NBER WORKING PAPER SERIES

IS THERE A FOSTER CARE-TO-PRISON PIPELINE? EVIDENCE FROM QUASI-RANDOMLY ASSIGNED INVESTIGATORS

E. Jason Baron
Max Gross

Working Paper 29922
http://www.nber.org/papers/w29922

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2022, revised October 2024

We are grateful to Brian Jacob, Michael Mueller-Smith, Joseph Ryan, Charlie Brown, and Kevin Stange, who provided invaluable support and advice. The study also benefited immensely from discussions with Anna Aizer, Peter Arcidiacono, Patrick Bayer, Jesse Bruhn, Lindsey Bullinger, Eric Chyn, Allan Collard-Wexler, Mark Courtney, Ashley Craig, Jennifer Doleac, Joseph Doyle, Susan Dynarski, George Fenton, Sarah Font, John Friedman, Ezra Goldstein, Matthew Gross, Sara Heller, Peter Hull, Joshua Hyman, Katrine Løken, Parag Mahajan, Matt Masten, Sam Norris, Emily Oster, Stephanie Owen, Matthew Pecenco, Roman Rivera, Evan Rose, Adam Rosen, Jonathan Roth, Andrew Simon, Juan Carlos Suárez Serrato, Elizabeth Weigensberg, and Fred Wulczyn, as well as feedback from seminar and conference participants at Abt Associates, the America Latina Crime and Policy Network, AEFP, APPAM, Brown University, Grinnell College, the NBER Children Spring Meeting, Mathematica, the Triangle Applied Microeconomics Conference, the Texas Economics of Crime Workshop, São Paulo School of Economics, the University of Calgary, the University of Wisconsin, and the University of Michigan. This research used data structured and maintained by the MERI-Michigan Education Data Center (MEDC). MEDC data are modified for analysis purposes using rules governed by MEDC and are not identical to those data collected and maintained by the Michigan Department of Education (MDE) and/or Michigan's Center for Educational Performance and Information (CEPI).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, nor of the National Bureau of Economic Research.

Is There a Foster Care-To-Prison Pipeline? Evidence from Quasi-Randomly Assigned Investigators
E. Jason Baron and Max Gross
NBER Working Paper No. 29922
April 2022, revised October 2024
JEL No. H75,I38,J13,K42

ABSTRACT

Foster care placement is strongly associated with crime—for example, close to one fifth of the prison population in the U.S. is comprised of former foster children—yet there is little evidence on whether this relationship is causal. Leveraging the quasi-random assignment of investigators and administrative data from Michigan, we show that placement substantially reduced the chances of adult arrests, convictions, and incarceration for children at the margin. Exploring mechanisms, we find evidence that children's birth parents made positive changes following placement. We show that most children in our setting reunified with their parents after being in foster care for one to two years, and that parents themselves were less likely to have criminal justice contact after placement. Considering recent historic federal policy which prioritizes keeping children with their families, our analysis indicates that safely reducing foster care caseloads will require improving efforts to ensure child wellbeing in the home.

E. Jason Baron
Economics Department
Duke University
228B Social Sciences Building
Durham, NC 27708
and NBER
jason.baron@duke.edu

Max Gross
600 Alexander Park Dr
Princeton, NJ 08540
mgross@mathematica-mpr.com

A non-technical summary of the paper is available at http://www.nber.org/data-appendix/w29922

Involvement with the child welfare and foster care systems is surprisingly common in the United States. Children are placed in foster care when government authorities find that they were abused or neglected and determine it is not safe for them to continue living in their homes. By age 18, more than one third of children are subject to a formal child welfare investigation for alleged maltreatment and 5 percent of children enter foster care (Kim et al., 2017; Yi, Edwards and Wildeman, 2020). Child welfare contact is even more common for Black children; more than half are subject to an investigation and up to 9 percent enter foster care at some point during their childhoods.

Children placed in foster care are particularly likely to be involved in the criminal justice system as adults. Close to one fifth of the prison population in the United States is comprised of former foster children (BJS, 2016) and about 70% of youth who exit foster care as legal adults are arrested at least once by age 26 (Courtney et al., 2011). Decades of research also show a positive association between foster care placement and criminality, and the media often cites a "foster care-to-prison pipeline" (Amon, 2021; Trivedi, 2020).

Despite these dismal descriptive statistics, there is little evidence for a causal relationship between foster care and later-in-life crime. Ex ante, it is unclear whether foster care would reduce or increase adult crime. On the one hand, keeping children in a harmful home environment could lead to worse adult outcomes. On the other hand, separating children from their families could lead to trauma and instability in their lives. Seminal work in Doyle (2008) provides the only causal evidence on the relationship between foster care and adult crime in the United States, and finds that placement tripled arrest, conviction, and imprisonment rates for children investigated in Illinois nearly three decades ago. However, in light of substantial changes to federal child welfare policy over time and the dramatic increase in foster care placements attributed to the opioid epidemic (Dallman, 2020; Evans, Harris and Kessler, 2022; Hou, 2022), it is critical to understand the effectiveness of foster care interventions as currently practiced.

This paper provides new evidence on the causal relationship between foster care and adult crime. We assembled a rich administrative dataset linking child welfare and adult criminal justice records in Michigan. Furthermore, to explore mechanisms, we linked child welfare records to administrative records from Michigan's K-12 public school system, juvenile detention spells, and nationwide postsecondary enrollment information. We also linked the birth parents involved in child welfare investigations to criminal justice records.

We study nearly 120,000 child welfare investigations involving children ages 6 to 16 between 2008 and 2016. Our research design exploits plausibly exogenous variation in foster care placements created by the rotational assignment of child welfare investigators. Investigators are assigned to cases based on who is next up on a list, not for reasons specific to the child or family, and they have discretion over whether to recommend placement. These decisions are in part subjective and some investigators are stricter than others. Using investigator stringency as an instrument, we compare the outcomes of children who by chance are assigned a strict investigator and placed in foster care to those who are not placed only because they are assigned a more lenient investigator. This research design allows us to identify local average treatment effects, which are impacts for children at the margin of placement: those for whom investigators might disagree about foster care placement. These children, or compliers, are a relevant population for policy because they are exactly the children for whom investigators wield discretion over placement (Berrick, 2018).

We find that foster care placement substantially reduces the chances of later-in-life criminal involvement for children at the margin. Children placed in foster care are 25 percentage points less likely to be arrested by age 19 relative to a control complier mean of 37%, a decrease of 68%. We find even larger reductions in the likelihood of being convicted (28 percentage points or 81%) and incarcerated (21 percentage points or 80%). We find similar declines through age 21 for the subset of youth that we observe at older ages. We show that these results are not driven by differential out-of-state migration and are robust to alternative designs and samples. These effect sizes are substantial, as the removal of a child is perhaps the most sweeping child and family intervention. For example, they are larger than those found for other interventions targeting youth that directly aim to reduce crime, such as cognitive behavioral therapy in the Becoming a Man program, which reduced violent-crime arrests by 45-50% (Heller et al., 2017). That we find such sizeable effects on crime is consistent with the large effect sizes in Doyle (2008), albeit opposite in sign.

Exploring mechanisms, we present evidence that children's birth parents make improvements in their own lives while children are temporarily in foster care. Having a child removed substantially decreased the likelihood of birth parents being arrested (8 percentage points or 85%) and convicted (7 percentage points or 86%) in the years after the investigation. There are institutional reasons to suspect that birth parents would make these substantial improvements in their own lives following child removal. Parents have strong incentives to make improvements. Caseworkers (who are different from the initial investigators) and judges monitor and review parents' progress before approving family reunification. As examples, they may review progress on strengthening parenting skills, overcoming substance abuse challenges, finding stable employment, or securing housing. We find less evidence for other mechanisms such as children moving to more advantaged neighborhoods following reunification or parental incarceration.

This study provides new evidence on whether or not there is a foster care-to-prison pipeline. Consistent with the idea of a pipeline, we can reproduce descriptive statistics showing that children who are placed in foster care are substantially more likely to be involved in the adult criminal justice system. Indeed, we find that this relationship holds even after controlling for a range of observable characteristics. However, our instrumental variables strategy reveals that, in our data, this relationship is the result of selection bias and does not represent the causal effects of foster care. Failing to isolate exogenous variation in foster care can lead to incorrect conclusions about the foster care-to-prison pipeline.

This study's findings run counter to the prior literature on the relationship between foster care and adult crime. Notably, they contrast those in Doyle (2008), which used the same examiner assignment research design to study placements nearly three decades ago and provides the only other causal estimates on this relationship. National changes in child welfare policy and practice in the 30 years since the beginning of the Doyle (2008) sample likely explain the difference in findings (see Online Appendix C for a detailed discussion).

First, child welfare policy has changed over time in ways that have likely improved foster care. The federal government has enacted several key policies after the end of the Doyle (2008) sample period focusing on reducing placement length, improving the quality of placement settings, and promoting the wellbeing of children while in foster care. At the same time, broader social policies have changed in ways that have likely improved the counterfactual to placement (e.g., child poverty has declined by nearly 60% (Thomson et al., 2022)). Absent improvements to foster care over time, we would therefore expect to find even more detrimental effects of placement more recently. That we instead find favorable effects suggests that the foster care system itself has likely improved over time.

Second, child welfare practice has also changed in ways that may have made children at the margin experience less harmful effects of placement. The share of children referred as possible victims of maltreatment in the United States who are placed in foster care has dropped by 50% over the last twenty years (USDHHS, 2000, 2019). Because a larger share of referrals resulted in placement during the time of the Doyle (2008) study, children at the margin of placement may have faced less risk in the home than marginal children today. We present evidence for this hypothesis below. In contrast, we find less evidence that differences in sample restrictions or outcome measures can account for the differences in findings.

Understanding the efficacy of current foster care systems is crucial in light of these nationwide policy changes. Compliers in our setting tended to have the types of experiences in foster care that recent policies have pushed for, which bolsters the generalizability of our findings. Most compliers were placed in a family home as opposed to a group home, spent a relatively short amount of time in care, and reunified with their birth parents. Further supporting generalizability to current-day systems, the complier population in our study is disproportionately likely to have been investigated for parental substance use and neglect, which is consistent with national trends in the types of cases that enter foster care (Bald et al., 2022*b*). Moreover, we extend Doyle (2008) by exploring the specific mechanisms through which foster care influences later-in-life crime, which is crucial for policy purposes. By showing that birth parents make improvements while their children are temporarily in foster care, we also highlight that foster care is a family intervention and that family structure and parental behavior contribute to criminal development. Particularly given the prevalence of foster care placements nowadays, in part due to the opioid epidemic, the findings in this study offer important insight into how present-day foster care systems are likely to influence children at the margin and the channels through which these effects operate.

This study builds upon Gross and Baron (2022), which found that removal improved children's safety and academic outcomes such as test scores and daily school attendance, by showing that foster care reduces children's later-in-life crime. There is little reason to expect that the findings in Gross and Baron (2022) would translate into the large declines in adult crime that we document in this study. For example, the reduction in crime is over six times larger than what would be expected based on the estimates in Gross and Baron (2022).[1] This is consistent with a growing body of evidence suggesting that effects on childhood outcomes alone may not be predictive of effects on later-in-life crime (Anders, Barr and Smith, 2022; Chetty et al., 2011; Deming, 2009, 2011; Gray-Lobe, Pathak and Walters, 2021; Heckman, Pinto and Savelyev, 2013). Indeed, consistent with Gross and Baron (2022), we show that removal improves test scores and attendance and protects children from subsequent maltreatment in our sample. We additionally provide new evidence that foster care increases the likelihood of high school graduation and college enrollment, and reduces the likelihood of being held in a juvenile detention center. However, we then show that, even under very conservative assumptions, improvements in all of these intermediate outcomes explain only about 40% of the estimated reduction in adult crime. Thus, the magnitude of the

[1]This result is based on the non-experimental relationships between the outcomes studied in Gross and Baron (2022) and adult criminal justice contact in our data, as well as the point estimates in Gross and Baron (2022); see Online Appendix B for more information.

later-in-life crime reduction documented in this study is not well-predicted by improvements in intermediate outcomes alone, including those documented in Gross and Baron (2022).

Furthermore, our focus on crime in this study is particularly important. Reducing crime generates large positive externalities for society that often drive benefit-cost ratios in program evaluation, whereas improvements in children's cognitive skills yield mostly private benefits. As such, we provide the first estimate of the Marginal Value of Public Funds (MVPF) of foster care, and show that its social benefits from reducing children's future crime alone are greater than its costs. Moreover, the close connection between child maltreatment, foster care, and crime is well established (Currie and Tekin, 2012), meaning that crime is an especially relevant outcome for this population.

Finally, we build on Gross and Baron (2022) by providing direct evidence that foster care had positive effects on birth parents. Gross and Baron (2022) suggested this as a likely mechanism in light of improvements in children's outcomes following reunification, yet lacked data on the outcomes of birth parents to directly test this hypothesis. In this study, we show that foster care reduces birth parents' contact with the criminal justice system. Our findings on the removal of a child are similar to recent evidence that "turning points" such as having a child can substantially influence parents' lives.[2]

We similarly complement other studies in this literature including Bald et al. (2022*a*) and Roberts (2019), who estimate the effects of foster care for young children in Rhode Island and South Carolina, respectively.[3] Bald et al. (2022*a*) focus on academic outcomes such as test scores and grade repetition and find substantial gains for girls younger than 6 years old but null effects for other gender-age groups. Roberts (2019) finds positive impacts on on-time grade progression, yet noisy estimates on daily school attendance and test scores.

Our findings also contribute to the economics of crime literature identifying effective crime

[2]For example, Massenkoff and Rose (2022) and Eichmeyer and Kent (2022) show that childbirth can reduce subsequent criminal behavior for mothers by as much as 50%.

[3]Two related studies estimate the causal effects of foster care outside of the U.S. Warburton et al. (2014) study the effects of foster care on a range of outcomes for 16- to 18-year-old male youth in Canada and find mixed results. Helénsdotter (2022) finds adverse impacts on mortality driven by suicides in Sweden.

reduction strategies. Although many studies have focused on policing or tougher sanctions as strategies to reduce crime (Bell, Jaitman and Machin, 2014; Chalfin et al., 2021; Drago, Galbiati and Vertova, 2009; Evans and Owens, 2007; Katz, Levitt and Shustorovich, 2003; Mello, 2019), we add to a growing literature emphasizing the efficiency gains of early policy interventions that prevent the development of offenders in the first place. For example, studies have shown that increasing access to mental healthcare (Jácome, 2020), increasing the quality of public schools (Baron, Hyman and Vasquez, 2024; Cullen, Jacob and Levitt, 2006; Deming, 2011; Dobbie and Fryer, 2015), moving out of disadvantaged neighborhoods (Chyn, 2018), and limiting lead exposure (Billings and Schnepel, 2018; Grönqvist, Nilsson and Robling, 2020) can reduce adult crime. Our study shows that preventing child maltreatment is another effective way to reduce later-in-life crime.

The results of this study are especially because of the historic changes to federal policy introduced in the Family First Prevention Services Act of 2018. A main goal of this legislation is to keep families intact and reduce foster care caseloads. We find that maltreated children who were not placed in foster care are more likely to be involved in the adult criminal justice system relative to those who were placed. Thus, our results indicate that safely reducing foster care caseloads will require improving efforts to ensure child wellbeing in the home.

I Overview of the Child Welfare System

Child welfare involvement is a surprisingly common experience for children in Michigan. Using the linked child welfare and public education data described below, we calculate that roughly one third of all public school children—and nearly half of Black children—are the subject of a maltreatment investigation by age 18. This section describes the maltreatment investigation process and the foster care system in Michigan.

I.A Child Maltreatment Investigations

In Michigan and across the country, a child maltreatment investigation is triggered by a report of suspected child abuse or neglect to a centralized hotline (Figure A1). Child abuse includes harm from non-accidental physical injuries and sexual abuse; child neglect includes harm from negligence, such as inadequate supervision due to parental substance abuse. Anyone can call the hotline to report suspected maltreatment and the most common reporters are education and law enforcement personnel (Fitzpatrick, Benson and Bondurant, 2022). Hotline employees screen and transfer reports to the child's local child welfare office.

Upon receiving a report of suspected maltreatment, the 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 in the rotation and investigators are not assigned based on their specific characteristics or skill sets, with two exceptions. Sexual abuse reports tend to be assigned to more experienced investigators, and repeat reports involving a child who was recently investigated are often re-assigned to the initial investigator. We exclude these exceptions from our analysis sample by dropping reports of sexual abuse and those involving children who had been the subject of an investigation in the year before the report. Typically, each county has its own local office but some large counties have multiple offices and some offices split investigators into geographic-based teams. As such, to compare children who could have been assigned to the same investigator, we include zip code by investigation year fixed effects in our analyses.

Investigators have about 30 days to make two decisions that jointly determine the outcome of the investigation. Investigators first decide whether there is enough evidence to substantiate the allegation by interviewing the people involved, visiting the home, and reviewing any police or medical reports. About 78% of cases are unsubstantiated in our sample, compared to three quarters of cases nationwide (AECF, 2017). In these cases, the investigation ends without any further follow-up. In the roughly 22% of substantiated cases in our sample, investigators then determine the level of risk the child faces in their home. To determine risk levels, investigators complete a 22-question risk assessment and cases deemed at the highest risk may result in removal from the home. Many of the questions on the risk assessment are subjective (e.g., one question asks whether the caretaker "views the incident less seriously than the department") and investigators often manipulate responses based on their priors about the child's risk level (Bosk, 2015; Gillingham and Humphreys, 2010). Consequently, investigators have immense discretion over child removal.

Investigator judgment over both evidence and risk determines the outcome of the investigation. As described earlier, unsubstantiated cases require no further follow-up. If investigators substantiate a report and the risk level is low, they must refer the family to community-based services like food pantries, support groups, or other local nonprofits. After the investigator refers the family to community-based services, the child welfare office does not follow up further. Roughly 12% of cases in our sample end up on this track. If investigators substantiate the allegation and the risk level is high, the family also receives more intensive, targeted services based on their needs, which could include substance abuse treatment, parenting classes, and counseling. These cases make up 8% of all investigations in our sample. Importantly, families are not typically compelled by the court to take up either community or targeted services. Lastly, substantiated allegations with particularly high risk not only trigger targeted and community services but also require investigators to file a court petition for child removal. The 2% of investigations that lead to removal take roughly 10 days, on average.

