARS) FY 2018 Estimates.

A hotline employee screens in the call and transfers the relevant reports to the child's local child welfare office. Once the office receives the reports, it assigns the case to a child maltreatment investigator based on a rotational assignment system rather than their particular skill set or characteristics. The investigator makes two key decisions that jointly determine the outcome of the investigation. First, the investigator interviews the people involved, reviews any relevant police or medical reports, and decides whether there is enough evidence to substantiate the allegation. In Michigan, 74 percent of investigations went unsubstantiated between 2010 and 2018. In these cases, the child welfare office simply concludes the investigation and, if warranted, connects families with services aimed at preventing future child welfare contact.

Conditional on a substantiated investigation, the investigator must also decide how much risk the child is in by continuing to live in the home. Investigators have immense discretion over foster placement. While there is technically a standardized system in place in which investigators complete a 22-question risk assessment that is used to determine whether removal is appropriate, many of the questions are inherently subjective and investigators often manipulate responses to match their priors (Bosk, 2015; Gillingham and Humphreys, 2010).

If the investigator substantiates and determines that the child's living arrangement does not place the child's well-being in imminent risk, then the investigator refers the family to community-based services (e.g., food pantries or support groups) or more intensive targeted services (e.g., substance abuse or parenting classes). However, if the investigator substantiates and believes that the child is in imminent risk, she not only refers the family to services but also requests to her supervisor that a court petition be filed to remove the child. Anecdotally, it is rare for either the supervisor or the judge to disagree with the investigator's recommendation.

Panel B of Figure 1 describes the foster care system in Michigan, which is similar to the rest of the country. Foster care is a temporary and family-oriented intervention. Children are temporarily removed from their homes and placed with either an unrelated foster family, relatives, or in a group home, while their birth parents receive services aimed toward increasing the likelihood of reunification. In Michigan, children spent roughly 17 months in foster care, after which most were reunified with their parents.

II.B Background on Blind Removals

Child welfare leaders and scholars have long expressed concerns regarding racial inequities in child welfare outcomes. Black children are over-represented in Michigan's foster care system: while they make up just 18 percent of the public school population, they account for nearly a quarter of all children in foster care. Such over-representation is common nationwide: 24 percent of children in the foster care system are Black despite their making up just 13.8 percent of the overall child population. This "racial disproportionality"—or the disproportional over-representation of Black children in foster care—has led to calls for reform ranging from policies that focus more on prevention to complete abolition (Barth et al., 2020).[8] While a number of initiatives such as diversity and anti-racism training for investigators have been piloted, blind removals have recently seen a dramatic surge in popularity (Reddy, Williams-Isom and Putnam-Hornstein, 2022).

Blind removals were pioneered over a decade ago in Nassau County, NY. In an effort to reduce racial disproportionality, the New York Office of Children and Family Services awarded 14 counties, including Nassau, a Disproportionate Minority Representation Grant. The grant provided funding to develop and implement strategies that would reduce disproportionality. Child welfare staff in Nassau chose to focus the grant on the removal of children—what they considered to be the most crucial point in child welfare's involvement. The idea behind blind removals is straightforward, and is borrowed from experiments that mask demographic characteristics of musicians auditioning for orchestras (see, for example, Goldin and Rouse (2000)).

[8]For example, the upEND Movement, in partnership with the University of Houston's Graduate College of Social Work promotes abolishing the child welfare system. See `https://cssp.org/our-work/project/upend/` for details.

Description of the Program

Under blind removals, the child welfare investigation process is the same as the one shown in Figure 1. However, whereas prior to blind removals investigators could file a petition to remove after getting their request approved by their supervisors, the decision to remove now faces increased scrutiny. Once the investigator substantiates the case, deems the child is in imminent harm, and gets approval from her supervisor to file a petition to remove, all relevant paperwork is provided to clerks who redact any information that could possibly reveal the child's race and socio-economic status. Specifically, the clerk redacts the child's name, race/ethnicity, zip code, income, the school district the child attends, and the names of any public safety departments involved in the case. The clerk does not, however, redact other relevant information including the child's sex, age, or previous contact with child welfare.

The investigator who conducted the initial assessment of risk then presents the redacted case to a committee of child welfare professionals at a "blind-removal meeting." These meetings are typically scheduled soon after the investigator determines removal should occur. How soon after depends on the level of risk the child may face. If the investigator determines the child must be removed immediately, an emergency blind-removal meeting may be held.

The committee of child welfare professionals typically consists of 10–12 individuals such as the investigator's supervisor, a mental health counselor, a domestic violence liaison, the agency's attorney, and blind removal supervisors—the individuals in charge of implementing the blind-removal process in the county.[9] During the meeting, attendees discuss the risk to the child and potential programs to expedite reunification in the event of removal. At the conclusion of the meeting, those in attendance, including the initial investigator, come to a consensus and determine whether or not a court petition will be filed to remove. Therefore, blind removal meetings serve to not only redact the child's demographic information, but they also add increased scrutiny to each removal decision.

Current Evidence of the Effectiveness of Blind Removals

The popularity of blind removals has drastically increased in the last few years. Driven largely by the intuitive appeal of the program and growing calls to reduce disproportionality, agencies across the country have expressed a marked interest in adopting blind removals. For instance, a bill before the California Legislature would adopt the program in select counties and former New York Governor Andrew Cuomo issued an administrative order in October 2020 calling for blind removals to be implemented statewide, though implementation has not yet occurred. Similarly, agencies in Los Angeles County, Austin, Chicago, and Baltimore have expressed interest in the program as well.[10]

[9] The specific committee members at each meeting can differ from case to case, but depend solely on schedule availability. Anecdotally, there is no evidence that cases with lower or higher risk or with particular characteristics schedule systematically different committee members. Importantly, the blind removal supervisors are always present at these meetings.

[10] See, for example Loudenback, J (2021), *Color-Blind Ambition*, The Imprint. Accessed at: `https://www.imprintnews.org/` (June 17, 2021), Cosgrove, J (2021), *Why are Black children removed from homes at high rate?*

Notably, while virtually all proposed implementations of the program closely mirror Kent County's, there is some variation in the proposed versions across the country: for instance, the New York directive would redact all demographic information of the child, expanding the aforementioned list to include the child's sex.

Despite its promise to address disproportionality, there is no causal evidence on the effectiveness of blind removals. A qualitative study of Nassau County's program used document analysis, in-depth interviews, focus groups and field visits to examine the perceived success of blind removals among child welfare workers (Pryce et al., 2019). The study documented that the Nassau County's commissioner, directors, supervisors, and caseworkers all reported that blind removals had contributed to decreasing the number of Black children being removed from their homes due to abuse or neglect.[11] This study's findings have been instrumental in the rising popularity of blind removals.

While this detailed qualitative evidence can begin to shed light on the blind removals program, examining the causal effects of Nassau County's program on foster care placement outcomes has been difficult for two main reasons. The first reason is data availability and quality concerns: The State of New York does not provide foster care placement data to NCANDS, a widely used data source to study child welfare systems across the U.S. There are also concerns regarding the quality of foster care placement data in Nassau County, where questions about whether or not racial categories have been consistently coded over time have been raised.[12] A second reason is that Nassau County implemented a variety of additional practices alongside blind removals aimed at reducing disproportionality (e.g., efforts to promote a racially and culturally diverse workforce, and school-based initiatives offering after-school care, medical, behavioral, and mental health treatment to families).

Due to these challenges with studying Nassau County's reform, we focus our analysis on the only other county to have implemented blind removals, Michigan's Kent County.[13] In response to perceived racial inequities in its child welfare system, Kent County's Child Welfare Director began exploring policies aimed at reducing disproportionality. Impressed with the program in Nassau County, she tasked child welfare personnel with implementing blind removals in Kent County. The resulting program closely mirrored the original pilot program in Nassau County, and was implemented in

L.A. County plans 'blind removal' pilot, Los Angeles Times. Accessed at: `https://www.latimes.com/` (July 14, 2021), Palmer, P (2022), *'Blind removal' program seeks to remedy racial disproportionality in LA County child welfare system*, ABC Los Angeles. Accessed at: `https://abc7.com` (August 10, 2022), and *How did the blind removal process in Nassau County, N.Y. address disparity among children entering care?*, Casey Foundation. Accessed at: `https://casey.org` (August 10, 2022).

[11] This study, coupled with a claim that the share of Black children in foster care declined from roughly 54 percent prior to blind removals to 35.5 percent in 2016, has contributed to the growing popularity of the program. However, data from the New York Office of Children and Families show that the time series of the share of Black children in foster care in Nassau County is erratic and shows no consistent evidence of declines following blind removals. In fact, 2016 represents an outlier: the share was as high as 62 percent in 2014 and 53 percent in 2019. For more details, see Loudenback, J (2021), *Color-Blind Ambition*, The Imprint. Accessed at: `https://www.imprintnews.org/` (June 17, 2021).

[12] See Riley, NS (2021), *Blinders on*, American Enterprise Institute. Accessed at: `https://www.aei.org/` (July 17, 2021).

[13] Minnesota's Ramsey County began training child welfare workers on blind removals in 2020, but has yet to implement the process.

August 2019.

III Sources of Racial Disproportionality in Foster Care

The goal of blind removals is to reduce racial disproportionality—the over-representation of Black children—in foster care systems. The expected effects of blind removals on this outcome therefore depend on how much disproportionality is due to racial disparities—or differences in probabilities—at the removal decision.

There are at least two, non-mutually exclusive reasons why the foster care system could exhibit disproportionality (Barth et al., 2020). The first reason relates to factors external to the child welfare system. In 2019 the poverty rate in the U.S. was 10.5 percent. However, the poverty rate for Black families was 18.8 percent.[14] These income disparities coupled with the fact that child maltreatment is positively related to economic hardship (Brown and De Cao, 2018; Lindo, Schaller and Hansen, 2018) could contribute to racial disproportionality in child welfare systems. Furthermore, child maltreatment is positively correlated with lower levels of parental educational attainment, higher levels of parental criminal justice contact, and the likelihood of living in a single-parent household (Norris, Pecenco and Weaver, 2021; Oliver, Kuhns and Pomeranz, 2006), and Black families tend to be over-represented along these dimensions as well.[15]

The second reason relates to factors internal to the child welfare system. Implicit biases in decisions made by child welfare personnel could lead to racial disparities within the system. For instance, child welfare decision-makers weighing allegations of abuse and neglect could judge parents of color more harshly and may be more likely to remove their children relative to White parents, even when their circumstances are similar.

Understanding the role of each of these explanations in shaping disproportionality in foster care is crucial to assess the potential effectiveness of blind removals. If disproportionality primarily stems from biases in decision-making within the system, and particularly at the removal decision, then programs designed to address biases within the system such as blind removals may hold promise. If instead disproportionality primarily stems from factors external to the child welfare system, then policies that target the decision to remove may have limited effects.

This section assesses the relative importance of these explanations. Specifically, we examine racial disparities at each step of the decision-making process within the child welfare system using detailed microdata from Michigan and Kent County. We use administrative data from MDHHS which consist of the universe of maltreatment investigations in Michigan between January 2010 and March 2021. The data include details of each investigation, including the date of the allegation, allegation type as coded by the investigator (e.g., abuse or neglect), demographics of the alleged child victim as coded

[14]These estimates come from the 2019 Current Population Survey Annual Social and Economic Supplement.

[15]For instance, Black students are 10 percentage points (50 percent) less likely to graduate high school on time (AECF, 2019), 1.5 times more likely to have a parent exposed to the criminal justice system (Finlay, Mueller-Smith and Street, 2022), and almost twice as likely to live in single-parent households (AECF, 2019).

by the investigator (e.g., race and sex), and indicators for substantiation and removal.

III.A Racial Disparities at Each Step of the Child Welfare Process

Panel A of Table 1 describes the shares by race in each of four populations prior to 2019, the year the blind removals program was implemented in Kent County: Column 1 describes the population of all public school students in Michigan in 2015-16; Column 2 describes the population of all child maltreatment investigations in Michigan from 2010–2018; Columns 3 and 4 describe the populations of all substantiated investigations and removals during the same time period, respectively. Panel B of the table describes the same populations for Kent County.

Table 1: Disproportionality in Child Welfare

	(1) Population	(2) All Allegations	(3) Substantiated	(4) Removed
Panel A: Michigan				
White	67.1	55.7	54.3	53.6
Black	18.2	25.1	26.1	25.4
Hispanic	7.4	6.6	7.2	7.4
Multi-Racial	3.4	8.6	9.2	13.0
Other	3.9	3.9	3.2	0.7
Panel B: Kent County				
White	60.6	42.8	41.1	38.8
Black	13.6	28.0	29.0	30.9
Hispanic	17.4	15.2	16.0	14.8
Multi-Racial	4.6	9.9	10.6	14.6
Other	3.8	4.1	3.3	0.9

Notes: The table shows population shares by race (×100) in four distinct populations. Column 1 shows the population of all public school students in Michigan (Panel A) and in Kent County (Panel B). This information comes from the Michigan Department of Education for the 2015-16 academic year. Column 2 of Panel A shows the population of 1,357,402 investigations in Michigan for alleged victims of child abuse or neglect between January 2010 and December 2018. Panel B shows the population of 93,934 investigations in Kent County. Column 3 in Panel A shows the population of 355,067 substantiated investigations in Michigan, while Panel B shows the population of 27,647 substantiated investigations in Kent. Finally, Column 4 in Panel A shows the population of 50,288 investigations in Michigan that resulted in removal, while Panel B shows the population of 3,437 such investigations in Kent.

