Foster Families, Group Homes, and Foster Child Outcomes

Cameron Taylor*

September 14, 2021

Abstract

This paper uses exogenous exits of children from foster families as an instrumental variable to estimate causal effects of non-kin foster family placements relative to group home placements on the outcomes of older children entering foster care. Non-kin foster families substantially reduce children's chances of being incarcerated, homeless or having substance abuse referrals at ages 20 and 21 relative to group homes. The difference in family and group home placement can explain a substantial share of the causal effects estimated in the literature on the overall effect of being placed in foster care. The paper then builds and estimates a structural model of foster child placement into families and group homes incorporating the instrument where foster family preferences over child characteristics influence placement and foster child outcomes. Family preferences reduce average foster child outcomes in equilibrium, and policies that subsidize the placement of children such as boys can achieve similar gains in average foster child outcomes as policies that increase the availability of families. These results nuance the policy discussion by showing that while increasing non-kin families can improve foster child outcomes, policies that consider changing the allocation of foster children to non-kin families can achieve quantitatively comparable gains.

*Stanford Graduate School of Business. Email contact: cntaylor@stanford.edu. I am very grateful to Rebecca Diamond and Paulo Somaini for their close guidance and advice on this project. I thank Luis Armona, Nick Bloom, Greg Buchak, Jose Ignacio Cuesta, Matthew Gentzkow, Neale Mahoney, Paul Oyer, and Isaac Sorkin for their helpful comments and feedback. The data used in this paper were made available by the National Data Archive on Child Abuse and Neglect, Cornell University, Ithaca, NY, and have been used with permission. Data from the National Youth in Transition Database (NYTD) were originally collected by the states and provided to the Children's Bureau. Funding for the project was provided by the Children's Bureau, Administration on Children, Youth and Families, Administration for Children and Families, U.S. Department of Health and Human Services. The collector of the original data, the funder, the Archive, Cornell University and their agents or employees bear no responsibility for the analyses or interpretations presented here.

1 Introduction

Every year, child protective service agencies in the U.S remove nearly 200,000 children from their homes and place them into state-sponsored foster care (Children's Bureau, 2016). These children represent some of the most disadvantaged children in the US: 1/3 of older children in foster care will end up homeless at some point in their life and 1/5 will end up incarcerated.[1] Understanding how to improve these children's outcomes is a first-order research question (Doyle and Aizer, 2018) and research on the impacts of removing abused and neglected children from their families and placing them into foster care has found mixed results. Recent research has found positive effects for young girls on test scores (Bald, Chyn, Hastings and Machelett, 2019) and for other children on education and maltreatment outcomes (Gross and Baron, Forthcoming), but most previous focused on older children has found consistent negative results on criminal behavior and earnings (Doyle, 2007b, 2008).

Why does foster care make some children, on average, worse off? With so many abused and neglected children in the US and around the globe, how can we improve foster care? The literature and policymakers have identified that child placement types are a crucial part of foster care and may contribute to its overall effectiveness. While in foster care children can be placed with substitute care families, aimed at simulating a loving and traditional home setting, or in larger group home settings with many other children and professional caretakers. The literature discusses the potential consequences of foster family and group home placements for foster children, with the general perspective that foster families provide better places for abused and neglected children to grow up.[2] Placement types in foster care are also a central part of foster care policymaking throughout the US. The largest child welfare reform in recent years in California, the Continuum of Care Reform, places foster care placements as the central theme of the welfare reforms and most welfare reforms attempt to keep children out of restrictive group home settings.[3] Despite a perceived importance of placing children in foster families relative to group homes for children's welfare, there is a notable shortage of foster families for children across the US (Adams, 2020; Child Welfare Monitor, 2020; DeGarmo, 2019), and many current policies focus on strategies for adding more foster families to the system.[4] Work that looks at how children differ in their outcomes by placement type either focus on cognitive outcomes (Nelson III, Zeanah, Fox, Marshall, T. and Guthrie, 2007) or assess

[1] Author's calculations from the 2011 and 2014 National Youth in Transition Database survey using this paper.

[2] I.e. Barth (2002); Nelson III, Zeanah, Fox, Marshall, T. and Guthrie (2007); Ryan, Marshall, Herz and Hernandez (2008).

[3] The CCR website states:" The Continuum of Care Reform draws together a series of existing and new reforms to our child welfare services program designed out of an understanding [foster children] do best when they are cared for in committed nurturing family homes." (California Department of Social Services, 2021)

[4] For example, the CCR website also states "Resources are being provided to counties to support the development and implementation of creative strategies for supporting, retaining and recruiting quality relative and non-relative resource families." (California Department of Social Services, 2021)

the impacts of these placements on children's outcomes using propensity score methods that rely on selection on observables assumptions (Ryan, Marshall, Herz and Hernandez, 2008). Moreover, there is no work that evaluates the effectiveness of commonly discussed policies, and the effectiveness of these polices may also depend on the type of foster children that foster families choose to care for in equilibrium if foster children differentially benefit from placement with families.

This paper combines an instrumental variables analysis with a structural model of foster care placement to study how older foster children's placement experiences in group homes and non-kin family placements while in foster care influence later incarceration, homelessness, substance abuse, employment and enrollment outcomes, and to evaluate the effect of different foster care policies. The first goal of this paper is to to relax the assumptions used in the literature to measure how families and group homes differentially affect children's outcomes by allowing for correlated unobservables that affect foster children's placements and later outcomes. I utilize administrative foster care data on placements from 2010-2015 linked to survey outcomes given to foster children at ages 17, 19 and 21 to examine how placements affect children's outcomes. I develop a new instrumental variable in the foster care context that shifts around whether a child is placed with a foster family or a group home, and is plausibly exogenous to a child's potential outcomes. In particular, inspired by methods from the market design literature that look for exogenous variation across markets (Agarwal, Hodgson and Somaini, 2020) this paper utilizes the fact that the number of available family slots coming from exits of other children from families is exogenous to a foster child's entry. Exits of other foster children vary due to the timing of a foster child's reunification with their birth family which can be driven by birth family and social worker factors, or being emancipated due to aging out of foster care. I confirm that this instrument strongly shifts placement across counties, accounting for the size of the foster care market in the county, and including county and month by year fixed effects to control for secular differences between counties and over time. The main threat to this instrument is that exits of other children may be correlated with factors that predict entries of children that are more likely to be placed. I develop tests using rich observables in the data and find no evidence to support this threat. I examine the instrument under the setting of heterogeneous treatment effects and provide evidence that the data are consistent with the other assumptions needed to interpret the instrument as providing a valid local average treatment effect estimate (Imbens and Angrist, 1994).

The IV results show that foster families cause better outcomes than group homes for foster children at age 21. On an outcome index that combines employment, enrollment, incarceration, homelessness and substance abuse, foster children gain between 0.97 and 0.99 standard deviations improvement from being placed with families relative to group homes. Children placed in families are estimated to be 24.9 percentage points less likely to experience incarceration, 39.5 percentage points less likely to experience homelessness, and 24.6 percentage points less likely to experience a substance abuse referral. These represent a local average treatment effect and are larger than the corresponding OLS estimates. I show that this is likely due to treatment effect heterogeneity, where children with pre-existing risks are most likely to be compliers and also get the largest treatment effects, and also likely due to measurement error in the endogenous variable.

I undertake a variety of robustness exercises to check the robustness of these IV results. In particular, I examine alternative definitions of the endogenous variable, check robustness to survey attrition, alternative samples of older children, and alternative regression specifications including dropping outlier observations and alternative outcome index definitions. The conclusions from the main specifications are not changed in these alternative analyses.

I further investigate possible mechanisms and correlates of these outcomes by looking at IV estimates on other outcomes in the survey data and examining more details on the foster care experiences of children in my sample. In particular, I show that it is unlikely that the differences in outcomes between children placed in families and group homes are due to incapacitation and the fact that these children remain in families until age 21, suggesting potential behavioral changes. I also show that, as hypothesized by the foster care literature, adult connections are much stronger for children placed in families, suggesting a role for family support. Family placement also causes foster children to take up social services less, suggesting ways that family placement saves the government money while still improving children's outcomes. I quantify these dollar benefits more directly using estimates from the literature on the personal and social costs of incarceration, homelessness and substance abuse treatment, finding that the annual benefits to an extra year of family placement for foster children is between \$5,797.15 and \$6,417.23. The analysis also finds that the cost of placing children in non-kin foster families is substantially less than the cost of placing children in group homes, suggesting substantial overall cost-savings (Barth, 2002).

What are the implications of these results for the previous literature on the treatment effects of the entire foster care experience for an abused or neglected child? I make an explicit comparison with Doyle (2008), who estimates the effects of foster care on incarceration and criminal behavior of older children in Cook County. I set up a simple framework in which the treatment effect of foster care is heterogeneous and depends on the placement setting of a child (family or group home). Combining Doyle (2008)'s estimates, my estimates and data on Cook County foster care placement I find that family placement can explain a large percentage of the negative effects found in Doyle (2008): if all children were placed in foster families as opposed to group homes, the negative treatment effects of foster care on incarceration outcomes would be reduced by 81.3%, though they would still remain negative. Future research could further explore other explanations for some of the negative effects found in foster care.

I also estimate heterogeneous treatment effects for foster children from family placement, looking at how the treatment effects vary with different child demographics and reasons for removal into foster care. The most notable heterogeneity is that boys benefit substantially more from placement with foster families than girls. This is consistent with a large literature on early childhood interventions (Bertrand and Pan, 2013).

To complement the instrumental variable analysis this paper moves beyond the local average treatment effects to estimate the effects of commonly discussed policies while also bringing new policies into the debate. To do this, I develop a model of foster child placement and outcomes. The model incorporates the instrument and allows for treatment effects and selection to depend on both observable and unobservable characteristics of children (Kline and Walters, 2016; Brinch, Mogstad and Wiswall, 2017). The model treats the decentralized process of how children are matched to foster families and group homes as a matching market where foster families have preferences over foster child characteristics and determine the allocation of foster children to families. The model adapts methods from the centralized matching market literature (e.g. Agarwal (2015); Agarwal, Hodgson and Somaini (2020)) to the decentralized foster care context with an eye towards simulating commonly discussed foster care policies.

I estimate the model by using simulated methods of moments (McFadden, 1989; Pakes and Pollard, 1989) to estimate preferences and the control function approach used in papers like Kline and Walters (2016) incorporating the instrument from the instrumental variable analysis. The model results do not find strong evidence for selection on unobservables, either on gains or levels, but does find evidence for selection on observable gains and levels. Families tend to prefer children, such as girls, that are likely to have better outcomes but do not benefit as much from being placed with foster families. Overall, family preferences make average foster child incarceration and homelessness about 2% worse than if children were allocated randomly.

After estimating the model I simulate an increase in the availability of non-kin families. This represents the first quantitative assessment of the benefits to foster children on incarceration, homelessness and other outcomes from an increase in available families in foster care in the US. I find that a 20% increase in the supply of available families in each county in the US would improve older foster children's outcomes by about 5.7% on an aggregate outcome index that includes incarceration, homelessness, substance abuse, employment and enrollment around the age of 21.

Foster care policy allows subsidies paid to depend on children's observable characteristics. For example, older children generally receive a higher subsidy. Motivated by this feature, I also study the effectiveness of a policies that shift the allocation of foster children to foster families based on observables, holding fixed the number of available families. This policy represents a different policy lever to adding families, and can be achieved without recruiting more foster families to the system.I simulate this policy by subsidizing the utility of different foster children on observable characteristics. I find that if older foster boys were subsidized to equate the rate of placement of older foster boys and older foster girls, the average foster child outcome would increase by 2%.

Moreover, this subsidy result is not specific to boys, but is true on observable demographics such as age as well, where subsidizing older children would lead to better average outcomes.

Outside of the foster care context, the results in this paper are also relevant to an extensive economics literature examining how a child's family and home circumstances affect their later social and economics outcomes. This has included looking at neighborhoods (Chyn, 2018; Chetty, Hendren, Kline and Saez, 2014; Kling, Liebman and Katz, 2007; Katz, Kling and Liebman, 2001) and family environments (Sacerdote, 2007; Fagereng, Mogstad and Rønning, 2021). This paper is most related to papers focusing on how family environments influence children's later outcomes. Most of the studies in this literature measure treatment effects on children's later outcomes by family characteristics such as parental wealth or education using adopted children and the characteristics of their adopted parents. The results in this paper add a new perspective on how family environments affect a child's outcomes by looking at the causal effects of families relative to institutionalized settings, a setting tens of thousands of children in the US grow up in every year.

This paper is related to papers child welfare literature that address how different foster care placements affect children's later outcomes. As previously mentioned, the two most similar papers in this literature that focus on group homes are Ryan, Marshall, Herz and Hernandez (2008) and Nelson III, Zeanah, Fox, Marshall, T. and Guthrie (2007). This paper adds an instrumental variable analysis to these papers and focuses on new outcomes for older children in the US, a policy-relevant group at high risk of emancipation, homelessness and incarceration. Another novelty relative to these studies is that this study is the first to the best of my knowledge to study the implications of foster family preferences on child outcomes in equilibrium, and provide an explicit quantitative policy analysis of commonly discussed foster care policies.

This paper is also related to a small but growing literature in economics on foster care, much of which is cited and discussed above (Doyle, 2007b, 2008; Bald, Chyn, Hastings and Machelett, 2019; Robinson-Cortes, 2019; Gross and Baron, Forthcoming). This literature is reviewed by Doyle and Aizer (2018). Most notably this literature focuses mainly on the overall treatment effect of being placed in foster care, while this paper provides causal evidence that the specific type of placement while in foster care matters while also incorporating other equilibrium considerations into policy simulations. One exception is Robinson-Cortes (2019) who builds a structural matching market model of foster care. A key difference between this paper and that study is the outcomes studied, as this paper incorporates welfare relevant economic and social outcomes for children.[5]

Methodologically, the structural model in this paper builds on and combines program evaluation methods (Kline and Walters, 2016) and tools from matching markets (Agarwal, 2015; Agarwal, Hodgson and Somaini, 2020). To my knowledge, the only other papers that look at these types

[5]That paper also studies different policies such as allowing children to be matched across larger geographic regions and increasing the share of certain types of foster homes.

of program evaluation questions in the context of matching markets are van Dijk (2019), Abdulkadiroğlu, Pathak, Schellenberg and Walters (2020) and Agarwal, Hodgson and Somaini (2020), and the only other paper that builds a model of foster care matching is Robinson-Cortes (2019). One difference with this first set of papers is that the instrument used in the analysis here varies at the market level and does not use detailed choice data to identify unobservables in the potential outcome equations. The way that I adapt these methods to an instrument that shifts overall market conditions in an equilibrium context might be useful in other contexts where researchers do not have access to detailed choice data or individual level shifters that affect treatment such as Doyle (2007b, 2008); Bald et al. (2019); Gross and Baron (Forthcoming).

The rest of this paper is organized as follows: Section 2 provides a description of foster care and the data used in the analysis. Section 3 describes the instrumental variable strategy and results. Section 4 describes the model setup. Section 5 describes the model estimation and results. Section 6 discusses the policy counterfactuals. Section 7 concludes.

2 Setting and Data

2.1 Overview of Foster Care and Foster Care Placement in the U.S.

Child protective services are administered at the county-level in the U.S. County officials receive reports of abuse, neglect, or a caretakers inability to care for a child and are tasked with investigating. Officials investigate over 4 million of these allegations every year (Children's Bureau, 2016).