I.B Foster Care System

Foster care is intended to be a temporary and family-oriented intervention. Children are placed with either an unrelated foster family, relatives, or in a group home. Child welfare typically tries to place children with relatives or, if no suitable relatives are available or willing, an unrelated family. Group home placements are considered a last resort for most children. In 2015, 41% of all foster children in Michigan were placed with an unrelated family, 35% lived with relatives, 9% lived in group homes or institutions, and 14% lived in other settings, such as pre-adoptive homes or supervised independent living (AECF, 2017). It is common to switch placement settings while in the foster system: 60% of all foster children in Michigan in 2015 lived in more than one setting, and 17% lived in at least four.

Birth parents work to regain custody in nearly all foster care cases. Child welfare caseworkers (who are different from the maltreatment investigator) meet with birth parents to create a reunification plan which stipulates how they can regain custody. Reunification plans might require birth parents to complete drug tests, secure housing, or get a job. Birth parents receive targeted services aimed at addressing the challenges in their own lives, which can range from counseling and parenting classes to job training and substance abuse treatment. Caseworkers monitor birth parents' progress and update the reunification plan as needed. Family reunification only occurs if the court decides that birth parents have made sufficient progress for their child to be safe in the home. About half of all foster children in Michigan reunify with their birth parents. Close to one third of children are adopted or have legal guardianship transferred, 9% exit as legal adults, and others enter into an informal guardianship with relatives.

II Data Sources and Sample Construction

II.A Administrative Data Sources

This study uses four sources of administrative data spanning: (1) child welfare, (2) adult crime, (3) K-12 education, and (4) postsecondary education.

Child welfare data come from the Michigan Department of Health and Human Services and consist of the universe of child maltreatment investigations in Michigan between August 1996 and July 2017. These data include information on the children involved in each investigation, and the details of each investigation, such as the allegation report date, allegation types as coded by the investigator, the child's zip code, and whether the investigator substantiated the allegation. The data also include placement records, such as whether a child was placed in foster care following an investigation, placement setting, and permanency outcome. We use this information to construct our main treatment variable: whether a child was placed in foster care due to a child welfare investigation. The data do not include the specific services to which families are referred or information about the take-up of services, nor do they include sibling identifiers. Crucial to our identification strategy, which leverages the quasi-random assignment of investigations, the files include unique investigator identifiers beginning in 2008.

The child welfare data also include information about the alleged perpetrator of maltreatment in each investigation. We use the term "birth parent" broadly to refer to this alleged perpetrator. This simplification helps us distinguish between the adult who had initial custody of the child before the investigation (e.g., a biological or stepparent) and the adult with custody of the child during the placement episode (e.g., foster care parents).[4] Thus, only "birth parents" listed as alleged perpetrators in the investigation are included in our data. For example, if a two-parent household is investigated for child maltreatment and both parents are alleged as perpetrators by the initial reporter, then both parents are included in our dataset. However, if only one parent in a two-parent household is listed as the alleged perpetrator in a child maltreatment investigation, then we only see that parent in our dataset.[5]

Criminal justice data come from the Michigan State Police and contain our three main adult crime outcomes: arrests, convictions, and incarceration. To measure adult arrests,

[4] In almost all cases in our sample (91%), the alleged perpetrator is the adult with legal custody of the child prior to the investigation. We show below that our main results are robust to dropping the 9% of investigations where the alleged perpetrator is not a parent (e.g., an uncle or cousin).

[5] It is possible that the investigator has some discretion in coding the alleged perpetrator. For example, if an initial reporter broadly describes a child's situation as neglectful but is unsure as to which parent is the alleged perpetrator, the investigator can decide which parent (or both) to list as the alleged perpetrator. To ensure that this discretion is not correlated with investigator stringency, we estimate a regression of our measure of investigator stringency (described below) on the characteristics of the alleged perpetrators listed in the investigation. These characteristics include the number of alleged perpetrators, as well as measures of their age, gender, and racial makeup. Conditional on zip code by investigation year fixed effects, these variables are not jointly predictive of investigator stringency ($p = 0.825$).

we use a dataset containing the universe of arrests in Michigan from January 2012–May 2020. For individuals who are arrested at age 17 or older (the age at which Michiganders are considered to be adults by the justice system during our sample period), these data include the date of the arrest, whether the arrest was for a misdemeanor or felony offense, and the type of crime, such as violent, property, or drug crime. We use a similar dataset that contains judicial information to measure whether an individual was convicted or incarcerated. To define outcomes consistently, our main analysis focuses on whether a person met each adult crime outcome before they turned 19 years old (for example, whether a person was ever arrested by age 19). In additional analyses, we examine outcomes through age 21 for the subset of children that we can observe at older ages, as well as the crime outcomes of the birth parents involved in each child welfare investigation.

K-12 education data come from the Michigan Department of Education and the Center for Educational Performance and Information. These data cover the universe of K-12 public school students in Michigan, including charter school students, between the 2002–2003 and 2020–2021 academic years. The K-12 education data include rich student-level demographic and socio-economic information, such as gender, race and ethnicity, and free or reduced-price lunch eligibility, which we use as covariates. We also use the K-12 education data to explore intermediate outcomes, including standardized test scores, school attendance, and whether a student enrolled in an educational program at one of Michigan's 23 juvenile detention centers. Enrollment in a juvenile detention center is a behavioral outcome that indicates youth contact with the juvenile justice system; youth younger than 17 years old may be held in a detention center after being arrested. We focus on enrollment in a juvenile detention center instead of other behavioral outcomes commonly available in administrative education data, such as school suspensions or expulsions, because school discipline is not reported consistently in the Michigan data. The K-12 data also include the census block where students live, which we link to publicly available census block group characteristics from the Census Bureau.

Postsecondary education data come from the National Student Clearinghouse, which covers enrollment at most two- and four-year colleges in the United States. Specifically, it includes enrollment information from 97% of degree-granting institutions in the United States (Dynarski, Hemelt and Hyman, 2015). We use these data to examine whether an individual ever attended any college.

The Michigan Education Data Center (MEDC) linked children from the child welfare, adult crime, and education administrative data using a probabilistic matching algorithm. These data sources do not contain a common identifier so MEDC staff linked the data based on first name, last name, date of birth, and gender using the Fellegi-Sunter model implemented via the *fastLink* R package (Enamorado, Fifield and Imai, 2019). Because MEDC manages the K-12 education data, the K-12 public school students serve as the base population. Staff linked the K-12 education data with the child welfare data and then matched the K-12 education data to the adult crime data. Both linkages performed well. For each of the matched records, the software rates the certainty level of the match using a posterior probability. Overall, 87.6% of records in the child welfare data matched to a public school student record and 92.4% of records in the adult crime data matched to a public school record with a high degree of certainty (over 99.6%).[6] We followed this same process to separately link birth parents involved in child welfare investigations to the adult crime data and found that 15% of birth parents matched to a record in the adult crime data with a high degree of certainty.

II.B Overview of Analysis Sample

Using the administrative data sources, we construct an analysis sample unique at the child by investigation level of Michigan children subject to a child welfare investigation between 2008 and 2016. Because the child welfare and adult crime data are linked separately via the K-12 education records, we restrict the sample to children who ever enrolled in a public

[6]This match rate is nearly identical for males and females, and MEDC closely validated the match by manually matching a randomly selected subset of 200 records. Furthermore, this rate is quite high, since some individuals investigated or arrested in Michigan could have gone to school in a different state, been enrolled in a private school, or been homeschooled.

school in Michigan. We also restrict the sample to children who were 16 years old or younger at the time of their investigation because 17-year-olds can be arrested as adults in Michigan. We further restrict the sample based on the years of available child welfare and adult crime data. For a child to be in our analysis sample, we must observe both: (1) who investigated the case and (2) their adult crime outcomes by age 19. Because the child welfare data first record investigator identifiers in 2008 and the adult crime data end in May 2020, we restrict our sample to children investigated at ages 6 and older. Children younger than 6 who were investigated in 2008 or later would not have turned 19 years old before May 2020. Lastly, we exclude cases from the analysis sample where investigators were unlikely to have been quasi-randomly assigned: allegations of sexual abuse and those involving children who had been the subject of an investigation in the year before the focal report. Overall, we focus on 118,273 investigations of 87,100 children.[7]

Table 1 compares the characteristics of all public school students in Michigan during the 2016-2017 school year (Column 1) to the children in our analysis sample—those subject to a child welfare investigation (Column 2). About half of the students in both groups are female and their average age is 12 years old. There are notable differences in terms of race and ethnicity and socio-economic status, however, as Black children and children from families with low income are disproportionately involved in the child welfare system. 81% of investigations involved children with low income, despite these children making up just half of the overall population.

About 2% of investigations in the analysis sample led to foster care placement (Column 3). Compared to all children subject to an investigation, those placed in foster care are slightly less likely to be female (46% compared to 50%) and considerably more likely to be Black (43% compared to 29%). Consistent with other research highlighting the association between foster care and crime, we find that children who are placed are disproportionately

[7]A potential concern is that structuring the data at the child by investigation level creates a violation of the Stable Unit Treatment Value Assumption if the outcome of one investigation influences the outcome of investigations in the future. We show below that the findings are robust to focusing only on a child's first investigation in the sample (Table 5).

likely to be arrested, convicted, and incarcerated. For example, children who are placed are 50% more likely to be arrested by age 19 than those subject to an investigation who are not placed (21% compared to 14%). Altogether, these descriptive statistics indicate that children placed in foster care differ in systematic ways from children who are investigated but not removed.

III Empirical Strategy

A naïve analysis of foster care might regress a measure of children's later-in-life criminal activity, such as an indicator for ever convicted as an adult, on a binary treatment variable equal to one if the child's investigation resulted in foster placement. Even with covariates to account for a wide range of observable characteristics, estimates of foster care from such a regression likely contain bias because foster children differ along unobservable dimensions from investigated children who are not removed. For example, children who are placed may be more severely abused or neglected than those who are not. Such unobserved features would bias OLS estimates to understate the benefits of foster care and overstate its costs.

III.A Research Design

To address omitted variable bias, we implement the examiner assignment research design, which has been used in other studies of foster care (Bald et al., 2022*a*; Doyle, 2007, 2008; Gross and Baron, 2022; Helénsdotter, 2022) as well as research on incarceration and prosecution (Agan, Doleac and Harvey, 2021; Aizer and Doyle, 2015; Bhuller et al., 2018, 2020; Garin et al., 2023; Kling, 2006; Mueller-Smith, 2015; Norris, Pecenco and Weaver, 2021), pretrial detention (Dobbie, Goldin and Yang, 2018; Leslie and Pope, 2017), disability insurance (Dahl, Kostøl and Mogstad, 2014), speeding fines (Goncalves and Mello, 2022), and evictions (Collinson et al., 2021), among others. Specifically, we instrument for placement using the removal tendencies of quasi-randomly assigned investigators. Children assigned by chance to particularly strict investigators—those with high propensities to remove—are more likely to be placed than they would have been had they been assigned to a more lenient investigator. To reliably measure investigator removal tendencies, we focus on children assigned to investigators who we observe work at least 50 cases in the child welfare data, inclusive of quasi-randomly assigned cases outside of the analysis sample. Overall, our analysis sample includes 3,011 investigators assigned to 235 cases each, on average.

Our main specifications take the following form:

$$FC_{iw} = \gamma_1 Z^R_{iw} + \gamma_2 X_{iw} + \Theta_r + \eta_{iw} \tag{1}$$

$$Y_{iw} = \beta_1 \hat{FC}_{iw} + \beta_2 X_{iw} + \theta_r + \epsilon_{iw} \tag{2}$$

where Y_{iw} is an outcome for child i assigned to investigator w (such as an indicator for ever arrested by age 19); Z^R_{iw} is the removal tendency instrument; X_{iw} is a vector of baseline covariates included in both the first and second stages.[8] Θ_r and θ_r represent child zip code by investigation year fixed effects to control for the level of investigator rotational assignment, which ensures that we only compare children who could have been assigned to the same investigator. There are 6,569 unique rotation groups, and the median rotation group consists of 15 investigators. We cluster standard errors at the child level to account for the mechanical correlation in outcomes that arises by including the same child more than once in the dataset.

As is common in the examiner assignment literature, we calculate the instrument as the fraction of all other investigations, both past and future (including quasi-randomly assigned cases outside of the analysis sample), that are assigned to the same investigator and result in foster care placement. Following Norris, Pecenco and Weaver (2021), we implement the unbiased jackknife instrumental variables estimator (UJIVE) approach in Kolesar (2013), which uses a leave-out approach to estimate Z^R_{iw}, conditional on the control

[8] See Table 2 for the full set of baseline covariates.

variables included in Equations 1 and 2. This approach ensures that covariates are handled consistently throughout the estimation.

We find that there is variation in investigator tendencies. A standard leave-out measure of removal tendencies in our data has a mean of 0.033 and a standard deviation of 0.025. This number is significantly smaller than the mean of 16% in Doyle (2008) and we discuss the extent to which this could explain the differences in findings across studies below. This average is also smaller than that in Bald et al. (2022*a*) (18%) and Roberts (2019) (17%). This is to be expected because these studies focus exclusively on substantiated investigations, and therefore the denominator is smaller by construction (since removal is conditional on substantiation and only about a quarter of all investigations are substantiated).

There is also variation in removal tendencies among investigators who work in the same local office. Figure 1 shows the distribution of the instrument, which varies from -0.04 to 0.07 after netting out child zip code by investigation year fixed effects. An investigator at the 10^{th} percentile removes at a rate 2.1 percentage points less than the average investigator in their local area, whereas an investigator at the 90^{th} percentile removes at a rate 2.3 percentage points greater. Relative to the average removal rate of 3%, this indicates that moving from the 10^{th} to the 90^{th} percentile represents an almost 150% increase in the likelihood of placement.

Under standard instrumental variables assumptions, which we discuss in the next section, β_1 represents the impact of foster care placement on outcomes for children at the margin (compliers). Compliers in this setting are children for whom investigators might disagree about removal. Because most policy debates surrounding foster care focus precisely on the complier population, the local average treatment effect (LATE) we identify in this study is for a particularly relevant population in child welfare policy.

III.B Identification Assumptions

Four assumptions must be satisfied to interpret our estimates as the causal effects of foster care for children at the margin of placement: relevance, exogeneity, monotonicity, and exclusion.

Relevance. This assumption requires that investigator removal stringency predicts foster care placement ($\gamma_1 \neq 0$). Figure 1 visually depicts the strong, positive relationship between investigator removal stringency and placement, and Table A1 reports the first-stage regression of foster placement on the removal stringency instrument. A one standard deviation (2.5 percentage point) increase in removal stringency increases the likelihood of placement by roughly one percentage point (Column 4), and the F-statistic of 225 indicates that the instrument is strong.

Exogeneity. This assumption requires that the unobserved determinants of children's later-in-life outcomes are independent of investigator removal stringency ($\text{Cov}[Z^R, \epsilon] = 0$). We test an implication of exogeneity: that observable child and case characteristics are uncorrelated with the removal tendencies of the assigned investigator. As expected due to the rotational assignment of child welfare investigators, a rich set of characteristics are not jointly predictive of the instrument despite being highly predictive of placement itself (Table A2). As further evidence of exogeneity, the first stage F-statistic in Table A1 is stable with the inclusion of covariates.

Monotonicity. Interpreting our estimates as a proper weighted average of complier treatment effects requires an average monotonicity assumption (Frandsen, Lefgren and Leslie, 2023). For average monotonicity to hold, the covariance between each child's investigator-specific treatment status and investigator stringency must be weakly positive. It follows from average monotonicity that removal stringency and placement should be positively correlated for all child subgroups. Table A3 shows that the first stage is positive and statistically significant across various subgroups of child and investigation characteristics.

Previous studies implementing examiner assignment research designs have made the stronger assumption of pairwise monotonicity, yet recent advances indicate that the pairwise monotonicity assumption is not necessary to estimate LATEs (Frandsen, Lefgren and Leslie, 2023). Unlike average monotonicity, pairwise monotonicity requires that children who are removed by a particularly lenient investigator must also have been removed by a stricter investigator, and vice versa.[9]

Exclusion. Our analysis requires an exclusion restriction in order for the estimates to be interpreted as LATEs. We discuss the exclusion restriction in detail in Section VI.

IV The Effects of Foster Care on Adult Crime

This section presents new evidence on the causal effects of foster care on adult crime. We first report our main findings on adult arrests, convictions, and incarcerations, and describe the complier population for whom these findings apply. We then explore sources of heterogeneity via subgroup analyses, as well as impacts on other indicators of child well-being.

IV.A Main Findings

A naïve OLS analysis suggests that placement increases later-in-life criminality (Panel A, Table 2). For example, we find that foster care is associated with a 4 percentage point increase in the likelihood of being arrested by age 19. This represents a 28% increase, relative to a control mean of 14.2%. We see similarly large increases in the likelihood of being convicted and incarcerated by age 19 (Columns 2 and 3). These results show that even controlling for detailed socio-demographic, school, and neighborhood characteristics, there is a strong positive association between foster care and adult crime. However, because of unobservable differences between children who are and are not placed (e.g., the severity of the

[9] We implement Frandsen, Lefgren and Leslie (2023)'s test for the null hypothesis that the pairwise monotonicity assumption and exclusion restriction jointly hold in Table A4. We reject the null hypothesis, which suggests that either the pairwise monotonicity assumption does not hold, the exclusion restriction is violated, or both. As described in detail in Section VI, we find evidence that the exclusion restriction may hold in our setting. Thus, we rely on the weaker assumption of average monotonicity throughout.

maltreatment), the OLS estimates likely conflate the impacts of placement with unobserved factors.

In contrast, UJIVE estimates show that placement substantially decreases future crime for children at the margin (Panel B, Table 2). We find that placement decreases the likelihood of being arrested by age 19 by 25 percentage points. Relative to a control complier mean of 37%, this suggests placement reduces the likelihood of an adult arrest by 68%. We find an even larger reduction in the likelihood of being convicted by age 19 of 28 percentage points, or 81% compared to a control complier mean of 35%. The estimated impacts on incarceration by age 19 are similarly large, equal to a 21 percentage point (or 80%) reduction. The point estimates are statistically significant and economically meaningful. For example, they are larger in magnitude than the estimated reduction in crime from cognitive behavioral therapy in the Becoming a Man program, which directly aims to reduce crime among vulnerable youth, and reduced violent-crime arrests by 45–50% (Heller et al., 2017). The estimates are substantively large, which is consistent with the large magnitudes found in Doyle (2008) for the effects of foster care, albeit in the opposite direction.