The table shows that Black children are over-represented both in Michigan's and Kent County's child welfare and foster care systems. Despite their making up just 18 and 14 percent of the overall student population in Michigan and Kent County, respectively, they account for roughly 25 and 31 percent of victims in foster care.[16]

[16]We focus on White and Black children in this paper for two main reasons: (1) The debate surrounding racial disparities in child welfare systems typically focuses on differential treatment between Black and White children. Specifically, a motivation for many agencies currently considering implementation of the program is to address perceived Black-White disparities in their child welfare outcomes. (2) It is well-known that the category "multi-racial" is not

Table 2 begins to examine whether the over-representation described above arises at particular points of the child welfare system. First, we match the universe of Michigan public education records to the universe of child maltreatment investigations to calculate the probability that a public school student in Michigan will ever be the alleged victim in a child maltreatment investigation from birth through age seventeen. In Michigan, nearly one quarter of public school students are ever the alleged victim in a child maltreatment investigation. However, there is substantial heterogeneity by race: while White children have a 20 percent chance of ever being an alleged victim, Black children have almost a 40 percent chance (Column 1). In Kent County, 17 percent of White children will ever be the alleged victim in a maltreatment investigation, while this share is nearly 50 percent for Black children.

Table 2: Racial Disparities At Each Point in the Child Welfare Process

	(1) $Pr(AllegedVictim)$	(2) $Pr(Sub \mid AllegedVictim)$	(3) $Pr(Rem \mid AllegedVictim)$	(4) $Pr(Rem \mid Sub)$
Panel A: Michigan				
White	20.5	25.5	3.2	11.7
Black	38.8	27.2	3.4	11.4
Hispanic	24.2	28.4	3.7	12.2
Other	14.1	22.3	0.6	2.3
Panel B: Kent County				
White	16.8	27.3	2.9	9.8
Black	47.5	29.4	3.5	10.8
Hispanic	28.1	29.8	3.1	9.7
Other	15.7	23.5	0.7	2.5

Notes: Column 1 shows the probability ($\times 100$) by race that a child will ever be the alleged victim in a child maltreatment investigation in Michigan (Panel A) and Kent County (Panel B). This calculation comes from MDHHS data linked to public education records from the Michigan Department of Education. Specifically, the sample consists of 717,363 first-time sixth-graders in 2008 through 2013. We measure whether a child was ever the alleged victim in a child maltreatment investigation as an indicator equal to one if the child ever appears as an alleged victim in Michigan at any point from birth through age 17. For more details on the match between child welfare and public education records, see Online Appendix D in Gross and Baron (2022). Columns 2–4 focus on the sample of child maltreatment investigations in Michigan between January 2010 and December 2018. This dataset comes from MDHHS. Column 2 shows the probability by race that an investigation is substantiated. Column 3 shows the probability by race that an investigation results in removal. Finally, Column 4 shows the probability by race that a substantiated investigation results in removal.

The table shows that most disproportionality is introduced at the initial point of contact: the probability that a child will be the alleged victim in a child maltreatment investigation. Conditional on being the alleged victim in a maltreatment investigation, substantiation and removal rates exhibit much smaller racial disparities. For instance, among alleged victims of child maltreatment in Kent County, substantiation rates for White and Black children differ by roughly two percentage points while removal rates differ by less than one percentage point. This pattern is also true in Michigan as a whole. Importantly, these patterns do not imply that differences in the substantiation and

consistently coded across investigators. Many investigators will default to this category when they are unsure about the child's race and are hesitant to ask the family directly. Given potentially systematic differences in these cases and the motivation behind the blind removals program, we choose to focus the analysis on White and Black children.

removal probabilities between races are negligible. For instance, the difference in substantiation rates in Kent County between White and Black children of two percentage points is not trivial when benchmarked to the baseline substantiation rate for White children (27.3 percent). The same is true of the difference in removal rates (1 percentage point relative to a baseline removal rate of 9.8 percent for White children). What these findings do imply, however, is that differences in substantiation and removal probabilities pale in comparison to the differences in the probability of being the alleged victim in a maltreatment investigation. These results begin to suggest that factors external to the child welfare system play an important role in shaping racial disproportionality within the system, and that policies that target the substantiation and removal decisions may have limited effects on overall disproportionality in foster care.

III.B Racial Disparities in Investigators' Decisions

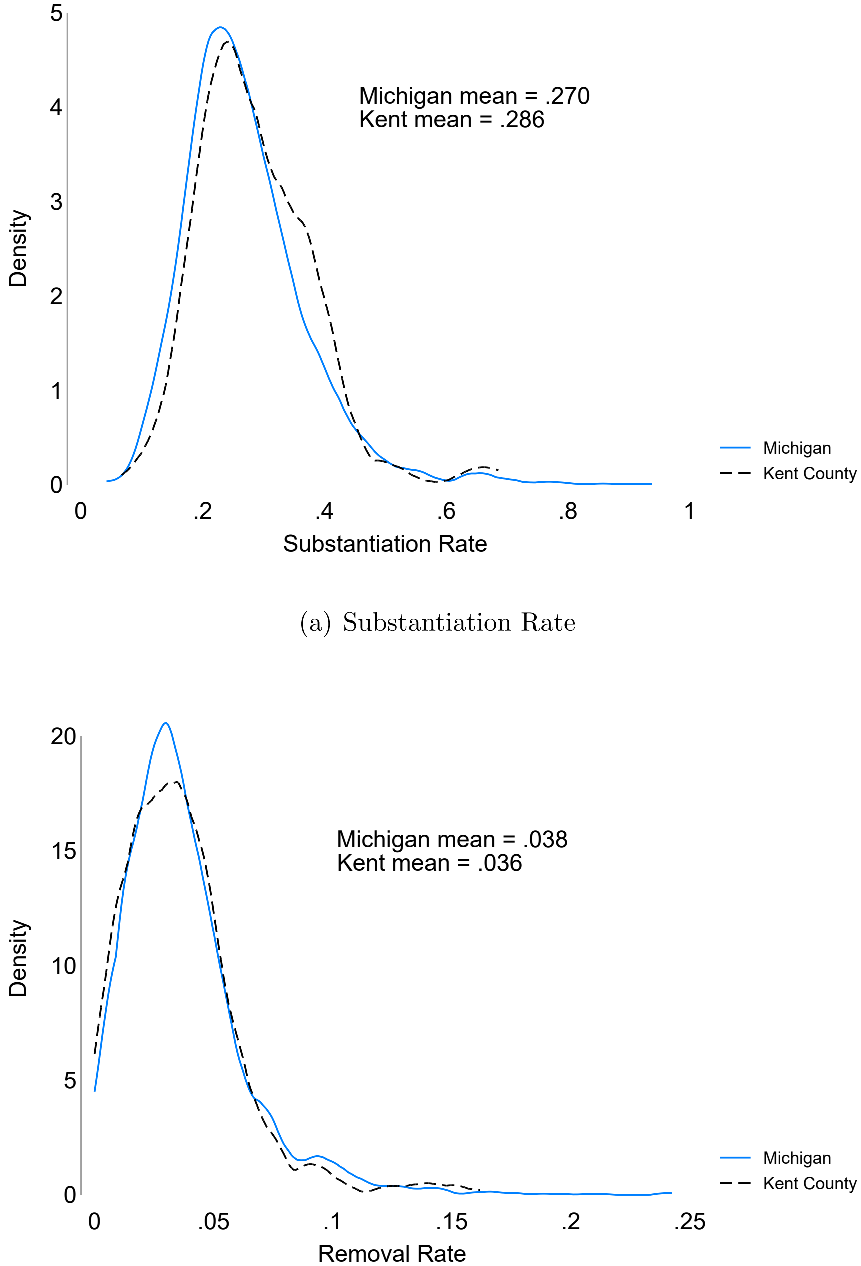
While racial disparities in substantiation and removal rates are much smaller than disparities in initial allegations, it is important to understand the sources of disparities at these later points in child welfare involvement. This section examines whether investigator biases play a role in shaping the small racial disparities at the substantiation and removal decisions.

We begin by comparing the investigator-specific substantiation and removal rates of White and Black children. In order to extract signal from noise in a measure of removal tendency, we restrict the analysis to Michigan's 2,066 child welfare personnel who investigated at least 25 Black and at least 25 White children from January 2010 to December 2018, and on Kent County's 248 workers who did the same.

Figure 2 shows the distribution of investigator-specific substantiation (S_i) and removal rates (R_i) separately for Michigan and Kent County. We define S_i and R_i as the fraction of total cases assigned to investigator i that were either substantiated or resulted in removal, respectively. Panel A shows that substantiation rates in Michigan appear to be slightly lower than those in Kent County, but the differences are small. On average, Michigan investigators substantiate at a rate of 27 percent, while Kent investigators substantiate at 28.6 percent. Panel B similarly shows the distribution of investigator removal rates. The figure shows that Kent County investigators remove at a lower rate relative to investigators in the rest of the state, but the averages are similar (3.6 percent versus 3.8 percent).

We next calculate, for each investigator, a White-Black substantiation and removal rate differential. In other words, for each investigator we calculate $S_i^* = S_i^W - S_i^B$ and $R_i^* = R_i^W - R_i^B$, the difference between investigator i's substantiation (removal) rate for investigations of White children and substantiation (removal) rate for investigations of Black children over the 2010–2018 time period.

Figure 2: Distribution of Substantiation and Removal Rates



(a) Substantiation Rate

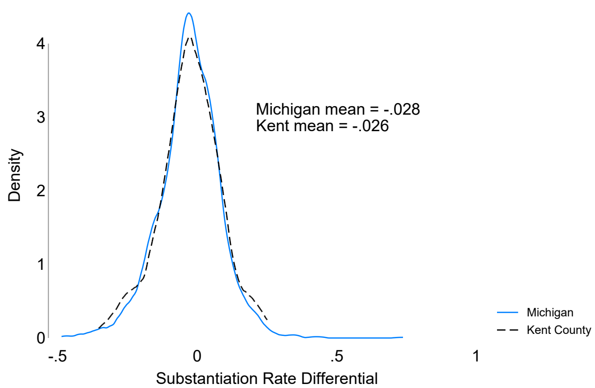
(b) Removal Rate

Notes: The figure shows the distributions of investigator-specific substantiation (S_i) and removal rates (R_i) separately for Michigan and Kent County. We define S_i and R_i as the fraction of total cases assigned to investigator i that were either substantiated or resulted in removal, respectively. This analysis consists of the 2,066 child welfare personnel who investigated at least 25 Black and at least 25 White children from 2010–2018 in Michigan, and the 248 workers in Kent County who did the same.

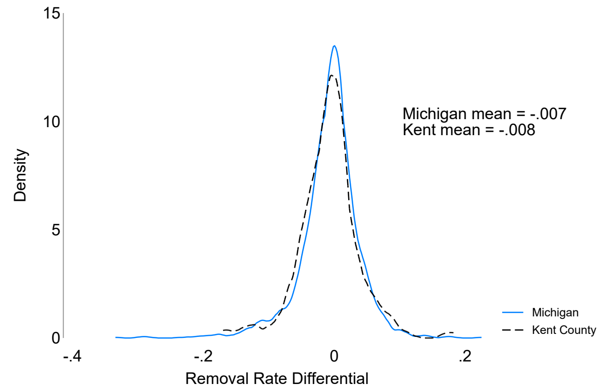
If child welfare investigators substantiate and remove Black and White children at similar rates, then the distributions of S_i^* and R_i^* should be centered around zero. However, if investigators substantiate and remove Black children at a systematically higher rate than White children, the average substantiation and removal rate differentials will be negative. Panels A and B of Figure

3 show that the distributions appear to be centered around zero, though slightly left-skewed. On average, Michigan investigators substantiate Black children at a rate 2.8 percentage points higher than White children, and they remove Black children at a rate 0.7 percentage points higher. Kent County's substantiation and removal rate differentials are similar to the rest of the state (-2.6 and -0.8 percentage points, respectively).