The general investigation process consists of social workers and judges working together to determine whether a child should be removed from their current birth family or guardian. This process can last up to a few weeks but can also be completely sidestepped if a police officer or other official believes the child must be removed from their current circumstances immediately. The literature generally uses detailed data on how foster children are randomly assigned to social workers who then use discretion in deciding whether a child should be removed from their home as an instrument for identifying the overall causal effects of foster care (i.e. Doyle (2007b); Bald, Chyn, Hastings and Machelett (2019); Gross and Baron (Forthcoming)). This paper focuses on the placement step of the process.

When a child enters foster care, there are, broadly, three different options for placement. The first is kin foster family placement. This placement consists of a child being placed with a relative while being formally or informally in foster care. This paper does not focus on kin families and drops children placed with kin families in the analysis. The second is non-kin foster family placement. These placements consist of placements with a family (which can include single adults, unmarried couples) that volunteer their time and house for foster children to remain. They receive basic training and go through an approval process that varies by state. While caring for children they are given a stipend that ranges between $500 and $1000 a month (WeHaveKids.com, 2020). These placements are the most common with 46% of children in foster care recently with non-kin foster families (compared to 32% in kin) (Children's Bureau, 2020).[6] The final placement option is group homes or institutions. In general and in this paper, they are distinguished from "family placements" by the number of children that are present in the placement. In the data analyzed in this paper, placement settings with 6 or more children are considered group homes or institutions. Group homes and institutions generally provide 24-hour care for children and are staffed with adults that maintain care for the children in a professional role. Some examples include residential treatment facilities and maternity homes.

Foster care placement generally works in two steps.[7] The first step, mandated by law in almost all states, is that social workers look for an appropriate relative caregiver to care for the child. This search may fail if the agency or social worker cannot locate a relative that lives close enough, or the relative is unwilling to care for the child. The second step occurs if a child is not placed with a relative. Social workers then attempt to contact an approved non-relative foster family. If one of these foster families has an available slot and is willing to care for the child, then the child is placed with that foster family. If both of these steps fail, then the child is placed in a group home or institution. Group homes are generally seen as an option of last resort.

Children exit foster care in three main ways. The first, and most common, is reunification with their parent or primary caretaker. When their child enters foster care, birth parents work with social workers on a plan for eventual reunification. For example, if a child is removed from their birth parents because they are abusing drugs, the social worker may ask the parents to undergo drug rehabilitation before the child reunites with them. The second is that children are adopted by another family, often their foster parents. The last is that the child is emancipated from the system once they turn old enough and lose eligibility for foster care funding. The first two outcomes are often described as preferred by practitioners and researchers and are referred to as "permanency", though there is some mixed evidence on how these intermediate outcomes translate to longer term outcomes (Unrau, Font and Rawls, 2012).

[6] A large literature looks at how kin and non-kin family placements differ (Font, 2014), and some recent work applies an instrumental variable analysis with payments to assess how in the same dataset used in this paper (Lovett and Xue, 2020).

[7] Details in this section are based on talks with foster care social workers and placement managers in Santa Clara County. These rules are relatively consistent across counties in California. While the data analyzed here encompass a wider range of counties and agencies than just those located in California, these rules generally seem to be consistent nationally.

2.2 Non-Kin Foster Families

This section focuses on documenting facts on non-kin foster families that are important for the empirical strategy used in this paper. Cherry and Orme (2013) document that in foster care there are two types of foster parents. There is a set of "vital few" foster mothers: foster mothers that account for a small proportion of foster parents in the system, and provide a disproportionate amount of care for children. Their analysis finds that 21% of foster mothers cared for 73% of foster children. In their sample, these foster parents fostered on average 104 children over almost 16 years of care. They adopt only 1.6 children on average. Other foster parents foster less but are more likely to adopt, caring for 11 children on average and adopting 0.8 children. It is thus conceivable that the availability of these foster parents that foster over many years could drastically impact a foster child's chances of being placed with a foster family, and that foster children's exits could affect availability of these foster parents.

Foster parents that serially foster may differ in important ways from other families, and these differences may be correlated with differences in treatment effects at the family level. Cherry and Orme (2013) show that these serial fosterers are less likely to work outside the home and have more time to foster, along with more professional support for fostering. These characteristics may make them more or less effective foster parents, and lead to better or worse outcomes for the children, and affect the interpretation of the treatment effects identified in this paper.

2.3 Main Data and Sample

There are two main datasets that are linked for the analysis in this project. Both come from the National Data Archive on Child Abuse and Neglect (NDACAN), funded by the Children's Bureau. The first is the Adoption and Foster Care Analysis and Reporting System (AFCARS) foster care file and the second is the National Youth in Transition Database (NYTD) outcomes file.

The AFCARS data is a federally mandated data collection system maintained to provide case specific information on all children covered by the protections of Title IV-B/E of the Social Security Act. This dataset covers all counties and states in the US, and all children in foster care for whom child welfare agencies in those states have responsibility for care.

The AFCARS data used in this paper contains foster care placement data for every foster child every 6 months between 2010 and 2015. This placement data includes the type of placement (kin, non-kin, group home, institution) every 6 months for children. It also includes demographics for each child (age, sex, and race) and reasons for removal, such as whether the child entered because their parents are in jail, they were abused, neglected, or had a behavioral problem.[8]

[8]Some of these observables, especially removal reasons, are known to be noisy indicators of services provided, but still useful proxies that can predict family placement and subsequent outcomes. To address this, the Appendix includes

The NYTD data contains results of a survey administered to eligible children at the ages of 17, 19 and 21. This paper uses two NYTD cohorts, those 17 in 2011 and those 17 in 2014. Children are eligible for the NYTD survey if they turn 17 years old while in foster care and remain in foster care within the 45-day period following their birthday. The survey asks about outcomes such as incarceration, homelessness, and substance abuse in the past two years. It also asks about whether the child is currently enrolled in school, whether they have a part-time or full-time job, whether they feel connection to an adult and their use of social services such as social security. How the survey results are obtained (electronic surveys, phone calls, etc.) is up to a state's own discretion. The effective survey response rate is 60%. While there is non-response this represents an increase over previous studies that survey adopted children's outcomes across the US (Sacerdote (2007) is able to get a response rate of 34%), but is lower than other studies that utilize county-level administrative data. This non-response rate necessitates addressing non-response bias and attrition, and analysis that assesses non-response bias and attrition is done in the Appendix. .

The main measure for placement used is an indicator for whether the initial placement of the child is with a non-kin family. The alternative is a group home or institutional placement. There are other ways one can measure a child's placement experience in foster care. I choose this as the primary measure for two reasons. First, the instrumental variable relies on market conditions when a foster child enters the system to exogenously shift their placements, and thus should have the most power for initial placements. Second, foster care placements are quite "sticky": initial placements into families are highly predictive of continued placement in a family. Robustness of this analysis to this choice of endogenous variable is assessed in the Appendix where I repeat the main analysis using endogenous variables of the percentage of time in a non-kin placement and months in a non-kin placement.

The main outcome measure is constructed following the Moving to Opportunity literature, which also analyzes survey results from a wide variety of outcomes. In particular, I follow Kling, Liebman and Katz (2007) in creating an economic and social outcome index for each foster child. This index combines whether a child is enrolled or employed, has been incarcerated at ages 20 or 21, has been homeless at ages 20 or 21, and has had a substance abuse referral at ages 20 or 21 into a single index. The summary index is defined to be the equally weighted average of z-scores of its components, where the sign of each component is set up so that more beneficial outcomes have higher scores (i.e. it is increasing in enrollment/employment, decreasing in incarceration). The z-scores are calculated by subtracting the control group mean and dividing by the control group standard deviation. This summary index is a useful way to maximize power and focus on a single outcome in the face of many different outcomes. The main results also break out the results

analyses that only include demographic child controls. Waldfogel (2000) discusses the benefits of the new AFCARS data and how it should assist in understanding important issues in child welfare and foster care through data.

into individual outcomes. Cost-benefit analyses that put a rough dollar value amount based on these outcomes are also done below to provide a concrete benchmark for the benefits of family placement.

The analysis is conducted at the child entry level in the AFCARS data with outcomes measured at age 21 in the NYTD data so that each observation is a unique child-entry and outcome at 21 pair. Only children placed in non-kin families, group homes or institutions upon entry are considered. If a child has multiple entries, I take only their latest entry. One challenge in the analysis is that children eligible for the survey are those that are in care at age 17, and children that make it to age 17 in foster care are a selected sample among all foster children. To make my analysis sample as representative of a foster care population of interest and reduce sample selection among this population as much as possible, I only consider children whose latest entry occurred at age 14 or older. This makes the sample more representative of "older" foster children and removes children that enter very young but linger in foster care for a long time. Those children may be substantially different on unobservables than other older children in the sample. Robustness of the main results to different age cutoffs (ages 12, 13, and 15) are included in the Appendix and show that the choice of the cutoff is immaterial to the main results. Finally, because the instrumental variable strategy used in the analysis in this paper requires knowing a child's county of removal, children without an identified county of removal are dropped. Some small counties are not included in AFCARS because of privacy concerns.

The final analysis sample, the "Outcome sample" consists of 5,113 foster children. To allow for a better understanding of potential attrition issues in the sample and allow for a better understanding of how the effects measured in this paper generalize to other foster care populations I compare this analysis sample to two other samples. The first is the set of children eligible to be included in the first wave of NYTD survey that also follow the same requirements described above i.e. initially placed in a non-kin family or group home, between ages 14 and 17 when entering, and having an identified county of removal. I call this the "Eligible sample".[9] I also include analysis of a sample of all foster children entering into non-kin family homes or group homes between ages 14 and 17 with an identified county of removal. This sample represents the general population of older foster children and is called the "Old Children sample".

Some descriptive statistics on the endogenous variables, main demographic covariates, and outcome variables for each sample are provided in Table 1. Half of the children with valid outcomes are placed with non-kin families while in foster care. This proportion is higher than the corresponding proportion in both the eligible sample and old children sample. The difference with

[9] Note that the difference in the sample sizes between the Eligible and Outcome sample shown in Table 1 does not reflect the effective response rate since some states further randomly sample some percentage from the set of eligible children. The response rate reported in the text does not reflect this extra random sampling, reflecting only children who states attempted to reach and did not respond.

the eligible sample highlight the importance of correcting for response bias and attrition in the survey. The difference with the old children sample is mainly driven by age differences between the samples. The old children sample is far more balanced on the age distribution, while most children in the outcome and eligible samples enter at age 16. The only other notable comparison between the samples is that there are substantially less boys with outcomes than in the eligible sample or old children sample. This is likely because boys are more likely to be incarcerated or homeless, and so are harder for each state to survey between ages 20 and 21. This type of bias will be accounted for in the attrition analysis below.

The main outcomes are also shown for the outcome sample. The index created has a mean of 1.01 and a standard deviation of 2.08. About 69% of foster children surveyed are employed or enrolled at the time of the survey, 23% have been incarcerated at some point between the ages of 20 and 21, 32% have been homeless at some point between the ages of 20 and 21, and 13% have had a substance abuse referral between the ages of 20 and 21. The children studied in this paper are likely to have poor economic and social outcomes.

2.4 Supplemental Data

I supplement the main AFCARS and NYTD data with NYTD services data which provides information on the services provided to foster children such as academic support, career preparation services and room and board financial assistance, and also measures their education at different points in time.

3 Placement Instrument and Regression Analysis

3.1 Research Design: Children Exiting Non-Kin Families as a Placement Instrument

Consider the following model for outcomes of older foster children:

$$Y_i = \beta \cdot \text{Non-Kin Foster Fam Place}_i + X_i\gamma + \epsilon_i \tag{1}$$

where Y_i is an outcome at age 21, Non-Kin Foster Fam Place$_i$ is an indicator variable for whether a foster child is initially placed with a non-kin family (a 0 means they are placed with a group home or institution), and X_i are child demographics and entry reasons. If family placement is correlated with other unobserved determinants of child outcomes ϵ_i then estimates of this equation will be biased.

Table 1: Summary Statistics: Sample Means

	Outcome Sample (1)	Eligible Sample (2)	Old Children Sample (3)
Initial placement with non-kin family	0.500	0.432	0.379
Sex: male	0.420	0.507	0.524
Race: black	0.300	0.324	0.321
Race: white	0.443	0.426	0.418
Race/ethnicity: hispanic	0.206	0.200	0.198
Age at entry: 14	0.120	0.0980	0.215
Age at entry: 15	0.284	0.247	0.268
Age at entry: 16	0.532	0.556	0.295
Age at entry: 17	0.0630	0.0986	0.222
Economic and social outcome index	1.01 (SD = 2.08)	-	-
Currently employed or enrolled	0.687	-	-
Incarceration ages 20-21	0.225	-	-
Homeless ages 20-21	0.321	-	-
Substance abuse referral ages 20-21	0.127	-	-
Number observations	5,113	18,461	209,075

Notes: This table provides means of variables across three different samples. The outcome sample is defined as children that have either valid incarceration, homelessness, substance abuse referral, employment or enrollment outcomes in the survey at age 21, are placed in a group home or non-kin family home for their first placement, and have their latest entry between ages 14 and 17. The eligible sample is defined as all children that were eligible for the survey at age 17, are placed in a group home or non-kin family for their first placement, and have their latest entry between ages 14 and 17. Note that the difference in the number of observations of the outcome and eligible sample does not reflect true attrition, since children surveyed at age 21 must have responded at age 17. The old children sample is all foster children that are placed in a group home or non-kin family home for their first placement and entering between ages 14 and 17.

To address this issue I utilize an instrumental variable (IV) strategy. Children may not be placed with a non-kin family in foster care if there are not enough families available in the same county around the time of their arrival. In over 95% of the counties that I study, some older children are placed in groups homes and institutions. This paper proposes that one source of availability that is exogenous to a child's potential outcomes is the exits of other foster children from non-kin foster families. If children exit placements of families that continue to foster, as the "Vital Few" described above do (Cherry and Orme, 2013), then those families may now be willing and able to care for entering foster children. If these exits satisfy certain assumptions, outlined below, then they can serve as an IV to measure a Local Average Treatment Effect (LATE) of placement into a non-kin foster family relative to a group home or institution (Imbens and Angrist, 1994).

In order to measure these types of exits, for every county-month-year t, I count the number of exits from non-kin placements that don't end in the child being adopted or the foster parent taking guardianship of the child, Exits_t. Adoptions and guardianship are less likely to represent true slots opening up int hese families. To account for between county differences in exits and secular seasonal trends, county and month-year fixed effects are included as controls in the IV model. Further the main specification normalizes the number of exits by the log population to account for the fact that some counties will have higher deviations of non-kin exits due to their population.

This instrument is then utilized in a two-stage least squares (2SLS) framework to estimate a causal effect:

$$Y_i = \beta \cdot \text{Non-Kin Foster Fam Place}_i + X_i\gamma + \delta_{c(i)} + \delta_{m(i)} + \epsilon_i \quad (2)$$

$$\text{Non-Kin Foster Fam Place}_i = \alpha \cdot \text{Exits}_{c(i),m(i)} + X_i\Gamma + \Delta_{c(i)} + \Delta_{m(i)} + \nu_i \quad (3)$$

where i is a child index, $c(i)$ is the county that child i enters into, $m(i)$ is the month-year (ex: December 2013) that child i enters. This framework includes child controls X_i, county fixed effects $\delta_{c(i)}$, $\Delta_{c(i)}$ and month-year fixed effects $\delta_{m(i)}$,$\Delta_{m(i)}$.