Figure A2 shows a reduced-form version of our approach, where we plot a child's likelihood of adult criminal justice contact against the removal stringency of the investigator assigned to their case. For this analysis and the remainder of the paper, we study convictions as the adult crime outcome to focus on incidents with a higher likelihood that a crime occurred and for ease of exposition. The figure shows a clear negative relationship between a child's probability of an adult conviction and their investigator's removal stringency. In other words, children assigned to particularly strict investigators had a lower chance of being convicted in adulthood.

Next, we conduct a more detailed analysis of the impacts of foster care placement on criminal justice contact. We begin by examining heterogeneity in the effects of placement by crime type (violent, property, drug, and public order) and severity (felony versus misdemeanor). Table A5 shows that placement reduces the likelihood of a conviction for a violent crime by 14 percentage points, or 78%. This estimate is statistically significant at the 5% level. Although we observe negative point estimates for the other three crime types, including a 70% reduction in the likelihood of being convicted for a property crime, these estimates are less precise. Furthermore, foster care placement leads to a 26 percentage point (86%) reduction in the probability of being convicted of a misdemeanor offense. While there is also an economically significant decline of 63% in the likelihood of a felony conviction, this estimate is imprecise. Finally, when examining the effects of foster care placement on the *number* of convictions, we find that foster care leads to a decline of nearly 0.4 convictions (81%) by age 19.

Altogether, the results in this section demonstrate that foster care led to a decline in adult crime for children at the margin of placement. In Online Appendix C, we discuss potential reasons for why these findings may differ from those in Doyle (2008). In particular, we present evidence that national changes in child welfare policy and practice over the 30 years since the beginning of the Doyle (2008) sample likely explain the differences in findings across the two studies. For instance, the proportion of children referred as possible victims of maltreatment who are placed in foster care has dropped by 50% over the last twenty years in the United States (AECF, 2017; USDHHS, 2000, 2019). Consequently, children on the margin of placement during the earlier study may have faced less risk at home compared to marginal children today, suggesting that more recent studies might find less harmful effects of placement. To test this hypothesis, we estimate Marginal Treatment Effects (MTEs) and show that foster care placement reduces adult crime the most for children on the margin of removal for more lenient investigators. The benefits of foster care placement sharply decline for children at the margin of placement for relatively more strict investigators. These estimates show that the decline in removal rates since the Doyle (2008) sample period could have contributed to more positive estimates of the effects of foster care in the current period.

IV.B Characteristics of Marginal Children

The UJIVE estimates show that foster care placement reduces later-in-life crime for compliers: children at the margin of placement, or those for whom investigators might disagree over whether removal from the home is appropriate.

Table 3 describes the baseline characteristics of compliers. Compared with the full sample, compliers are more likely to be Black and from families with low income. For example, nearly 40% of compliers are Black compared to 29% of all children subject to an investigation in our sample. Compliers are also more likely to live in an urban county and a low income neighborhood. The reason for the child welfare investigation also differs between compliers and the overall sample. Compliers are more likely to be investigated for neglect (67% compared to 52%) and parental substance abuse (20% compared to 16%), and less likely to be investigated for physical abuse (13% compared to 31%). This could reflect that investigators may wield greater discretion in cases involving parental substance abuse and neglect, and less discretion in physical abuse cases. The last four rows of the table present characteristics of the birth parents listed as the alleged perpetrators in the investigations in our analysis sample. Investigations involving compliers are significantly less likely to involve a female birth parent, and are more likely to involve a Black and relatively younger birth parent (under 38 years old, the average age of birth parents in our analysis sample).[10] Additionally, the average investigation in the sample lists 1.4 alleged perpetrators, compared to 1.2 for compliers.

After removal, most compliers are initially placed in a family home and reunify with their birth parents after about 1.5 years. Table 4 compares the average experience of all children placed in foster care to the experiences of treated compliers. On average, 56% of foster children are initially placed with relatives, 30% with an unrelated family, and 14% in a group home. Compliers are equally likely to be placed with relatives (55%), slightly more likely to be placed with an unrelated family (33%), and less likely to be placed in a group

[10] Recall that the average age of children in our sample is 12 years old.

home (12%). About 40% of all foster children in our sample live in just one or two different placements while in foster care and 60% live in three or more. Placements for compliers are more stable; over half (52%) of compliers live in one or two placements. Compliers also had relatively shorter stays in foster care, spending about three fewer months in the system (18 months compared to 21 months). Conditional on exiting, nearly four in five compliers reunify with their birth parents (83%). Fewer compliers are adopted (7%), have legal guardianship transferred (6%), or are emancipated as legal adults (4%). These permanency outcomes are similar for the average foster child.

Our research design cannot identify the causal effects of certain placement experiences on adult crime, such as placement in family homes or more stable placements. However, the experiences of compliers suggest that foster care reduces crime in a setting where most are initially placed in a family home, have stable placements, are in foster care for a relatively short period of time, and ultimately reunify with their birth parents. This is consistent with a large body of research showing that these experiences correlate with improved outcomes (Rubin et al., 2007, 2004; Ryan and Testa, 2005; Ryan et al., 2008).[11]

IV.C Heterogeneity by Child and Investigation Characteristics

We test whether the effects of foster care on adult criminality differ by child and investigation characteristics (Table 5). Studies show that male children are often more vulnerable to disruptions or disadvantages than female children (Autor et al., 2019; Kling, Ludwig and Katz, 2005). We find that the reduction in convictions is larger for male children. The impacts of foster care on adult crime for female children are economically meaningful (roughly 70%), but statistically insignificant. This could be due to much lower baseline conviction rates for females. Previous research also shows that younger children benefit more from changes to their environment than older children (Chetty, Hendren and Katz, 2016; Chyn, 2018). Consistent with this research, our results suggest that the reduction

[11] For example, Font and Mills (2022) find evidence of negative peer effects in congregate care settings.

in crime is driven by children ages 6 to 11 at the time of the investigation. Yet, it is also possible that these effects are due to the effectiveness of foster care placement during the child's first investigation in the sample, which seems to be more impactful than placements during subsequent investigations.

There is also policy interest in understanding whether the causal effects of placement vary by child race and ethnicity (Barth et al., 2020). Our results suggest that foster care reduces crime for White, Black, and Hispanic children. The estimates are similar in magnitude for White and Black children and are more negative, yet less precise for Hispanic children given the smaller sample size. However, in percent terms, the effects are significantly larger for White children. Finally, we find that—while foster care placement leads to economically meaningful reductions in adult crime across all main types of investigations in our sample (neglect, physical abuse, and substance abuse)—the estimates are only statistically significant for investigations involving neglect, likely due to the larger sample size in this category.

IV.D Other Indicators of Child Well-being

In addition to later-in-life crime, we examined the effects of foster care on a range of other indicators of child well-being, including academic, behavioral, and safety outcomes. We replicated the analyses on school absences, test scores, and subsequent maltreatment from Gross and Baron (2022) for the sample in the current study. We also added new analyses on high school graduation and college enrollment, which were unavailable for most children in Gross and Baron (2022), and a more comprehensive measure of juvenile delinquency.

Across all of these other indicators, we found that foster care improves well-being. Compared with the close to half of complier children who are not placed that are chronically absent from school in the years after the investigation (meaning they are absent for over 10% of school days in a given year), foster care reduces chronic absenteeism by 21 percentage points (Table 6). We also see gains in standardized math test scores of about 0.4 standard deviations. These improvements appear to translate to increased educational attainment, as we find large, yet somewhat imprecise, positive impacts on high school graduation and college enrollment. We also find that placement reduces the likelihood of being held in a juvenile detention center by 17 percentage points or 80%.[12] Lastly, we find that placement improves children's safety. It decreases the chances that children are the subject of a subsequent maltreatment investigation by 22 percentage points, or 58% relative to a control complier mean of 36%, and reduces the likelihood that children are confirmed as victims in a subsequent investigation by 8 percentage points (67%).

Interestingly, the effects of foster care on these other indicators of well-being explain a modest share of the effect on later-in-life crime. A rough estimate of the implied effects on adult convictions can be obtained by pairing estimates of the causal effects of foster care on these intermediate outcomes with the non-experimental relationship between each indicator and the probability of an adult conviction in our data. Table A7 shows that, under very conservative assumptions, effects operating through intermediate outcomes alone yield an implied effect on adult convictions of -12 percentage points, or 40% of the estimate documented in Table 2.

V Mechanisms

The results presented so far show that foster care placement leads to large reductions in adult crime. This section explores potential mechanisms driving these effects. We begin by presenting new evidence that a key reason for the improvements in children's outcomes is that birth parents make positive changes in their own lives while their children are temporarily in foster care. We then show that there is less evidence for possible alternative

[12] Juvenile detention centers are secure facilities generally used to hold youth after an arrest but before a court hearing and the median youth is held for about two weeks (Baron, Jacob and Ryan, 2022). This measure is more comprehensive than the juvenile delinquency outcome in Gross and Baron (2022). The earlier study used juvenile court petitions, which were not available for 8 counties in Michigan, including the counties that include 3 of the 6 most populated cities in Michigan: Grand Rapids, Lansing, and Ann Arbor. Because the data for that outcome were incomplete, the study could not rule out positive or negative effects on juvenile crime. Juvenile detention center information is available statewide.

mechanisms, including a short-lived reduction in criminality that ends after youth complete their schooling, neighborhood effects, and the removal of a parent.

V.A Evidence That Birth Parents Made Improvements

Gross and Baron (2022) suggested that improvements made by children's birth parents were a likely mechanism through which placement influenced children's outcomes. The study showed that gains in children's safety and academic outcomes improved only after most children reunified with their birth parents. However, it lacked data on adult outcomes to directly test this mechanism. In this section, we use the administrative adult crime data to examine how having a child removed influences the criminality of birth parents up to 12 years after the focal child welfare investigation.

We find that for cases on the margin, having a child placed in foster care substantially decreases the probability that birth parents have subsequent contact with the criminal justice system (Table 7). Placement reduced the likelihood of birth parents being arrested in the years after the investigation by 8.3 percentage points, or 85% relative to a control complier mean of 9.8%. We see a similar decrease in the likelihood of being convicted after the investigation. Both of these effects are statistically significant and, interestingly, are similar in magnitude to the observed reduction in children's later-in-life arrests and convictions.[13] In Table A9, we examine heterogeneity in these effects across various subgroups of birth parent, child, and investigation characteristics. While somewhat imprecise, the estimates suggest that the effects of foster care on the reduction in the likelihood that birth parents are subsequently convicted are larger for relatively younger birth parents and investigations involving parental substance abuse.[14]

These findings offer direct evidence that having a child placed in foster care causes birth

[13] These findings contrast with Grimon (2020), who finds no effects of child removal on jail or court involvement for mothers in Allegheny County.

[14] When examining effects by crime type, we find negative point estimates for most types, driven by large and statistically significant reductions in violent and public order offenses. Additionally, we find economically significant declines in the likelihood of convictions for both felony and misdemeanor offenses, though the point estimate for felonies is more imprecise (Table A8).

parents to make positive changes in their own lives. Criminal justice contact is only a proxy for the many ways in which child removal may influence birth parents. Birth parents may also make changes in ways that we cannot observe in the data available for this study; for example, they may gain employment in a better job or enter a healthier romantic relationship. Reduced criminal justice contact is indicative of any number of changes in birth parents' lives that could lead to family reunification and positively impact children's subsequent life outcomes. These findings are also broadly consistent with Finlay, Mueller-Smith and Street (2022), who show that intergenerational exposure to the criminal justice system is associated with many indicators of children's adult outcomes, including increases in criminal justice contact.

Dynamics in the intermediate outcomes of children are also consistent with foster care having positive effects on birth parents. Close to half of complier children exit foster care within one year and nearly all exit within two years (Panel A of Figure 2); more than 80% of compliers reunify with their parents when they exit (Table 4). Children's intermediate outcomes, such as subsequent reports of maltreatment, chronic absenteeism, and math test scores, tend to improve only after children exit foster care, when most have reunified. For example, Panel B of Figure 2 shows that the effects of placement on the likelihood of being the subject in a subsequent investigation are smaller in the first few years after the initial investigation and are larger in later years. We see similar trends for absenteeism and test scores (Panels C and D). That these outcomes appear to improve only after most foster children reunify offers further evidence that removal had positive effects on parents.

There are specific features of the child welfare system which support the idea that foster care leads birth parents to make positive changes in their lives. First, a judge must confirm that it is safe for children to return home before they can be reunified with their birth parents. Consistent with this process, evidence of declines in the likelihood that birth parents are arrested begins to appear during the year following removal, especially around the time of reunification in the second year after the investigation (Panel E of Figure 2). The reduction in arrests persists years later. Second, after their children are removed, birth parents work closely with social workers to address challenges in their own lives, such as confronting drug addiction and other substance abuse, finding stable employment, securing housing, or strengthening parenting skills. Birth parents receive fully funded services to help with these challenges, such as substance abuse treatment, parenting classes, or counseling, and have strong incentives to engage with these services in order to regain custody of their children. These findings are also consistent with recent studies showing that "turning points" for parents such as childbirth can substantially reduce their criminal justice contact (Dustmann and Landersø, 2021; Eichmeyer and Kent, 2022; Massenkoff and Rose, 2022).[15]

V.B Examining Alternative Explanations

We find less evidence for three alternative mechanisms. First, as foster care appears to increase educational attainment, placement could reduce crime by age 19 only because youth spend more time in high school and mechanically have less time to commit crimes. Such an "incapacitation effect" would imply a short-lived reduction in criminality that ends after youth complete their schooling (Bell, Costa and Machin, 2021). We estimate the impacts of foster care using the subset of youth for whom we observe crime at older ages and find sustained crime reductions through age 21, a time when youth would have either graduated or dropped out of high school (Table A10).

Second, it is possible that, because children by definition move when they are placed in foster care, they move to more advantaged neighborhoods which could directly improve their long-term outcomes. In Gross and Baron (2022), we showed that children experienced modest improvements in their neighborhood and school environment while temporarily in foster care (during the first year after the investigation). However, two years after the investigation (when most compliers had exited the system and reunified with their birth parents), there were no discernible differences in the neighborhood and school environments

[15] As further evidence of parental improvement, birth parents whose initial child victim entered foster care were less likely to be alleged perpetrators in subsequent child maltreatment investigations, even years after the initial investigation (Panel F of Figure 2).

of treated and control compliers.

While it is possible that a one-year exposure to a more advantaged neighborhood and school could generate the large declines in adult crime documented in this study, there are at least three reasons why we believe this is not the primary mechanism. First, studies of mobility often find much smaller and even mixed results on later-in-life crime. For instance, Chyn (2018) finds a 14% decline in arrests for violent crimes, but an 18% increase in arrests for property crimes. Similarly, Kling, Ludwig and Katz (2005) find that males relocating to lower poverty neighborhoods in Moving to Opportunity experienced only a short-term reduction in violent arrests, as well as increases in arrests for property crimes. We find much larger and persistent declines in violent offenses, as well as suggestive declines in property crimes (Table A5). Second, studies of mobility often find that such effects tend to increase with dosage (Chetty and Hendren, 2018; Chetty, Hendren and Katz, 2016; Chyn, 2018), whereas exposure in our context was only temporary. Third, long-run benefits of moving to more advantaged neighborhoods do not appear to run through schooling channels (Chyn, 2018; Jacob, 2004), and foster care had large impacts on educational outcomes.

A third alternative explanation is that foster care caused the removal of a parent (e.g., through parental incarceration), which could influence children's outcomes (Arteaga, 2021; Billings, 2019; Norris, Pecenco and Weaver, 2021). We find no evidence that foster care increased the likelihood of parental incarceration (Column 3, Table 7).

VI Threats and Robustness Checks

VI.A Exclusion Restriction

The exclusion restriction requires that the removal stringency instrument only influences children's later-in-life outcomes through foster care placement. This section explores two potential concerns about exclusion in our context.

Potential Influence on Children's Experiences in Foster Care

Exclusion would be violated if investigators influence children's experiences in foster care, yet there is institutional and empirical support that this is not a first-order concern in our setting. As an example, suppose that children who happened to be assigned to stricter investigators were both more likely to be placed in foster care and to have particular experiences while in the system, such as short and stable stays. If this were the case, then it could be children's experiences in foster care that influence their outcomes, and not foster care placement alone. This is similar to concerns in studies that apply the examiner assignment design to estimate the effects of incarceration, in that the judge who decides whether to incarcerate also plays a role in the sentencing decision.

Unlike the criminal justice setting, however, there is little institutional reason to suspect that investigators would influence children's experiences in foster care. The original quasi-randomly assigned investigators do not work with the child or their family after the investigation is completed, and they are not involved in decision-making at that point. If a child is placed in foster care, their case is transferred to a different agency staff member who works in a separate "foster care" department. These workers are often also assigned to cases according to a rotation, but this rotation is distinct from the one for the initial investigators as they include different types of workers. Consequently, investigators are not involved in decisions related to children's foster care experiences: where they are placed, how long they remain in foster care, or the stability of their placements.

Several empirical tests support the notion that investigators do not influence children's experiences in foster care in our setting. First, conditional on being in foster care, the instrument is uncorrelated with the number of days in foster care, the number of different settings while in the system, and the initial placement type (Table A11). For this test, we set the indicators of foster care experiences to missing for children who were not placed, which may yield biased estimates. Thus, in a complementary analysis, we use the full sample and estimate our main specification, but holding constant variation in these indicators across investigators. We calculate leave-out means using UJIVE for each indicator of foster care experiences; e.g., for each investigator, we calculate the average length in foster care in all other cases assigned to the investigator, where length is equal to zero for cases that did not result in foster care placement. We do the same for four other indicators of foster care experiences: the number of different foster placements, and indicators for initial placement with relatives, an unrelated family, or in a group home. When controlling for these five measures, we find point estimates that are statistically indistinguishable from our baseline estimates and nearly identical in percent terms (Table A12).

Discretion During the Investigation Other Than for Placement

Having established that investigators do not influence children's experiences in foster care, we next turn to areas of the investigation other than placement that investigators have discretion over. Investigators make decisions around prevention services at the end of the investigation that could influence children's outcomes. As shown in Figure A1, investigators place families on one of four tracks based on the strength of evidence that maltreatment occurred and the child's risk of future harm: (1) no services (78% of all cases), (2) community-based services (12%), (3) both community-based and targeted services (8%), and (4) child removal plus community-based and targeted services (2%).

The exclusion restriction would be violated if investigators who are more likely to remove children are also more likely to recommend prevention services, and these services impact children's outcomes (Mueller-Smith, 2015). However, we argue that there is little empirical or institutional reason to expect that prevention services *without removal* would have meaningful impacts on children's outcomes. That is, the potential outcomes under treatments (1), (2), and (3) are likely to be approximately equal to each other. First, community-based services are light-touch and could include simply leaving a pamphlet that provides information on local food pantries, support groups, or other community organizations. These services require no follow-up by the child welfare agency. Thus, it would be quite surprising if these "information-only" treatments had large effects on children.