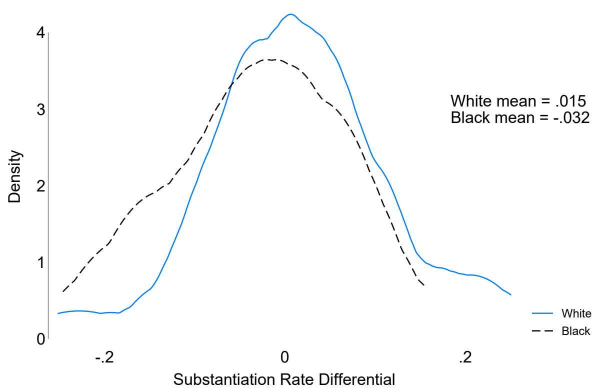
Figure 3: Substantiation and Removal Rate Differentials



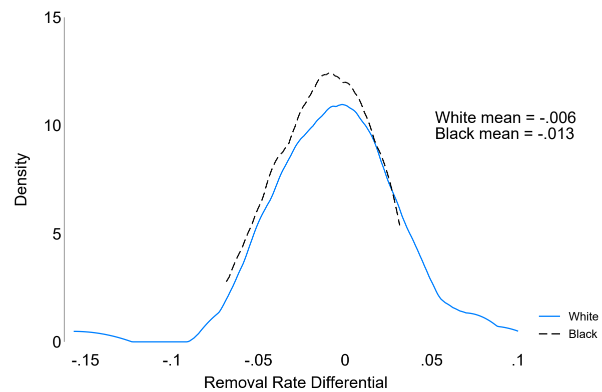
(a) Substantiation Rate Differential



(b) Removal Rate Differential



(c) Substantiation Differential by Race



(d) Removal Differential by Race

Notes: The figures present the distribution of substantiation and removal rate race differentials for child welfare investigators in Michigan and Kent County between 2010 and 2018. Specifically, Panels A and B show the distributions of S_i^* and R_i^* separately for Michigan and Kent County, while Panels C and D plot the distributions of S_i^* and R_i^* in Kent County separately by investigator race. This analysis focuses on Michigan's 2,066 CPS workers who investigated at least 25 Black and at least 25 White children between 2010 and 2018, and on Kent County's 248 workers who did the same.

What explains these differences in substantiation and removal rates? The fact that investigators in Michigan are quasi-randomly assigned allows us to narrow in on the sources of racial disparities at each decision point in child welfare involvement. This is the case because the characteristics of cases investigated, including the racial composition of alleged victims, should be unrelated to investigator characteristics such as experience, sex, race, and stringency.

For instance, in a setting in which investigators can select into cases with particular characteristics such as the child's race, risk levels, and so on, it would be virtually impossible to interpret differences in the substantiation and removal rates of White and Black children. Quasi-random assignment of investigators to cases allows us to examine the extent to which differences in an investigator's removal rate for Black and White children are attributed to investigators' racial prejudice or broad systematic differences in the case characteristics of cases involving White and Black children.

As described in Section II.A, upon receiving a report of suspected maltreatment, the child's local child welfare office assigns the report to a child maltreatment investigator according to a rotational assignment system. Reports cycle through investigators based on who is next up in the rotation. Typically, each county has its own local office. As such, investigators tend to be quasi-randomly assigned conditional on county and year fixed effects (to account for the set of case workers at any given time and in any given agency).

We test an implication of the conditional random assignment of investigators in Kent County: that observable child and case characteristics are uncorrelated with investigator characteristics. Because we observe only a very limited set of investigator characteristics in our data, we construct a measure of investigator "stringency" or "strictness." Specifically, for each investigation, we calculate the removal tendency of the investigator assigned to that investigation as the fraction of all other investigations, both past and future, assigned to the same investigator that resulted in foster care placement. We refer to investigators with higher removal tendencies as "stricter." In other words, for each investigation j assigned to investigator i, we calculate:

$$Z_{ji}^{R} = (\frac{1}{n_i - 1}) \sum_{k \neq j}^{n_i - 1} (FC_{ki}) \tag{1}$$

We then assign Z_{ji}^{R} to each of the 85,441 investigations in Kent County from 2010 to 2018 that was handled by one of the 248 investigators in Kent who investigated at least 25 Black and 25 White children during this period, and regress Z_{ji}^{R} on a set of child and investigation characteristics. Column 1 of Table 3 shows that child characteristics such as age, sex, and race, and investigation characteristics such as whether the allegation was for neglect (versus abuse) are not correlated with the investigator's strictness, despite these characteristics being highly predictive of foster care placement itself (Column 2). Column 3 shows that cases are well-balanced across investigators within Kent, even without controlling for year fixed effects. As a result, in the analyses that follow we do not control for year fixed effects.

Having established that investigators may be quasi-randomly assigned within Kent County, we first examine the role of racial prejudice on the part of the investigator in explaining the differences in substantiation and removal probabilities between White and Black children. While this question is difficult to answer, we propose a simple framework to examine this hypothesis. We obtained demographic information for the subset of investigators who are currently employed by Kent County to test whether differences in substantiation and removal rates by race are similar between White and Black investigators. The intuition behind this approach is simple: if differences in the substantiation and removal rates by race are entirely driven by one group of investigators, then this may suggest that racial prejudice on the part of that group could play a role in the observed racial disparities. If rate differentials are similar for White and Black investigators, then this finding would be inconsistent with most models of racial prejudice.

Table 3: Balance Tests for the Conditional Random Assignment of Investigators

	(1)	(2)	(3)	(4)
Dependent Variable:	Removal Tendency	Foster Care	Removal Tendency	Foster Care
	Year FEs		No Year FEs	
Age	-0.000**	-0.002***	-0.000***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Female	0.000	0.001	0.000	0.001
	(0.000)	(0.001)	(0.000)	(0.001)
White	0.000	0.025***	0.000	0.024***
	(0.001)	(0.004)	(0.001)	(0.003)
Black	0.001	0.037***	0.000	0.035***
	(0.001)	(0.004)	(0.001)	(0.003)
Hispanic	0.000	0.031***	-0.000	0.029***
	(0.001)	(0.004)	(0.001)	(0.004)
Neglect	0.000	-0.018***	0.000	-0.017***
	(0.000)	(0.002)	(0.000)	(0.002)
Dependent Variable Mean	0.038	0.039	0.038	0.039
N	85,441	85,441	85,441	85,441

Notes: The table reports the results from regressions of the dependent variable (either the investigator removal stringency or foster care placement) on a set of child and investigation covariates. The specification is estimated on the sample of 85,441 unique investigations in Kent County from January 2010 to December 2018 that were assigned to one of the 248 investigators in Kent County who investigated at least 25 Black and at least 25 White children during this time period. Columns 1 and 2 additionally include year fixed effects. Standard errors are clustered at the investigator level. The reference racial group is "other," which is equal to one if the child is not White, Black, or Hispanic.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Panels C and D of Figure 3 plot the distributions of S_i^* and R_i^* in Kent County separately by investigator race. Panel C shows that Black investigators substantiate investigations of Black children at a higher rate than those of White children, while the reverse is true for White investigators. However, Panel D shows that Black and White investigators both remove Black children at higher rates than White children. These findings suggest no clear evidence that racial prejudice on the part of investigators drive racial disparities at the substantiation and removal stages.

A second explanation is that White and Black investigations are systematically different. For instance, if investigations of Black children are more likely to be severe on average, then one would expect the removal rate for Black children to be larger than that of White children, even in the absence of investigator prejudice. Table 4 shows the characteristics of investigations for Michigan and Kent County prior to blind removals and separately by the child's race. The table shows that White and Black investigations tend to be similar along gender and allegation type dimensions (abuse versus neglect). However, the table shows that investigations of Black children in Michigan and Kent County tend to be for younger children relative to those of White children. The difference, however, is modest (0.6 years or roughly 7 percent). Because investigations of younger children may be more likely to both be substantiated and result in removal, these estimates suggest that the relatively small disparities at the substantiation and removal points could be driven by small differences in the case characteristics of White and Black children.[17]

Table 4: Systematic Differences in Case Characteristics by Race

	(1) White Investigations	(2) Black Investigations	(3) Difference
Panel A: Michigan			
Female	0.50	0.50	0.006**
	[0.50]	[0.50]	(0.003)
Age at Investigation	8.03	7.54	0.586***
	[5.12]	[5.33]	(0.084)
Investigation for Alleged Abuse	0.61	0.58	0.035***
	[0.49]	[0.49]	(0.010)
Panel B: Kent County			
Female	0.50	0.49	0.009*
	[0.50]	[0.50]	(0.005)
Age at Investigation	8.01	7.43	0.582***
	[5.12]	[5.19]	(0.117)
Investigation for Alleged Abuse	0.64	0.61	0.024***
	[0.48]	[0.49]	(0.004)

Notes: Columns 1 and 2 of the table report the means and standard deviations (in brackets) of case-level characteristics for investigations of White and Black children, respectively, averaged in the time period prior to blind removals (2010–2018). Column 3 reports the point estimates and clustered standard errors (in parentheses) of tests for equality of means. Panel A shows case characteristics for the state of Michigan as a whole, while Panel B focuses on Kent County. In Panel A, we two-way cluster standard errors at the county and investigation-year level, while in Panel B we cluster at the investigation-year. The case-level characteristics include whether the investigation was for a female child, the child's age at the time of the investigation, and whether or not the investigation was for any alleged abuse (as opposed to neglect).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

III.C Counterfactual Disproportionality Under Equalized Removal Rates

So far, we have shown that while racial disparities exist at later points in the child welfare process, they are much smaller than disparities in the initial rates of allegation and do not appear to be driven by investigator prejudice. This suggests that factors external to the child welfare system (e.g., differential reporting patterns or underlying rates of maltreatment) play an outsized role in shaping racial disproportionality within the system, and that any intervention that targets disparities at the

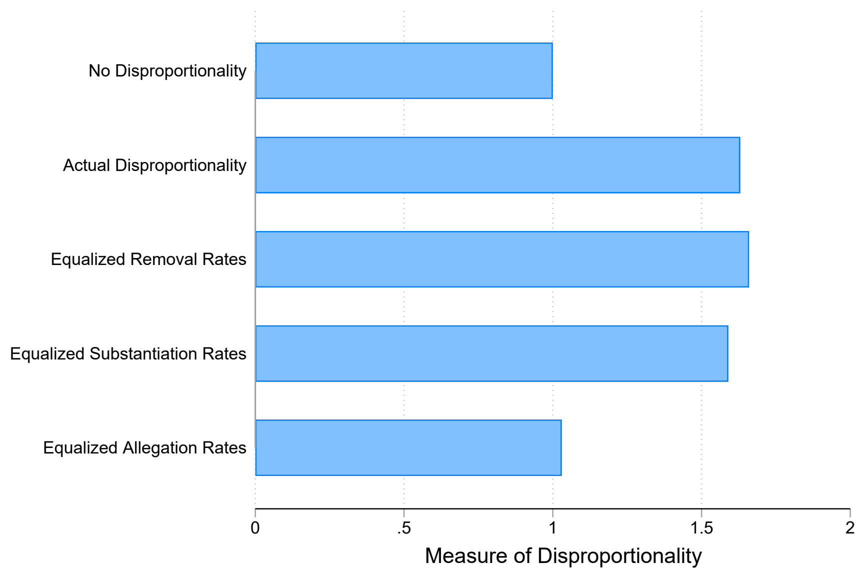
[17] For instance, in 2017 the rate of children aged three and younger entering foster care was double that of older children and youth (Child Trends, 2019).

decision to remove may have limited effects on racial disproportionality.

To more formally quantify the relative importance of each step of child welfare involvement in shaping racial disproportionality, we conduct simple counterfactual exercises. Specifically, we ask: how much would disproportionality in Michigan's foster care system be reduced by if racial disparities at the several points in child welfare involvement were completely eliminated?

Figure 4 shows actual and counterfactual disproportionality measures in Michigan's foster care system, based on the numbers reported in Tables 1 and 2. For the detailed calculations underlying this figure, see Table A1. We measure disproportionality as the ratio of the share of Black children in foster care to the share of Black children in the population. In other words, in a world with no disproportionality, this measure would be equal to one (Bar 1). According to the numbers reported in Table 1, and assuming only White and Black children in the population, Black children make up roughly 21 percent of children in Michigan. Based on the probabilities in Table 2 of ever being the alleged victim in a child maltreatment investigation, having that investigation substantiated, and being removed and placed in foster care, Black children make up 35 percent of the foster care population. Thus, Michigan has an actual disproportionality measure of roughly 1.63 (Bar 2).

Figure 4: Counterfactual Measures of Disproportionality



Notes: The figure shows actual and counterfactual disproportionality measures in Michigan's foster care system, based on the numbers reported in Tables 1 and 2. For the detailed calculations underlying this figure, see Table A1.

If removal rates (conditional on a substantiated investigation) were completely equalized across White and Black children, then disproportionality in Michigan's foster care system would slightly increase (Bar 3). This is because, conditional on a substantiated investigation, removal rates for White children are actually slightly higher than those of Black children (see Table 2). The fourth bar shows that even if *both* substantiation and removal rates were completely equalized across White and Black children, disproportionality would remain largely unchanged in Michigan. However, in a counterfactual where White and Black children are equally likely to be the alleged victims in an allegation of maltreatment, disproportionality would be 1.03, 96 percent lower than actual disproportionality, and extremely close to a world in which there is no disproportionality.