3.2 Instrument Variation

Figures 1 and 2 provide visualizations of the raw variation in the instrument. Figure 1 plots the residual non-kin exits variation after controlling for county and month by year fixed effects within the four largest counties in the data. Because the instrument is normalized by log county population there is not an obvious interpretation of these differences. The deviations from 0 in each of these figures for each county provide the identifying variation of the instrument.

Where does the identifying variation from this instrument come from? This identifying variation is best described as driven by idiosyncratic variation in non-kin exits within counties away from general seasonal trends that could be due to variation in birth families finishing court-mandated rehabilitation or other action items in their case plan,[10] or aging of children in non-kin exits out of the foster care system.

Figure 2 plots the raw instrument against the raw endogenous variable at the county-by-month-year unit. This figure shows that first, in the largest county, Los Angeles, there is a clear positive

[10]More information on how families can expect to be reunified with a child placed in foster care can be found here: https://www.childwelfare.gov/pubPDFs/reunification.pdf (accessed August 3, 2021).

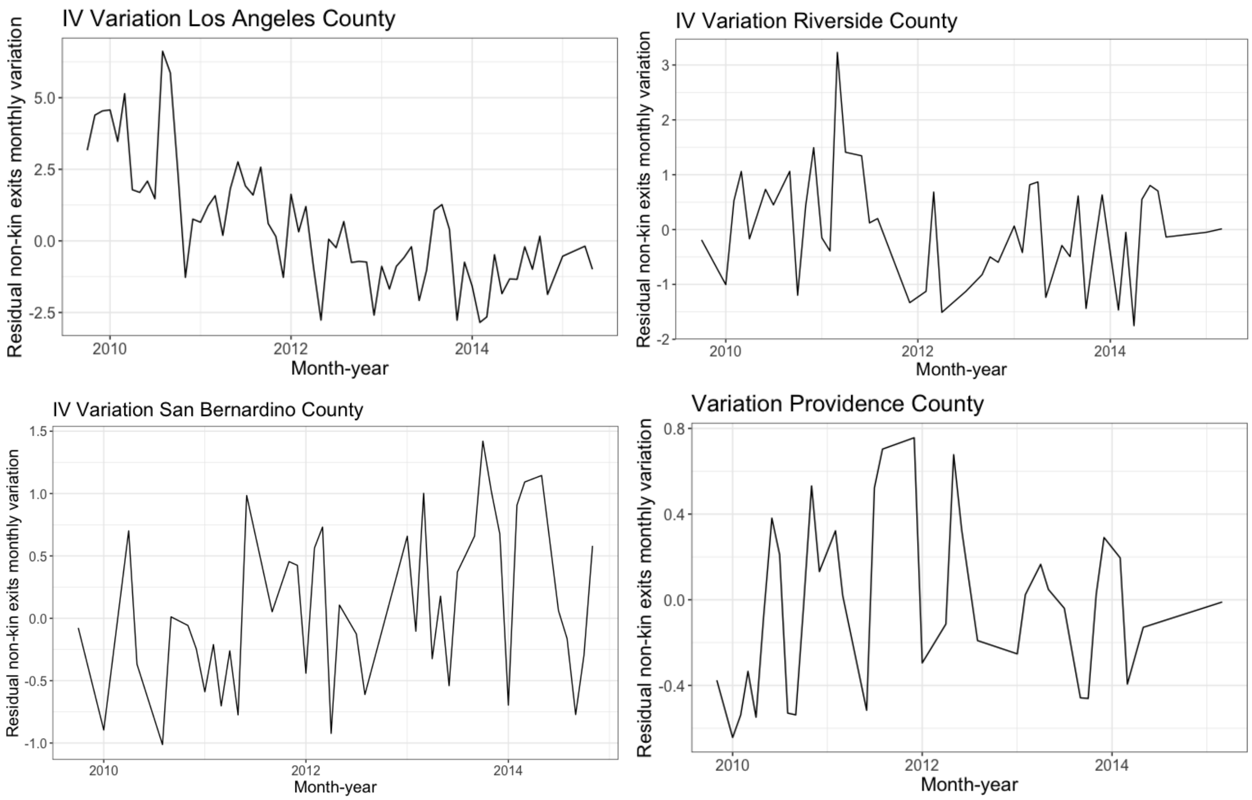

Figure 1: Time-series for IV within County

Notes: These figures plot $\widetilde{\text{Exits}}_m$ which looks at the residual of the Exits instrument defined in the text on county and month-year fixed effects defined over month-years m for the four largest counties in the outcomes data.

relationship between placements and exits in the raw data. Second, when generalized to all counties in the data, this relationship continues to hold.

3.3 Instrument Validity: Conditional Independence, Exclusion Restriction and Monotonicity

For non-kin exits to be a valid instrument in the presence of heterogeneous treatment effects it must satisfy four requirements: relevance, conditional independence, exclusion, and monotonicity. I now discuss each of the latter three assumptions and associated tests to provide evidence that each is satisfied.

The validity of the instrument depends on whether non-kin exits in the month of entry are uncorrelated with child characteristics that could affect a children's future outcomes, conditional on county and month-year fixed effects. This assumption is not testable, but I discuss two main ways in which this assumption is likely to be violated, and discuss how the methodology and data provide evidence that the instrument satisfies this conditional independence.

One potential source of endogeneity is seasonality. Exits from families may be predictable

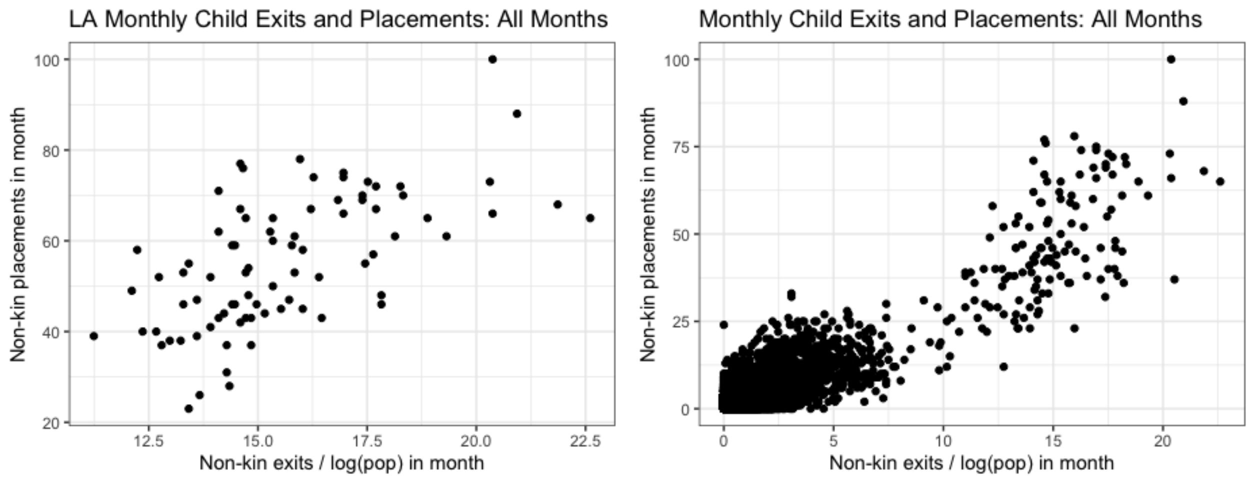

Figure 2: Raw Correlations for IV

Notes: These figures plot the raw instrument Exits$_t$ at the county-month-year level against total non-kin placements at the same level. The left panel shows only Los Angeles County and the right panel includes all counties included in the outcome sample.

based off of seasonal patterns such as the beginning and end of school year, or transitions in the fiscal year. Unobservable characteristics of entering children may also be correlated with this timing, threatening validity. This strategy accounts for secular seasonal patterns with month-year fixed effects.

However, even controlling for seasonality, exits might be correlated with types of children entering for other reasons. While my main models control for a rich amount of child level observables, it is possible that there are unobservables of children that are correlated with outcomes and the level of non-kin exits in the county at the month-year time horizon. For example, there may be county-specific trends in exits in which more exits signify "good times" for a county, and so children entering may generally be more likely to have good outcomes. I develop two tests to examine whether there is evidence in the data of such correlated unobservables.

The first test examines the correlation between the instrument and the set of child-level observables in the data. This test regresses the instrument on these observables, including the county and month-by-year fixed effects. I also compare these correlations with the correlations between family placement and the same observables to show that these observables are indeed informative of family placement.

The results from this test are in Table 2. This test is done in all three of the main samples constructed. Table 2 also contrasts the instrument regression with regressions of the endogenous variable (family placement) on the same variables. This table shows that there is little correlation between the instrument and child observables in all three samples considered. In particular the p-values for the F-test testing the null hypothesis that all coefficients in the regression are zero are above 0.05. Moreover, the coefficient sizes in columns (1) - (3) are very small compared to the

Table 2: Instrument and Endogenous Variable Correlation with Observables

	Instrument: Non-Kin Exits Month / log(Population)			Endogenous Variable: Initial Placement with Non-Kin Family		
	Outcome Sample	Eligible Sample	Old Children Sample	Outcome Sample	Eligible Sample	Old Children Sample
	(1)	(2)	(3)	(4)	(5)	(6)
Sex: male	−0.022	−0.014	−0.011	-0.114^{***}	-0.129^{***}	-0.114^{***}
	(0.018)	(0.011)	(0.007)	(0.012)	(0.007)	(0.004)
Race: white	−0.053	0.027	−0.151	−0.038	-0.042^{**}	-0.013^{*}
	(0.032)	(0.023)	(0.161)	(0.034)	(0.018)	(0.007)
Race: black	−0.039	0.032	−0.162	-0.071^{**}	-0.053^{***}	-0.020^{***}
	(0.035)	(0.024)	(0.184)	(0.035)	(0.018)	(0.007)
Race: hispanic	−0.053	0.020	−0.145	−0.021	−0.020	0.013^{*}
	(0.034)	(0.023)	(0.143)	(0.034)	(0.018)	(0.007)
Age: 15	−0.049	-0.027^{*}	−0.002	−0.014	-0.037^{**}	-0.028^{***}
	(0.051)	(0.016)	(0.003)	(0.030)	(0.015)	(0.003)
Age: 16	−0.062	−0.008	−0.004	−0.018	-0.045^{***}	-0.039^{***}
	(0.050)	(0.019)	(0.003)	(0.030)	(0.016)	(0.004)
Age: 17	−0.004	−0.030	-0.009^{*}	−0.039	-0.064^{***}	-0.049^{***}
	(0.040)	(0.021)	(0.005)	(0.038)	(0.016)	(0.006)
Physical abuse	0.002	0.026	0.015	0.107^{***}	0.076^{***}	0.093^{***}
	(0.029)	(0.038)	(0.013)	(0.020)	(0.015)	(0.009)
Sexual abuse	0.068	0.001	−0.004	0.038	0.034^{**}	0.064^{***}
	(0.046)	(0.030)	(0.015)	(0.027)	(0.016)	(0.008)
Neglect	0.063	0.034	0.031	0.132^{***}	0.109^{***}	0.112^{***}
	(0.045)	(0.038)	(0.026)	(0.020)	(0.013)	(0.011)
Parent alcohol abuse	0.041	0.0003	-0.029^{*}	0.022	0.086^{***}	0.072^{***}
	(0.044)	(0.024)	(0.016)	(0.036)	(0.020)	(0.008)
Parent drug abuse	−0.028	-0.037^{**}	-0.030^{*}	0.083^{***}	0.076^{***}	0.081^{***}
	(0.026)	(0.015)	(0.017)	(0.029)	(0.013)	(0.010)
Child alcohol abuse	−0.044	0.003	-0.037^{*}	-0.092^{*}	-0.043^{*}	-0.036^{***}
	(0.045)	(0.026)	(0.021)	(0.047)	(0.024)	(0.009)
Child drug abuse	0.043	0.009	−0.008	-0.061^{*}	-0.085^{***}	-0.095^{***}
	(0.039)	(0.028)	(0.022)	(0.036)	(0.015)	(0.009)
Child disability	−0.020	−0.013	−0.035	−0.048	-0.068^{***}	-0.069^{***}
	(0.048)	(0.026)	(0.024)	(0.037)	(0.025)	(0.015)
Child behavior problem	−0.054	-0.053^{*}	-0.047^{*}	-0.265^{***}	-0.242^{***}	-0.253^{***}
	(0.046)	(0.028)	(0.026)	(0.025)	(0.016)	(0.014)
Parent(s) died	−0.105	-0.081^{**}	−0.017	0.063	0.115^{***}	0.159^{***}
	(0.083)	(0.040)	(0.012)	(0.061)	(0.037)	(0.013)
Parent(s) jail	−0.066	−0.011	0.0005	0.025	0.053^{**}	0.063^{***}
	(0.051)	(0.018)	(0.010)	(0.039)	(0.021)	(0.007)
Inability to cope	0.014	−0.016	0.004	0.049^{**}	0.067^{***}	0.074^{***}
	(0.024)	(0.017)	(0.011)	(0.020)	(0.011)	(0.007)
Abandonment	0.001	0.009	−0.004	0.037	0.031^{*}	0.037^{***}
	(0.024)	(0.019)	(0.010)	(0.029)	(0.016)	(0.008)
Relinquished	0.042	0.064	0.015	0.127^{***}	0.079^{***}	0.081^{***}
	(0.049)	(0.040)	(0.018)	(0.040)	(0.021)	(0.017)
Housing problem	−0.019	−0.001	−0.015	0.084^{***}	0.059^{***}	0.063^{***}
	(0.049)	(0.022)	(0.028)	(0.030)	(0.015)	(0.008)
Number observations (children)	5,113	18,461	208,808	5,113	18,461	209,075
Mean outcome variable	1.98	1.83	1.74	0.5	0.432	0.379
R^2	0.976	0.968	0.954	0.440	0.331	0.296
F-statistic (p-value)	0.741	1.14	1.409	21.86	46.97	106
	(0.799)	(0.2955)	(0.0989)	(<0.001)	(<0.001)	(<0.001)
County, month x year fes	Y	Y	Y	Y	Y	Y

Notes: Columns (1)-(3) report OLS regression results from regressing the instrument, normalized by log population, on all child demographics and entry reasons. Columns (4)-(6) report OLS regression results from regressing the endogenous variable, initial placement in a non-kin family, on all child demographics and entry reasons. F-statistics are for statistical tests where the null hypothesis is that all coefficients on observables are 0. See Table 1 and the text of the paper for descriptions of the different samples. The instrument is not defined for some very small counties in the old children sample, explaining the discrepancy between the number of observations in columns (3) and (6). Standard errors are clustered at the county level. *p<0.1; **p<0.05; ***p<0.01

Table 3: Instrument Correlation with Non-Kin and Kin Placement

	Placement with Non-Kin Family			Placement with Kin Family		
	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome Sample	Eligible Sample	Old Children Sample	Outcome Sample	Eligible Sample	Old Children Sample
Instrument: non-kin exits / log(pop)	0.022***	0.014***	0.012***	0.005	0.002	−0.002
	(0.005)	(0.003)	(0.002)	(0.006)	(0.002)	(0.002)
Mean outcome variable	0.420	0.368	0.313	0.160	0.147	0.175
Number observations (children)	6,088	21,638	252,960	6,088	21,638	252,960
County, month x year fes	Y	Y	Y	Y	Y	Y
Child demographic, entry controls	Y	Y	Y	Y	Y	Y

Notes: Columns (1)-(3) give coefficient estimates on the instrument for a regression of placement with non-kin family on the instrument, demographic and entry reason controls and county and month-year fixed effects. Columns (4)-(6) do the same with a regression of placement with kin family. The samples in all columns are the same as in Table 2 but also include foster children whose initial placement is with a kin family. Standard errors are clustered at the county level. *p<0.1; **p<0.05; ***p<0.01

coefficient sizes in columns (4) - (6) although the mean of the instrument is much larger than the mean of the endogenous variable, suggesting the economic relationships between the observables and the instrument are very small.