Targeted services may include substance abuse treatment, parenting classes, and counseling, which are more intensive. However, studies show that these types of programs tend to have small, if any, impacts on the outcomes of adult participants, let alone large intergenerational effects on children's adult outcomes (Barnow and Smith, 2015; Wood et al., 2014). One potential reason for these findings is that, while birth parents have a strong incentive to engage with post-removal services to regain custody of their children, parents are not typically compelled by the courts to take up prevention services without removal. Consequently, engaging these families tends to be difficult. While our dataset does not include information on the take-up of services, Baron et al. (2024) review data from the Michigan Department of Health and Human Services on the take-up rates of SafeCare, a targeted service commonly offered to families and a program that the State of Michigan is currently evaluating. They find that take-up rates for this program range from 5% to 10%, while completion rates are significantly lower (on the order of 1% to 2%). Because take-up and completion rates are extremely low conditional on the service being offered, it is unlikely that the typical targeted services to which families are referred substantially influence children's outcomes.[16] Taken together, the results in this section suggest that the exclusion restriction is plausible in our context.

VI.B Robustness Checks

Table A13 presents a variety of alternative specifications that probe the robustness of our main estimates. To assess whether certain sample restrictions influence the results, we conduct the analysis using three other samples: (1) excluding children assigned to investigators who worked fewer than 75 cases (the main analysis uses a threshold of 50);

[16]Indeed, even among the highly selected sample of child welfare prevention programs that states submit to the Title IV-E Prevention Services Clearinghouse precisely because they are thought to have some level of evidence supporting their efficacy, only about half were determined to show some evidence of impacts, and only 14% received the highest evidence rating. Those that have been determined to show evidence were rarely used in Michigan during our sample period (MDHHS, 2021).

(2) excluding investigators with a removal stringency above the 90th percentile in our sample—since Figure 1 shows the first-stage relationship is particularly strong for these investigators; and (3) dropping the 9% of investigations in our sample in which the alleged perpetrator is not the adult with legal custody of the child (e.g., an uncle or cousin)—since we showed that improvements made by birth parents may explain the impacts on their children's outcomes.

To assess robustness to other reasonable measures of investigator removal tendencies, we recreate the instrument in five ways: (1) randomly splitting the sample in half and defining the instrument as the investigator's removal rate from the other half of the sample, (2) allowing tendencies to vary over time by creating a leave-out-other-years measure, (3) constructing a leave-out-same-year measure to account for removal decisions occurring around the same time potentially being correlated, (4) constructing the instrument using only cases in our analysis sample, and (5) constructing the instrument without using the UJIVE approach of Kolesar (2013), which simplifies to using a two-stage least squares estimator. We also examine sensitivity to the definition of investigator rotational assignment by including fixed effects for the county by investigation year instead of zip code by investigation year. Because some of the local offices in Michigan divide investigators into teams based on small regions, the main analysis includes zip code by investigation year fixed effects. However, a small share of zip codes in Michigan span more than one county, which could generate measurement error in our main analysis. Lastly, even though the F-statistics in Table A1 are large, we follow the recommendation in Andrews, Stock and Sun (2019) and construct Anderson-Rubin confidence intervals, which are robust to weak identification and are efficient in the just-identified case. Across all of these robustness checks, the estimates of foster care on our main three outcomes remain negative and large in magnitude.

We also assess robustness to the specific covariates included in the UJIVE regressions and to the level of clustering standard errors. Our main analysis includes a variety of covariates in addition to the rotation group fixed effects and a potential concern is that the main findings are unique to that particular specification. However, the estimated impact of placement on later-in-life crime is nearly identical even if we do not include additional covariates beyond zip code by investigation year fixed effects (Table A14).

In addition, although our main specification clusters standard errors at the child level to account for the correlation in outcomes that arises mechanically by including the same child more than once in the panel, we assess sensitivity to alternative levels of clustering. We find that the results are robust to the following alternative levels of clustering standard errors: investigator level, zip code-by-year level, two-way clustering at the child and investigator level, and two-way clustering at the child and zip code-year level (Table A15).

Finally, one may be concerned that foster placement causes children to leave Michigan and commit crimes in other places, which we do not observe in the Michigan adult crime data. That is, the observed reduction in crime could be driven by increased out-of-state migration rather than an actual decrease in crime. We explore the extent to which out-of-state migration may influence our findings in Table A16. We find that foster care does not impact the probability that children leave the state during grades K–12 (Column 1) or for college (Column 2). We also estimate our main specification while excluding children who left Michigan in K–12 (Columns 3, 5, and 7) and children who ever attended college outside of Michigan (Columns 4, 6, and 8). Estimates from these restricted samples are very similar to our baseline estimates in Table 2.

VII Marginal Value of Public Funds

The MVPF is a benefit-cost framework that produces a common metric for the relative effectiveness of spending on different programs. It compares the benefits that a policy provides to society, or society's willingness to pay, to the net cost to the government of implementing the policy (Hendren and Sprung-Keyser, 2020). In this section, we calculate the first estimate of the MVPF for a policy that promotes children's safety.

The first step to calculate the MVPF is to estimate society's willingness to pay for foster care. To do so, we measure the social benefit as the reduction in social costs from foster care's effects on children's later-in-life crime. We pair our detailed criminal justice records which show which types of crimes (if any) children committed with social cost estimates for each type in Chalfin (2015). We construct a variable equal to each child's social cost of adult crime, equal to zero for children who were never convicted by age 19 and equal to the sum of the cost of each conviction for children who were convicted. For example, if a child was convicted twice, first for homicide and then for carjacking, then we define the child's social cost as the sum of the social costs of homicide and carjacking. Because foster care placement occurs years before an adult conviction, we discount the social cost of each crime using a 3 to 5% rate (Anders, Barr and Smith, 2022). Using this variable as the outcome in our main specification, we show that the reduction in social costs ranges from about $84,000 to $95,00 depending on the discount rate (Panel A of Table 8).

The next step is to calculate the net cost of foster care to the government which includes both the direct costs of each placement as well as the cost savings from less criminal activity (for example, savings from fewer incarcerations). This improves on typical cost-benefit analyses which do not consider long-run government savings as part of a program's cost. The direct cost of each placement in Michigan is about $50,000 (ChildTrends, 2021) (Panel B).[17] For the cost savings associated with less crime, we use estimates from Heckman et al. (2010) for the police and court costs associated with each arrest and the incarceration costs for a given incarceration spell. Similar to calculating the reduction in social costs, we estimate

[17]Information on the cost of foster care placements in Michigan is sourced from Child Trends' national survey of child welfare agency expenditures. The survey indicates the total amount Michigan spends annually on out-of-home placements. This category encompasses expenditures for both family foster care and congregate care placements and includes any services or activities associated with providing foster care in an approved setting. Examples of such expenditures include the initial assessment of the child's needs, case planning and management, periodic case reviews, and the recruitment and licensing of out-of-home placements. This total does not include expenditures on child protective services investigations, since both treated and control compliers experience an investigation. It also excludes costs associated with adoption and guardianship, as most children in our context reunify with their birth parents. Note that the direct cost of marginal placements in our setting may be even lower because compliers spend relatively less time in foster care than the average placement and are somewhat less likely to be placed in institutional settings (Table 4), which tend to be more expensive.

our baseline specification where the dependent variable is the sum of the cost of each arrest and incarceration for each child in our sample. Panel C shows that cost savings range from $12,000 to $14,000 depending on the discount rate. Combining the direct cost of each placement with these cost savings, the net cost of foster care to the government is between $36,000 and $38,000.

We calculate the MVPF as the reduction in the social cost of crime divided by the net cost of foster care to the government. The MVPF ranges from 2.22 to 2.63 depending on the discount rate, which means that society receives more than $2 in benefits for every $1 in costs (Panel D). These estimates are larger than the MVPF of other social policies for children and adults. For example, Hendren and Sprung-Keyser (2020) calculate an average MVPF of 1.78 across four health insurance expansions to children and an upper bound of 1.20 for adult policies such as housing vouchers, tax credits, and cash welfare programs.

It is important to note that our estimates of the MVPF are limited in scope. For society's willingness to pay, birth parents presumably have a high willingness to pay to avoid child removal, and this is not taken into account in the above calculations. At the same time, there are several reasons why our estimates could be understated. First, we only measure reductions in criminal justice contact through age 19. Second, our measure of "adult criminality" is almost surely a lower bound. Specifically, there are far fewer convictions than there are offenses due to both the under-reporting of crimes to law enforcement and the low rate of clearance rates among reported crimes. Third, the benefits we consider include only the children's subsequent crime reductions and not the other benefits we estimate, such as reductions in crime for parents and increases in children's educational attainment, which could increase earnings. Fourth, as is common in the economics of crime, the social cost of crime reductions excludes relatively minor crimes, such as traffic or drug offenses (Chalfin, 2015). Lastly, the cost savings to the government exclude savings from fewer subsequent child welfare investigations or juvenile detentions.

VIII Conclusion

This study provides new evidence on whether or not there is a foster care-to-prison pipeline. Studying nearly 120,000 child welfare investigations in Michigan between 2008 and 2016, and leveraging the quasi-random assignment of investigators, we find that placement for children at the margin reduces the likelihood that children are arrested, convicted, and incarcerated as adults.

We find that a likely explanation for the reduction in crime is that birth parents make improvements while their children are in foster care. Parents whose children were removed are less likely to have criminal justice contact in the years following the investigation. This pattern could be explained by the fact that parents work closely with social workers after removal and receive fully funded services to address challenges in their lives. As a result, one might suspect that the improvements in children's outcomes are driven by adult services and not placement. That is, that we would observe similar impacts for children if their parents had received services while they remained in the home. However, the families of complier children who are not placed often also receive services, indicating that there is something different about the experiences of parents whose children are placed. It is challenging to determine why the combination of foster care placement and services leads to improved parental outcomes, while prevention services alone may not. One interpretation is that the combination of services and removal results in improvements because parents are compelled to engage with the services if they wish to reunify with their children. Conversely, when their children are not removed, parents do not face as strong an incentive to utilize these services, and evidence shows that they often do not. Another possibility is that parents benefit from the time and space away from their children, which allows them to focus on themselves.

Because abused and neglected children who are not placed are more likely to be involved in the criminal justice system as adults, our results indicate that current efforts to protect vulnerable children in their homes are falling short. The findings in this study are particularly important in light of the Family First Prevention Services Act of 2018, bipartisan legislation which prioritizes keeping children with their families. For the first time, this historic policy allows states to use federal Title IV-E funding on evidence-based programs and services that aim to prevent foster care placement. Given the federal push to keep families intact, our analysis indicates that safely reducing foster care caseloads will require improving efforts to ensure child wellbeing in the home. Child welfare programs and services funded through Title IV-E represent one path to strengthen prevention, yet policies outside of child welfare can also promote child safety. For instance, an extensive literature has found that broader social policies, such as the social safety net, can reduce child abuse and neglect (Aizer et al., 2016; Berger et al., 2017; Raissian and Bullinger, 2017; Rittenhouse, 2023). The fact that foster care passes cost-benefit analyses even when accounting only for its effects on children's later-in-life crime suggests that policies that promote child safety are likely to yield substantial positive externalities. To meet the federal goal of safely reducing foster care caseloads, identifying and scaling effective policies and programs to ensure child wellbeing in the home is a crucial frontier for future research.

References

Abadie, Alberto. 2003. "Semiparametric instrumental variable estimation of treatment response models." *Journal of Econometrics*, 113(2): 231–263.

AECF. 2017. "KIDS COUNT Data Center." The Annie E. Casey Foundation. https://datacenter.kidcount.org.

Agan, Amanda Y, Jennifer L Doleac, and Anna Harvey. 2021. "Misdemeanor prosecution." *NBER Working Paper #w28600*.

Aizer, Anna, and Joseph J Doyle. 2015. "Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges." *The Quarterly Journal of Economics*, 130(2): 759–803.

Aizer, Anna, Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney. 2016. "The long-run impact of cash transfers to poor families." *American Economic Review*, 106(4): 935–71.

Amon, Elizabeth. 2021. "New Washington Laws Aim to Interrupt Foster Care-to-Prison Pipeline." *The Imprint*. Accessed at https://imprintnews.org/ on 10-10-2021.

Anders, John, Andrew Barr, and Alexander Smith. 2022. "The Effect of Early Childhood Education on Adult Criminality: Evidence from the 1960s through 1990s." *American Economic Journal: Economic Policy*, Forthcoming.

Andrews, Isaiah, James H Stock, and Liyang Sun. 2019. "Weak instruments in instrumental variables regression: Theory and practice." *Annual Review of Economics*, 11(1).

Arteaga, Carolina. 2021. "Parental Incarceration and Children's Educational Attainment." *Review of Economics and Statistics*, Forthcoming.

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.

Barnow, Burt S., and Jeffrey Smith. 2015. "Employment and Training Programs."

Baron, E. Jason, Brian Jacob, and Joseph Ryan. 2022. "Pretrial Juvenile Detention." *NBER Working Paper #29861*.

Baron, E. Jason, Joseph J. Doyle, Natalia Emanuel, Peter Hull, and Joseph Ryan. 2024. "Discrimination in Multi-Phase Systems: Evidence from Child Protection." *The Quarterly Journal of Economics*, Forthcoming.

Baron, E. Jason, Josh Hyman, and Brittany Vasquez. 2024. "Public School Funding, School Quality, and Adult Crime." *The Review of Economics and Statistics*.

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.

Bayer, Patrick, Randi Hjalmarsson, and David Pozen. 2009. "Building criminal capital behind bars: Peer effects in juvenile corrections." *The Quarterly Journal of Economics*, 124(1): 105–147.

Bell, Brian, Laura Jaitman, and Stephen Machin. 2014. "Crime deterrence: Evidence from the London 2011 riots." *The Economic Journal*, 124(576): 480–506.

Bell, Brian, Rui Costa, and Stephen J Machin. 2021. "Why does education reduce crime?" *Journal of Political Economy*, Forthcoming.

Berger, Lawrence M, Sarah A Font, Kristen S Slack, and Jane Waldfogel. 2017. "Income and child maltreatment in unmarried families: Evidence from the earned income tax credit." *Review of Economics of the Household*, 15(4): 1345–1372.

Berrick, Jill Duerr. 2018. *The Impossible Imperative: Navigating the competing principles of child protection.* Oxford University Press.

Bhuller, Manudeep, Gordon B Dahl, Katrine V Loken, and Magne Mogstad. 2018. "Intergenerational effects of incarceration." *AEA Papers and Proceedings*, 108: 234–40.

Bhuller, Manudeep, Gordon B Dahl, Katrine V Løken, and Magne Mogstad. 2020. "Incarceration, recidivism, and employment." *Journal of Political Economy*, 128(4): 1269–1324.

Billings, Stephen B. 2019. "Parental Arrest and Incarceration: How does it impact the children?" Working paper.

Billings, Stephen B, and Kevin T Schnepel. 2018. "Life after lead: Effects of early interventions for children exposed to lead." *American Economic Journal: Applied Economics*, 10(3): 315–44.

BJS. 2016. *Bureau of Justice Statistics Survey of Prison Inmates.*

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

Carneiro, Pedro, James J Heckman, and Edward J Vytlacil. 2011. "Estimating marginal returns to education." *American Economic Review*, 101(6): 2754–81.

Chalfin, Aaron. 2015. "Economic costs of crime." *The encyclopedia of crime and punishment*, 1–12.

Chalfin, Aaron, Benjamin Hansen, Emily K Weisburst, Morgan C Williams, et al. 2021. "Police Force Size and Civilian Race." *American Economic Review: Insights*, Forthcoming.

Chetty, Raj, and Nathaniel Hendren. 2018. "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects." *The Quarterly Journal of Economics*, 133(3): 1107–1162.

Chetty, Raj, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How does your kindergarten classroom affect your earnings? Evidence from Project STAR." *The Quarterly journal of economics*, 126(4): 1593–1660.

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.

ChildTrends. 2018. "Foster Care." Child Trends Databank.

ChildTrends. 2021. "Child Welfare Agency Spending in Michigan in SFY 2018."

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.

Chyn, Eric, Brigham Frandsen, and Emily C Leslie. 2024. “Examiner and Judge Designs in Economics: A Practitioner’s Guide.” National Bureau of Economic Research.

Collinson, Robert, John Eric Humphries, Nicholas Mader, Davin Reed, Daniel Tannenbaum, and Winnie van Dijk. 2021. “Eviction and Poverty in American Cities: Evidence from Chicago and New York.” *Working Paper.*

Courtney, M, A Dworsky, A Brown, C Cary, K Love, V Vorhies, and Chapin Hall. 2011. “Midwest Evaluation of the Adult Functioning of Former Foster Youth: Outcomes at Age 26.”

Cullen, Julie Berry, Brian A Jacob, and Steven Levitt. 2006. “The effect of school choice on participants: Evidence from randomized lotteries.” *Econometrica*, 74(5): 1191–1230.

Currie, Janet, and Erdal Tekin. 2012. “Understanding the cycle childhood maltreatment and future crime.” *Journal of Human Resources*, 47(2): 509–549.

Dahl, Gordon B, Andreas Ravndal Kostøl, and Magne Mogstad. 2014. “Family welfare cultures.” *The Quarterly Journal of Economics*, 129(4): 1711–1752.

Dallman, Scott. 2020. “The Effect of Opioid Abuse on Child Out-of-Home Placements.” *Working Paper.*

Deming, David. 2009. “Early childhood intervention and life-cycle skill development: Evidence from Head Start.” *American Economic Journal: Applied Economics*, 1(3): 111–34.

Deming, David J. 2011. “Better schools, less crime?” *The Quarterly Journal of Economics*, 126(4): 2063–2115.

Dobbie, Will, and Roland Fryer. 2015. “The Medium-Term Impacts of High-Achieving Charter Schools.” *Journal of Political Economy*, 123(5): 985–1037.

Dobbie, Will, Jacob Goldin, and Crystal S Yang. 2018. “The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges.” *American Economic Review*, 108(2): 201–40.

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.

Drago, Francesco, Roberto Galbiati, and Pietro Vertova. 2009. “The deterrent effects of prison: Evidence from a natural experiment.” *Journal of political Economy*, 117(2): 257–280.

Dustmann, Christian, and Rasmus Landersø. 2021. “Child’s gender, young fathers’ crime, and spillover effects in criminal behavior.” *Journal of Political Economy*, 129(12): 3261–3301.