This simple counterfactual exercise reinforces that differential reporting patterns or underlying rates of maltreatment play an outsized role in shaping racial disproportionality within the system, and that even if blind removals were to completely eliminate racial disparities stemming from the decision to remove, it would have a negligible impact on overall racial disproportionality in foster care.

Importantly, one could argue that equalizing removal rates may not go far enough and that, in fact, if biases in child welfare reports are present and a differentially higher number of minor allegations are reported for Black children relative to White children, then equal removal rates conditional on initial investigations would be a sign of bias in the removal decision. In other words, optimal removal rates for Black children conditional on investigation should be lower than those of White children. However, this scenario is unlikely as both case severity and substantiation rates for White and Black children tend to be quite similar in both Kent County and Michigan more broadly. Specifically, as mentioned above, in Kent County the share of investigations for White children that are for child abuse (versus neglect) is roughly 64 percent relative to 61 percent for Black children (see Table 4). Similarly, conditional on an investigation, the substantiation rate is roughly 27.3 percent for White children relative to 29.4 percent for Black children (see Table 2).

IV Comparing Removal Rates to Other Counties in the U.S.

In Table 2, we showed that conditional on being the alleged victim in a child maltreatment investigation, substantiation and removal rates exhibit relatively small racial disparities in both Kent County and the average Michigan county. Using a counterfactual exercise, we then showed that policy interventions that target the decision to remove may not meaningfully impact racial disproportionality in such settings. Still, one may wonder whether other counties across the U.S. display similarly small differences in removal rates by race. In other words, are there counties in the U.S. in which blind removals (or other interventions that target the removal decision) could hold promise?

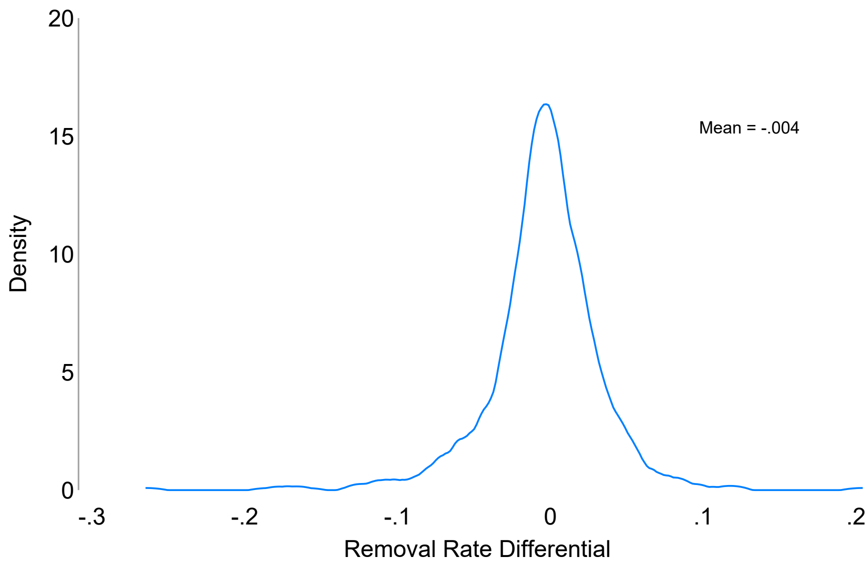
To approach this question, we collected the NCANDS Child File for 2018 which is comprised of all referrals to state child protective services agencies in that year.[18] These data represent a census of child protective services cases during a federal fiscal year. We aggregate case-level data to obtain county-level rates of investigations that result in removal. We then compute the county's average removal rate by race, conditional on an investigation, and calculate the White - Black difference in this rate (so that negative numbers indicate that investigators in the county tend to remove Black children at higher rates, relative to White children, and conditional on an investigation). Since NCANDS only reports geographic identifiers for typically populous counties that received over 1,000

[18] For a handful of states, referrals do not contain information regarding the eventual outcome of the case (e.g., services provided or foster care placement). As a result, we do not include New York, Pennsylvania, and North Carolina in this analysis. We focus only on the remaining states in the lower 48.

unique investigations during the federal fiscal year, it is important to caveat that the results below may not be representative of less populous counties in the U.S.

As in Panel B of Figure 3, Figure 5 plots the distribution of differences in removal rates between White and Black children across the country. Similar to both the average Michigan county and Kent County, conditional on an investigation, investigators in the average county represented in NCANDS remove Black children at a rate nearly identical to that of White children (-0.4 percentage points). Moreover, the bulk of the distribution is centered around zero.

Figure 5: Differences in the Conditional Probability of Removal (NCANDS)

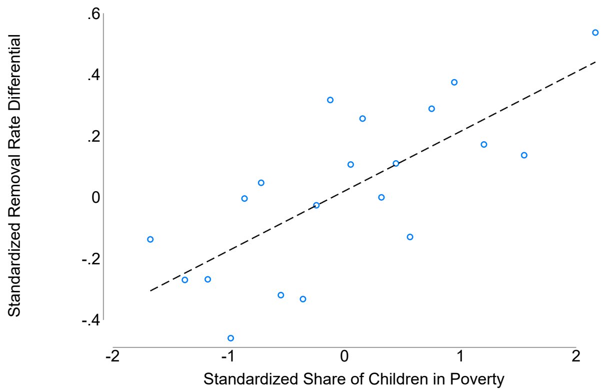


Notes: The figure plots the distribution of differences in removal rates (conditional on an investigation) between cases involving White and Black children among counties represented in NCANDS during 2018. Given reporting censoring in NCANDS, the analysis focuses on the 579 counties in the U.S. that had more than 1,000 unique investigations during the year and had at least 100 cases involving White children and 100 cases involving Black children.

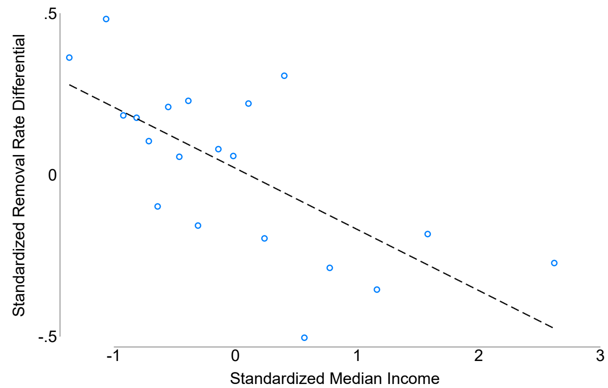
Still, the figure shows that about 9 percent of counties in the U.S. remove Black children at a rate greater than 5 percentage points relative to White children even conditional on an investigation, and about 5 percent of counties do the opposite. In these counties, interventions that target the decision to remove may have a larger scope to impact disproportionality, and it is important for policy purposes to describe the characteristics of such counties.

To this end, we collected county-level data from the U.S. Census' Small Area Income and Poverty Estimates Program and Population Estimates Program on county-level characteristics such as the percent of children in poverty, median household income, and the share of the population who is Black. Figure 6 shows binned scatter plots relating the county's White-Black removal rate differential to each of the three county-level characteristics. The figure shows that counties with a smaller share of children in poverty, with a higher median income, and with a smaller share of the population who is Black tend to have lower values of the removal rate differential (i.e., tend to remove Black children at higher rates than White children, conditional on an investigation). While we examine only a limited set of county-level characteristics, this analysis begins to shed light on the types of counties in which policy interventions targeting the decision to remove may have more scope to affect disproportionality. We leave it to future research to fully characterize such counties.

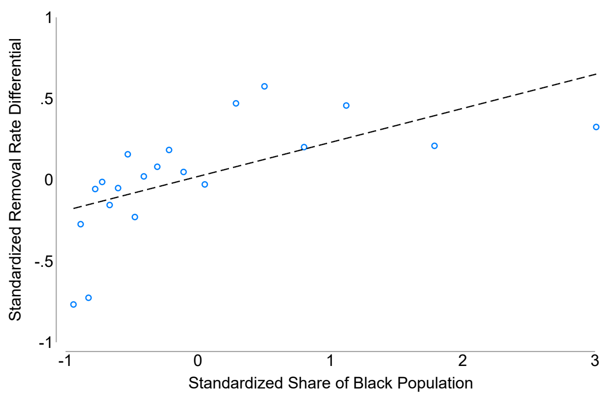
Figure 6: Correlations between Removal Rate Differentials and County Characteristics (NCANDS)



(a) Percent of Children in Poverty



(b) Median Income



(c) Share of the Population who is Black

Notes: The figure shows binned scatter plots of the relationship between a county's White-Black removal rate differential and county characteristics. All measures are standardized to have mean zero and standard deviation one in the population of 579 counties in our sample. The county's percent of children in poverty and median income come from the U.S. Census' Small Area Income and Poverty Estimates Program. The share of the population who is Black comes from the U.S. Census' Population Estimates Program.

V Causal Effects of Blind Removals

We have shown so far that most of the observed racial disproportionality in Michigan's foster care system is largely due to racial disparities in the probability of ever being the alleged victim in a child maltreatment investigation. Thus, policies that target specific points of the child welfare decision making process such as blind removals are likely to have limited effects. Specifically, even if blind removals were to completely eliminate racial disparities stemming from the decision to remove, it would have a limited impact on overall racial disproportionality in foster care. We view establishing this descriptive fact as the main contribution of this study.

Nevertheless, one may wonder whether Kent County's blind removal program had any impact on racial disparities at the removal decision. While this section attempts to shed light on this question, we note that our setting is not ideal for rigorous program evaluation. The treatment group in our setting consists only of one county with few investigators in both the pre- and post-period. Furthermore, the post-period is extremely short: Kent implemented blind removals in August 2019 and the COVID-19 pandemic began shortly thereafter. As a result, the main results in this section should be interpreted with caution.

For our main analysis, we restrict the sample to investigators in Kent County and in control counties who had at least 25 investigations of both Black and White children in both the pre- and post-blind removal periods. Our control group consists of Michigan's four other most populous counties: Wayne, Oakland, Macomb, and Genesee. We also restrict our analysis to before March 2020 to avoid conflating any differential effects of the COVID-19 pandemic.

Recall that in Section III.B, we calculated a within-investigator removal rate race differential. In other words, for each investigator we calculated $R_i^* = R_i^W - R_i^B$, the difference between investigator i's removal rate for White and Black children over the 2010–2018 time period. We begin by computing this rate pre- and post-August 2019—the first month in which Kent County implemented blind removals.

The intuition behind our identification strategy is simple: We compute $\bar{R}^{*K}_{post}$, $\bar{R}^{*K}_{pre}$, $\bar{R}^{*C}_{post}$, and $\bar{R}^{*C}_{pre}$—where $\bar{R}^{*K}_{pre}$ is the average investigator removal rate differential in Kent County from January 2010 to July 2019 and $\bar{R}^{*K}_{post}$ is the average investigator removal rate differential in Kent County from August 2019 to March 2020. We similarly compute these rates for investigators in counties in the control group. We use the estimated $\bar{R}^{*K}_{post}$, $\bar{R}^{*K}_{pre}$, $\bar{R}^{*C}_{post}$, and $\bar{R}^{*C}_{pre}$ to calculate a classic 2×2 DID estimate. Specifically, we compute $\hat{\delta} = (\bar{R}^{*K}_{post}$ - $\bar{R}^{*K}_{pre})$ - $(\bar{R}^{*C}_{post}$ - $\bar{R}^{*C}_{pre})$ using the following equation:

$$Y_{it} = \alpha + \beta Kent_i + \theta Post_t + \delta(Kent_i \times Post_t) + \varepsilon_{it} \tag{2}$$

where Y_{it} is either R^*_{it}, the removal rate race differential of investigator i in period t, or R_{it}, investigator i's overall removal rate in t; $Kent_i$ is an indicator equal to one if the investigator works in Kent County and zero otherwise; $Post_t$ is an indicator equal to one if the time period is during or after August 2019.[19]

Statistical inference in our setting is complicated by the fact that treatment occurs only for one cluster, and models with few treated units can lead to improper inference (Cameron, Gelbach and Miller, 2008; Ferman and Pinto, 2019; MacKinnon and Webb, 2017, 2018).[20] We perform statistical

[19] The goal of blind removals is to reduce racial disproportionality by eliminating racial disparities in investigators' decisions to remove. As a result, our analysis directly tests the effects of blind removals on investigators' removal behavior. An alternative way to test the causal effects of blind removals is at the child-by-investigation level. In Online Appendix B, we show that the main conclusions of the paper are unchanged if we instead model the probability that a particular investigation results in removal as a function of the blind removals program.

[20] It is common to assume in regression models that the error term is correlated within clusters but uncorrelated between them. Inference under this assumption can be achieved by a cluster-robust variance estimator. However, test

inference in a number of ways, all of which yield similar conclusions. In our main results, we simply show standard errors that are robust to heteroskedasticity and compute p-values based on Ferman and Pinto (2019)'s alternative method for inference when there is a small number of overall clusters, only one treated cluster, and errors are heteroskedastic. In Online Appendix B, we also use synthetic control methods that are well-suited to handle inference in contexts with a single treated unit and find similar results.