The second test consists of seeing whether the instrument predicts more kin family placements. The logic of this test is the following: if the unobservable characteristics of children entering when non-kin exits are such that they are children that are easier to care for in general, i.e. the instrument is just a proxy for child quality, then these children should also be more likely to be accepted and placed in kin families.

The results from this test are shown in Table 3. The results show that while non-kin placement is highly correlated with the instrument, kin placement is not correlated with the instrument. The economic magnitude of the coefficient is small and it is not statistically significant.

The second assumption required to interpret the IV estimates as a causal effect of family placement is the exclusion restriction.[11] The exclusion restriction in this context requires that non-kin exits should affect children's outcomes only through family placement. One challenge to this assumption is that non-kin exits may also signify other changes in the foster care system such as changed attention or effort of social workers, or services available to the foster children. Another challenge is that non-kin exits may shift both whether a child is placed with a foster family, and the number of children in an average foster family placement.

To tackle the challenge that exits may predict other changes in the foster care system that make children better off I look in the NYTD services data to measure services received around a child's

[11] If this assumption is not satisfied, I can still interpret the reduced form impact of changes in other foster children's exits on a foster child's outcomes as causal.

Table 4: Correlations between Instrument and Services Received at Entry

	Coefficient on instrument (1)	p-value (2)	Outcome mean (3)	Number observations (children) (4)
Special education services	-0.0042 (0.0027)	0.116	0.188	28,589
Independent living needs assessment	-0.0207 (0.0112)	0.0642	0.487	28,589
Academic support services	-0.0195 (0.0124)	0.115	0.501	28,589
Career services	-0.0274 (0.0193)	0.156	0.295	28,589
Employment vocational services	-0.0301 (0.0211)	0.156	0.144	28,589
Financial management services	-0.0278 (0.0162)	0.0876	0.283	28,589
Housing education and management	-0.0265 (0.0210)	0.207	0.329	28,589
Health education	-0.0262 (0.0133)	0.0492	0.364	28,589
Mentor services	-0.0305 (0.0160)	0.0565	0.168	28,589
Educational financial assistance	-0.0457 (0.0268)	0.0887	0.0858	28.589
Other financial assistance	-0.0432 (0.0290)	0.137	0.167	28,589
Instrument	non-kin exits / log(pop)			
County, month x year fes	Y			
Child demographic, entry reason controls	Y			

Notes: Each row of this table is associated with a separate regression of a different service outcome on a child entry. Each of these regressions includes demographic, entry reason controls, and county and month by year fixed effects.The sample for each regression is all children entering between 14 and 17 years old receiving any services.

entry and correlate services received with the instrument. The results from these correlations are in Table 4. There are no statistically significant (at the 5% level) correlations between these services and the instrument. Moreover, the coefficient sizes are also small relative to the outcome means, and estimated to be negative, suggesting that if anything, non-kin exits suggest less service delivery. Thus I do not find evidence that service delivery is significantly higher when there are more exits, and the data seem inconsistent with the fact that other parts of the foster care system work in foster children's favor when there are more exits.

To tackle the challenge that the instrument effects could be driven by the intensive margin instead of the extensive margin I directly measure in the data the number of children per placement. This is a noisy exercise because there are no dedicated family identifiers in the AFCARS. Instead, I rely on identifying unique families according to their county, their family structure, the age of the primary caretaker and secondary caretaker, and the race of the primary caretaker and secondary caretaker. This still leaves families with an invalid number of children cared for at a time in larger counties and so I only look at families that are estimated to have at most 6 children at a time,

Table 5: Correlations between Instrument and Number of Children in Family Placement

	Number Children in Family Placement (1)	Indicator for More Than 1 Child in Family Placement (2)
Non-kin exits / log(pop)	0.1327 (0.0501)	0.0432 (0.0218)
County, month x year fes	Y	Y
Child demographic, entry reasons	Y	Y
Children placed with families only	Y	Y
Mean outcome	2.25	0.553
Number observations (children)	2,071	2,071

Notes: Column (1) provides the coefficient estimate on the instrument for a regression of number of children estimated in a child's initial placement for children from the outcome sample placed with a family who also have a valid measure of number of children in placement. A family has a valid number of children in their placement if, after accounting for the sequential arrival and exit of foster children in the AFCARS data, they have 8 or less children in their care. A family is identified by a unique sequence of county, family structure, age of primary caretaker, age of secondary caretaker, race of primary caretaker and secondary caretaker. Column (2) provides the coefficient estimate on the instrument for a regression of an indicator of having more than 1 child in a placement. Standard errors are clustered at the county level.

as this is the cutoff for group homes given by AFCARS. I then measure how many children are estimated to be in a family placement for entering children with outcomes in my data and look at the correlation between the number of children in these placements and my instrument, for those children that are placed in families only.

The results for this exercise are in Table 5. Table 5 shows that the instrument is positively correlated with the number of children in family placement given placement in a family. This is consistent with heterogeneous foster families if the average number of children in serial foster families captured by the instrument is higher than other non-kin foster families. These results show that the data are inconsistent with children being pushed into families that are smaller and thus better for that reason. Instead, it suggests that the instrument is truly capturing an extensive margin effect of shifting children in and out of families in the presence of heterogeneous families.

The final assumption to test for the validity of the instrument with heterogeneous treatment effects is monotonicity. I follow the literature in testing this assumption by confirming that the first stage is positive and significant in various subsamples in the data (Bhuller, Dahl, Løken and Mogstad, 2020; Dobbie, Goldin and Yang, 2018). Appendix Tables A2 - A9 include first stage coefficients, standard errors and cluster robust F-statistics for 32 different subsamples of the outcome sample based off child demographics and entry reasons.[12] In all subsamples except for 2, the estimated coefficient is positive. The negative coefficients in the 2 differing subsamples are not estimated to be statistically significant. Moreover, most of the positive coefficients are statistically

[12] When the sample sizes are too small, less than 250, for a subgroup, they are left out of this exercise.

significant with cluster robust F-statistics above 10. Theses tables also includes first stage coefficients in the larger eligible sample. These findings generalize to this sample as well. I address the relevance condition and assess the first stage in the full sample in the next subsection.

3.4 Instrument Relevance and the First Stage

I now analyze the first stage of the model equation (3). To make the first stage interpretable, I first show the variation and coefficients at an aggregated county-month-year form. Figure 3 plots the a regression spline model of the first stage in a sample of county-month-years. The figure also shows the weighted density of the instrument, where the weights are the number of children in the corresponding county, so that counties with more children are weighted more highly. This figure shows a strong relationship between the (residualized) instrument and the endogenous variable at the county-month level. Moving across the support of the residualized instrument, the probability of placement rises from about 0.35 to about 0.75.

I also give the coefficients of this county-month-year level regression in Table 6. I show both a non-weighted and weighted version of this regression, where weights are determined by the number of children in the county. Both regressions show strong cluster robust F-statistics. These coefficients are also easily interpretable. The 0.0033 coefficient in column (1) of Table 6 can be interpreted as saying that if there are 10 extra non-kin exits than normal in a county-month-year, then the percent of entering foster children that is expected to be matched with non-kin families as opposed to group homes increases by 3.3 percentage points.

Estimates of the coefficient α on the instrument corresponding to the details of the first stage equation (3) are provided in Appendix Table A1. Table A1 shows that non-kin exits in the same county and month as a child entry is strongly correlated with the placement of that child in a non-kin foster family. In the outcome sample, the preferred instrument specification (exits / log(pop)) has an F-statistic well over 10 in the outcome sample, the eligible sample, and the old children sample. The old children sample first stage is closer to the outcome sample first stage when weighted by county representation in the outcome sample due to heterogeneous survey response rates and sampling of survey participants across counties.

This strong correlation is robust across changes in the instrument specification, and the samples in which the instrument is defined. I include specifications that do not normalize by population, that control for the total entries of children in the same month (an alternative way to control for the market size), and that look at log of one plus the raw exits. Of these instrument specifications in the outcome and eligible samples, the only one that does not have a strong first stage is log non-kin exits. Table A1 shows though that this is due to county representation, as the old children sample has a strong first stage, which is diminished if weighted towards the outcome sample.

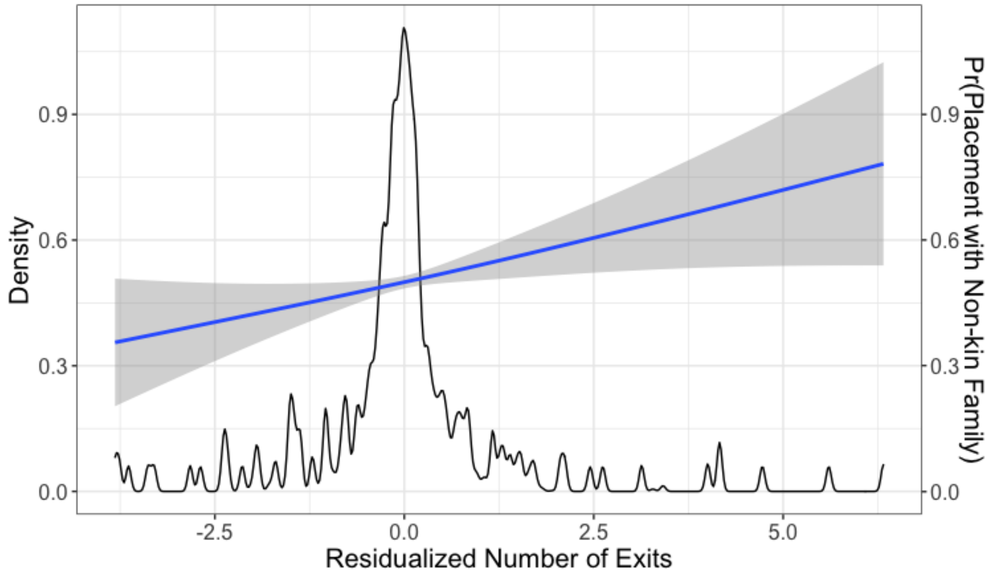

Figure 3: First Stage Variation

Notes: This figure shows the first-stage of non-kin foster family placement (vs. group home placement) on exits from non-kin families on the aggregated county-month-year sample (4,129 total observations). The x-axis plots the residualized number of exits divided by log population, residualized on county and month by year fixed effects. The y-axis on the right gives the probability of placement in a non-kin family. A generalized additive model with penalized regression splines is plotted along with 95% confidence bands. The density plot with y-axis on the left is a weighted density of the residualized number of exits divided by log population, where weights are given by the number of children in the corresponding county.

Table 6: First Stage Coefficients and F-Statistics

	% Children Placed in Non-Kin Families	
	(1)	(2)
Number non-kin exits	0.0033	0.0031
	(0.0008)	(0.0005)
Cluster robust F-statistic	16.2	40.7
Weighted by number children in county	N	Y
County, month x year fes	Y	Y
Mean dep var	0.500	0.514
Number observations (county - month - year)	4,129	4,129

Notes: This table shows OLS regressions of the endogenous variable of percent of children placed in non-kin families on the raw instrument, number of exits, across county-month-year cells. Column (2) further weights these regression results by the number of total children in the corresponding county. Standard errors are clustered at the county level.

Overall, Figure 3, Table 6, and Table A1 suggest that the first stage of the instrument is strong enough to satisfy the relevance condition of the instrument. In particular, the F-statistics are well over the standard thresholds cited in the literature for weak instruments including Stock and Yogo (2005) and Olea and Pflueger (2013).

3.5 Compliers

If the assumptions of the empirical strategy discussed hold, then our IV estimates a LATE of non-kin family placement for children relative to placement in a group home or institution for foster children who would have received a different placement in foster care if there were more or less slots with non-kin families available due to more or less non-kin exits at the time of their entry. To get some insight into this population based off of observables, I follow previous literature by comparing the first stage coefficients of different subsamples to the first stage coefficient of the full sample. Abadie (2003) shows that with a binary instrument, the ratio of these coefficients provides the relative likelihood that a complier has a certain observable characteristic. This procedure is performed and extended in Dahl, Kostol and Mogstad (2014). For simplicity, I keep the instrument in the continuous form here and compare the coefficients. The results are qualitatively similar when the instrument is transformed to a binary version.

Tables A2 - A9 provide the first stage coefficients for subsamples by demographics and entry reasons. Table 7 provides a condensed form of this table that focuses on the subsamples with the strongest first stage (the largest first stage coefficient) with column (1) providing the first stage coefficients and F-statistics. By the method described above, compliers have the highest likelihood of belonging to these subsamples. In particular, it suggests that the relative likelihood that a complier is a child that is removed due to housing problems (homelessness or other inadequate living conditions) or drug abuse problems is extremely high. The relative likelihoods that a compliers is age 15 and that a child does not enter for neglect are also high. Tables A2 - A9 also show that there is not a substantial difference in the likelihood of a complier being a boy relative to a girl, and that compliers are more likely to be Hispanic or White than Black. First stage results across other entry reasons such as abuse type are relatively uniform.

Thus it appears that compliers are substantially more likely to be homeless or have drug abuse problems at the time of entry, more likely to be age 15, and more likely to not have neglect as a reason for entry. These characteristics can help us interpret the LATE and will be useful in comparing the LATE estimate to the OLS estimates in the heterogeneous treatment effect framework.

Table 7: Condensed Complier Table

	First Stage	Reduced Form	IV	OLS	First Stage Eligible Sample
	(1)	(2)	(3)	(4)	(5)
Full Sample					
Coefficient and s.e.	0.0454 (0.006)	0.0916 (0.022)	2.016 (0.513)	0.886 (0.067)	0.0193 (0.003)
Cluster robust F-statistic	64.6	-	-	-	36.1
Number of children		5,113			18,461
Subgroup: Drug Abuse Child					
Coefficient and s.e.	0.256 (0.109)	1.093 (0.799)	4.274 (1.834)	0.916 (1.011)	0.0554 (0.018)
Cluster robust F-statistic	5.5	-	-	-	9.8
Number of children		214			1,103
Subgroup: Housing Problems					
Coefficient and s.e.	0.215 (0.038)	0.594 (0.301)	2.768 (1.127)	2.234 (0.923)	-0.0477 (0.016)
Cluster robust F-statistic	31.7	-	-	-	8.6
Number of children		252			979
Subgroup: Age 15					
Coefficient and s.e.	0.061 (0.015)	0.083 (0.055)	1.357 (0.815)	0.959 (0.131)	0.0435 (0.007)
Cluster robust F-statistic	17.0	-	-	-	38.7
Number of children		1,454			4,560
Subgroup: No Neglect					
Coefficient and s.e.	0.0504 (0.0076)	0.147 (0.035)	2.917 (0.754)	0.891 (0.095)	0.0334 (0.00498)
Cluster robust F-statistic	44.4	-	-	-	44.8
Number of children		3,109			11,688

Notes: This table presents first stage, reduced form (ITT), instrumental variable and OLS regression results for different subsamples of the outcome sample. It also includes the first stage regressions in the eligible sample. All models include county and month by year fixed effects. Standard errors are clustered at the county level. The full sample is the entire outcome sample. The drug abuse child subsample is children that enter at least in part due to their use of narcotics. The housing problems subsample is children that enter at least in part due to inadequate housing, including homelessness. The age 15 subgroup is children whose entry is at age 15. The no neglect subsample is children who do not enter because of a failure to provide adequate food, clothing shelter or care.