Dynarski, Susan, Steven Hemelt, and Joshua Hyman. 2015. “The Missing Manual: Using National Student Clearinghouse Data to Track Postsecondary Outcomes.” *Educational Evaluation and Policy Analysis*, 37(1S): 53S–79S.

Eichmeyer, Sarah, and Christina Kent. 2022. “Parenthood in Poverty.” *Working Paper.*

Enamorado, Ted, Benjamin Fifield, and Kosuke Imai. 2019. “Using a probabilistic model to assist merging of large-scale administrative records.” *American Political Science Review*, 113(2): 353–371.

Evans, Mary F, Matthew C Harris, and Lawrence M Kessler. 2022. “The Hazards of Unwinding the Prescription Opioid Epidemic: Implications for Child Maltreatment.” *American Economic Journal: Economic Policy*, Forthcoming.

Evans, William N, and Emily G Owens. 2007. "COPS and Crime." *Journal of public Economics*, 91(1-2): 181–201.

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.*

Fitzpatrick, Maria D, Cassandra Benson, and Samuel Bondurant. 2022. "Beyond reading, writing, and arithmetic: The role of teachers and schools in reporting child maltreatment." *Journal of Human Resources*, 0319–10084R2.

Font, Sarah, and Christopher Mills. 2022. "Safe from Harm? Peer Effects and Criminal Capital Formation in Foster Care." *Working Paper.*

Frandsen, Brigham. 2020. "TESTJFE: Stata module to perform test for instrument validity in the judge fixed effects design."

Frandsen, Brigham, Lars Lefgren, and Emily Leslie. 2023. "Judging judge fixed effects." *American Economic Review*, 113(1): 253–77.

Garin, Andrew, Dmitri Koustas, Carl McPherson, Samuel Norris, Matthew Pecenco, Evan K Rose, Yotam Shem-Tov, and Jeffrey Weaver. 2023. "The Impact of Incarceration on Employment, Earnings, and Tax Filing." *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.

Goncalves, Felipe, and Steven Mello. 2022. "Should the Punishment Fit the Crime? Discretion and Deterrence in Law Enforcement." *Working Paper.*

Gray-Lobe, Guthrie, Parag A Pathak, and Christopher R Walters. 2021. "The long-term effects of Universal Preschool in Boston." *NBER Working Paper #28756.*

Grimon, Marie-Pascale. 2020. "Effects of the Child Protection System on Parents." Working paper.

Grönqvist, Hans, J Peter Nilsson, and Per-Olof Robling. 2020. "Understanding how low levels of early lead exposure affect children's life trajectories." *Journal of Political Economy*, 128(9): 3376–3433.

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): 170–199.

Heckman, James J, and Edward J Vytlacil. 2007. "Econometric evaluation of social programs, part II: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments." *Handbook of econometrics*, 6: 4875–5143.

Heckman, James J, and Edward Vytlacil. 2005. "Structural equations, treatment effects, and econometric policy evaluation 1." *Econometrica*, 73(3): 669–738.

Heckman, James J, Seong Hyeok Moon, Rodrigo Pinto, Peter A Savelyev, and Adam Yavitz. 2010. "The rate of return to the HighScope Perry Preschool Program." *Journal of public Economics*, 94(1-2): 114–128.

Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes." *American Economic Review*, 103(6): 2052–86.

Heller, Sara, Anuj Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold Pollack. 2017. "Thinking fast and slow? Some field experiments to reduce crime and dropout in Chicago." *Quarterly Journal of Economics*, 132(1): 1–54.

Helénsdotter, Ronja. 2022. "Health Effects of Removing a Child From Home." Unpublished.

Hendren, Nathan, and Ben Sprung-Keyser. 2020. "A unified welfare analysis of government policies." *The Quarterly Journal of Economics*, 135(3): 1209–1318.

Hou, Claire. 2022. "The Opioid Crisis and Foster Care Dynamics." *Working Paper.*

Jacob, Brian A. 2004. "Public housing, housing vouchers, and student achievement: Evidence from public housing demolitions in Chicago." *American Economic Review*, 94(1): 233–258.

Jácome, Elisa. 2020. "Mental Health and Criminal Involvement: Evidence from Losing Medicaid Eligibility." *Job Market Paper, Princeton University.*

Katz, Lawrence, Steven D Levitt, and Ellen Shustorovich. 2003. "Prison conditions, capital punishment, and deterrence." *American Law and Economics Review*, 5(2): 318–343.

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. 2006. "Incarceration length, employment, and earnings." *American Economic Review*, 96(3): 863–876.

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.

Kolesar, Michal. 2013. "Estimation in an instrumental variables model with treatment effect heterogeneity." *Unpublished Working Paper.*

Leslie, Emily, and Nolan G Pope. 2017. "The unintended impact of pretrial detention on case outcomes: Evidence from New York City arraignments." *The Journal of Law and Economics*, 60(3): 529–557.

Lochner, Lance. 2004. "Education, work, and crime: A human capital approach." *International Economic Review*, 45(3): 811–843.

Massenkoff, Maxim N, and Evan K Rose. 2022. "Family formation and crime." *NBER Working Paper #w30385.*

MDHHS. 2021. "Michigan Title IV-E Prevention Plan." *Michigan Department of Health and Human Services.*

Mello, Steven. 2019. "More COPS, less crime." *Journal of Public Economics*, 172: 174–200.

Mogstad, Magne, Alexander Torgovitsky, and Christopher R Walters. 2021. "The causal interpretation of two-stage least squares with multiple instrumental variables." *American Economic Review*, 111(11): 3663–98.

Mueller-Smith, Michael. 2015. "The Criminal and Labor Market Impacts of Incarceration." Working paper.

Mueller-Smith, Michael, Benjamin Pyle, and Caroline Walker. 2022. "Estimating the Impact of the Age of Criminal Majority: Decomposing Multiple Treatments in a Regression Discontinuity Framework." *Working Paper.*

Norris, Samuel. 2019. "Examiner inconsistency: Evidence from refugee appeals." *Working Paper.*

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.

Olea, José Luis Montiel, and Carolin Pflueger. 2013. "A robust test for weak instruments." *Journal of Business & Economic Statistics*, 31(3): 358–369.

Raissian, Kerri M, and Lindsey Rose Bullinger. 2017. "Money matters: Does the minimum wage affect child maltreatment rates?" *Children and youth services review*, 72: 60–70.

Rittenhouse, Katherine. 2023. "Income and Child Maltreatment: Evidence from a Discontinuity in Tax Benefits." *Working Paper.*

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

Rubin, David M, Amanda LR O'Reilly, Xianqun Luan, and A Russell Localio. 2007. "The impact of placement stability on behavioral well-being for children in foster care." *Pediatrics*, 119(2): 336–344.

Rubin, David M, Evaline A Alessandrini, Chris Feudtner, David S Mandell, A Russell Localio, and Trevor Hadley. 2004. "Placement stability and mental health costs for children in foster care." *Pediatrics*, 113(5): 1336–1341.

Ryan, Joseph P, and Mark F Testa. 2005. "Child maltreatment and juvenile delinquency: Investigating the role of placement and placement instability." *Children and youth services review*, 27(3): 227–249.

Ryan, Joseph P, Jane Marie Marshall, Denise Herz, and Pedro M Hernandez. 2008. "Juvenile delinquency in child welfare: Investigating group home effects." *Children and Youth Services Review*, 30(9): 1088–1099.

Thomson, Dana, Renee Ryberg, Kristen Harper, James Fuller, Katherine Paschall, Jody Franklin, and L Guzman. 2022. "Lessons from a historic decline in child poverty." *Child Trends.*

Trivedi, Shanta. 2020. "Police feed the foster care-to-prison pipeline by reporting on Black parents." *NBC News.* Accessed at https://www.nbcnews.com/ on 02-05-2021.

USDHHS. 2000. "Child Maltreatment 2000." United States Department of Health and Human Services.

USDHHS. 2015. "Child Maltreatment 2015." United States Department of Health and Human Services.

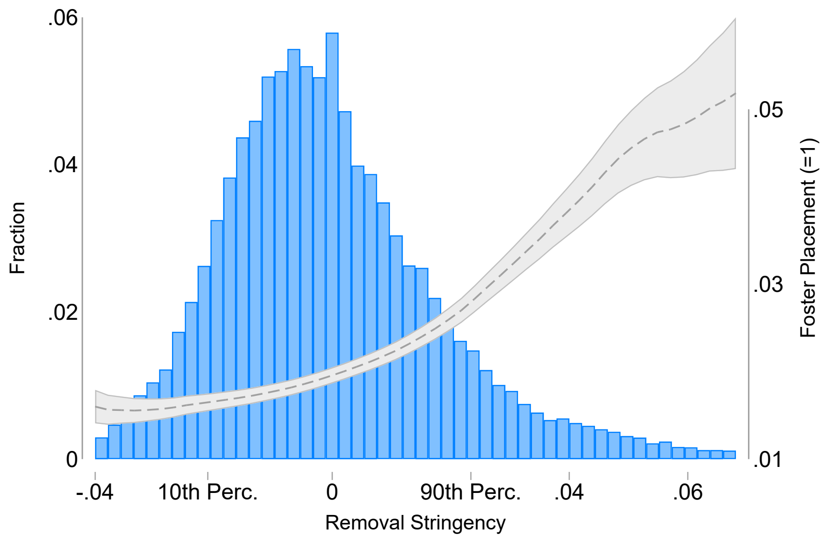
USDHHS. 2019. "Child Maltreatment 2019." United States Department of Health and Human Services.

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.

Wood, Robert G, Quinn Moore, Andrew Clarkwest, and Alexandra Killewald. 2014. "The long-term effects of building strong families: A program for unmarried parents." *Journal of Marriage and Family*, 76(2): 446–463.

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.

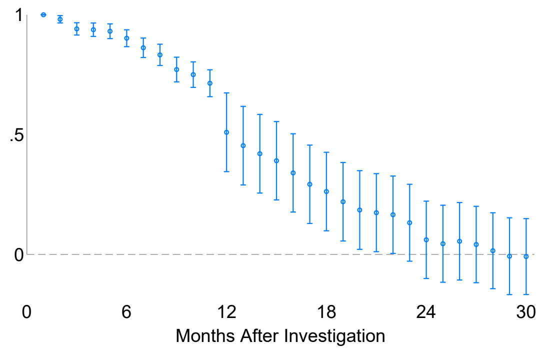
Figure 1: Distribution of Investigator Removal Stringency Instrument



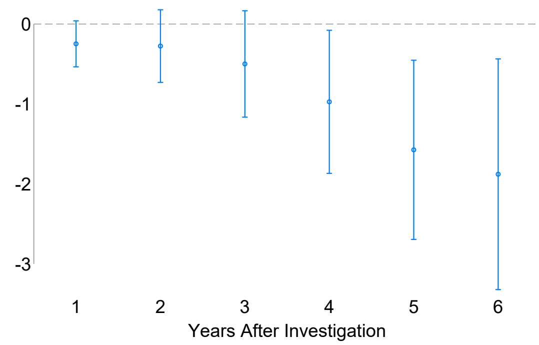
Notes. This figure shows the distribution of the removal stringency instrument. The dashed line shows point estimates from a non-parametric regression of placement on Z^R (net of zip code by investigation year fixed effects) and the shaded region shows the 95 percent confidence interval.

Figure 2: Effects of Foster Care Over Time

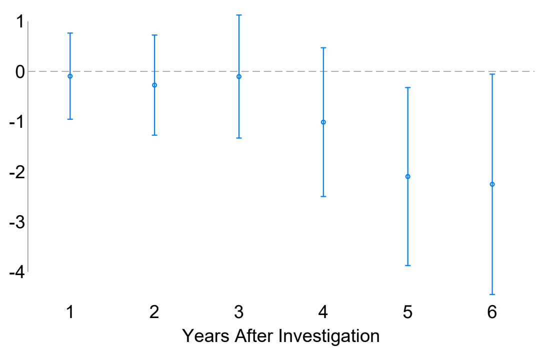
(a) Child is in foster care in month t



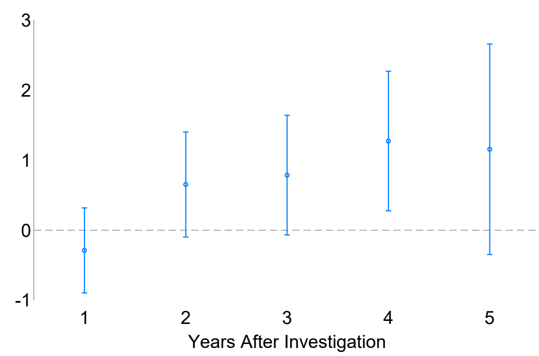
(b) Ever the subject of another maltreatment investigation by the end of year t (cumulative)



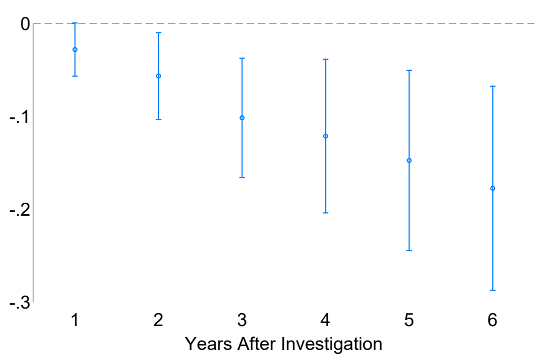
(c) Ever chronically absent by the end of year t (cumulative)



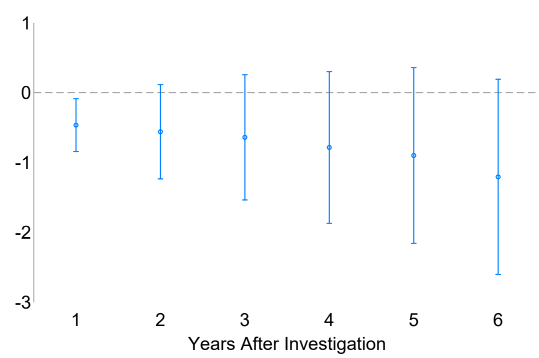
(d) Math test scores in year t



(e) Birth parent is ever arrested by the end of year t (cumulative)



(f) Birth parent is ever alleged for child maltreatment by the end of year t (cumulative)



Notes. The figures report UJIVE point estimates and 95 percent confidence intervals from separate regressions of the outcome variable on foster care in the focal year t. All regressions in Panels A through D include zip code by investigation year fixed effects and the covariates listed in Table 2. Panel C shows outcomes up to five years after the investigation (as opposed to six) because students in Michigan are only tested in grades 3–8. As a result, the sample size drops dramatically for this outcome after year five. The estimates in Panels E and F come from a sample at the birth parent × investigation level for the birth parents listed as the alleged perpetrators in the investigations in our analysis sample. Regressions in these panels include zip code by investigation year fixed effects as well as controls for the birth parent's gender, race and age, the child's gender and free or reduced price lunch receipt, and whether the investigation was for abuse or neglect. Standard errors are clustered at the child level in Panels A through D. We cluster standard errors at the birth parent level in panels E and F since a birth parent can appear in the birth parent × investigation sample more than once. The outcomes in Panels B, C, E, and F are cumulative, meaning the point estimates represent the effect of child removal on the cumulative number of times the child or birth parent experienced the outcome at least once in a given year by the end of year t.

Table 1: Summary Statistics

	(1)	(2)	(3)
		Analysis Sample	
	All MI Students	All	Foster Care
Socio-demographic Characteristics			
Female	0.49	0.50	0.46
White	0.67	0.63	0.49
Black	0.21	0.29	0.43
Hispanic	0.08	0.06	0.06
Other Race	0.05	0.02	0.02
Age	11.70	11.85	12.08
Grade in School	6.15	6.22	6.36
Low Income	0.49	0.81	0.86
Prior Schooling Characteristics			
Attendance Rate	0.95	0.83	0.75
Std Math Score	0.00	-0.39	-0.52
Std Reading Score	0.00	-0.36	-0.49
Adult Crime Outcomes			
Arrested by Age 19		0.14	0.21
Convicted by Age 19		0.08	0.11
Incarcerated by Age 19		0.06	0.09
Observations	1,262,665	118,273	2,595

Notes. Column 1 reports the characteristics of Michigan public school students enrolled in grades 1 through 11 in the 2016-17 school year. Column 2 includes all investigations in the analysis sample and Column 3 includes the subset of investigations that resulted in foster placement. For Columns 2 and 3, the socio-demographic variables are measured in the school year of the investigation and the prior schooling variables are measured in the school year before the investigation. Low income represents free or reduced-price lunch eligibility and math and reading test scores are normalized for the entire state to have a mean of zero and a standard deviation of one within every subject by grade by year cell.

Table 2: Effects of Foster Care on Adult Crime

	(1) Arrested by Age 19	(2) Convicted by Age 19	(3) Incarcerated by Age 19
		Panel A: OLS Estimates	
Foster Care	0.040***	0.021***	0.025***
	(0.008)	(0.006)	(0.006)
	{0.142}	{0.076}	{0.056}
		Panel B: UJIVE Estimates	
Foster Care	-0.252**	-0.281***	-0.210**
	(0.126)	(0.095)	(0.083)
	{0.370}	{0.346}	{0.262}
Observations	118,273	118,273	118,273

Notes. All regressions include zip code by investigation year fixed effects, baseline controls (the child's gender, race, and indicators for grade in school), investigation controls (the number of investigations prior to the focal case and an indicator for whether the allegation included physical abuse), academic controls measured in the year prior to the investigation (attendance rate, special education status, an indicator for free or reduced price lunch eligibility, an indicator for whether the student was expelled, and math and reading test scores), controls for the child's school in the year before the investigation (indicators for urbanicity and charter status, and the share of students by race and by free or reduced price lunch eligibility), and controls for the child's census block in the year prior to the investigation (household median income, employment rate, share of the population with a bachelor's degree or higher, share of the population by race, and an indicator for whether the student was ever homeless in that year). Standard errors clustered at the child level are in parentheses and control means (Panel A) or control complier means (Panel B) are in curly brackets. For a given outcome Y, control complier means come from a regression of $(1 - FC)Y$ on $(1 - FC)$ (instrumented by Z) and zip code by investigation year fixed effects (Abadie, 2003).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Baseline Complier Characteristics

	(1) Full Sample	(2) All Compliers
Child Characteristics		
Female	0.500	0.506
White	0.629	0.448
Black	0.287	0.394
Low income	0.811	0.956
Young	0.440	0.430
Urban/suburban county	0.648	0.850
Low income neighborhood	0.500	0.539
Investigation Characteristics		
Neglect allegation	0.522	0.668
Physical abuse allegation	0.314	0.131
Substance abuse allegation	0.164	0.202
Birth Parent Characteristics		
At least one female	0.812	0.733
At least one White	0.760	0.542
At least one Black	0.251	0.427
At least one young	0.531	0.636
Number listed per investigation	1.436	1.210

Notes. Column 1 contains all investigations in the sample. Column 2 displays average complier characteristics. For a given covariate X, control complier means come from a regression of $(1 - FC)X$ on $(1 - FC)$ (instrumented by Z) and zip code by investigation year fixed effects. Similarly, complier means in the treated state come from a regression of $FC \times X$ on FC (instrumented by Z) and zip code by investigation year fixed effects (Abadie, 2003). We follow Chyn, Frandsen and Leslie (2024) and report average complier characteristics across the treated and untreated states in Column 2. Young children are defined as those under the age of 12 at the time of the investigation, reflecting the average age of children in our sample. The last four rows present characteristics of the birth parents listed as the alleged perpetrators in the investigations in our analysis sample. Because there can be multiple alleged perpetrators in any given investigation, we report whether the investigation included at least one birth parent with that particular characteristic. Young birth parents are individuals under the age of 38 at the time of the investigation, the average age of birth parents in the sample.