The parameter of interest is δ, which—under identification assumptions that we probe below—measures the causal effects of the blind removal program on the removal decisions of child welfare professionals. It is important to be clear about the blind removals treatment: the program could change removal outcomes because (1) it changes the decision-making process to include more decision-makers, (2) it increases scrutiny to each decision, and/or (3) it masks children's demographic information to decision makers. Our study can shed light on the net effects of the policy, but it is impossible to isolate the relevant importance of each of these channels since they are all part of the blind removals program. However, δ is likely the parameter of interest for policymakers deciding whether or not to implement the program, as currently proposed implementations throughout the country would include each of these three components as well.

V.A Identification Assumptions

The identification assumption of this approach is that ($\bar{R}^{*C}_{post}$ - $\bar{R}^{*C}_{pre}$) constitutes a good counterfactual for ($\bar{R}^{*K}_{post}$ - $\bar{R}^{*K}_{pre}$) in the absence of blind removals. In order to probe this assumption, we test whether investigators in Kent and in control counties were on similar trends in their removal rate differentials prior to blind removals.

Figure A1 begins by showing the distributions of $R^*_{i,(2010-2014)}$ and $R^*_{i,(2015-2018)}$ separately for Kent and the four control counties.[21] The figure shows that the removal rate race differential became slightly less negative in Kent County from 2010–2014 to 2015–2018 (by roughly 0.7 percentage points). In other words, the differential rate at which Black children were removed relative to White children shrank. However, this trend was strikingly similar in control counties, which suggests that Kent and control counties followed similar trajectories in the removal rate race differential prior to blind removals.

Table A2 formalizes this argument. Specifically, we compute the difference between $\bar{R}^{*K}_{2015-2018}$ and $\bar{R}^{*K}_{2010-2014}$ and test whether it is statistically different from the difference between $\bar{R}^{*C}_{2015-2018}$ and

statistics based on cluster-robust standard errors tend to over-reject the null hypothesis when the number of clusters is small. While the wild cluster bootstrap proposed by Cameron, Gelbach and Miller (2008) can lead to more reliable inference, it can lead to improper inference if the number of *treated* clusters is small (MacKinnon and Webb, 2018).

[21] As in Section III.B, this analysis focuses on investigators with at least 25 investigations of both Black and White children in either the 2010–2014 or the 2015–2018 time periods. Given these sample restrictions, we cannot examine differences in annual trends due to a small number of observations in each year. Specifically, this approach would require that we observe a large enough number of investigators who had at least 25 cases of both Black and White children in a given year. Instead, our approach simply requires that we identify investigators who had at least 25 cases of both Black and White children in either 2010–2014 or 2015–2018.

$\bar{R}^{*C}_{2010-2014}$. The table shows estimates of ($\bar{R}^{*K}_{2015-2018}$ - $\bar{R}^{*K}_{2010-2014}$), ($\bar{R}^{*C}_{2015-2018}$ - $\bar{R}^{*C}_{2010-2014}$), and the difference between the two in bold. Robust standard errors are shown in parentheses below the point estimates, while Ferman and Pinto (2019) p-values are shown in curly brackets. Consistent with the results in Figure A1, the DID estimate in Panel A is small (0.265 percentage points) and statistically insignificant. Panel B of the table additionally shows that the removal rates of investigators in Kent and in control counties were also trending similarly prior to blind removals. Furthermore, these conclusions are unchanged when performing alternative inference procedures (Columns 1 and 2, Table A3). Altogether, the results in this section suggest that the assumption of parallel trends in our setting is plausible.

V.B Effects of Blind Removals on Removal Rates

We begin to examine the causal effects of blind removals by showing the distributions of R^*_i and R_i separately for Kent County and counties in the control group, before and after the implementation of the program. Panel A of Figure A2 overlays the pre- and post-blind removal distribution of R^*_i for Kent County, while Panel B shows the distribution of R_i instead. Panels C and D show the same distributions for control counties.

Panels A and C show that, in both Kent and control counties, the distribution of R^*_i remained largely unchanged pre- and post-blind removals. If anything, the distributions in both Kent and control counties shifted slightly to the left: investigators removing Black children at a higher rate relative to White children in the post period compared to the pre-blind removals period—though the differences between the distributions are not statistically significant. Panel A of Table 5 formalizes this comparison. Specifically, it shows the estimated $\hat{\delta}$ as well as its four main components. The table shows that the average removal rate race differential in both Kent County and in the control counties became more negative after blind removals. However, the relative shift toward removing Black children at a higher rate was similar between Kent and the control counties: close to -1 percentage points in both Kent and the control group. As a result, the DID estimate is small (0.142 percentage points) and statistically insignificant either when performing inference using simple robust standard errors or Ferman and Pinto (2019) p-values.

As mentioned above, these results should be interpreted with caution: the standard error of nearly 2 percentage points does not allow us to rule out meaningful changes in either direction, and the standard errors only become larger when we use alternative methods. For instance, Column 3 of Table A3 shows that the point estimate of 0.142 is also statistically insignificant when deriving p-values from either heteroskedasticity-robust standard errors, standard errors clustered at the investigator level, standard errors clustered at the county level, or wild-clustered bootstrapped standard errors. Thus, even though the point estimate is small in magnitude, given the size of the standard errors, we take this as suggestive evidence that blind removals had a limited impact on removal rate race differentials in Kent County.

Table 5: DID Estimates of Blind Removals on Placement Outcomes

	(1) Post	(2) Pre	(3) Difference	(4) DID
Panel A: Removal Rate Differential				
Kent	-0.93	-0.06	-0.869	
	[6.77]	[4.44]	(1.653)	
N	24	24	48	
Control Counties	-2.22	-1.21	-1.011	**0.142**
	[7.43]	[3.45]	(0.849)	(1.906)
				{0.946}
N	93	93	186	234
Panel B: Removal Rate				
Kent	3.17	3.97	-0.803	
	[2.43]	[2.59]	(0.724)	
N	24	24	48	
Control Counties	3.14	2.86	0.288	**-1.091**
	[3.39]	[2.30]	(0.424)	(0.832)
				{0.157}
N	93	93	186	234

Notes: In Panel A, Column 1 presents estimates of $\bar{R}^{*K}_{post}$ and $\bar{R}^{*C}_{post}$ while Column 2 presents estimates of $\bar{R}^{*K}_{pre}$ and $\bar{R}^{*C}_{pre}$. Column 3 presents estimates of ($\bar{R}^{*K}_{post}$ - $\bar{R}^{*K}_{pre}$) and ($\bar{R}^{*C}_{post}$ - $\bar{R}^{*C}_{pre}$). We present estimates of δ in Column 4 in bold. For expositional purposes, we multiply all rates by 100 and present them as percentage points. Standard deviations are shown in brackets in Columns 1 and 2, while robust standard errors are shown in parentheses in Columns 3 and 4. Ferman and Pinto (2019) p-values are shown in curly brackets, as discussed in the main text. Panel B shows the same information but for overall removal rates (R_{it}) instead.

While blind removals do not appear to have impacted removal rate differentials by race, they may have impacted the overall removal rate. As mentioned above, blind removal meetings place increased scrutiny on every petition to remove, so that these meetings may decrease overall removal rates in Kent County. In addition, during blind-removal meetings investigators are connected to other investigators and child welfare specialists that may be able to provide alternative viewpoints regarding removal decisions. This may mean that investigators who learn about new viewpoints and resources to deal with specific cases may remove at lower rates in the future.

Panels B and D of Figure A2 show that investigators in Kent County remove at lower rates following the implementation of blind removals, though this difference is not statistically significant. Panel B of Table 5 confirms this observation: the average removal rate in Kent County decreased from almost 4 percent prior to blind removals to roughly 3.2 percent, a difference of 0.8 percentage points. In contrast, counties in the control group saw an increase in the removal rate during this period, such that the DID estimate suggests a one percentage point decline in the removal rate. However, even though this estimate is modest in magnitude, the estimate is statistically insignificant at the 5 percent level both when computing Ferman and Pinto (2019) p-values and when conducting the alternative inference procedures in Table A3 (see Column 4). Still, given the magnitude of the estimate and the relatively small number of investigators in our sample, we take this as suggestive evidence that blind removals led to a decline in overall removal rates.

Why did the blind removals program not impact racial disparities at the removal decision? While the exact mechanism is difficult to empirically pin down, we propose three, non-mutually exclusive explanations for this finding. First, we showed in Section III.B that there is little evidence that investigator prejudice is responsible for the relatively small racial disparities at the removal decision. Thus, blind removals had limited scope to begin with in our setting. Second, as mentioned in Section II.B, while most demographic and socio-economic information has technically been concealed from nearly all decision-makers in the blind-removal committee, the investigator—who knows the race of the child—prepares the relevant documentation, presents the case to the blind removal committee, and has a vote in the committee. Therefore, there is still an unblinded decision-maker, and one that plays a substantial role in the direction of the outcome. If the initial investigator's prejudice influences the way in which she writes and presents the case, blind removals will have limited effects. Finally, the blind removal program is only focused on cases where the investigator would recommend removal. If racial disparities at the removal decision are driven by investigators removing fewer than optimal White children, then the program cannot achieve its intended outcome. This is a limitation of the program as a tool to impact racial disparities at the removal decision.

V.C Potential Costs of Blind Removals

The fact that blind removals have a limited impact on overall disproportionality is of little concern if the program is costless to implement. However, even beyond the additional monetary costs of personnel at both the extensive (e.g., clerks) and intensive margins (e.g., additional hours worked), there are a number of potential unintended consequences of the program that should be considered when evaluating whether or not to implement it. While a full examination of the costs of blind removals is beyond the scope of this paper, this section highlights and tests a potential cost of the program: its impact on time to removal.

Child welfare investigators usually cite concerns about the time that the blind removal process takes and its potential impact on emergency situations.[22] We test this concern directly by estimating the effect that Kent County's blind removal program had on the median number of days from the beginning of an investigation to removal. Panel A of Table 6 shows that the median time to removal increased in Kent County from roughly 14 days to 27 days after the implementation of blind removals. However, in control counties, the median time to removal increased by just 3 days. The DID estimate shows that blind removals increased the median time to removal by roughly 9 days, an increase of roughly 30 percent relative to the pre-blind removal median in control counties. This estimate is statistically significant at the five percent level per Ferman and Pinto (2019) p-values.

[22]See, for example, *The Idea of Removing Race from Child Removal Decisions*, Children's Bureau. Accessed at `https://cbexpress.acf.hhs.gov` (July 17, 2021).

Table 6: Potential Unintended Consequences of Blind Removals

	(1) Post	(2) Pre	(3) Difference	(4) DID
Panel A: Median Time to Removal				
Kent	26.52	14.24	12.286	
	[40.22]	[14.91]	(9.361)	
	21	21	42	
Control Counties	33.16	29.86	3.299	**8.987**
	[49.83]	[34.10]	(7.376)	(11.842)
				{0.032}
N	67	67	134	176
Panel B: Placements in a Family Setting				
Kent	0.80	0.81	-0.014	
	[0.24]	[0.15]	(0.059)	
	21	21	42	
Control Counties	0.83	0.86	-0.032	**0.018**
	[0.25]	[0.10]	(0.036)	(0.069)
				{0.498}
N	67	67	134	176

Notes: The table presents estimates of the effects of blind removals on Y_i, where Y_i is either investigator i's median time to removal (Panel A), or investigator i's share of removals in which the child was placed in a family setting—as opposed to a group home (Panel B). Column 1 presents estimates of $\bar{Y}^K_{post}$ and $\bar{Y}^C_{post}$ while Column 2 presents estimates of $\bar{Y}^K_{pre}$ and $\bar{Y}^C_{pre}$. Column 3 presents estimates of ($\bar{Y}^K_{post}$ - $\bar{Y}^K_{pre}$) and ($\bar{Y}^C_{post}$ - $\bar{Y}^C_{pre}$). We present estimates of δ in Column 4 in bold. Standard deviations are shown in brackets in Columns 1 and 2, while robust standard errors are shown in parentheses in Columns 3 and 4. Ferman and Pinto (2019) p-values are shown in curly brackets, as discussed in the main text. The number of investigators in this table is lower than that of Table 5 due to the small number of investigators who had zero removals in the post period (3 in Kent County and 26 in control counties).

Importantly, it is not entirely clear that an increase in the median time to removal is necessarily a cost of the program. On the one hand, this result could imply that children at the margin of removal spend more time in potentially unsafe conditions. Though we cannot empirically test this hypothesis directly, we note that child welfare personnel remain involved with the family while the case goes through the removal process. On the other hand, the increase in time to removal may reflect that more careful consideration is given to each petition to remove, which could lead to more effective placement decisions. For instance, delayed placements could help investigators secure more effective matches between the child and potential foster care settings. Similarly, blind removals could increase worker morale if the psychological burden of the removal decision is now shared among multiple individuals as opposed to the worker alone. Numerous investigators have highlighted this as a salient benefit of the program thus far.[23]

We directly test whether blind removals impacted placement match quality by examining the

[23]Investigators often mention other benefits of blind removal meetings such as improvements in teamwork, accountability, and shared historical child welfare knowledge.

effects of the blind removals program on the proportion of first foster care placements in a family setting (kinship or unrelated caregiver) as opposed to a congregate care setting (group homes or residential facilities). While there is limited empirical evidence on the relative effectiveness of each placement type, in practice, placements in family settings are generally preferred to placements in congregate care.[24] Panel B of Table 6 shows no evidence that blind removals impacted the proportion of first placements in a family setting. The share of first placements in a family setting (roughly 80 percent) stayed relatively constant in both Kent and the control group before and after blind removals.