3.6 Main Reduced Form Results: Effects of Non-Kin Foster Family Placement vs. Group Homes on Outcomes

Table 8 contains the LATE estimates of the effect of family placement for older foster children on outcomes measured at age 21. It also compares the estimates to the OLS estimates of the same treatment effects. Columns (1) to (4) include OLS and IV estimates of family placement on the index previously discussed. These columns represent the estimates with the highest power. They also include estimates with and without demographic and entry reason controls. Columns (5) and (6) compare OLS and IV estimates of the effects on current employment or enrollment with controls, columns (7) and (8) compare OLS and IV estimates of the effects on incarceration between ages 20-21, columns (9) and (10) compare OLS and IV estimates of the effects on homelessness between ages 20-21, and columns (11) and (12) compare OLS and IV estimates of the effects on substance abuse referrals between ages 20-21.

Columns (3) and (4) show that the IV estimate gives a statistically significant and substantially large effect of placement on economic and social outcomes. Initial placement with a non-kin family relative to a group home or institution increases outcomes by 0.97 or 0.99 standard deviations of the index. The comparison of the IV estimate with and without controls provide further evidence supporting conditional independence assumption and validity of the estimate, as adding a large set of child-level controls does not alter the coefficient substantially.

When the index is broken out into individual indices in columns (6), (8), (10) and (12), a statistically significant effect is identified for both homelessness and substance abuse. Incarceration is marginally statistically significant ($p = 0.069$) and employment or enrollment is not statistically significant, though the estimated coefficient is very close to the precise and positive OLS estimate.

One way to interpret the magnitudes for each outcome is to use the regression model to get a predicted probability of the outcome of an average child when placed in a group home vs. a non-kin foster family. For homelessness, this method predicts that if half of the children are placed in non-kin foster families, then the estimated probability of homelessness for a child in a group home is 0.51, while the estimated probability for a child in a non-kin foster family home is 0.132.[13] Thus, placement in a group home almost quadruples the chance that a child ends up homeless. Similar calculations give that group homes triple the chance that a child ends up incarcerated (prob 0.1025 vs. 0.3515) and increase the chances that a child ends up with a substance abuse referral by more than 10 times. The results on homelessness and substance abuse do not have easy-to-compare benchmarks in the literature, but the results on incarceration have a benchmark in Ryan, Marshall, Herz and Hernandez (2008) who finds that the risk of delinquency associated with group homes is 2.5 times that associated with other foster care settings.[14] Another way to interpret these coefficients is to assign dollar values in a cost-benefit analysis. This is done below.

[13]Mathematically the method uses two equations:

$$\bar{y} = (0.5)(\bar{y}(1)) + (0.5)(\bar{y}(0))$$
$$\bar{y}(1) - \bar{y}(0) = \beta_{IV}$$

to solve for the two unknowns where $\bar{y}$ is the overall mean, $\bar{y}(1)$ is the predicted outcome for children receiving treatment and $\bar{y}(0)$ is the predicted outcome for children receiving the control.

[14]While the difference between the propensity score results from Ryan, Marshall, Herz and Hernandez (2008) and the OLS results in Table 8 seem quite large, the raw differences in incarceration rates by placement type for children in my sample are similar to Ryan, Marshall, Herz and Hernandez (2008). Thus, the differences in the propensity score results of Ryan, Marshall, Herz and Hernandez (2008) is likely due to data differences, the fact that they undertake a proportional hazards survival analysis, or a differences in samples (they consider younger children, and only focus on Los Angeles).

Table 8: Impact of Non-kin Family Placement on Outcomes of Foster Children

	Economic and Social Outcome Index				Employment or Enrollment		Incarceration		Homelessness		Substance Abuse	
	OLS		IV		OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Initial non-kin family placement	0.886	0.646	2.016	2.056	0.0941	0.107	-0.115	-0.249	-0.078	-0.395	-0.048	-0.246
	(0.067)	(0.067)	(0.513)	(0.727)	(0.016)	(0.193)	(0.014)	(0.148)	(0.016)	(0.182)	(0.011)	(0.110)
County, month x year fes	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Child demographic, entry controls	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Number observations (children)	5,113				5,113		5,039		5,036		5,011	
Mean of outcome	1.01				0.687		0.227		0.321		0.128	
Sd of outcome	2.08				0.464		0.419		0.467		0.334	
Instrument for IV specifications	Non-kin exits / log(population)											

Notes: This table presents OLS and IV results for β, the coefficient on initial non-kin family placement, in equation (2) for different outcome variables and with different specifications. Columns (1)-(4) present results with the economic and social outcome index, described in Section 2.3 which includes variables on employment, enrollment, incarceration, homelessness and substance abuse referrals. They include OLS results with and without the set of demographic and entry reason controls, and IV results with and without the set of demographic and entry reason controls. Columns (5)-(6) present OLS and IV results for an indicator variable for whether a child is employed or enrolled at age 21 (at the time of the survey). These only include specifications with full controls. Columns (7)-(8) present OLS and IV results for an indicator variable for whether a child has experienced incarceration in the past two years since the survey, surveyed at age 21. These only include specifications with full controls. Column (9)-(10) present OLS and IV results for an indicator variable for whether a child has experienced homelessness in the past two years since the survey, surveyed at age 21. These only include specifications with full controls. Columns (11)-(12) present OLS and IV results for an indicator variable for whether a child has had a substance abuse referral in the past two years, surveyed at age 21. The set of controls include demographics with age of entry categories, sex (male or female), and race (white, black, hispanic, other). The set of controls also includes a set of 15 indicator variables indicating the reasons a child was removed from their family.

These results suggest that family placement substantially improves outcomes, consistent with the general OLS relationships found in previous studies (Ryan, Marshall, Herz and Hernandez, 2008) and much of the practical intuition that is used in current foster care policy.

The differences between the OLS and IV estimates are substantial and noticeable for all outcomes except for Employment or Enrollment. In all cases the IV estimates are larger than the OLS estimates. This might be surprising given the expected direction of selection effects: one might expect that children with unobservables that make them more likely to be placed with families would also have unobservables that make them more likely to have positive outcomes if foster families are able to choose which children they foster. However, the the evidence in Table 8 is not consistent with this. The next section is dedicated to exploring this difference under the heterogeneous treatment effects framework.

3.7 LATE and OLS

Table 8 shows that the estimated LATE is much larger than OLS. For example, the coefficient on the overall index in column (4) is more than three times the size of the OLS coefficient on the overall index in column (2). These large differences warrant careful discussion.

The literature mainly distinguishes two reasons for discrepancies between OLS and IV treatment effects when a LATE is estimated: heterogeneous treatment effects and selection bias (see, for example, Dahl, Kostol and Mogstad (2014)). However, another important possibility in the difference between any IV estimate and OLS is measurement error. If the endogenous variable is measured with some error, then OLS estimates will be attenuated while IV estimators can sweep out this error (Reiersøl, 1941).

First, as noted briefly above, it seems dubious that selection bias can explain these differences. For them to explain these differences would require that families choose to care for foster children on unobservables that also predict worse outcomes at age 20-21. Since adding observables to the OLS regressions in 8 make the estimated effects smaller and anecdotal evidence that families prefer children that are easier to care for suggest the opposite, it seems unlikely that this explains the differences.

Can heterogeneous treatment effects explain these differences? The IV estimates in Table 8 are LATEs with certain compliers. Table 7 shows that of all the observable subgroups considered, the group for which compliers have the highest likelihood of belonging is children that have drug abuse problems and are homeless or have housing problems. Table 7 also shows the IV and OLS estimates in this subsample. The OLS coefficient in the housing problem subsample is large and comparable to the IV coefficient in the full sample. The OLS coefficient in the drug abuse subsample is noisy, but the IV coefficient is estimated to be very large. Moreover, the results show that

Table 9: Complier Adjusted OLS Results

	OLS (1)	OLS Weighted (2)	OLS Housing Problem Subsample (3)	IV (4)
Outcome: Economic and Social Outcome Index				
Non-kin family placement	0.886 (0.067)	0.903 (0.126)	2.234 (0.923)	2.016 (0.513)
Outcome: Incarceration				
Non-kin family placement	-0.189 (0.013)	-0.187 (0.135)	-0.278 (0.137)	-0.351 (0.110)
Outcome: Homeless				
Non-kin family placement	-0.087 (0.016)	-0.095 (0.017)	-0.396 (0.167)	-0.270 (0.119)
Outcome: Substance Abuse				
Non-kin family placement	-0.068 (0.010)	-0.074 (0.011)	-0.528 (0.141)	-0.210 (0.074)
Outcome: Employment or Enrollment				
Non-kin family placement	0.108 (0.016)	0.107 (0.018)	0.015 (0.144)	0.136 (0.135)
County, month-year fixed effects	Y	Y	Y	Y
Child demographic, entry reason controls	N	N	N	N
Number observations (children)	5,113	5,113	252	5,113

Notes: This table presents various OLS specifications and IV results across the outcome index and the outcomes that make up the outcome index. Column (1) presents OLS results. Column (2) presents OLS results where the sample is weighted according to first stage coefficient of the housing subsample following Dahl, Kostol and Mogstad (2014) and Bhuller, Dahl, Løken and Mogstad (2020). Column (3) presents OLS results only looking at the subsample of children that enter at least partly due to inadequate housing or homelessness. Column (4) presents IV results. All specifications include county and month by year fixed effects, but do NOT include demographic or entry reason controls, following closely the procedure in Bhuller, Dahl, Løken and Mogstad (2020). Standard errors are clustered at the county level throughout.

one cannot reject that the IV and OLS coefficient in the housing subsample are the same. Table 9 shows that the main differences in OLS estimates in the full sample (columns (1) and (3)) are due to homelessness and substance abuse outcomes. The increase in effects on substance abuse and homelessness is consistent with the possibility that children previously abusing drugs and homeless are more at risk of future abuse of drugs and homelessness, and family environments more strongly mitigate this risk.

If all compliers had housing problems, then the housing problem OLS treatment effect could explain the entire difference between the OLS and IV differences. To investigate this possibility, Table 9 performs an additional exercise following Dahl, Kostol and Mogstad (2014) and Bhuller, Dahl, Løken and Mogstad (2020) that reweights the OLS sample according to the ratio of the first

Table 10: Measurement Error: OLS Results on More Precise Subsample

	Economic and Social Outcome Index		
	OLS	OLS Precise Measurement Subsample	IV
	(1)	(2)	(3)
Initial non-kin family placement	0.886	1.281	2.016
	(0.0673)	(0.234)	(0.513)
Number observations (children)	5,113	752	5,113
County, month-year fes	Y	Y	Y
Child entry, demographics	N	N	N
% IV - OLS difference explained		29.3%	

Notes: This table presents results from OLS and IV regressions of the outcome index on an indicator for a child's initial placement being with a non-kin family estimated in different subsamples. All regressions include county and month-by-year fixed effects but do not include child-level controls. Column (1) gives OLS results for the full outcome sample. Column (2) gives OLS results for children that enter foster care in the same month as the reporting period for the data, or the precise measurement subsample. Column (3) gives IV results for the full outcome sample. Standard errors are clustered at the county level throughout.

stage coefficients and reruns OLS. This is given in column (2) of Table 9. The table shows that there is an increase in the estimated treatment effect, though it is very modest and not nearly the same magnitude as the subsample coefficient in housing. This is because while compliers are relatively likely to have housing problems, the proportion of children with housing problems is very small, approximately 250 out of 5,000 children or 5%. However, I take this as suggestive evidence that if I were to add other unobservables of compliers and reweight on those unobservables, I could explain at least part of the difference between the OLS and IV estimates as treatment effect heterogeneity of the compliers.

An alternative but not mutually exclusive reason for the difference between the OLS and IV estimates is measurement error. The source of measurement error in this data comes from the way initial placements are measured. Placements are reported every 6 months and children may change placements between the time of entry and the report time introducing measurement error.

To see whether this measurement error can account for some of the difference between the OLS and IV estimate, I look at OLS estimates in a subsample for which measurement error should be minimized in the period - the subsample of children whose entries occur in the same month as the reporting period. These children should have the most precise measurement of placement. Table 10 provides results from this exercise. When I estimate on the 750 or so children with these entries, I see that OLS estimate increases substantially by almost 50%. This increase can explain about 29% of the difference between OLS and IV difference. This also suggests that measurement error is likely an important contributor to the difference between the OLS and IV.

Overall, the evidence in this section suggests that the difference in the OLS and IV can potentially be explained by two sources: heterogeneous treatment effects, in which compliers are children like those with pre-existing conditions who receive especially large benefits from placement with a family, and measurement error in the endogenous variable.

3.7.1 Alternate Explanations for LATE and OLS Differences: Foster Family Treatment Effect Heterogeneity

An alternate and related explanation for the difference between the LATE and OLS estimates is heterogeneity on the treatment side: the types of non-kin foster families that children shifted by the instrument end up in may be particularly good at caring for children relative to the average non-kin foster family. This would be consistent with these families "specializing" in fostering children and having high human capital in the occupation of fostering children.

It is challenging to test this hypothesis under the constraints of the data. As mentioned previously, there are no family identifiers in the AFCARS data and there are only a limited set of observables in the data related to non-kin and kin foster families.

One piece of evidence related to this hypothesis is contained in Table 13 below which shows that the instrument predicts that children placed in non-kin foster families are substantially more likely to be adopted by age 18. These types of specializing families are less likely to adopt (Cherry and Orme, 2013) and so this finding suggests the instrument may be picking up other types of families, too. Thus, while it is possible that family treatment effect heterogeneity could explain the differences, the available data do not necessarily point to this specific type of specialized foster family driving this effect.

3.8 Mechanisms: Incapacitation, Family Connections, Public Welfare Takeup and Intermediate Foster Care Outcomes

The main results in Table 8 leave some important questions unanswered regarding how foster families make foster children better off than group homes. First, since children are entering between the ages of 14 and 17, and outcomes are measured at ages 20-21, it is possible that some of these effects reflect incapacitation effects of being in the foster care system as opposed to actual behavioral or other changes. For instance, if children placed with families remain in foster care until age 21 on average, while children placed in group homes remain in foster care until age 18 on average, it is possible that these effects simply reflect that children in family homes are still with their families. Thus, for example, children would not be avoiding homelessness after leaving a family due to changes in their behavior or the support they receive, but instead because they are still in that family home. Ryan et al. (2008) finds that associations between group home placement and arrests are immediate after placement, and occur while a child is still in foster care.