Table 4: Foster Care Experiences of Compliers

	(1) All Placements	(2) Treated Compliers
Initial Placement		
With Relatives	0.556	0.550
With Unrelated Family	0.304	0.334
In Group Home	0.140	0.116
Placement Stability		
Number of Different Placements	3.468	3.272
One or Two Different Placements	0.397	0.523
Three or More Different Placements	0.603	0.477
Placement Length		
Months in Foster System	21	18
Permanency Outcomes		
Reunified	0.802	0.832
Adopted	0.087	0.065
Guardianship Transferred	0.064	0.060
Emancipated	0.047	0.044

Notes. Column 1 reports the mean outcome among all placements, and Column 2 reports the mean outcome among treated compliers. For a given outcome Y, complier means in the treated state come from a regression of $FC \times Y$ on FC (instrumented by Z) and zip code by investigation year fixed effects (Abadie, 2003). Permanency outcomes are conditional on having exited foster care by the last available day in the child welfare data.

Table 5: Effects of Foster Care on Adult Convictions by Child and Investigation Characteristics

	(1) Male	(2) Female	(3) Young	(4) Old	(5) First Inv.	(6) Subseq. Inv.	(7) White	(8) Black	(9) Hispanic	(10) Neglect	(11) Physical Abuse	(12) Substance Abuse
Foster Care	-0.496***	-0.114	-0.446***	-0.095	-0.366***	-0.089	-0.243*	-0.253*	-0.282	-0.429***	-0.194	-0.052
	(0.189)	(0.094)	(0.135)	(0.140)	(0.120)	(0.166)	(0.144)	(0.133)	(0.346)	(0.146)	(0.289)	(0.144)
	{0.593}	{0.149}	{0.530}	{0.138}	{0.440}	{0.143}	{0.283}	{0.424}	{0.358}	{0.529}	{0.227}	{0.119}
Observations	59,148	59,125	52,076	66,197	87,100	31,173	74,375	33,999	7,084	61,777	37,112	19,384

Notes. Each column reports estimates from a separate UJIVE regression of whether a child was convicted by age 19 on foster care. All regressions include zip code by investigation year fixed effects and the covariates listed in Table 2. Control complier means are in curly brackets. Standard errors are clustered by child. Young children are defined as those under the age of 12 at the time of the investigation, reflecting the average age of children in our sample. Column 5 restricts the sample to only the child's first investigation in our sample. The distinction between first and subsequent investigations is made within our final analysis sample, rather than across the entire spectrum of investigations in Michigan. On average, children in our sample had 1.36 investigations each.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Effects of Foster Care on Academic, Behavioral, and Safety Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Academic				Behavioral	Safety	
	Chronically Absent	Math Test Score	Graduated High School	Ever Attended College	Ever Detained as Juvenile	Subject in Maltreatment Investigation	Confirmed Victim of Maltreatment
Foster Care	-0.211**	0.428*	0.184	0.268*	-0.172*	-0.215***	-0.081**
	(0.101)	(0.256)	(0.177)	(0.154)	(0.094)	(0.080)	(0.037)
	{0.463}	{-0.685}	{0.558}	{0.304}	{0.207}	{0.364}	{0.123}
Observations	118,189	93,764	114,601	118,273	118,273	118,273	118,273

Notes. This table reports the results from UJIVE regressions of the dependent variable on foster care. All regressions include zip code by investigation year fixed effects and the covariates listed in Table 2. Standard errors in parentheses are clustered by child and the control complier means are reported in curly brackets. Because the outcomes in the first and last two columns are time-varying, we construct an unbalanced investigation by school year panel and follow students in the years after their investigation. The point estimates in these columns come from a specification where we pool all available years following the focal investigation. The outcomes in the remaining columns are measured at a single point in time and we use our main approach to estimate effects for these outcomes. Certain outcomes are not available for the full sample: (1) chronically absent, due to a small amount of missing absences data; (2) math test scores, because some children were investigated after tested grades (3–8) or exempt from state tests; and (3) high school graduation, because some students leave Michigan public schools before grade 9 and their graduation status is unknown. Math test scores and chronic absenteeism in a given academic year are conditional on public school enrollment in that year.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Effects of Foster Care on Birth Parents' Criminal Justice Contact

	(1) Ever Arrested Following the Investigation	(2) Ever Convicted Following the Investigation	(3) Ever Incarcerated Following the Investigation
Foster Care	-0.083**	-0.067**	-0.038
	(0.036)	(0.033)	(0.030)
	{0.098}	{0.078}	{0.046}
Observations	165,829	165,829	165,829

Notes. This table reports the results from UJIVE regressions of the dependent variable on foster care. We construct a sample at the birth parent × investigation level for the birth parents listed as the alleged perpetrators in the investigations in our analysis sample. The sample consists of 108,205 unique birth parents and 165,829 unique birth parent × investigation observations. We then estimate our main specification in Equations 1 and 2 to examine how removal of one's child impacts the probability of subsequent contact with the criminal justice system. All regressions include as controls the birth parent's gender, race and age, the child's gender and free or reduced price lunch receipt, and whether the investigation was for abuse or neglect. Standard errors clustered at the birth parent level are in parentheses and control complier means are in curly brackets.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Marginal Value of Public Funds

	(1) 3% Discount Rate	(2) 4% Discount Rate	(3) 5% Discount Rate
		Panel A: Society's Willingness to Pay	
Foster Care	-95,319***	-89,374***	-83,854***
	(33,667)	(31,619)	(29,715)
Observations	118,273	118,273	118,273
		Panel B: Direct Cost to the Government	
Direct Cost	$49,920	$49,920	$49,920
		Panel C: Cost Savings to the Government	
Foster Care	-13,742**	-12,937**	-12,188**
	(6,992)	(6,588)	(6,212)
Observations	118,273	118,273	118,273
		Panel D: Estimates of the MVPF	
Willingness to Pay	$95,319	$89,374	$83,854
Net Cost	$36,178	$36,983	$37,732
MVPF	2.63	2.42	2.22

Notes. All monetary amounts are inflated to 2012 dollars. Panel A reports the results from UJIVE regressions of the total social cost of the convictions for each child on foster care. The dependent variable is equal to zero for children who were never convicted by age 19. We discount the social cost of each crime using a 3 to 5% rate from the age at conviction to age 12, the average age at investigation in our sample. Panel B reports the cost of each out-of-home placement in Michigan in 2018 to the federal, state, and local governments (ChildTrends, 2021). Using the same methods as in Panel A, Panel C reports the results from UJIVE regressions of the child's total police, court, and incarceration costs on foster care. Panel D presents estimates of the MVPF, equal to society's willingness to pay divided by the net cost of foster care to the government. All regressions include zip code by investigation year fixed effects and the covariates listed in Table 2. Standard errors clustered at the child level are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Is There a Foster Care-To-Prison Pipeline?

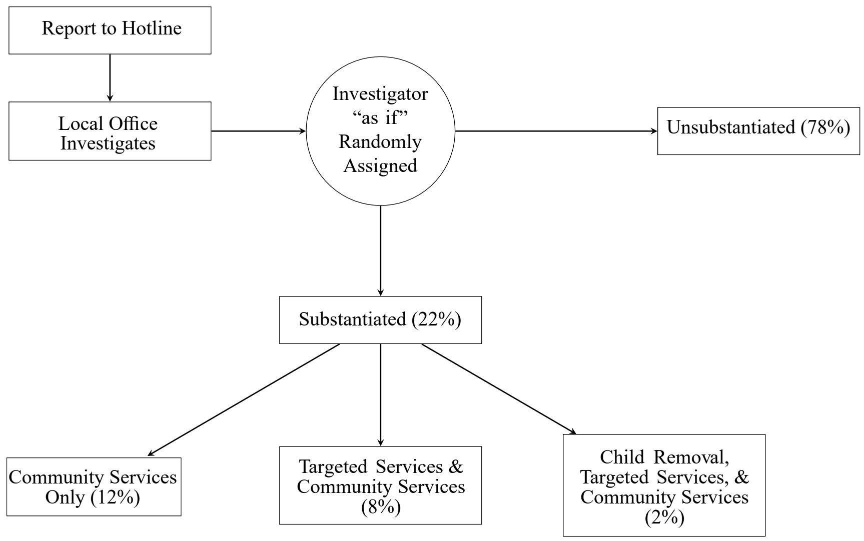
Evidence from Quasi-Randomly Assigned Investigators

E. Jason Baron and Max Gross

Online Appendix

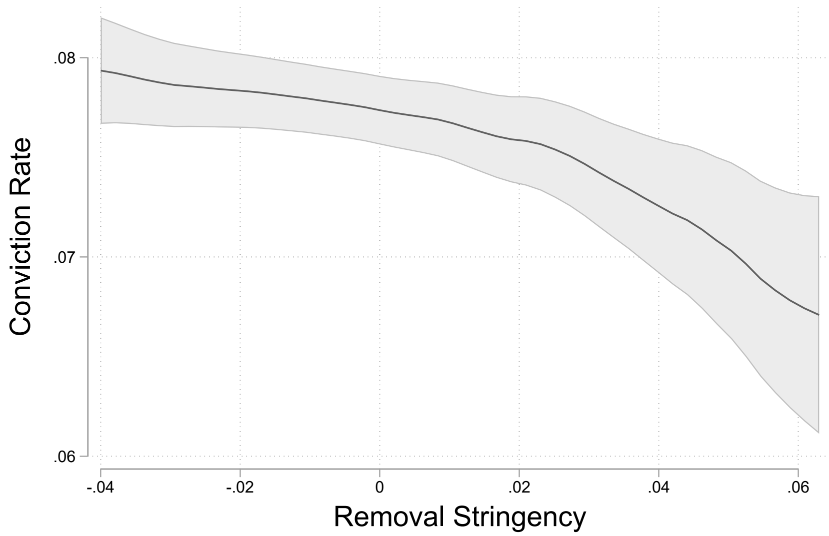
A Supplemental Online Figures and Tables

Figure A1: Child Welfare Investigation Process in Michigan



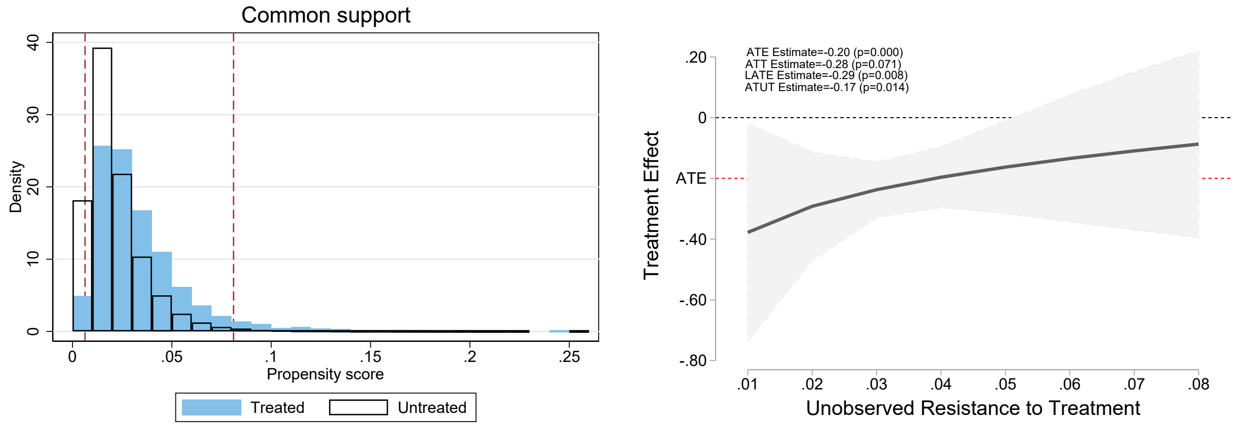
Notes. This figure describes the child maltreatment investigation process in Michigan. "Substantiated" means that investigators found enough evidence to support the abuse or neglect allegation. The shares of investigations in our analysis sample that result in each outcome are included in the boxes.

Figure A2: Reduced Form Relationship Between the Probability of an Adult Conviction and Investigator Removal Stringency



Notes. This figure plots the point estimates and 95 percent confidence intervals from a non-parametric regression of conviction by age 19 on investigator removal stringency, net of zip code by investigation year fixed effects. Standard errors are clustered at the child level.

Figure A3: Common Support and MTEs



Notes. Panel A shows the distribution of the propensity score for treated (those that resulted in removal) and non-treated (those that did not result in removal) investigations. The dashed vertical red lines indicate the upper and lower points of the propensity score distribution with common support (based on one percent trimming). The MTE estimates in Panel B are based on a local instrumental variables approach using a global quadratic polynomial specification for the trimmed sample with common support. The shaded area represents 95% confidence intervals. Standard errors are constructed based on 100 bootstrap replications.

Table A1: First Stage Relationship

	(1) Foster Care	(2) Foster Care	(3) Foster Care	(4) Foster Care
Removal Stringency	0.462***	0.435***	0.436***	0.435***
	(0.028)	(0.029)	(0.029)	(0.029)
Observations	118,273	118,273	118,273	118,273
F-Statistic	281.796	224.746	225.993	225.17
Zip Code by Year FE		✓	✓	✓
Socio-demographic Controls			✓	✓
Academic Controls				✓

Notes. This table reports estimates of γ_1 from Equation 1. Socio-demographic controls include gender, race and ethnicity, indicators for grade in school, an indicator for free or reduced price lunch eligibility in the school year prior to the investigation, an indicator for whether the child was the subject of a prior investigation, and an indicator for whether the allegation was for abuse (versus neglect). Academic controls are measured in the year prior to the investigation and include an indicator for special education receipt, an indicator for grade repetition, daily attendance rate, and standardized math and reading test scores. We report robust Kleibergen-Paap F-statistics, which in the just-identified case are equivalent to the effective F-statistics of Olea and Pflueger (2013) (Andrews, Stock and Sun, 2019). Standard errors are clustered by child.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Balance Tests

Dependent Variable:	(1) Foster Care Placement	(2) Investigator Stringency
Female	-0.003***	0.000
	(0.001)	(0.000)
White	-0.002	0.000
	(0.003)	(0.000)
Black	0.007**	0.000
	(0.003)	(0.000)
Hispanic	0.003	-0.000
	(0.003)	(0.000)
Grade 2	-0.002	0.001*
	(0.004)	(0.001)
Grade 3	-0.001	0.000
	(0.003)	(0.001)
Grade 4	-0.003	0.000
	(0.004)	(0.001)
Grade 5	-0.002	0.000
	(0.004)	(0.001)
Grade 6	-0.003	0.001
	(0.004)	(0.001)
Grade 7	0.000	-0.000
	(0.004)	(0.001)
Grade 8	0.002	0.000
	(0.004)	(0.001)
Grade 9	0.002	0.000
	(0.004)	(0.001)
Grade 10	0.004	0.000
	(0.004)	(0.001)
Grade 11	-0.003	0.000
	(0.005)	(0.001)
Had a Prior Investigation	0.005***	-0.000
	(0.001)	(0.000)
Allegation was for Physical Abuse	-0.003***	0.000
	(0.001)	(0.000)
Free or Reduced Price Lunch	0.004***	-0.000*
	(0.001)	(0.000)
Attendance Rate	-0.037***	-0.001*
	(0.006)	(0.001)
IEP	0.002	0.000
	(0.001)	(0.000)
Repeated a Grade	0.001	-0.000
	(0.001)	(0.000)
Math Test Score	-0.001**	-0.000
	(0.001)	(0.000)
Reading Test Score	-0.000	-0.000
	(0.001)	(0.000)
F Stat from Joint Test	9.865	1.223
P-Value from Joint Test	0.000	0.193
Observations	118,273	118,273

Notes. This table reports the results from regressions of the dependent variable on zip code by investigation year fixed effects and a set of child and investigation covariates. Standard errors are clustered at the child level. The variables in the last five rows are measured in the year before the investigation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Testable Implications of Monotonicity: First Stage by Subgroup

	(1) Male	(2) Female	(3) Young	(4) Old	(5) First Inv.	(6) Subseq. Inv.	(7) White	(8) Black	(9) Hispanic	(10) Neglect	(11) Physical Abuse	(12) Substance Abuse
Removal stringency	0.491***	0.385***	0.398***	0.481***	0.379***	0.643***	0.371***	0.550***	0.651***	0.399***	0.283***	0.824***
	(0.042)	(0.041)	(0.039)	(0.045)	(0.032)	(0.075)	(0.034)	(0.059)	(0.146)	(0.040)	(0.051)	(0.096)
Observations	59,148	59,125	52,076	66,197	87,100	31,173	74,375	33,999	7,084	61,777	37,112	19,384

Notes. Separately for each child subgroup, this table reports the first-stage relationship between investigator removal stringency and foster placement. All regressions include zip code by investigation year fixed effects and the covariates listed in Table A1. Standard errors are clustered at the child level. Young children are defined as those under the age of 12 at the time of the investigation, reflecting the average age of children in our sample. Column 5 restricts the sample to only the child's first investigation in our sample. The distinction between first and subsequent investigations is made within our final analysis sample, rather than across the entire spectrum of investigations in Michigan. On average, children in our sample had 1.36 investigations each.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Frandsen, Lefgren and Leslie (2023) Joint Test of Pairwise Monotonicity and Exclusion

Outcome:	Frandsen, Lefgren and Leslie (2023) Combined P-Value
Arrested by Age 19	0.000
Convicted by Age 19	0.000
Incarcerated by Age 19	0.000

Notes. The table presents results from the semi-parametric test proposed in Frandsen, Lefgren and Leslie (2023) for the null hypothesis that the pairwise monotonicity assumption and exclusion restriction jointly hold. We implemented the test via the Stata package *testjfe* (Frandsen, 2020), using the default number of knots in the spline function (3) and a relative weight of 0.8 on the fit component of the test.