Understanding all potential costs of blind removals is beyond the scope of this paper, but has important consequences for the general equilibrium effects of this program. Nevertheless, it is important to remember that the explicit goal of blind removals is to reduce racial disproportionality. While the program could lead to unintended costs and benefits in the longer-run, our analysis shows that it is likely to have limited effects on its intended outcome.

VI Conclusion

Child welfare leaders and scholars have long expressed concerns regarding racial disproportionality in child welfare outcomes. Calls to reform the system have grown more urgent in recent months, as the nationwide push to examine structural racism in institutions has reached the child welfare system. While a number of smaller initiatives such as diversity and anti-racism training for investigators have been piloted, a salient reform in the last decade has been blind removals.

The popularity of blind removals has drastically increased in the last few years. Driven largely by the intuitive appeal of the program, a perception that an early implementation in Nassau County, NY was successful, and growing calls to reduce disproportionality in child welfare, agencies across the country from Los Angeles to the state of New York have expressed a marked interest in adopting blind removals. However, until this paper, there had been no quantitative analysis of the program.

Our study addresses this critical gap in the literature and is the first to shed light on the broad effects of blind removals on foster care placement outcomes. We derived two main findings. First and foremost, we showed that most racial disproportionality in foster care systems in Michigan is driven by disparities in the initial rates of child maltreatment allegations. Specifically, we showed that even if a policy were to completely eliminate any disparity in the decision to remove, it would have a limited effect on overall racial disproportionality in foster care systems.

Second, we set out to understand whether blind removals could narrow the already small racial disparity in the decision to remove. We find suggestive evidence that the program led to a decline

[24] In recent years states have prioritized moving away from congregate care placements. In fact, the 2018 Family First Prevention Services Act incentivizes the reallocation of children away from congregate care. Kinship care is thought to be less disruptive to children's lives because it allows them to live with someone they know and who shares their culture (Lovett and Xue, 2020); these placements also exhaust fewer state resources. Furthermore, it has long been posited that negative peer effects in congregate-care placements are a critical mechanism that may enhance risk: by grouping children who may have behavioral problems or mental trauma together, these behaviors may be reinforced (Font and Mills, 2020).

in removals, but that this decline was similar for both White and Black children. In other words, while the program may have led to an overall decline in removal rates, we find limited evidence that it had an effect on racial disparities at the removal decision. However, we showed that the program significantly increased the time to removal for the median investigation.

These findings have important implications for policy. First, blind removals are not well-suited to reduce racial disproportionality in settings where most of the disparity is introduced at the initial rates of allegations since it targets a decision late in child welfare involvement. In other words, this policy is likely to have the most promise in settings in which disparities in the decision to remove are larger. Still, using NCANDS data, we showed that such settings tend to be uncommon. Understanding whether the program can be successful in settings where the removal rate race differential is large is an important question for future research.

The second policy implication is that states and local child welfare agencies should more carefully consider whether implementation of blind removals is appropriate in their context, and should keep in mind that the existing empirical evidence does not suggest it is an effective strategy at reducing racial disproportionality in the average foster care system. Specifically, our results suggest that policies that target the disparities in the initial rates of allegations are likely to have substantially larger impacts on racial disproportionality.

Blind removals cannot address some of the reasons that Black families and other underserved populations are brought into the attention of child welfare in the first place. We still know very little about why racial disparities are so large at the initial rates of allegations. Policymakers and scholars focused on this question should seek to understand whether disparities are driven by either implicit (or explicit) biases in the initial reporting of child maltreatment or by factors external to the child welfare system such as conditions linked to poverty and barriers to access resources that may help mitigate risk factors. The answer is likely a combination of the two explanations and we still do not know the relative importance of each.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program." *Journal of the American statistical Association*, 105(490): 493–505.

Abadie, Alberto, and Javier Gardeazabal. 2003. "The economic costs of conflict: A case study of the Basque Country." *American economic review*, 93(1): 113–132.

AECF. 2019. "KIDS COUNT Data Center." The Annie E. Casey Foundation. https://datacenter.kidcount.org.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2019. "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes." *American Economics Journal: Applied Economics*, 11(3): 338–381.

Bald, Anthony, Eric Chyn, Justine S. Hastings, and Margarita Machelett. 2022*a*. "The causal impact of removing children from abusive and neglectful homes." *Journal of Political Economy*, Forthcoming.

Bald, Anthony, Joseph Doyle Jr., Max Gross, and Brian Jacob. 2022*b*. "Economics of Foster Care." *Journal of Economic Perspectives*, Forthcoming.

Baron, E Jason, and Max Gross. 2022. "Is There a Foster Care-To-Prison Pipeline? Evidence from Quasi-Randomly Assigned Investigator." *NBER Working Paper #w29922*.

Baron, E Jason, Ezra G Goldstein, and Cullen T Wallace. 2020. "Suffering in silence: How COVID-19 school closures inhibit the reporting of child maltreatment." *Journal of public economics*, 190: 104258.

Bartholet, Elizabeth, Fred Wulczyn, Richard P Barth, and Cindy Lederman. 2011. "Race and child welfare." *Chapin Hall Issue Brief, June*, 11–27.

Barth, Richard P, Melissa Jonson-Reid, Johanna KP Greeson, Brett Drake, Jill Duerr Berrick, Antonio R Garcia, Terry V Shaw, and John R Gyourko. 2020. "Outcomes following child welfare services: what are they and do they differ for black children?" *Journal of Public Child Welfare*, 14(5): 477–499.

Bertrand, Marianne, and Jessica Pan. 2013. "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior." *American Economics Journal: Applied Economics*, 5(1): 32–64.

Bosk, Emily Adlin. 2015. "All Unhappy Families: Standardization and Child Welfare Decision-Making." *Doctoral Dissertation, University of Michigan.*

Brown, Dan, and Elisabetta De Cao. 2018. "The impact of unemployment on child maltreatment in the United States." ISER Working Paper Series.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics*, 90(3): 414–427.

Chetty, Raj, Nathaniel Hendren, and Lawrence Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review*, 106(4): 855–902.

Child Trends. 2019. "Infants and toddlers are more likely than older children to enter foster care because of neglect and parental drug abuse." Child Trends.

Chyn, Eric. 2018. "Moved to Opportunity: The Long-Run Effect of Public Housing Demolition on Labor Market Outcomes of Children." *American Economic Review*, 108(10): 3028–3056.

Doyle, Joseph J. 2007. "Child protection and child outcomes: Measuring the effects of foster care." *American Economic Review*, 97(5): 1583–1610.

Doyle, Joseph J. 2008. "Child protection and adult crime: Using investigator assignment to estimate causal effects of foster care." *Journal of political Economy*, 116(4): 746–770.

Doyle, Jr, Joseph J. 2013. "Causal effects of foster care: An instrumental-variables approach." *Children and Youth Services Review*, 35(7): 1143–1151.

Edwards, Frank, Sara Wakefield, Kieran Healy, and Christopher Wildeman. 2021. "Contact with Child Protective Services is pervasive but unequally distributed by race and ethnicity in large US counties." *Proceedings of the National Academy of Sciences*, 118(30).

Ferman, Bruno, and Cristine Pinto. 2019. "Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity." *Review of Economics and Statistics*, 101(3): 452–467.

Finlay, Keith, Michael Mueller-Smith, and Brittany Street. 2022. "Measuring Intergenerational Exposure to the US Justice System: Evidence from Longitudinal Links between Survey and Administrative Data." *Working Paper.*

Font, Sarah, and Christopher Mills. 2020. "Safe from Harm? Peer Effects and Criminal Capital Formation in Foster Care." *Working Paper.*

Gillingham, Philip, and Cathy Humphreys. 2010. "Child protection practitioners and decision-making tools: Observations and reflections from the front line." *British journal of social work*, 40(8): 2598–2616.

Goldin, Claudia, and Cecilia Rouse. 2000. "Orchestrating impartiality: The impact of" blind" auditions on female musicians." *American economic review*, 90(4): 715–741.

Grimon, Marie-Pascale. 2021. "Effects of the Child Protection System on Parents." *Working Paper.*

Gross, Max, and E. Jason Baron. 2022. "Temporary Stays and Persistent Gains: The Causal Effects of Foster Care." *American Economic Journal: Applied Economics*, 14(2): 170–99.

Kim, Hyunil, Christopher Wildeman, Melissa Jonson-Reid, and Brett Drake. 2017. "Lifetime Prevalence of Investigating Child Maltreatment Among US Children." *American Journal of Public Health*, 107(2): 274–280. PMID: 27997240.

Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz. 2005. "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment." *Quarterly Journal of Economics*, 120(1): 87–130.

Lindo, Jason M, Jessamyn Schaller, and Benjamin Hansen. 2018. "Caution! Men not at work: Gender-specific labor market conditions and child maltreatment." *Journal of Public Economics*, 163: 77–98.

Lovett, Nicholas, and Yuhan Xue. 2020. "Family first or the kindness of strangers? Foster care placements and adult outcomes." *Labour Economics*, 65: 101840.

MacKinnon, James G, and Matthew D Webb. 2017. "Wild Bootstrap Inference for Wildly Different Cluster Sizes." *Journal of Applied Econometrics*, 32(2): 233–254.

MacKinnon, James G, and Matthew D Webb. 2018. "The Wild Bootstrap for Few (Treated) Clusters." *The Econometrics Journal*, 21(2): 114–135.

Norris, Samuel, Matthew Pecenco, and Jeffrey Weaver. 2021. "The Effects of Parental and Sibling Incarceration: Evidence from Ohio." *American Economic Review*, 111(9): 2926–63.

Oliver, William J, Lawrence R Kuhns, and Elaine S Pomeranz. 2006. "Family structure and child abuse." *Clinical pediatrics*, 45(2): 111–118.

Pryce, Jessica, Wonhyung Lee, Elizabeth Crowe, Daejun Park, Mary McCarthy, and Greg Owens. 2019. "A case study in public child welfare: County-level practices that address racial disparity in foster care placement." *Journal of public child welfare*, 13(1): 35–59.

Putnam-Hornstein, Emily, Barbara Needell, Bryn King, and Michelle Johnson-Motoyama. 2013. "Racial and ethnic disparities: A population-based examination of risk factors for involvement with child protective services." *Child abuse & neglect*, 37(1): 33–46.

Putnam-Hornstein, Emily, Eunhye Ahn, John Prindle, Joseph Magruder, Daniel Webster, and Christopher Wildeman. 2021. "Cumulative Rates of Child Protection Involvement and Terminations of Parental Rights in a California Birth Cohort, 1999–2017." *American journal of public health*, 111(6): 1157–1163.

Reddy, Julia, Anne Williams-Isom, and Emily Putnam-Hornstein. 2022. "Racial sensitivity training: An inadequate solution to systemic racial disparities in child protection systems." *Child Abuse & Neglect*, 128: 105584.

Roberts, Kelsey V. 2019. "Foster Care and Child Welfare." Working paper.

Ryan, Joseph P, Brian A Jacob, Max Gross, Brian E Perron, Andrew Moore, and Sharlyn Ferguson. 2018. "Early exposure to child maltreatment and academic outcomes." *Child maltreatment*, 23(4): 365–375.

Warburton, William P, Rebecca N Warburton, Arthur Sweetman, and Clyde Hertzman. 2014. "The impact of placing adolescent males into foster care on education, income assistance, and convictions." *Canadian Journal of Economics*, 47(1): 35–69.

Wildeman, Christopher, and Natalia Emanuel. 2014. "Cumulative risks of foster care placement by age 18 for US children, 2000–2011." *PloS one*, 9(3): e92785.

Yi, Youngmin, Frank R Edwards, and Christopher Wildeman. 2020. "Cumulative prevalence of confirmed maltreatment and foster care placement for US children by race/ethnicity, 2011–2016." *American journal of public health*, 110(5): 704–709.