Table 11: Effects of Family Placement and Child Still in Foster Care

	Outcome Index × 1{In FC at Age 18} (1)	Outcome Index × 1{In FC at Age 19} (2)	Outcome Index × 1{In FC at Age 20} (3)
Initial non-kin family placement	1.470 (0.439)	1.381 (0.450)	-0.0455 (0.414)
County, month x year fes	Y	Y	Y
Child demographic, entry reason controls	N	N	N
NYTD 2011 only	N	N	Y
Number observations	5,113	5,113	1,808

Notes: This table reports results from IV regressions of the main outcome index interacted with the maximum age that a child is observed in the AFCARS data. For the regression looking at age 20 in column (3) I only include children that were surveyed at age 17 in 2011, since I do not observe children surveyed at age 17 when they are 20 because I use placement data that ends in the fall of 2016. Throughout county and month by year fixed effects are included. Child demographic and entry reason controls are not included. Standard errors are clustered at the county level.

To tease out these mechanisms, I estimate models that incorporate the length of time that the child is in foster care. I estimate the parameters in IV regressions where the outcome variable is the outcome index featured in Table 8 interacted with an indicator for whether the child is still in foster care at a certain age. Interacting the outcome variable with these indicator variables allows us to see whether the variation in outcomes that is estimated to come from the family placement is driven by variation in outcomes for children that remain in foster care until a certain age.

Table 11 shows the results from this exercise. Columns (1) and (2) show that most of the treatment effects estimated through the instrument on the outcome index in Table 8 come from children that remain in foster care until age 18 or 19. Column (3) shows that this is not the case for children that remain in foster care until age 20. Since outcomes are measured at age 21 and refer to the outcomes in the previous 2 years, these results suggest that these outcomes may not simply be due to incapacitation or a child remaining with a family at the times outcomes are measured. A sharper and more conclusive test would require outcomes that are measured with more time precision.

In Appendix Table A10 we further investigate incapacitation by looking at OLS models that compare exit times for children in non-kin family placements vs. group home placements. Those results are consistent with the results in Table 11 discussed previously. The results show that placement with a family increases the probability of being in foster care at ages 18 and 19. The results for age 20 suggest a marginally statistically significant increase in the probability of still being in care, though conventional confidence intervals contain 0. Moreover, the outcome means suggest that by age 20, very few children in the sample are still in foster care: over 95% have exited the system. This is consistent with many state laws that cap foster care at ages 18 or 19 as shown in Figure A1. For incapacitation to drive the IV estimates would require that the 5% of children still in family foster care placements at ages 20 and onward account for the entire IV estimate. The mean economic and social outcome index for children still in family foster care placements at age 20 is 1.30 while for children still in group home placements at age 20 is actually higher at 1.52. Thus, it seems unlikely that this specific and small subgroup could be driving these effects. Taken together, the evidence suggests it is unlikely that incapacitation plays an important role in the estimated impacts of family placements versus group homes.

Table A10 presents the results from this exercise, estimating whether family placement predicts whether children are still in placement. Due to precision, I focus on OLS estimates for this exercise.

I now turn to other outcomes that may offer more insight into what drives the boost in outcomes for children placed in families. One potential pathway studied in the foster car literature is a meaningful sense of connection to an adult or family. This has been hypothesized to be an important component of a foster child's successful transition to adulthood (Freundlich and Avery, 2006).[15] However, achieving these connections can be challenging in practice, and little causal evidence has been found to suggest that foster children more easily develop these support systems and connections through family placements.

Table 12 Panel A columns (1) and (2) includes IV and OLS estimates of placement with a family on connections with an adult at age 21.[16] The IV estimate suggests a statistically significant 49 percentage point increase in the probability of developing a connection, or 57 percent on the mean outcome of 0.896. While methods to more formally test whether connection to an adult is an important mediator of the economic and social outcomes considered above are not appropriate in this setting (Dippel, Gold, Heblich and Pinto, 2020), the evidence is consistent with this connection to adult being correlated with these outcomes and potentially being an important mediator.[17]

The other results in Panel A of Table 12 show that the IV estimates do not estimate precise strong effects for other outcomes such as having children or receiving payments. The IV estimates do suggest that placement with a family leads to a large decrease in the probability of participating in an apprenticeship or on-the-job training during age 20. This could be consistent with families shifting children into more enrollment as opposed to employment to invest in human capital to increase lifetime earnings, but I lack the power to precisely test this hypothesis.

One important question about how children achieve better outcomes through placement with

[15]Biehal (2014) also studies what belonging means in substitute foster families.

[16]The wording of the question involves that the adult is someone "who he or she can go to for advice or guidance when there is a decision to make or a problem solve, or for companionship when celebrating personal achievements. The adult must be easily accessible to the youth, either by telephone or in person. This can include, but is not limited to adult relatives, parents or foster parents." (NYTD Outcomes Codebook p. 37).

[17]Interestingly the OLS coefficient estimates a precise 0 on connection to an adult for children. This is quite drastic and different, but consistent with the treatment effect heterogeneity found elsewhere, where family effects are amplified for the complier population.

Table 12: Connection to Adult, Public Welfare Outcomes and Other Economic and Social Outcomes

Panel A: Other Economic and Social Outcomes

	Connection to Adult		Had Children		Private Financial Payments: Family, Child Support, Legal		Apprenticeship, Internship, On-the-Job Training	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Initial non-kin	-0.006	0.490	-0.046	0.103	-0.002	-0.247	0.026	-0.553
family placement	(0.012)	(0.158)	(0.015)	(0.180)	(0.011)	(0.166)	(0.016)	(0.247)
Child demographic, entry controls					Y			
County, month x year fes					Y			
Mean outcome	0.896		0.275		0.115		0.315	
Number children	5,097		5,063		5,052		5,099	

Panel B: Social Services

	Total Public Aid		Social Security		Educational Aid		Food Stamps		Housing Vouchers		Other Cash Welfare	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Initial non-kin	-0.168	-1.114	-0.063	-0.013	0.091	-0.028	-0.056	-0.612	-0.017	-0.352	-0.029	-0.296
family placement	(0.034)	(0.490)	(0.012)	(0.158)	(0.014)	(0.238)	(0.020)	(0.366)	(0.010)	(0.186)	(0.011)	(0.194)
Child demographic, entry controls							Y					
County, month x year fes							Y					
Mean outcome	0.592 (sd = 0.831)		0.103		0.203		0.315		0.0752		0.0989	
Number children	4,122		5,064		5,048		4,241		4,228		4,228	

Notes: This table presents OLS and IV results from other economic and social outcomes and public welfare use outcomes. The other economic and social outcomes contained in the NYTD data include whether the child has a connection to an adult they feel comfortable going to for advice, they have mothered or fathered children in the past 2 years, they receive financial payments from a family, child support or other legal source. The public welfare use source includes an index of total public aid which adds together indicators for social security, food stamps, housing vouchers and other cash welfare. These are also broken out separately, with the addition of an outcome on whether the child receives financial aid. All regressions include child demographic and entry reason controls, and county and month by year fixed effects. Standard errors are clustered at the county level.

families is whether they rely on social services to achieve these gains. If so, this might dampen the overall monetary benefit of family placement, as this benefit comes with a social cost of welfare take-up. Panel B of Table 12 hows OLS and IV estimates of the effect of family placement on take-up of social services. It includes a measure of total public aid, which sums the social security, food stamps, housing vouchers and other cash welfare measures. The IV estimate suggests that placement in families leads children to take-up less public aid, with the results seeming especially strong (and marginally statistically significant) for food stamps and housing vouchers. The point estimate for educational aid take-up is negative though with wide confidence intervals.

The final set of results in this subsection look at potential mechanisms and mediators in intermediate outcomes in foster care including placement stability and permanency. These are closely studied in the literature (Becker, Jordan and Larsen, 2007; Koh and Testa, 2008; Andersen and Fallesen, 2015) but focus more on the differences in achieving stability and permanency in kin and non-kin placements. These outcomes are of first order importance to foster care policy makers as short-term markers of how well the foster care system is working. I contribute to this literature by looking at differences contributed by group homes and foster family placements. These could also be important mediators for the effects on social and economic outcomes estimated.

Table 13: Intermediate Foster Care Outcomes

Panel A: IV	Outcome Sample		Eligible Sample		Old Children Sample (Weighted)	
	Adopt or Guardian by 18	Number Placements after Entry	Adopt or Guardian by 18	Number Placements after Entry	Adopt or Guardian by 18	Number Placements after Entry
	(1)	(2)	(3)	(4)	(5)	(6)
Initial non-kin family placement	0.0918	-3.495	-0.0599	-1.122	0.0795	-0.550
	(0.1470)	(2.611)	(0.1562)	(1.806)	(0.0383)	(0.636)
Instrument	Non-kin exits / log(county population)					
County, month x year fes	Y	Y	Y	Y	Y	Y
Child demographic, entry reason controls	Y	Y	Y	Y	Y	Y
Mean outcome	0.045	4.46	0.043	4.42	0.025	2.46
Number observations (children)	3,619	4,454	13,840	15,731	143,409	151,372

Panel B: OLS						
	(1)	(2)	(3)	(4)	(5)	(6)
Initial non-kin family placement	0.0459	-0.734	0.0413	-0.747	0.0349	-0.439
	(0.109)	(0.169)	(0.0052)	(0.096)	(0.0094)	(0.076)
County, month x year fes	Y	Y	Y	Y	Y	Y
Child demographic, entry reason controls	Y	Y	Y	Y	Y	Y
Mean outcome	0.045	4.46	0.0431	4.42	0.0250	2.46
Number observations (children)	3,619	4,454	13,840	15,731	143,409	151,372

Notes: This table presents OLS and IV regression results of adoption or guardianship indicator variables and number of placement numeric variables on the initial placement with non-kin indicator variable. It does this across the outcome, eligible and old children sample, where observations in the old children sample are weighted according to (obs weight = percent observations with same county in outcome sample) to ensure a stronger first stage. The samples for adoption and guardian by 18 models exclude children who do not exit by age 18. Smaller sample sizes for number placements are smaller because missing value in the number placements variable. Throughout models include child demographic and entry controls, and county and month by year fixed effects. Standard errors are clustered by county.

Table 13 shows IV and OLS estimates of adoption and guardianship by age 18 and the total number of placements after entry. Because these outcomes are observed in the AFCARS data, I examine the results in all three analysis samples, but the preferred specifications in columns (5) and (6) use the larger older children sample. The IV and OLS estimates in columns (5) and (6) both suggest that adoption and guardianship is shifted by a large and statistically significant percentage. The number of placement estimates are consistent but the IV is less precise and cannot reject 0 effects or even positive effects. These results show that placement with a foster family significantly boosts the probability of adoption or guardianship and they are consistent with placement increasing placement stability, though there is less precision for this result.

3.9 Longitudinal Analysis

So far the analysis has focused on the longest term outcomes, outcomes at ages 20-21. However, the NYTD data include outcomes for children at ages 17 and before, and at age 19. This section brings these outcome measures into the analysis.

Table A13 shows results for the IV strategy on outcomes measured at age 17 or before and outcomes measured at ages 18-19. To ensure the analysis attributes these outcomes to the specific foster care placement and not circumstances before, sample restrictions are required. For these results two separate samples are estimated on: children with early enough entries (for whom these outcomes happen reasonably far enough after placement) and children with age 21 outcomes who I estimate outcomes for in the main specifications. In all cases, the results show relatively consistent OLS effects but the IV is noisy for outcomes at age 19. The confidence intervals do contain the IV estimates. I take this to mean that my empirical strategy and the data only have enough power to identify effects that are more likely to be revealed later at ages 20-21.

Table A14 further analyzes the longitudinal data. Panel A presents correlations between the different outcomes. The coefficients and R-squareds show that, while there is some correlation, outcomes at earlier ages do not have much explanatory power for predicting outcomes at age 21. Panel B presents an alternative lagged dependent variable empirical strategy that controls for outcomes at age 17 and earlier in the OLS regression equation. The results show coefficient estimates similar to those produced in the older sample, though the samples are quite different because less children have outcomes measured at both age 17 and age 21. Panel C provides a validity test of the empirical strategy, correlating the instrument with outcomes at age 17 or before and comparing this to the correlation with outcomes at age 21. If outcomes in a child's life up until age 17 predict later outcomes, are observed by families, and correlated with the instrument, one might worry that the instrument may reflect something about underlying unobservables that also predict child outcomes. This panel shows no correlation between outcomes in life before 17 and the instrument. Finally, Panel D supplements the empirical analysis in Table 8 by also including outcomes at age 19 as a proxy for later life outcomes. I define the outcome variable to be outcomes at age 21 if they exist, and if not, then outcomes at age 19 if they exist. This boosts the sample size and alters the outcome in Table 8. The results are similar to the main results, with a noticeable difference being a significant boost in enrollment or employment coming from including 19 year olds. Thus there is some evidence that family placement increases enrollment or employment for children when considering earlier ages, too.

3.10 Cost-Benefit Analysis

This section undertakes a cost-benefit analysis of family vs. group home placement in foster care using the IV strategy. The goal of this exercise is to produce a rough dollar estimate of the benefits of family placement. This helps put the results in more concrete terms and provide some back-of-the-envelope numbers that could be cited in policy-making. A more extensive policy analysis accounting for the equilibrium placement of children due to family preferences, unobservable differences between children, and heterogeneous treatment effects is undertaken below.

I measure two types of benefits from outcomes in the NYTD. The first is personal benefits in terms of a foster child's earnings. For simplicity, I consider a single year of earnings. Panel A of Table 14 considers estimates contained in the literature on the earnings costs to incarceration, substance abuse, and homelessness. The second type of benefit is social benefits in terms of the cost to the government and taxpayers for providing the service. For incarceration I consider the cost of imprisoning someone for one year; for substance abuse I consider the cost of methadone treatment, a common treatment for opioid abuse, for one year; and for homelessness I consider the cost of providing a bed in a homeless housing facility. Panel B considers estimates of these costs in the literature.

The costs of placing children with foster families and group homes are computed in Panel B. All sources I could find pointed towards the cost of placing children in group homes as more expensive than placing children in families. This table also contains an estimate in the literature on the elasticity of supply for foster families from Doyle (2007a).

Finally, Panel C uses the benefits numbers in Panel A and the instrument to produce IV estimates of all the benefits with the endogenous variable as years in non-kin family placements to match the estimated costs. The total benefits have a lower bound of around \$5,800 and an upper bound of around \$6,400. The annual cost-savings of placing children in families is also estimated to be between \$25,000 and \$30,000 per year. These numbers suggest that family placement more than compensates for its cost relative to group homes in dollar terms and can save on the order of \$30,000 to \$35,000 per year per child.