Table A5: Heterogeneity by Crime Type and Impacts on the Number of Convictions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Ever convicted for a . . .						
	Violent Offense	Property Offense	Drug Offense	Public Order Offense	Felony	Misdemeanor	Number of Convictions
Foster Care	-0.140**	-0.059	-0.036	-0.064	-0.054	-0.264***	-0.380**
	(0.059)	(0.062)	(0.042)	(0.064)	(0.055)	(0.088)	(0.184)
	{0.179}	{0.084}	{0.045}	{0.099}	{0.086}	{0.308}	{0.468}
Observations	118,273	118,273	118,273	118,273	118,273	118,273	118,273

Notes. This table reports the results from UJIVE regressions of the dependent variable on a foster care indicator, using removal stringency to instrument for foster care. Standard errors clustered by child are shown in parentheses and control complier means are shown in curly brackets. All regressions include zip code by investigation year fixed effects and the covariates listed in Table 2. The first four columns present heterogeneity by crime type, while Columns 5 and 6 present heterogeneity by crime severity. See Table A6 for a tabulation of offenses by crime type and severity. The dependent variable in Column 7 represents the number of convictions a child has by age 19. It is zero for individuals with no convictions by that age.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Tabulation of Offenses, by Crime Type and Severity

(1)	(2)	(3)	(4)	(5)	(6)
Crime Type				Crime Severity	
Violent	Property	Drug	Public	Felony	Misdemeanor
Assault/batt.	Retail fraud	Poss. marijuana	Op. impaired	Assault police	Retail fraud
Dom. violence	Larceny < \$200	Use marijuana	Purchase/consump. by minor	Carrying concealed weap.	Poss. control. subst.
Home invasion	Breaking/entering	Deliv./manuf. marijuana	Op. w/o license	Home invasion	Op. impaired
Armed robbery	Larceny > \$200	Poss. control. subst.	Disturb. peace	Breaking/entering	Assault/batt.
Assault w/ dang. weap.	Unarmed robbery	Poss. cocaine/heroin	Carrying concealed weap.	Armed robbery	Dom. violence

Notes. This table displays a breakdown of offenses in our analysis sample, categorized by crime type (violent, property, drug, or public order) and severity (felony or misdemeanor). The table highlights the five most common offenses, ranked by frequency in the analysis sample.

Table A7: Implied Estimates for Adult Convictions Based on Impacts on Other Indicators of Child Well-being

	(1) Convicted by Age 19	(2) Estimates in Table 6	(3) Implied Effect (pps)
Chronically Absent	0.068***	-0.211	-1.43
	(0.002)		
Math Test Score	-0.031***	0.428	-1.33
	(0.001)		
Graduated High School	-0.096***	0.184	-1.77
	(0.002)		
Ever Attended College	-0.063***	0.268	-1.69
	(0.002)		
Ever Detained as Juvenile	0.233***	-0.172	-4.01
	(0.007)		
Subject in Maltreatment Investigation	0.053***	-0.215	-1.14
	(0.003)		
Confirmed Victim of Maltreatment	0.044***	-0.081	-0.36
	(0.004)		
Total Implied Effect (pps)			-11.72
Actual Effect (pps)			-28.1

Notes. The estimates in Column 1 come from bivariate regressions in our main analysis sample, where the dependent variable is adult conviction and the independent variable is the specific intermediate outcome. Standard errors are clustered at the child level. Column 2 summarizes the estimates of the effects of foster care on the particular intermediate outcome, as previously reported in Table 6. To be as conservative as possible, we include each of the intermediate outcomes in Table 6 (including the effect on high school graduation, even though the point estimate is not statistically significant at conventional levels). The implied percentage point effects in each row are multiplications of the non-experimental relationships in the first column and the estimates in Column 2. For example, the estimate in Column 3, Row 1 is the result of multiplying $0.068 \times -0.211 \times 100$. We calculate the total implied effect in the second to last row by (conservatively) summing up each of the individual implied effects. The estimate in the last row comes from Table 2.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Effects of Foster Care on Birth Parents' Criminal Justice Contact by Crime Type and Its Impacts on the Number of Convictions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Ever convicted for a ...						
	Violent Offense	Property Offense	Drug Offense	Public Order Offense	Felony	Misdemeanor	Number of Convictions
Foster Care	-0.036*	-0.007	0.001	-0.072***	-0.023	-0.056*	-0.062
	(0.022)	(0.018)	(0.017)	(0.027)	(0.021)	(0.030)	(0.101)
	{0.039}	{0.012}	{0.001}	{0.083}	{0.025}	{0.067}	{0.102}
Observations	165,829	165,829	165,829	165,829	165,829	165,829	165,829

Notes. This table reports the results from UJIVE regressions of the dependent variable on foster care. We construct a sample at the birth parent × investigation level for the birth parents listed as the alleged perpetrators in the investigations in our analysis sample. The sample consists of 108,205 unique birth parents and 165,829 unique birth parent × investigation observations. We then estimate our main specification in Equations 1 and 2 to examine how removal of one's child impacts the probability of subsequent contact with the criminal justice system. All regressions include as controls the birth parent's gender, race and age, the child's gender and free or reduced price lunch receipt, and whether the investigation was for abuse or neglect. Standard errors clustered at the birth parent level are in parentheses and control complier means are in curly brackets. The first four columns present heterogeneity by crime type, while Columns 5 and 6 present heterogeneity by crime severity. The dependent variable in Column 7 represents the number of convictions the birth parent had after the investigation. It is zero for birth parents with no convictions following the investigation.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Heterogeneity in the Effects of Foster Care on Birth Parents' Criminal Justice Contact

	(1) Young Birth Parent	(2) Old Birth Parent	(3) Young Child	(4) Old Child	(5) Child's First Inv.	(6) Child's Subseq. Inv.	(7) Neglect Inv.	(8) Phy. Abuse Inv.	(9) Subs. Abuse Inv.	(10) Male Child	(11) Female Child	(12) White Child	(13) Black Child	(14) Hisp. Child
Foster Care	-0.111*	-0.004	-0.045	-0.067	-0.066	-0.083	-0.064	0.071	-0.141**	-0.058	-0.089**	-0.086*	-0.021	-0.163
	(0.057)	(0.008)	(0.048)	(0.043)	(0.041)	(0.059)	(0.051)	(0.085)	(0.057)	(0.053)	(0.039)	(0.047)	(0.051)	(0.113)
	{0.122}	{0.011}	{0.047}	{0.097}	{0.077}	{0.091}	{0.071}	{0.001}	{0.153}	{0.068}	{0.115}	{0.094}	{0.059}	{0.179}
Observations	88,339	77,490	73,582	92,247	121,790	44,039	87,436	49,593	28,800	82,944	82,885	107,562	44,186	10,060

Notes. Each column reports estimates from a separate UJIVE regression of whether the birth parent was ever convicted following the focal investigation. We construct a sample at the birth parent $\times$ investigation level for the birth parents listed as the alleged perpetrators in the investigations in our analysis sample. The sample consists of 108,205 unique birth parents and 165,829 unique birth parent $\times$ investigation observations. We then estimate our main specification in Equations 1 and 2 to examine how removal of one's child impacts the probability of subsequent contact with the criminal justice system. Standard errors clustered at the birth parent level are in parentheses and control complier means are in curly brackets. Young children are defined as those under the age of 12 at the time of the investigation, reflecting the average age of children in our sample. Young birth parents refer to individuals under the age of 38 at the time of the investigation, reflecting the average age of birth parents in our sample.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Effects of Foster Care on Adult Convictions Through Age 21

	(1) Convicted by Age 19	(2) Convicted by Age 20	(3) Convicted by Age 21
Foster Care	-0.281***	-0.302**	-0.221
	(0.095)	(0.127)	(0.200)
	{0.346}	{0.367}	{0.311}
Observations	118,273	93,279	68,067

Notes. This table reports the results from UJIVE regressions of the dependent variable on foster care. Standard errors clustered by child are shown in parentheses and control complier means are shown in curly brackets. All regressions include zip code by investigation year fixed effects and the covariates listed in Table 2.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Testable Implications of the Exclusion Restriction

	Removal Stringency
Days in Foster Care	0.000
	(0.000)
# Different Foster Placements	-0.000
	(0.001)
Initial Placement with Relatives	0.006
	(0.005)
Initial Placement with Unrelated Family	0.007
	(0.005)
Observations	2,595
F Stat from Joint Test	0.727
P-Value from Joint Test	0.574

Notes. This table reports the results from a regression of the removal stringency instrument on indicators of the child's experience while in foster care. All regressions include zip code by investigation year fixed effects and the covariates listed in Table 2. Standard errors are clustered at the child level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Effects of Foster Care on Adult Crime, Controlling for Children's Experiences in Foster Care

	(1) Arrested by Age 19	(2) Convicted by Age 19	(3) Incarcerated by Age 19
Foster Care	-0.366**	-0.322***	-0.290***
	(0.150)	(0.113)	(0.101)
	{0.548}	{0.424}	{0.369}
Observations	118,273	118,273	118,273

Notes. This table reports the results from UJIVE regressions of the dependent variable on foster care. All regressions include zip code by investigation year fixed effects, the covariates listed in Table 2, and the five measures of children's foster care experiences, as described in the main text. Standard errors clustered at the child level are in parentheses and control complier means are in curly brackets. For a given outcome Y, control complier means come from a regression of $(1-FC)Y$ on $(1-FC)$ (instrumented by Z), zip code by investigation year fixed effects, and the five measures of children's foster care experiences (Abadie, 2003).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Robustness to Alternative Samples and Specifications

	(1) Arrested by Age 19	(2) Convicted by Age 19	(3) Incarcerated by Age 19
	Panel A: Alternative Samples		
Investigator Assigned > 75 Investigations	-0.215*	-0.224**	-0.145*
N = 113,661	(0.130)	(0.098)	(0.086)
Excluding the Strictest Investigators	-0.290	-0.232*	-0.210*
N = 106,447	(0.181)	(0.138)	(0.121)
Excluding Non-Parent Perpetrators	-0.201*	-0.265***	-0.166**
N = 107,629	(0.118)	(0.089)	(0.077)
	Panel B: Alternative Stringency Measures		
Split Sample	-0.327**	-0.256**	-0.221**
N = 118,273	(0.155)	(0.116)	(0.103)
Leave-Out Other Years	-0.193	-0.264***	-0.148*
N = 118,273	(0.118)	(0.090)	(0.077)
Leave-Out Same Year	-0.330	-0.244	-0.220
N = 118,273	(0.347)	(0.255)	(0.227)
Only Cases in the Analysis Sample	-0.244	-0.335**	-0.212*
N = 69,551	(0.315)	(0.143)	(0.123)
2SLS	-0.176	-0.262***	-0.181**
N = 118,273	(0.126)	(0.096)	(0.084)
	Panel C: Alternative Level of Rotation		
County by Year	-0.251**	-0.296***	-0.217***
N = 118,273	(0.125)	(0.095)	(0.083)
	Panel D: Anderson-Rubin Confidence Intervals		
Standard Errors Clustered by Child	-0.252**	-0.281***	-0.210**
N = 118,273	(0.126)	(0.095)	(0.083)
	[-0.504,-0.017]	[-0.470,-0.104]	[-0.377,-0.055]

Notes. Panel A reports the results from UJIVE regressions using alternative sample definitions, Panel B uses alternative measures of removal stringency to instrument for foster care, and Panel C reports the results using the main stringency instrument but replaces zip code by investigation year fixed effects with county by investigation year fixed effects. Except for Panel C, all regressions include zip code by investigation year fixed effects and the covariates listed in Table 2. In Panel B, the split sample instrument is the removal rate of the assigned investigator from a random half of the sample. The leave-out other years measure is the leave-out removal rate of the assigned investigator from other children who had investigations in the same year. The leave-out same year measure is the leave-out removal rate of the assigned investigator from other children who had investigations in different years. The number of observations drops in the fourth row of Panel B because, as for our main analysis, we focus on investigators who were assigned to at least 50 cases—but in this case in our analysis sample. As with the main removal instrument, we calculate each of these four alternative instruments using UJIVE.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: Robustness to Control Selection

	(1) Arrested by Age 19	(2) Arrested by Age 19	(3) Convicted by Age 19	(4) Convicted by Age 19	(5) Incarcerated by Age 19	(6) Incarcerated by Age 19
Foster Care	-0.236*	-0.252**	-0.269***	-0.281***	-0.200**	-0.210**
	(0.129)	(0.126)	(0.096)	(0.095)	(0.084)	(0.083)
Baseline Controls						
Grade 2		0.024***		0.015***		0.015***
		(0.006)		(0.005)		(0.004)
Grade 3		0.046***		0.027***		0.026***
		(0.006)		(0.005)		(0.004)
Grade 4		0.044***		0.023***		0.023***
		(0.008)		(0.006)		(0.005)
Grade 5		0.070***		0.030***		0.032***
		(0.008)		(0.006)		(0.005)
Grade 6		0.100***		0.047***		0.048***
		(0.008)		(0.006)		(0.005)
Grade 7		0.130***		0.060***		0.065***
		(0.009)		(0.007)		(0.006)
Grade 8		0.167***		0.077***		0.079***
		(0.009)		(0.007)		(0.006)
Grade 9		0.200***		0.095***		0.095***
		(0.010)		(0.007)		(0.006)
Grade 10		0.224***		0.106***		0.105***
		(0.010)		(0.008)		(0.007)
Grade 11		0.230***		0.094***		0.105***
		(0.012)		(0.009)		(0.008)
Std Math Score		-0.020***		-0.013***		-0.008***
		(0.003)		(0.002)		(0.002)
Std Reading Score		-0.014***		-0.011***		-0.008***
		(0.002)		(0.002)		(0.002)
Female		-0.092***		-0.067***		-0.056***
		(0.003)		(0.002)		(0.002)
Black		0.014**		0.007		-0.000
		(0.007)		(0.005)		(0.005)
White		-0.044***		-0.030***		-0.029***
		(0.006)		(0.005)		(0.004)
Investigation Controls						
# of Prior Investigations		0.010***		0.008***		0.005***
		(0.001)		(0.001)		(0.001)
Allegation was for Physical Abuse		0.026***		0.017***		0.013***
		(0.002)		(0.002)		(0.002)
Prior Academic Characteristics						
Attendance Rate		-0.193***		-0.123***		-0.097***
		(0.015)		(0.011)		(0.010)
IEP		-0.003		-0.001		0.003
		(0.003)		(0.003)		(0.002)
Ever Expelled		0.208***		0.167***		0.128***
		(0.029)		(0.026)		(0.024)
Free or Reduced Price Lunch		0.010***		0.004		0.003
		(0.003)		(0.002)		(0.002)

Std Math Score X Std Reading Score		0.004*		0.004*		0.004**
		(0.003)		(0.002)		(0.002)
Std Math Score Squared		-0.004**		-0.003***		-0.002*
		(0.002)		(0.001)		(0.001)
Std Reading Score Squared		0.001		0.000		-0.000
		(0.001)		(0.001)		(0.001)
Std Math Score Cubed		0.002***		0.001***		0.001*
		(0.001)		(0.000)		(0.000)
Std Reading Score Cubed		0.000		0.000		0.000**
		(0.000)		(0.000)		(0.000)
School Controls						
Urban		-0.002		-0.003		-0.001
		(0.004)		(0.003)		(0.003)
Charter		-0.016***		-0.008**		-0.003
		(0.005)		(0.003)		(0.003)
% White		0.012		0.010		0.017
		(0.020)		(0.015)		(0.014)
% Black		0.038**		0.022		0.025*
		(0.019)		(0.015)		(0.013)
% FRPL		0.015		0.013*		0.011
		(0.010)		(0.007)		(0.007)
Neighborhood Controls						
# Neighborhoods Lived in Before Investigation		0.003***		0.002***		0.002***
		(0.001)		(0.000)		(0.000)
Household Median Income		0.000		0.000		0.000
		(0.000)		(0.000)		(0.000)
Employment Rate		-0.015		0.004		0.002
		(0.015)		(0.011)		(0.010)
% Bachelor's Degree or Higher		-0.013		0.002		0.004
		(0.014)		(0.011)		(0.009)
% White		-0.002		-0.008		0.011
		(0.024)		(0.017)		(0.015)
% Black		-0.020		-0.022		-0.000
		(0.024)		(0.018)		(0.016)
Homeless in SY Before Investigation		0.202		0.370		0.160
		(0.579)		(0.468)		(0.404)
Observations	118,273	118,273	118,273	118,273	118,273	118,273
Rotation Group FE	✓	✓	✓	✓	✓	✓
Baseline Controls		✓		✓		✓
Investigation Controls		✓		✓		✓
Academic Controls		✓		✓		✓
School Controls		✓		✓		✓
Neighborhood Controls		✓		✓		✓

Notes. Our preferred UJIVE specification throughout the paper includes the covariates listed in the leftmost column of this table. This table shows that the main UJIVE results in Table 2 are robust to not including any additional control variables. Columns 1, 3, and 5 include zip code by investigation year fixed effects and no other control variables. Columns 2, 4, and 6 replicate our preferred specification by including the full set of controls as well as indicators for any missing covariates. Standard errors are clustered by child.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A15: Robustness to Alternative Levels of Clustering Standard Errors

	(1) Arrested by Age 19	(2) Convicted by Age 19	(3) Incarcerated by Age 19
Baseline (by child)			
Foster Care	-0.252**	-0.281***	-0.210**
	(0.126)	(0.095)	(0.083)
By Investigator			
Foster Care	-0.252*	-0.281***	-0.210***
	(0.131)	(0.095)	(0.081)
By Rotation			
Foster Care	-0.252**	-0.281***	-0.210**
	(0.125)	(0.094)	(0.083)
By Child and Investigator			
Foster Care	-0.252*	-0.281***	-0.210***
	(0.130)	(0.094)	(0.080)
By Child and Rotation			
Foster Care	-0.252**	-0.281***	-0.210**
	(0.125)	(0.094)	(0.083)