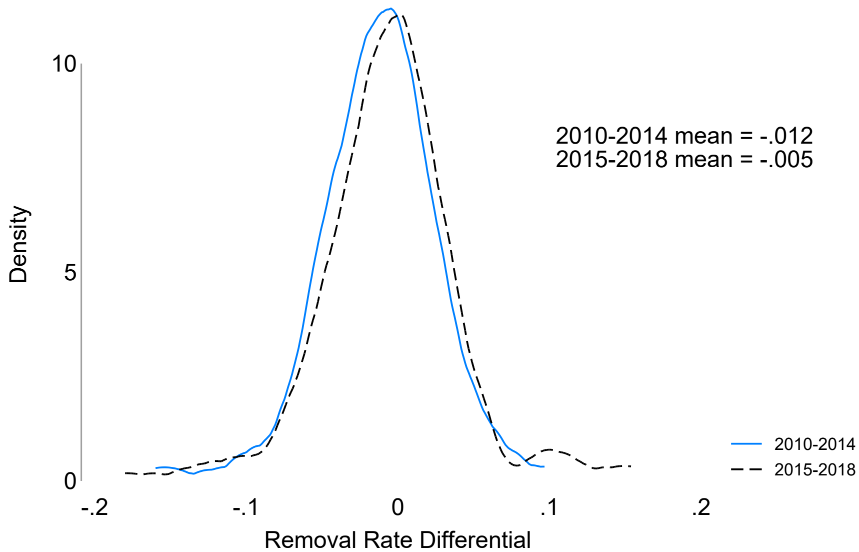
The Push for Racial Equity in Child Welfare: Can Blind Removals Reduce Disproportionality?

E. Jason Baron, Ezra Goldstein, and Joseph Ryan

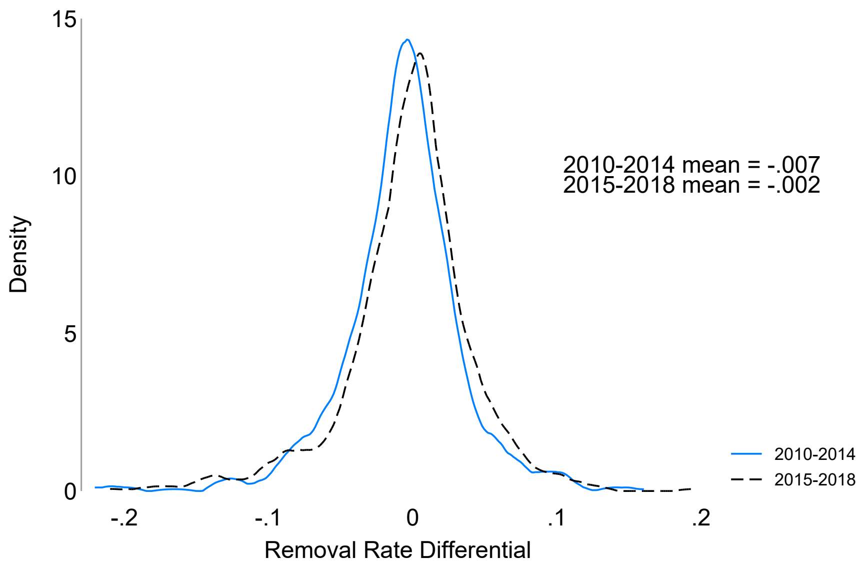
Online Appendix

A Supplemental Figures and Tables

Figure A1: Trends in the Distribution of Removal Rate Race Differentials Pre-Blind Removals



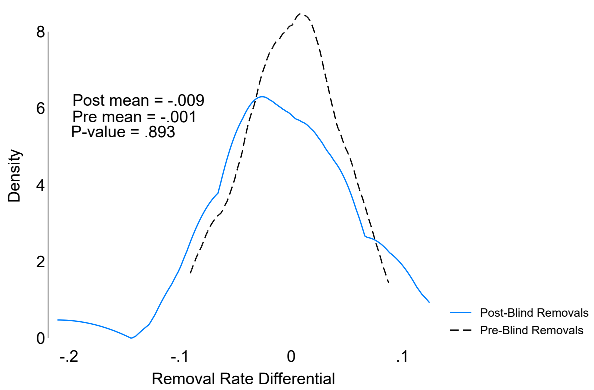
(a) Kent County



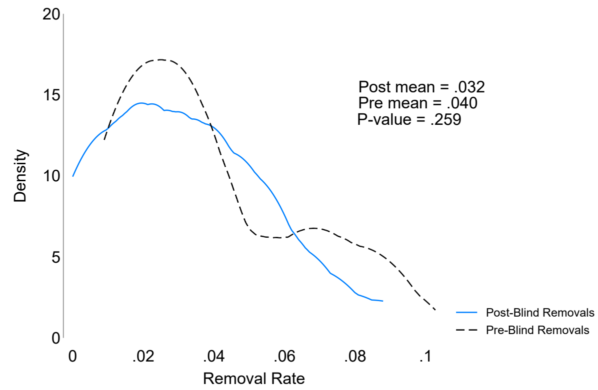
(b) Control Counties

Notes: Panel A of the figure overlays the distribution of R_i^* for Kent County from 2010 to 2014 and from 2015 to 2018, while Panel B shows the same distributions but for the four control counties: Wayne, Oakland, Macomb, and Genesee.

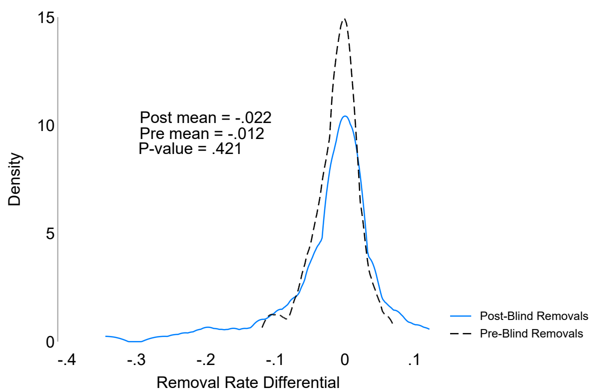
Figure A2: Distributions of Removal Rates Pre- and Post-Blind Removals



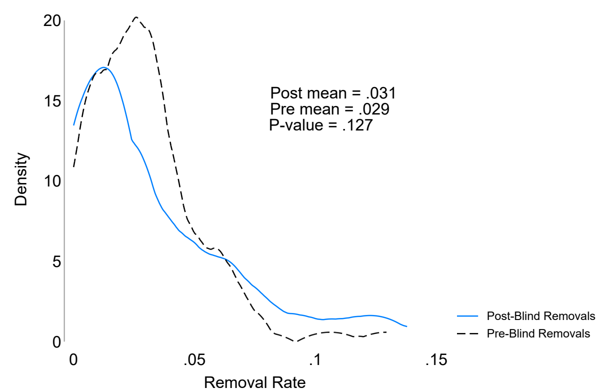
(a) Removal Rate Differential in Kent



(b) Removal Rate in Kent



(c) Removal Rate Differential in Control Counties



(d) Removal Rate in Control Counties

Notes: Panel A of the figure overlays the pre- and post-blind removal distribution of R_i^* for Kent County, while Panel B shows the distribution of R_i instead. Panels C and D show the same distributions but for the four control counties: Wayne, Oakland, Macomb, and Genesee. The figure reports the p-value for a two-sample Kolmogorov-Smirnov test for the equality of distribution functions.

Table A1: Counterfactual Disproportionality Under Equalized Removal Probabilities

	(1) Number MI	(2) Pr(Alleged Victim)	(3) Number Inv.	(4) Pr(Sub| Alleged Victim)	(5) Number Inv & Sub	(6) Pr(Rem|Sub)	(7) Number Rem
Panel A: Actual Disproportionality							
White	78.65	0.20	16.09	0.26	4.11	0.12	0.48
Black	21.35	0.39	8.29	0.27	2.26	0.11	0.26
Share Black	0.21		0.34		0.35		0.35
Measure of Disproportionality	**1.63**						
Panel B: Equal Rem							
White	78.65	0.20	16.09	0.26	4.11	0.11	0.47
Black	21.35	0.39	8.29	0.27	2.26	0.11	0.26
Share Black	0.21		0.34		0.35		0.35
Measure of Disproportionality	**1.66**						
Panel C: Equal Sub&Rem							
White	78.65	0.20	16.09	0.26	4.11	0.11	0.48
Black	21.35	0.39	8.29	0.26	2.12	0.11	0.13
Share Black	0.21		0.34		0.34		0.34
Measure of Disproportionality	**1.59**						
Panel D: Equal Inv							
White	78.65	0.20	16.09	0.26	4.11	0.12	0.48
Black	21.35	0.20	4.37	0.27	1.19	0.11	0.14
Share Black	0.21		0.21		0.22		0.22
Measure of Disproportionality	**1.03**						

Notes: The table shows actual and counterfactual disproportionality measures in Michigan's foster care system, based on the numbers reported in Tables 1 and 2. Panel A of the table shows the actual disproportionality in Michigan's child welfare system, as well as the racial disparities that shape disproportionality at each point in child welfare involvement. The first column shows the "number" of children in Michigan who are White and Black, based on the population shares in Table 1 and assuming only White and Black children in the population. Column 2 shows the probability by race that a child in MI will ever be the alleged victim in a child maltreatment investigation (see Table 2, Column 1). Column 3 shows the number of children who will ever be the alleged victims in a child maltreatment investigation, based on the total number of children in Column 1 and the probabilities in Column 2. Column 4 shows the probability by race that a child who is the alleged victim in a child maltreatment investigation will have that investigation substantiated (see Table 2, Column 2). Column 5 shows the number of children who will ever be the victim in a substantiated maltreatment investigation, based on the number of children in Column 3 and the probabilities in Column 4. Column 6 shows the probability by race that a child who is the victim in a substantiated maltreatment investigation will be removed (see Table 2, Column 4). Finally, Column 7 shows the number of children in MI who will ever be removed, based on the total number of children in Column 5 and the probabilities in Column 6. The third row of each panel shows the share of Black children in each of the populations described above. The fourth row presents a measure of disproportionality, defined as the ratio of the share of Black children removed from their homes to the share of Black children in the population. For instance, in Panel A the measure of disproportionality is $1.63 = 0.35/0.2135$. Panel B presents a measure of disproportionality under equalized removal probabilities, Panel C presents a measure of disproportionality under equalized substantiation *and* removal probabilities, and Panel D presents a measure under equalized investigation rates.

Table A2: Probing the Parallel Trends Assumption

	(1) 2015-2018	(2) 2010-2014	(3) Difference	(4) DID
Panel A: Removal Rate Differential				
Kent	-0.55	-1.25	0.702	
	[4.58]	[3.86]	(0.472)	
N	155	170	325	
Control Counties	-0.25	-0.69	0.438	**0.265**
	[4.37]	[4.06]	(0.216)	(0.518)
				{0.623}
N	706	843	1,549	1,874
Panel B: Removal Rate				
Kent	3.42	4.30	-0.875	
	[2.47]	[2.46]	(0.274)	
N	155	170	325	
Control Counties	3.45	3.97	-0.513	**-0.362**
	[2.56]	[2.25]	(0.122)	(0.300)
				{0.332}
N	706	843	1,549	1,874

Notes: In Panel A, Column 1 presents estimates of $\bar{R}^{*K}_{2015-2018}$ and $\bar{R}^{*C}_{2015-2018}$ while Column 2 presents estimates of $\bar{R}^{*K}_{2010-2014}$ and $\bar{R}^{*C}_{2010-2014}$. Column 3 presents estimates of ($\bar{R}^{*K}_{2015-2018}$ - $\bar{R}^{*K}_{2010-2014}$) and ($\bar{R}^{*C}_{2015-2018}$ - $\bar{R}^{*C}_{2010-2014}$). We present DID estimates in Column 4 in bold. For expositional purposes, we multiply all rates by 100 and present them as percentage points. Standard deviations are shown in brackets in Columns 1 and 2, while robust standard errors are shown in parentheses in Columns 3 and 4. Ferman and Pinto (2019) p-values are shown in curly brackets, as discussed in the main text. Panel B shows the same information but for overall removal rates (R_{it}) instead.

Table A3: Robustness to Alternative Inference

	(1) Removal Rate Diff. Pre-Trend	(2) Removal Rate Pre-Trend	(3) Removal Rate Differential	(4) Removal Rate
Point Estimate	0.26	-0.36	0.14	-1.09
P-values from...				
Robust Ses	[0.610]	[0.228]	[0.939]	[0.191]
Clustered by Investigator Ses	[0.590]	[0.193]	[0.940]	[0.145]
Clustered by County Ses	[0.351]	[0.140]	[0.944]	[0.069]
Wild Cluster Bootstrapped Ses	[0.705]	[0.553]	[0.982]	[0.367]
Ferman and Pinto Ses	[0.623]	[0.332]	[0.946]	[0.157]

Notes: The table shows robustness of the main results of the paper to alternative statistical inference procedures. Specifically, rows 3, 4, and 5 show p-values derived from robust standard errors, standard errors clustered by investigator, and standard errors clustered by county, respectively. The fourth row shows p-values derived from wild-cluster bootstrapped standard errors. We follow MacKinnon and Webb (2018) and perform a subcluster bootstrap at the investigator level since our model has only one treated unit. Finally, the fifth row shows p-values derived using the method outlined in Ferman and Pinto (2019). For expositional purposes, we multiply all rates by 100 and present them as percentage points.

B Additional Research Designs

Case-Level Analysis

The goal of blind removals is to reduce racial disproportionality by eliminating racial disparities in investigators' decisions to remove. As a result, our analysis in the main body of the paper is done at the investigator level. Specifically, we directly test the effects of blind removals on investigators' removal behavior. An alternative way to test the causal effects of blind removals is at the child-by-investigation level. In this section, we show that the main conclusions of the paper are unchanged if we instead model the probability that a particular investigation results in removal as a function of the blind removals program.