Table 14: Cost-Benefit Analysis

Panel A: Benefits	Personal Benefits (1 year)				Social Benefits (1 year)			
	Lower Bound	Lower Bound Source	Upper Bound	Upper Bound Source	Lower Bound	Lower Bound Source	Upper Bound	Upper Bound Source
Incarceration	$618.75	Aizer and Doyle (2015), Card (1999)	$618.75	Aizer and Doyle (2015), Card (1999)	$31,978	Bailey et al. (2020), Prisons Bureau	$31,978	Bailey et al. (2020), Prisons Bureau
Substance Abuse	$0	Van Ours and Williams (2015)	$3,300	Van Ours (2007)	$4,700	National Institute on Drug Abuse (2021)	$4,700	National Institute on Drug Abuse (2021)
Homelessness	$3,000	Collinson and Reed (2018)	$5,310	Cohen (2021)	$13,470	Wong et al. (2006)	$13,470	Wong et al. (2006)

Panel B: Costs	Overall Cost (1 year)			
	Lower Bound	Lower Bound Source	Upper Bound	Upper Bound Source
Cost of non-kin foster family for 1 year	$500 * 12 months = $6,000	WeHaveKids.com (2020)	$1,000 * 12 months = $12,000	County (2020)
Cost of group home for 1 year	$2,762 * 12 months = $33,144	County (2020)	$42,654	Trends (2018), AFCARS
Elasticity of foster families	0.5	Doyle (2007a)	0.5	Doyle (2007a)

Panel C: IV Estimates of Overall Benefits	Personal Benefits		Social Benefits	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound
1 year placement with non-kin family	$504.06	$1,124.14	$5,293.09	$5,293.09

	Total Benefits	
	Lower Bound	Upper Bound
1 year placement with non-kin family	$5,797.15	$6,417.23

Notes: This table presents results from the cost-benefit analysis, placing US dollar values on the benefits derived. Panel A reports lower and upper bounds from the literature on the benefits related to incarceration, substance abuse, and homelessness, including both personal benefits related to one year of earnings in young adulthood and social benefits related to one year of government spending. Panel B reports lower and upper bounds from the literature and county and state-based foster care institutional details on the cost of placing a foster child in a non-kin foster family, the cost of placing a foster child in a group home, and the elasticity of foster families. Panel C uses the lower and upper bounds in Panel A to compute IV estimates of lower and upper bounds on personal benefits, social benefits and overall benefits (personal benefits + social benefits).

3.11 Comparison to Doyle (2008) and Implications for Overall Removal Effects on Older Children

This section makes an explicit comparison to the literature looking at the causal effects of entry into foster care on subsequent outcomes. This paper provides one way to think about heterogeneity in the treatment of entry into foster care and shows that there can be substantial heterogeneity in foster care impacts on subsequent outcomes through placement types. Quantitatively, I compare the estimates in this paper to those found in the literature and perform some back-of-the-envelope calculations.

Doyle (2008) estimates the causal effect of foster care placement for children of average age 11 on incarceration at ages 18 or older in Cook County. He finds that placement into foster care causes a 22.5 percentage point increase in the probability of incarceration (Table 4, Panel C, Column 4) on a mean of 0.066 (Table 4, Panel C, Column 1). This paper shows that it is possible that placement into foster care *and* placement in a group home could be an important part of these negative effects, which are also found for other outcomes in Doyle (2007b).[18]

This paper estimates that the effect of placement with a family relative to a group home for children in foster care causes a 24.9 percentage point decrease in the probability of incarceration. Moreover, between 2005 and 2015, the placement rate of children into families (kin and non-kin) in Cook County for children entering between ages 14 and 17 is 0.264. For simplicity I assume that treatment effects are the same for kin families as for non-kin families relative to group homes.[19]

Now suppose that the causal effect of placement into foster care estimated in Doyle (2008) can be written as

$$\beta_{overall} = \beta_0 + \beta_{family} F + e \tag{4}$$

where e is some random noise, so that the treatment effect is now a random coefficient that also depends on family placement. Using this setup and the numbers above, the expected treatment effect as a function of average family placement in Cook County can be written as

$$\mathbb{E}[\beta_{overall}] = 0.291 - 0.249\mathbb{E}[F]. \tag{5}$$

Equation (5) gives a rough and simple way to understand the implications of family placement for the overall effect of foster care. If all children were placed in families in Cook County, this method would estimate that the probability increase in incarceration would be reduced to 4.2 percentage points, and that if no children were placed with families, the probability increase would

[18]Note, though, that some recent studies have found positive effects on children. These include (Bald et al., 2019; Gross and Baron, Forthcoming).

[19]OLS estimates of these treatment effects on the outcomes considered in this paper are surprisingly similar using AFCARS and NYTD.

jump up to 29.1 percentage points. This suggests a large role for family placements and placement types in understanding the overall effects of foster care. However, this example shows that even with full placement policy, there is an expected increase in incarceration. This result might suggest future research on studying how foster care shapes child outcomes through channels other than family placement or institutionalization, such as the trauma of being separated from a birth family.

3.12 Heterogeneous Treatment Effects

This section explores heterogeneous treatment effects by child characteristics. Previously heterogeneity based on compliers was explored and in this section I further explore heterogeneity based on previous observables. These heterogeneity tests have important precedents in the literature. For example, many early childhood family and education interventions aimed at disadvantaged children have larger effects for boys (Bertrand and Pan, 2013). Do boys or girls benefit more from family placement in foster care? This is particularly interesting to the later analysis in this paper which examines consequences of the foster child allocation to families for the distribution of foster child outcomes.

Tables A2 - A9 provide a comprehensive set of heterogeneous treatment effects by child demographics and entry reasons observed in AFCARS. Due to precision challenges, it is difficult to confidently make comparisons between certain subgroups but these tables do contain a few notable results. Figure A2 provides a summary of the most precisely estimated heterogeneous treatment effects, plotting IV coefficients and 95% confidence intervals in different subsamples.

The first notable result is that boys benefit substantially more from family placement than girls, with an estimated benefit of almost four times the girls benefit on the economic and social outcome index. This is driven by reductions in the chances of incarceration and homeless. The second notable result is that children that enter foster care due to neglect appear to get less benefit from family placement than children than enter for a reason other than neglect. This is also consistent with the OLS coefficient difference. The same is true for children that enter because their caretaker has some disabling condition keeping them from properly caring for the child. The third notable result is that children that have drug abuse problems or housing or homeless problems appear to have strong treatment effects. As noted above, these children also have larger first stage coefficients, making them more likely to be compliers too and driving the difference between the OLS and IV coefficients. This last result suggests that children with pre-existing risk or conditions associated with a certain negative outcome benefit the most from placement with a family.

To further utilize the richness in the observable characteristics of children available, I apply the random forest method developed in Wager and Athey (2018) for estimating heterogeneous treatment effects as conditional average treatment effects (CATE). Table A15 gives estimates of summary statistics of the distribution of CATE and Figure A3 summarizes on different child observables (demographics and entry reasons) whether children with that observable characteristic are above or below median in the CATE distribution. The results are very similar to those in Tables A2 - A9.

3.13 Robustness

This subsection assesses robustness to the results in Table 8 by checking how the results change when we consider different variations of the endogenous variable, the possibility of non-random attrition, alternative child age cutoffs for defining the sample, and other changes to the analysis sample and outcome index considered.

3.13.1 Measurement of Family Placement

An important part of the analysis done so far is that it uses a foster child's measured initial placement as the endogenous variable. This was done to help motivate the instrument conceptually, but may not be the best measure of a child's overall experience with families while in foster care nor the exact policy variable of interest. Table A12 reruns the IV models on the economic and social index with two alternative endogenous variables. The first endogenous variable measures the months in non-kin family placement that a child is placed. The second endogenous variable measures the percent of time that a child spends with a non-kin foster family while in foster care before they exit. Both endogenous variables lead to larger statistically significant positive effects larger than the OLS results, consistent with the patterns in Table 8. Thus the results are robust to the measurement of the endogenous variable.

3.13.2 Attrition

As shown in Table 1, there are differences in family placement for children that have a measured outcome versus those that were eligible to be included in the survey. This type of attrition could bias our estimates.

Table A16 contains results that account for and measure the consequences of potential non-random attrition. Two methods are implemented to address attrition. First following Sacerdote (2007) and the method developed by Wooldridge (1999), Panel A looks at inverse propensity score weighted versions of the IV estimate. This panel shows that if the IV results are weighted on child observables that predict whether we observe a child's outcome at age 21, the IV estimates are larger. This is consistent with results in Table 1, where boys are more likely to be left out of the survey and also more likely to be incarcerated or homeless. This suggests those raw patterns in Table 1 are not cause for concern, if anything suggesting the estimates are a lower bound.

Panel B of Table A16 computes Lee (2009) bounds for OLS and intent to treat effects across different subsamples. One challenge with this exercise is that the attrition rates observed in the data do not reflect actual attrition since some states randomly sample from eligible children. Columns (3) - (6) adjust for this in two ways. Columns (3) and (4) look only at the sample of children that were eligible for the follow-up survey at age 21. These children are 17 year olds that are in foster care and also respond to the initial survey. When Lee bounds are computed on this subsample the reduced form yields positive coefficients for both the lower and upper bound. Columns (5) and (6) include only children that are in states that do not sample. Note that this includes most of the sample. The Lee lower bound is above 0 here, too. Thus while attrition may affect the results quantitatively, the conservative Lee bounds still give the same qualitative signs of the results, and weighting on observables leads in fact to stronger results.

3.13.3 Robustness to Child Age Cutoff

The analysis above utilized children that were eligible for the sample, and that enter between ages 14 and 17 in their latest entry. This section tests robustness to the lower age cutoff. Table A17 presents IV and OLS estimates of the coefficient on family placement for the economic and social outcome index model where the age cutoff is changed to be 12, 13 and 15. When the age cutoff is 12 or 13, a large statistically significant effect of family placement on outcomes is estimated. The estimate for the age cutoff of 15 is large, but the standard errors are larger. Similar to the main results, the IV estimates are consistently above the OLS estimates, presumably for the same reasons discussed above about the LATE.

3.13.4 Specifications and Outcome Indices

Table A18 Panel A provides further first stage and IV estimates under different specifications. Across all these specifications the qualitative results are very similar to the main results. A notable specification in Panel A is in Column (1) that adjusts the instrument to only look at exits from non-kin families where the exit is for a child that is 14 years or older. These exits may be more representative of available slots for older children, since some families may not be willing to care for older children. The IV results are stronger in this specification.

Panel B of Table A18 explores robustness to the outcome index utilized, looking at different mixtures of incarceration, homelessness, substance abuse, employment, enrollment, and educational attainment. All estimates are positive, large and statistically significant, with magnitudes usually around a 1 standard deviation increase in the outcome index.

4 Model of Equilibrium Foster Care Placement and Child Outcomes

I now build a model of foster care placement and foster child outcomes to study how family preferences determine the allocation of foster children to foster families and aggregate foster child outcomes in equilibrium. I then use the model evaluate counterfactual policies that could improve foster child outcomes.

Within the policy context discussed in the introduction, the main goal of the model is to simulate and evaluate policies that increase the supply of available families for foster children and also shift children that are placed in foster families based on observables while holding fixed the supply of available families.

The model has two main components: a matching market equilibrium across foster care markets in which foster family preferences, available foster families and available children determine the equilibrium allocation in each market, and an outcome model where children's outcomes depend on their placement and other characteristics that also determine placement. The following subsections describe the setup of the model and include relevant institutional or empirical details that inform the assumptions.

4.1 Market Definition and Available Foster Families

The model treats placement of foster children with foster families as a set of distinct markets. This choice is made due to the institutional details of foster care. Social workers and other stakeholders involved treat foster children's placement on a case-by-case basis due to the time constraints they face in placing children. Social workers are constrained by the law to find a placement for a child within a reasonable time frame of that child entering.[20]

Each market t is a county-month-year tuple (ex: Los Angeles December 2011). In each market there is a set of foster children I_t that must be placed and a set of available families J_t. Each child $i \in I_t$ can either be placed with one of the available families, in which case their placement is denoted $A_i = 1$, or in a group home, in which case their placement is denoted $A_i = 0$.

In the model, all foster families are the same, and the number of available families in market t is modeled by

$$|J_t| = \text{Number Available Families}_t = Z_t\lambda + \phi_{c(t)} + \phi_{m(t)} + \eta_t \tag{6}$$

where Z_t is the non-kin exits instrument from the previous section. Here η_t represents an unob-

[20]In California, for example, this time frame is 24 or 48 hours.

served supply shock to the number of families that are available in a given market. Equation (6) gives a model-based description of how the instrument will change a foster child's chances of being placed with a foster family. The second component of the model to pin down which foster children are placed based on observable characteristics, unobservable characteristics and the instrument is foster family preferences over child characteristics.

4.2 Foster Family Preferences

Every family has preferences over foster child characteristics. This setup is motivated by distinctive patterns in the types of children that are more often placed with foster families. It is also motivated by discussions with foster care social workers who say that the family preferences they elicit often have a strong influence on which children are placed with foster families. For example, boys and older children are anecdotally preferred less often by families, and these patterns will show up in the data here, too.

Descriptive evidence for foster family preferences over observable demographic characteristics of foster children is given in Table 15. This table presents results of OLS regressions of an indicator for whether a child is placed in a non-kin foster family (or group home) on child demographics. There are clear patterns in the types of children more likely to be placed with families: girls are predicted to be more likely, black children are predicted to be less likely, and older children are predicted to be less likely. These are generally consistent with previous work on the types of foster children placed in group homes (Ryan et al., 2008) and anecdotal evidence from social workers on the types of preferences that foster parents often state and other work on foster care placement.

To capture these patterns in the data and interpret them as foster family preference parameters, I assume that preferences are homogeneous and vertical (Lancaster, 1966; Berry and Pakes, 2007) over child characteristics and the preferences of family j for child i are given by

$$u_{ji}(X_i) = u(X_i) + v_i = X_i\alpha + v_i \tag{7}$$

where X_i is a set of child observable characteristics (i.e. demographics) and v_i is an unobservable taste shock for child i common to all foster families. Equation (7) describes how foster families "rank" foster children across different markets based on their characteristics, where all foster families have the same ranking criteria. Note that no outside option is purposefully defined: families in the market see all foster children as acceptable, but have a preference ordering over foster children based on observable and unobservable characteristics.[21]

[21] No outside option is defined because the data cannot identify both the number of available families in (6), and the outside option of not caring for any child. To see why, with both an outside option and a number of available families, we can rationalize any number of available families with any number of children being matched in. market by simply making the outside option of not caring for a foster child good enough in that market.

Table 15: Descriptive Evidence on Foster Family Preferences

	Placement with Non-Kin Foster Family (1)	(2)
(Intercept)	0.545***	0.527***
	(0.004)	(0.006)
Sex: male	−0.195***	−0.211***
	(0.002)	(0.003)
Race: black	−0.050***	−0.071***
	(0.004)	(0.006)
Race: white	−0.00001	−0.042***
	(0.004)	(0.006)
Race: hispanic	0.004	0.009
	(0.004)	(0.006)
Age: 15	−0.060***	−0.065***
	(0.003)	(0.004)
Age: 16	−0.084***	−0.090***
	(0.003)	(0.004)
Age: 17	−0.093***	−0.099***
	(0.003)	(0.004)
Observations	231,342	93,606
R^2	0.050	0.066

Notes: This table presents OLS regressions of an indicator variable for placement with non-kin foster family on entry (versus placement in a group home). Column (1) includes all child entries for children with non-missing demographics entering between the ages of 14 and 17. Column (2) includes child entries in county-month-years where at least 10 children entered in the same county-month-year. The reference group for race is asian pacific islander and native american, and the reference group for age is entering at 14 years old. Standard errors are given in parentheses.