Notes. Our preferred UJIVE specification throughout the paper clusters standard errors at the child level. This table shows that the main UJIVE results in Table 2 are robust to alternative levels of clustering standard errors. There are 87,100 unique children, 3,011 unique investigators, and 6,569 unique rotation groups in our sample.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A16: Robustness to Out-of-State Migration

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Arrested by Age 19		Convicted by Age 19		Incarcerated by Age 19	
	Ever Left State in K-12	Ever Attended College Outside the State	Excluding K-12 Leavers	Excluding Out-of-State College	Excluding K-12 Leavers	Excluding Out-of-State College	Excluding K-12 Leavers	Excluding Out-of-State College
Foster Care	-0.003	-0.072	-0.232*	-0.249*	-0.300***	-0.304***	-0.200**	-0.220**
	(0.112)	(0.076)	(0.139)	(0.129)	(0.105)	(0.098)	(0.091)	(0.086)
	{0.141}	{0.118}	{0.372}	{0.372}	{0.377}	{0.373}	{0.263}	{0.278}
Observations	118,273	118,273	105,709	112,746	105,709	112,746	105,709	112,746

Notes. Columns 1 and 2 report point estimates from UJIVE regressions of the dependent variable on foster care. All regressions include zip code by investigation year fixed effects and the covariates listed in Table 2. Standard errors are clustered at the child level and control complier means are reported in curly brackets. The dependent variable in Column 1 is an indicator variable equal to one if the student ever left the state while in grades K–12. We measure this outcome using exit codes that are assigned to students who leave the Michigan public school system. The dependent variable in Column 2 is an indicator variable equal to one if the student ever enrolled in postsecondary education outside of Michigan according to the National Student Clearinghouse. The remaining columns report the UJIVE effects of foster care on whether a child is arrested, convicted, and incarcerated by age 19 using a sample excluding children who ever left the state in K-12 (Columns 3, 5, and 7), and children who ever enrolled in postsecondary education outside Michigan (Columns 4, 6, and 8).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A17: Implied Estimates for Adult Criminal Justice Contact

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
					Implied Effects (pps)		
	Arrested by Age 19	Convicted by Age 19	Incarcerated by Age 19	Gross & Baron (2022) Estimates	Arrested	Convicted	Incarcerated
Alleged victim	0.053***	0.038***	0.024***	-0.132	-0.70	-0.50	-0.32
	(0.003)	(0.002)	(0.002)				
Confirmed victim	0.044***	0.032***	0.019***	-0.053	-0.23	-0.17	-0.10
	(0.004)	(0.003)	(0.003)				
Daily attendance rate	-0.003***	-0.001***	-0.001***	0.055	-0.02	-0.01	-0.01
	(0.000)	(0.000)	(0.000)				
Standardized math score	-0.053***	-0.031***	-0.024***	0.356	-1.89	-1.10	-0.85
	(0.002)	(0.001)	(0.001)				
Standardized reading score	-0.037***	-0.021***	-0.017***	0.175	-0.65	-0.37	-0.30
	(0.002)	(0.001)	(0.001)				
Juvenile Delinquency	0.275***	0.180***	0.146***	-0.028	-0.77	-0.50	-0.41
	(0.007)	(0.006)	(0.005)				
Total Implied Effect (pps)					-4.3	-2.7	-2.0
Actual Effect (pps)					-25.2	-28.1	-21.0

Notes. The estimates in Columns 1 through 3 come from bivariate regressions in our main analysis sample, where the dependent variable is the adult crime outcome of interest and the independent variable is the specific intermediate outcome. Standard errors are clustered at the child level. Column 4 summarizes the estimates of the effects of foster care on the particular intermediate outcome, as previously reported in Table 4 in Gross and Baron (2022). The implied percentage point effects in each row are simple multiplications of the non-experimental relationships in the first three columns and the estimates in Column 4. For example, the estimate in Column 5, Row 1 is the result of multiplying $0.053 \times -0.132 \times 100$. The estimates in the second to last row are obtained by (conservatively) summing up each individual implied effect. The estimates in the last row come from Table 2.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A18: Effects of Foster Care on Adult Crime for Sample Comparable to Doyle (2008)

	(1) Arrested by Age 19	(2) Convicted by Age 19	(3) Incarcerated by Age 19
Foster Care	-0.263*	-0.277***	-0.192**
	(0.134)	(0.101)	(0.089)
	{0.378}	{0.338}	{0.241}
Observations	95,541	95,541	95,541

Notes. This table reports the results from UJIVE regressions of the dependent variable on foster care. The analysis sample is restricted to mirror the sample in Doyle (2008) and includes only children who were eligible for free or reduced-price lunch in any year before the investigation. All regressions include zip code by investigation year fixed effects and the covariates listed in Table 2. Standard errors are clustered at the child level and control complier means are reported in curly brackets.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Implied Adult Crime Effects Based on the Estimates in Gross and Baron (2022)

The current study builds on Gross and Baron (2022), which used the same research design and found that foster care improved children's safety and academic outcomes, such as test scores and daily school attendance. This section discusses the extent to which the effects of foster care on later-in-life crime presented in this study could have been predicted by estimates in Gross and Baron (2022).

Panel B of Table 4 in Gross and Baron (2022) shows that foster care reduced the likelihood that children were alleged (or confirmed) as victims of maltreatment in the future by 13.2 percentage points (or 5.3 percentage points). The study also showed that foster care improved subsequent daily school attendance rates by 5.5 percentage points, standardized math test scores by 0.36 standard deviations, and reading test scores by 0.18 standard deviations. These estimates are statistically significant at conventional levels, except for reading test scores. The study also reported a 2.8 percentage point decline in juvenile delinquency using data on juvenile court petitions (an official document filed following juvenile arrests in cases where youth are not immediately diverted from the courts). However, as acknowledged in Gross and Baron (2022), data on juvenile petitions were only available for roughly half of the sample, for a limited number of years, and for certain counties in Michigan. As a result, the estimate on delinquency was imprecise (the standard error was 4 percentage points) and the study could not rule out large effects in either direction.

A rough estimate of the implied adult crime effect can be obtained by pairing the estimates in Gross and Baron (2022) with the non-experimental relationship between each of the intermediate outcomes and measures of adult crime. We first estimated separate bivariate regressions of the adult crime outcome (arrested, convicted, and incarcerated by age 19) on the intermediate outcome of interest (for example, standardized math test scores) using our main analysis sample. These estimates are reported in Columns 1 through 3 of

Table A17. Next, for each outcome in Gross and Baron (2022), we calculate the implied effect on the three adult crime outcomes by multiplying the non-experimental relationship by the estimates in Gross and Baron (2022), which are shown in Column 4.[18] Finally, we calculate the total implied effect as the sum of the implied effects for each intermediate outcome. To be as conservative as possible, we include each individual implied effect, including those in which the estimates in Gross and Baron (2022) were imprecise.

Columns 5, 6, and 7 of Table A17 show that the estimates in Gross and Baron (2022) predict only a modest share of the actual effects found in the current study. As an example, the implied effect on adult arrests is a decline of 4 percentage points, whereas we document a decline of 25 percentage points (Table 2). The implied effects for the other two adult crime outcomes are even smaller. These findings are consistent with recent work underscoring the importance of non-cognitive and behavioral measures, and showing that effects on childhood outcomes alone may not be predictive of effects on later-in-life crime. Altogether, the findings in this section provide compelling evidence that the magnitude of the later-in-life crime reduction documented in this study is not well predicted by estimates in Gross and Baron (2022) alone.

C Comparison to Doyle (2008)

The analysis in this study contrasts with the findings in Doyle (2008), the only other study to estimate the causal impacts of foster care on later-in-life crime in the United States. Using administrative data on Illinois children investigated between 1990 and 2003, the study finds that placement increases adult arrests by 39 percentage points and convictions by 41 percentage points (see Table 4 in Doyle (2008)). In this appendix, we investigate three differences between studies that could contribute to the divergent findings: the study sample period, sample restrictions, and outcomes.

[18]We acknowledge that the bivariate regressions of adult crime on the intermediate outcomes are not causal. However, estimates from these regressions are likely to overstate the relationship between the intermediate outcome and adult crime, leading to an upward bias estimate of the implied effect.

1. Differences between study time periods: We first discuss how several differences between the study time periods could explain the different findings.

First, child welfare policy has changed over time in ways that likely have improved foster care. The federal government has enacted several key policies over the last two decades that focused on improving the quality of foster care in three areas: reducing placement length, increasing placements with relatives and limiting placements in group homes, and promoting the wellbeing of children while in foster care. The Adoption and Safe Families Act of 1997 aimed to reduce the length of stay by generally requiring states to terminate parental rights for children who spent 15 of 22 consecutive months in foster care. Accordingly, the proportion of children in foster care with relatively shorter stays (between one and two years) increased from 18% to 30% from 1998 to 2017 (ChildTrends, 2018). This could have improved foster care because there is an extensive literature showing a negative association between placement stability and children's outcomes (Rubin et al., 2007, 2004; Ryan and Testa, 2005), and shorter stays tend to be more stable. There has also been a shift toward increasing placements with relatives and decreasing placements in institutional settings, in part driven by the Fostering Connections to Success and Increasing Adoptions Act of 2008. The share of children placed with relatives increased by 33% from 2008 to 2017 (24% to 32%) while the share placed in group homes decreased by roughly the same amount (16% to 11%) (AECF, 2017). Perhaps due to negative peer effects, research shows that group home placements are associated with worse outcomes (Bayer, Hjalmarsson and Pozen, 2009; Font and Mills, 2022; Ryan et al., 2008). Lastly, there have been efforts to ensure children's safety and wellbeing while in foster care, including requiring more regular visits from caseworkers (e.g., through the Child and Family Services Improvement Act of 2006), mandating criminal background checks of prospective foster parents (the Adam Walsh Child Protection and Safety Act of 2006), and coordinating mental health care services for foster children (the Child and Family Services Improvement and Innovation Act of 2011). To the extent that these policies have improved child welfare practice, one would expect more recent studies to find less detrimental effects of foster care.

At the same time, broader social policies have changed in ways that have likely improved the counterfactual to placement, meaning that we would have expected to find even more harmful effects than in Doyle (2008) without improvements to the foster care system itself. For example, child poverty has declined by about 60% between 1993 and 2019, in part due to the vastly expanded social safety net over that period (Thomson et al., 2022). Many other indicators of child well-being tell a similar story (e.g., child deaths, high school graduation rates, teen pregnancy rates). All else equal, this means that children are likely to be better off remaining in the home (as opposed to foster care placement) more recently than during the Doyle (2008) sample period. Absent improvements to foster care over time, we would therefore expect to find even more detrimental effects of placement more recently. That we instead find favorable effects offers additional evidence that the foster care system itself has likely improved over time.

Second, nationwide child welfare practice has also changed over time in ways that may have shifted the margin of who is placed in foster care (that is, shifted the complier population). In 2000, toward the end of the Doyle (2008) sample period, about 6% of the children referred as possible victims of abuse or neglect in the United States were placed in foster care (AECF, 2017; USDHHS, 2000). That number declined to 4% in 2015 (USDHHS, 2015), toward the end of our child welfare data, and dropped to 3% in 2019 (USDHHS, 2019). Consistent with these national trends, the mean of the removal instrument in Doyle (2008) was larger than in our study. Because a larger share of referrals resulted in placement during the time of the Doyle (2008) study, children at the margin of placement may have faced less risk in the home than marginal children today. As such, one might again expect more recent studies to find less harmful effects of placement.

To test this hypothesis, we estimate Marginal Treatment Effects (MTEs). MTEs in this setting are average treatment effects for children on the margin of foster care placement—where the margin varies across the distribution of the unobserved propensity to be removed (and thus investigator leniency). We follow the MTE framework described in Bhuller et al. (2020). We model the observed outcomes as $Y = I \times Y(1) + (1–I) \times Y(0)$. I is an indicator for foster care placement, $Y(1)$ is the potential outcome if children are placed and $Y(0)$ is the potential outcome if children are not placed. The assigned investigator will remove a child ($I = 1$) according to the following choice equation: $v(X, Z) - U > 0$, where $v()$ is an unknown function, X represents the child's observable characteristics, Z is the investigator's stringency, and U is an unobserved continuous random variable. In other words, the investigator's decision to remove is based on three factors: their own stringency, the child's characteristics observed in our data, and the child's characteristics not unobserved in our data. U represents a child's unobserved resistance to foster care placement, meaning children with lower values of U are more likely to be removed conditional on X and Z.

As Bhuller et al. (2020) show, after normalizing $U|X = x$ to be uniformly distributed over $[0, 1]$ for every value of X, then $v(X, Z)$ equals the propensity score of removal, defined as $p(X, Z) \equiv P[D = 1|X = x, Z = z]$. The MTE is defined as $E[Y(1)–Y(0)|U = u, X = x]$. That the MTE depends on U for a given X reflects unobserved heterogeneity in treatment effects. In addition to the assumptions of relevance and exogeneity, estimating the MTE requires the strong assumption of pairwise monotonicity, as well as separability between observed and unobserved heterogeneity in treatment effects. Under these assumptions, the MTE is point identified over the common support of the propensity score. Nevertheless, because the pairwise monotonicity assumption is strong and has recently been scrutinized (Frandsen, Lefgren and Leslie, 2023; Mogstad, Torgovitsky and Walters, 2021; Mueller-Smith, 2015; Norris, 2019), estimates of the MTE should be interpreted with caution.

To characterize the region of common support, we plot the distribution of propensity scores for removal separately for children in our sample who are and are not placed (Panel A; Figure A3). The region of common support for which we can estimate the MTE ranges from 0% to 9%. Because of the small number of observations at the extremes, as is standard in the literature, we trim the top and bottom 1 percentiles of the overlapping sample prior to estimating the MTEs. The vertical dashed lines in the figure show the trimmed region of common support. Within this region, we estimate MTEs using a local instrumental variables approach and a global quadratic polynomial specification. We construct confidence intervals based on 100 bootstrap replications.

Panel B of Figure A3 plots the MTE point estimates on adult convictions by the unobserved resistance to treatment. The estimates are most negative for children with a low unobserved resistance to treatment and rise as the unobserved resistance to treatment increases. This implies that foster care placement reduces adult crime the most for children on the margin of removal for more lenient investigators. At the other extreme, estimates of the MTE are smaller and statistically insignificant for children at the margin of removal for relatively strict investigators. While there is relatively limited support at high levels of the removal tendency, it is likely that children at the margin of removal for even more strict investigators would see increases in adult crime if removed. These estimates show that the decline in removal rates since the Doyle (2008) sample period could have contributed to more positive estimates of the effects of foster care in the current period.[19]

Third, the types of cases that are placed in foster care have also changed over time in ways that may have shifted the complier population. The share of placements due to neglect have increased and placements due to physical abuse have decreased (see Table 1 in Bald et al. (2022*b*)). Consistent with these trends, compliers in Doyle (2008) were more likely to be investigated for abuse, whereas those in our setting were more likely to be investigated

[19]We also use the MTEs to estimate other economically interesting treatment effect parameters. As shown in Heckman and Vytlacil (2007, 2005), all conventional treatment effect parameters can be expressed as weighted averages of the MTEs, including the average treatment effect (ATE), the average treatment effect on the treated (ATT), and the average treatment effect on the untreated (ATUT). Recovering these parameters for the entire population requires full support of the propensity score over the unit interval (Bhuller et al., 2020), which Panel A of Figure A3 shows is not the case in our setting. Therefore, we follow Carneiro, Heckman and Vytlacil (2011) and Bhuller et al. (2020) and rescale the weights so that they integrate to one over our region of common support. We report estimates of the LATE, ATE, ATT, and ATUT in the upper left corner of Panel B. The estimated ATT shows that foster care reduces adult convictions by 28 percentage points for children who are placed, which is nearly identical to the estimated LATE. The estimated ATE is also negative yet smaller in magnitude than the ATT and LATE. Given the relatively narrow region of common support, however, these estimates should be interpreted with caution.

for neglect. Placements due to parental substance use have also increased over time, in part due to the opioid epidemic, and compliers in our setting are disproportionately likely to be investigated for parental substance use.[20] For these reasons, the complier population in our setting appears to be representative of current trends.

2. Differences in sample restrictions: The sample in Doyle (2008) included children ages 4 to 16 who received Medicaid before their investigation. The sample in the current study includes investigated children ages 6 to 16 regardless of Medicaid receipt. To assess whether the differences in the sample selection could explain the differences in findings, we restrict our analysis to students who were eligible for free or reduced price lunch in any academic year prior to the investigation (a possible proxy for Medicaid receipt). We find estimates of foster care placement for this sample that are similar to our main analysis (Table A18). Although we cannot examine impacts for 4- and 5-year-olds because they are too young to appear in our adult crime data, it is extremely unlikely that excluding these children could explain the differences in findings because the reduction in crime in the current study is driven by younger children (Table 5). Therefore, differences in the analysis sample definition do not appear to contribute to the differences in findings.

3. Differences in the outcome measures: Doyle (2008) examines adult crime outcomes up to age 31, whereas our crime outcomes are measured by age 19. The difference in outcome measures is unlikely to explain the different findings. First, this would require a counterintuitive pattern that foster care substantially decreases crime through age 19, but then substantially increases crime thereafter. However, foster care in our setting decreases adult convictions through at least age 21 (Table A10). Second, the shape of the age-crime profile shows that criminal activity tends to peak at ages 17–20, after which criminal activity decreases (Lochner, 2004). Given the lower level of criminal activity in later ages, it is unlikely that an outcome measure which included crime after age 21 would reverse the sign of our estimated treatment effects. Third, there is a positive correlation between criminality

[20]The reported complier characteristics in Doyle (2008) do not include removals due to parental substance use.

earlier in adulthood and criminality later into adulthood (Mueller-Smith, Pyle and Walker, 2022), which provides further evidence against the different outcome measures explaining the divergent findings.

Takeaways: We conclude that policy and practice have changed foster care as an intervention over the last two decades in ways that could plausibly alter its effects compared to the earlier time period. Other differences between studies, such as the samples and outcomes, are far less likely to contribute to the divergent results.

Crucially, the makeup of compliers in our study mirrors foster care trends, which supports the generalizability of our results to current foster care systems. Compliers were more likely to be investigated for neglect than physical abuse, and disproportionately likely to be investigated for parental substance use. This aligns with trends in the last two decades in the type of cases that are placed in foster care (Bald et al., 2022*b*). Compliers in our setting also tended to have the types of experiences in foster care that recent policies have pushed for. For instance, most were placed in a family home as opposed to a group home, spent a relatively short amount of time in care, and reunified with their birth parents. These patterns of the complier population further bolster the generalizability of our study findings to present-day foster care systems.