We focus on the 321,093 substantiated investigations for either White or Black children in Michigan from January 2010 to March 2020 and estimate the following specification:

$$\begin{aligned} FC_{ict} =& \beta_0 + \beta(Kent_c \times Post_t) + \delta Black_i + \gamma(Kent_c \times Post_t \times Black_i) \\ & + \theta_c \times Black_i + \tau_t \times Black_i + \varepsilon_{ict} \end{aligned} \tag{B.1}$$

where FC_{ict} is an indicator equal to one if investigation i in county c in time period t resulted in a removal; $Kent_c$ is an indicator equal to one if the investigation occurred in Kent; $Post_t$ is an indicator equal to one if the investigation occurred during or after August 2019; $Black_i$ is an indicator equal to one if investigation i involved a Black child; θ_c and τ_t are county- and month-year-level fixed effects.

In this specification, β measures the effects of the blind removal program on the removal probability of investigations involving White children. In other words, it measures—for the subset of substantiated investigations involving White children—the effects of blind removals on the probability that an investigation results in removal from the home. γ is the parameter of interest and measures the differential effect of blind removals on the probability that an investigation involving a Black child resulted in removal. If blind removals decreased racial disparities in the decision to remove, then γ would be negative. If blind removals decreased the probability of removal for both White and Black children but the decline was similar, then β would be negative and γ would be close to zero.

Table B1 presents estimates of β and γ along with standard errors two-way clustered at the county and month-year levels in parentheses. The first two columns of the table present the estimates when using the four counties in the main text as a control group: Wayne, Oakland, Macomb, and Genesee. The last two columns show estimates when using every other county in Michigan as a control group. The second and fourth columns include investigator-specific fixed effects.

The table yields two main takeaways. First, there is some evidence that blind removals decreased the probability of removal for White children. Estimates of β are consistently negative and large: Relative to an average removal rate among substantiated investigations of 12.5 percent, blind removals decreased the probability of removal by roughly 3 percentage points, or 24 percent. Second, there is no evidence that blind removals differentially impacted the removal rate of Black children (relative to White children). Estimates of γ are mostly small in magnitude and statistically insignificant.

Table B1: Case-Level Analysis

	(1)	(2)	(3)	(4)
	Control Counties		All Counties	
β	-0.028	-0.039**	-0.034***	-0.031***
	(0.013)	(0.009)	(0.006)	(0.007)
γ	0.008	0.005	0.018*	0.002
	(0.017)	(0.021)	(0.010)	(0.011)
Control Mean	*0.125*	*0.125*	*0.128*	*0.128*
N	130,275	130,275	321,093	321,093
County FE	✓	✓	✓	✓
Month-Year FE	✓	✓	✓	✓
Investigator FEs		✓		✓

Notes: The table presents estimates of β and γ from Equation B.1 along with standard errors two-way clustered at the county and month-year levels in parentheses. The first two columns of the table present the estimates when using the four control counties in the main text as a control group: Wayne, Oakland, Macomb, and Genesee. The last two columns show estimates when using every county in Michigan as a control group. The second and fourth columns include investigator-specific fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

As discussed in the main body of the paper, inference in our setting is complicated by the fact that there is only one treated cluster. As a result, traditional clustered standard errors may yield improper inference. Indeed, Ferman and Pinto (2019) p-values for the estimates in the first row range from 0.102 to 0.25, so it is difficult to definitively conclude that blind removals reduced overall removal rates. Still, these results yield a remarkably similar story to our findings in the main body of the paper: While there is some suggestive evidence that blind removals reduced removal rates overall, there is limited evidence that the program did so differentially for White and Black children.

Heterogeneity by Child and Case Characteristics

Previous work highlights that how children respond to foster care and other environmental changes varies by both age and sex. For instance, young children benefit from moving to lower-poverty areas more than older youth (Chetty, Hendren and Katz, 2016; Chyn, 2018), and previous studies of foster care find that foster placement harms the outcomes of 16- to 18-year-old males (Warburton et al., 2014), significantly improves outcomes for young girls (Bald et al., 2022*a*), but has no impacts on young boys. In addition, previous work shows that males are often more vulnerable than females to childhood disadvantage or disruption (Autor et al., 2019; Bertrand and Pan, 2013; Kling, Ludwig and Katz, 2005).

Given these findings, it is important to understand whether the blind removals program differentially impacted the probability that children with differing demographics are placed in foster care. Table B2 shows estimates of β and γ from Equation B.1 and from our preferred specification in Column 4 of Table B1 on six distinct subgroups of investigations. Columns 1 and 2 compare the effects of blind removals on investigations that involved an allegation of abuse and investigations that involved only allegations of neglect, respectively. Columns 3 through 6 compare the effects of blind removals across four distinct demographic subgroups: older ($\geq$ 6 years old, the median age in our sample) males, younger males, older females, and younger females.

Table B2: Heterogeneity in Case-Level Analysis

	(1) Abuse	(2) Neglect	(3) Old Male	(4) Young Male	(5) Old Female	(6) Young Female
β	-0.063***	-0.027***	-0.026***	-0.029	-0.016	-0.047***
	(0.017)	(0.007)	(0.006)	(0.020)	(0.012)	(0.012)
γ	0.017	0.018	-0.060***	0.042***	0.032	0.007
	(0.024)	(0.013)	(0.014)	(0.016)	(0.034)	(0.023)
Control Mean	0.131	0.122	0.105	0.146	0.112	0.148
N	213,398	105,676	74,704	84,115	79,949	76,964
County FE	✓	✓	✓	✓	✓	✓
Month-Year FE	✓	✓	✓	✓	✓	✓
Investigator FEs	✓	✓	✓	✓	✓	✓

Notes: The table presents estimates of β and γ from Equation B.1 along with standard errors two-way clustered at the county and month-year level in parentheses. We estimate the specification in Column 4 of Table B1 separately on each of six subgroups: cases involving at least one allegation of abuse, cases involving no allegations of abuse (only neglect), old ($\geq$6 years old, the median age in our sample) and young males, and old and young females.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The first row of the table shows how the blind removals program impacted the probability of placement for different subgroups of White children. The table shows that the probability of removal decreased more for investigations involving an abuse allegation relative to investigations involving only allegations of neglect. The decline in the probability of removal was also larger for young females relative to other White subgroups such as older White females and younger White males.

The second row of the table tests whether the effects mentioned above differ for Black children relative to White children across each of the subgroups. The table shows some evidence that the decline in the removal probability was smaller (in magnitude) for some subgroups of Black children. Specifically, both Black victims of abuse and Black victims of neglect had a lower decline in the probability of removal relative to White victims of abuse and White victims of neglect. This pattern is also true for young Black males, young Black females, and older Black females. In fact, for some subgroups such as young Black males, the estimates actually indicate that the blind removals program *increased* the probability of removal.

Synthetic Control

Given the problems with inference in the analysis presented above, we also estimate the causal effects of blind removals using synthetic control methods proposed in Abadie and Gardeazabal (2003) and Abadie, Diamond and Hainmueller (2010). The synthetic control method chooses a weighted average of control counties in Michigan to best match Kent for the outcome of interest and other covariates prior to the implementation of blind removals. Specifically, we compare the evolution of our two main outcomes—the total number of children removed in a given year and the share of total children removed who are Black—in Kent County to a "synthetic" Kent County comprised of a weighted average of control "donor" counties that mirror Kent in the pre period. Table B3 shows, for each outcome, the weight assigned to each of the donor counties that comprise synthetic Kent.

Table B3: County Weights for Synthetic Kent County

County	Weight
Panel A: Number Removed	
Washtenaw	0.412
Oakland	0.300
Menominee	0.258
Wayne	0.030
Panel B: Black Share of Removals	
Allegan	0.632
Wayne	0.337
Kalamazoo	0.031

Notes: The table presents the combination of counties and weights that comprise the synthetic control for Kent County for the total number of children removed (Panel A) and the Black share of total removals (Panel B). The initial pool of potential donor counties contained the 70 counties in Michigan that had at least one child removal in every year from 2016 to 2020.

Table B4 compares Kent County to synthetic Kent in the years prior to blind removals. Since the post-blind removal period in our sample consists of August 2019 to March 2020, we defined each year by a similar window of months. For instance, the "year" 2016 consists of August 2015 to March 2016. Similarly, the "year" 2020 consists of August 2019 to March 2020. We define the pre-treatment period as the years from 2016 to 2019, and focus on a single post-period—2020—to avoid conflating differential effects of the COVID-19 pandemic.

For both outcomes, we used the following variables to construct a synthetic Kent County: the particular dependent variable, the county's total population, share of the population who is Black, share of the child population in poverty, median household income, and unemployment rate, all averaged over the pre-treatment period. The table shows that, in both cases, the main outcomes in synthetic Kent closely match those in Kent County.

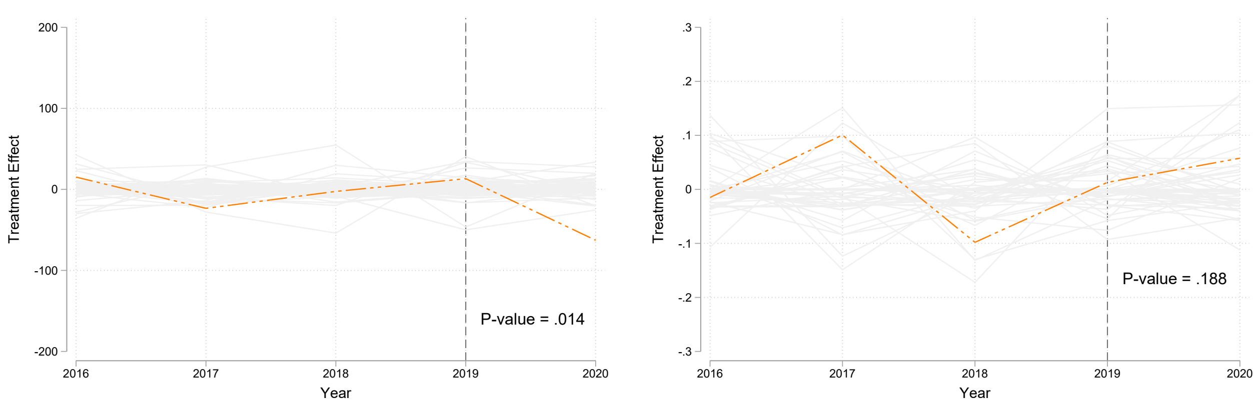
Table B4: Pre-Treatment Balance

	(1)	(2)	(3)
		Synthetic Control Outcome	
	Kent County	Number Removed	Black Share of Removals
Number Removed	246	246	-
Black Share of Removals	0.324	-	0.324
Total Population	175,919	197,045	147,031
Black Share of Population	0.120	0.162	0.125
Share of Child Population in Poverty	0.156	0.204	0.144
Median Household Income	59,255	54,278	63,379
Unemployment Rate	3.235	4.355	3.867

Notes: The table presents the average value of variables prior to the implementation of the blind removals program for Kent County and its synthetic control. County-level demographics were obtained from the Census Bureau's Population and Housing Unit Estimates Program. The share of children in poverty and the median household income come from the Census Bureau's Small Area Income and Poverty Estimates Program. The unemployment rate comes from the Bureau of Labor Statistics' Local Area Unemployment Statistics Program.

We first examine the total number of child removals in Kent County. Panel A of Figure B1 plots the "treatment effect" (the difference in the outcome between Kent County and synthetic Kent County) in orange before and after the implementation of blind removals, as well as 70 placebo synthetic control comparisons using all other counties as placebo treated counties in light gray. Using this empirical distribution of placebo treatment effects, we also calculate the "p-value" associated with the post-blind removal treatment effect—the proportion of placebo effects that are at least as large as the main effect for the post-treatment period.

Figure B1: Synthetic Differences, Kent vs. Placebo Counties



(a) Number Removed

(b) Black Share of Removals

Notes: The figure plots the results of a permutation test of the significance of the difference between Kent County and synthetic Kent County. The dashed orange line plots the difference between Kent County and synthetic Kent County. The light gray lines instead compute synthetic control differences using all other counties as placebo treated units.

Consistent with the evidence in the main body of the paper, the figure shows that, relative to its synthetic counterpart, Kent removed roughly 50 fewer children following the implementation of blind removals. The treatment effect appears to lie mostly outside of the range of placebo effects, yielding a "p-value" of 0.014. Panel B repeats this exercise but for the share of total removals in the county that consisted of Black children. For this outcome, the actual treatment effect between Kent County and synthetic Kent lies entirely inside the range of placebo effects, yielding a "p-value" of 0.188.

Altogether, the results in this section paint a similar picture to those in the main body of the paper. We find suggestive evidence that the blind removals program decreased the probability of removal, but we find no consistent evidence that the program had any impact on the share of removals by race, and as a result on racial disproportionality in foster care.