4.3 Foster Child Outcomes

Foster child outcomes depend on family placement (the treatment), and the same observable and unobservable characteristics that determine foster family preferences. In particular, I follow Heckman (1979), Kline and Walters (2016), and Walters (2018) and model the conditional expectation of an outcome for child Y_i as

$$E[Y_i(A)|X_i, Z_t, v_i] = \mu_A(X_i) + \gamma_A v_i = X_i\beta_A + \gamma_A v_i. \tag{8}$$

This model of outcomes includes the common assumption of separability between the observables and unobservables in determining outcomes (conditional on treatment) (Brinch, Mogstad and Wiswall, 2017). Here β_0 and β_1 allow for children with different characteristics X_i to vary in their average potential outcomes, and to vary in the average impact of the treatment of being placed with a family relative to a group home. γ_0 and γ_1 allow for unobservable selection on levels and unobservable selection on gains.

4.4 Market Equilibrium

This paper follows the literature on two-sided matching markets (e.g. Agarwal (2015)) in assuming that a market equilibrium in market t consists of a stable match between available families and entering foster children. Because children have identical and trivial preferences, and families have identical vertical preferences, stability in this case is equivalent to assigning children in each market to maximize family utility. This can be written as:

$$\sum_{j \in J_t} \sum_{i \in I_t} u_{ij} A_{ij} \text{ s.t. } \sum_{i} A_{ij} \leq 1 \; \forall j \in J_t, \; \sum_{j} A_{ij} \leq 1 \; \forall i \in I_t \tag{9}$$

where A_{ij} is indicator for assignment of child i to family j.

Together equations (6) and (7), and the stability condition in equation (9) provide a way for the model to predict whether a foster child is placed with a foster family, and serve the basis of the environment in which policy simulations will be conducted.

The model assumes that there are a certain number of available families in each market with preferences that rank children, but assumes that all children are acceptable. Because of this implicit individual rationality assumption, all families are matched to children in equilibrium and so

$$\text{Number Available Families}_t = \text{Number Children Matched}_t. \tag{10}$$

4.5 Model Discussion

This model emphasizes a few important aspects of the foster care market. The first is that families are scarce, and random supply shocks to the availability of families occur. These random supply shocks, which are modeled directly as coming from the instrument, are crucial to the identification of the model parameters. The second is that available families select their most preferred foster children based on observable and unobservable characteristics. This provides a method for predicting which foster children will be placed under different counterfactual scenarios. It also introduces an economic mechanism where foster family preferences may cause the allocation of children to families to not maximize the average outcomes of foster children, or satisfy other important distributional properties. These features distinguish it from previous work in the program evaluation space and allows us to directly study and compare policies that increase the supply of families and also affect the allocation of children to families holding constant the supply of families while taking into account that family preferences play an important role in any allocation policy.

4.5.1 Limitations

The model abstracts from a few important features of the foster care market. I discuss some of these below.

Dynamics and Timing: The model treats all children entering in the same month as being placed in a one-shot style market at the same time. This is an approximation of reality and is used to emphasize how foster family preferences allocate children into treatment more broadly within the constraints provided by the data. In reality, children are allocated more dynamically based on the relatively strict time requirements discussed previously. Social workers may be able to change children's placements over a longer time horizon (for example switching them to a different placement after staying in a group home temporarily for a few weeks) but my talks with social workers and the enormous case load many of them face with entering children suggested this was not a large phenomenon. Thus, I see the model assumptions as capturing the important economic mechanism whereby families have preferences over children's characteristics, and the characteristics of children entering around the same time have spillover effects on the placement of other children. Extensions to the model could allow for foster children to potentially change placements over time based on a priority system, and better data could allow for a sharper model of how family arrivals and exits to the system play a role in foster family placement.

Social Worker Discretion: In this model social workers play a trivial role. The stability concept and family preferences imply that they simply maximize outcomes for family's based on their preferences. There are no social worker identifiers in the AFCARS data as it is currently circulated, so it is difficult to separate out the role of social worker and family preferences. The literature has shown that in the context of some decision problems in foster care, heterogeneity in social worker discretion can have important consequences for foster children (Doyle, 2007b, 2008; Bald, Chyn, Hastings and Machelett, 2019; Gross and Baron, Forthcoming). Moreover, because social workers make offers to families of foster children, it is possible that how they view what placements are suitable for children or how they view their portfolio of children could affect a child's placement. However, under child welfare laws, social workers are generally expected to make the best possible effort to find a child "the least restrictive home possible", and in my talks with social workers, they emphasized more heavily how family preferences influence placement. It seems more likely that social workers provide an important function in matching types of children to types of families than in deciding whether a child is matched at all. Thus, I see the model assumptions as approximating the feature of foster care in which social workers are solely meant to facilitate family placements for all foster children, under the preferences provided to them by those families. However, important extensions could allow for the incorporation of social worker preferences or objectives in foster family placement.

5 Model Identification, Estimation, and Results

5.1 Identification

This section outlines how the model utilizes the variation in the data to estimate the model parameters with an emphasis on how the instrument is used to identify the parameters that determine foster children's outcomes.

Using the equilibrium condition in (10) the parameter λ can be identified in (6) by looking at the covariance between the instrument and the number of children matched across markets, conditional on the county and month-year fixed effects. The key assumption required to identify λ is that Z_t is independent of η_t unobservable shocks to the supply of available families. This identification mirrors in the first stage equation (3) in the IV regressions of Section 2, but is instead run at the market level instead of the child level.

To identify the preference parameters α in equation (7), the model implies that children are ranked according to $u_i(\alpha) = X_i\alpha + v_i$, and that the top $|J_t|$ ranked children are matched, where the number of children matched is determined by the number of available families. To see how the observables influence this ranking, the model looks at the covariances between each child's placement and the observables of that child. Since a child's rank is relative to their own market and thus depends on the characteristics of children in the same market, the estimation routine also includes moments that look at the covariance between the children's placement and the observables of other children in that market. Together, these moments identify α.

The last set of parameters are the outcome parameters, (β_0, β_1) and (γ_0, γ_1). To identify these parameters, I first discuss how the model identifies an estimate of v_i for each child. To identify an estimate of v_i for each child, the model adds the parametric assumption that $v_i \sim N(0, 1)$. Using the parametric assumptions on v_i, the child observables in each market and the preference parameters α, the model implies a way to estimate v_i for each child. The main challenge in estimating v_i is the potential correlation between η_t and v_i. To allow for that correlation, the estimate is formed based off of the instrument Z_t instead of the potentially endogenous number of children matched observed in equilibrium.

How does Z_t identify v_i? This is best illustrated by looking through a few examples. Consider a market with 3 children $(1, 2, 3)$. Suppose that the model for family supply and Z_t predicts that there is 1 available family and suppose it is observed that child 1 is matched. This implies that

$$X_1\alpha + v_1 \geq \max\{X_2\alpha + v_2, X_3\alpha + v_3\} \tag{11}$$

and using the parametric assumptions on the v_i's pins down a conditional expectation v_i based on Z_t and other market details. The model uses the equilibrium implications and predicted number of available families from the instrument to find expectations of the unobservables by treating these unobservables as order statistics in each market.

If instead Z_t predicts there are 2 available families and children 1 and 2 are observed to be matched, the equilibrium implication is

$$\min\{X_1\alpha + v_1, X_2\alpha + v_2\} \geq X_3\alpha + v_3 \tag{12}$$

which again pin down all the expectations of the v_i's.

The outcome parameters (β_0, β_1) and (γ_0, γ_1) are then identified by correlating outcomes between observables and this unobservable shock, separately, and for children placed with families and without families.

5.2 Estimation

To estimate the model parameters, I proceed in three steps.

5.2.1 Estimation of Family Supply

First, to estimate the parameters in (6) I use the equilibrium condition (10) and OLS to estimate

$$\text{Number Children Matched}_t = Z_t\lambda + \phi_{c(t)} + \phi_{m(t)} + \eta_t. \tag{13}$$

I include all children entering into identifiable counties at the ages of 14-17 between 2010 and 2015 in this equation.

5.2.2 Estimation of Family Preferences

To estimate the preference parameters α I utilize Simulated Method of Moments (McFadden, 1989; Pakes and Pollard, 1989). I minimize the criterion function

$$(\hat{\mathbf{m}} - \hat{\mathbf{m}}^S(\alpha))'\mathbf{W}(\hat{\mathbf{m}} - \hat{\mathbf{m}}^S(\alpha)) \tag{14}$$

where $\hat{\mathbf{m}}$ represents two sets of moments calculated separately for each market, and then averaged across markets. The simulated moments $\hat{\mathbf{m}}^S$ are also computed separately for each market, and are averaged across markets and across simulations.

For the data in market t ($\{X_i\}_{i\in I_t}, \{A_i\}_{i\in I_t}$), the two sets of moments are

1. The covariance between child characteristics and children placed in each match. For each observable characteristic X_i^k,

$$\hat{m}_{t,1,k} = \frac{1}{|I_t|}\sum_{i\in I_t} X_i^k \cdot A_i. \tag{15}$$

2. The covariance between characteristics of other children in a market and a child's own placement. For each observable characteristic X_i^k

$$\hat{m}_{t,2,k} = \frac{1}{|I_t|}\sum_{i\in I_t}\Big(\frac{\sum_{j\neq i, i\in I_t} X_j^k - X_i^k}{|I_t|-1}\Big)\cdot A_i. \tag{16}$$

Intuitively, the first set of moments is used to estimate which characteristics of children make them more likely to be placed, and the second set of moments is used to estimate which characteristics of children competing for slots make a child more likely to be placed. This second set of moments adds additional information on how a child's placement varies with the characteristics of other children in the market. This information might be important in estimating preferences since the placement of a child is dependent on his/her ranking among the set of children in the same market.

To simulate an equilibrium in market t the model utilizes the parametric assumption $v_i \sim N(0,1)$ and the fact that $\{X_i\}_{i\in I_t}$ are assumed to be exogenous characteristics of children. Let $i \in I_t$ to consider children in only one market. For a guess of the preference parameter α and a simulated draw v_i^s, compute

$$u_i^s(\alpha) = X_i\alpha + v_i^s.$$

Sort $i \in I_t$ according to $u_i^s(\alpha)$, $\{u_{i,l}^s(\alpha)\}_l$ where $u_{i,l}^s(\alpha) \geq u_{i,l'}^s(\alpha)$ if and only if $l \leq l'$. Then set $A_i^s = 1$ if the rank of i according to u_i^s is better than or equal to $|J_t|$, the number of available families in market t. For all other i set $A_i^s = 0$. This gives a simulated set of A_i^s. We can then average over these simulations for a market, and average over markets to compute a corresponding estimate of the moment.

For simplicity, I compute the estimator with the identity matrix $\mathbf{W} = \mathbf{I}$ and utilize $S = 25$ total simulations. The optimization routine underlying estimation keeps fixed the set of simulated draws of the unobservable v_i^s across parameter draws to speed up estimation. Because the optimization routine is time consuming, and is especially computationally expensive as a function of the number of markets used in estimation, I estimate the preferences based off of the 20 largest counties and focus only on children entering in 2010 and 2013. I choose larger counties and these years since they have the most overlap with children whose outcomes are measured in the NYTD survey. Leaving out some counties also leaves room for assessing model fit in a hold-out sample.

5.2.3 Estimation of Outcome Parameters

To estimate the parameters in (8) I follow standard two-step control function estimators (i.e. Kline and Walters (2016)). Using the law of iterated expectations (8) can be written as

$$\mathbb{E}[Y_i|X_i, Z_t, A_i = A] = X_i\beta_A + \gamma_A\mathbb{E}[v_i|X_i, Z_t, A_i = A]. \tag{17}$$

The parameters in this equation can be estimated by computing control function estimates $\hat{v}_i$ of $\mathbb{E}[v_i|X_i, Z_t, A_i]$. I compute these estimates directly from the simulated equilibria with v_i^s, α and (X_i, Z_t, A_i), and taking the conditional mean of the v_i^s when i's equilibrium assignment matches what is observed in the data.[22] With the control function estimates $\hat{v}_i$ a second step regression can be used to estimate the outcome parameters:

$$Y_i = \beta_{00} + X_i\beta_{0x} + \gamma_0\hat{v}_i + 1\{A_i = A\}\left[\beta_{10} + X_i\beta_{1x} + \gamma_1\hat{v}_i\right] + \omega_i. \tag{18}$$

To interpret the coefficients $\beta_{10} - \beta_{00}$ as an average treatment effect of being placed with a family, all covariates are normalized to have unconditional mean 0. To avoid overfitting and for reasons of precision, I estimate preferences and outcome parameters with respect only to key demographics: sex, age and race.

To compute standard errors for all parameters, I utilize a block bootstrap at the county level, with 100 bootstrap replications, drawing with replacement and repeating all three steps of the estimation procedure.

5.3 Parameter Results

Table 16 gives parameter estimates for the preference parameters α in (7) and λ in equation (6). It also provides information on the spread of the fixed effects in equation (6).

[22] Because I use 25 simulations in computing equilibrium, some children do not have any simulated equilibrium assignments that match their observed assignment for the estimated parameters. When this occurs, I can still estimate $\hat{v}_i$ using properties of a truncated normal distribution. For example, if a child is never observed to be matched in the simulated equilibria, but is matched in the data, and her highest draw of simulations is V_i^s, then I infer that her true v_i follows a standard normal distribution truncated from below at V_i^s, and thus I can compute a corresponding expectation.

Table 16: Preference and Family Availability Estimates

Family Preference Estimates	
Male	-0.716
	(0.143)
Age 15	-0.208
	(0.091)
Age 16	-0.281
	(0.090)
Age 17	-0.311
	(0.096)
Race: white	-0.006
	(0.101)
Race: black	-0.066
	(0.090)
Race: hispanic	0.060
	(0.089)
Number children	17,731

Family Availability Estimates	
Non-kin exits month	0.0707
	(0.0313)
Sd(county effects)	1.42
Sd(month-year effects)	0.205
Sd(non-kin exits month)	13.5
Number county-month-years	60,068
Number counties	1,208
Number month-years	86

Notes: These tables present estimates of the preference parameters α in (7) and the instrument shifting parameter λ in (6), along with some summaries of the fixed effects in (6). The family preference estimates are estimated according to the simulated method of moments procedure outlined in the text while the family available estimates are estimated according to the OLS regression procedure. Family preferences are estimated on the 20 largest counties for children entering in 2010 and 2013. Family availability estimates are estimated on all identified counties and children between ages 14 and 17. Standard errors for all parameters are computed using a block bootstrap where the blocks are counties with 100 bootstrap replications.

The preference parameters, while not directly quantitatively interpretable, show a few important patterns. First, boys are preferred to girls by families. Second, younger children are preferred to older children. Both of these patterns are consistent with the descriptive patterns in Table 15. The model's incorporation of choice sets and potential spillovers between children allows for a more direct interpretation of these patterns as reflecting preferences of families. The parameter estimates on race do not suggest strong preferences over race of foster children for foster families.

The family availability estimates show that the instrument, non-kin exits at a market level, strongly affects the number of available families. The coefficient estimate can be interpreted as saying that around 7.07% of exits from non-kin families in a market convert to a placement slot for an entering foster child in that market. The sign of this coefficient is consistent with the first stage regressions examined previously. This table also shows that there are sizable differences in family supply across counties with an estimated standard deviation of about 1.42 available families. Variation in seasonality across counties appears to be less important, with a standard deviation of only about 0.2 available families.

Table 17 provides the selection corrected estimates of parameters $(\beta_0, \beta_1, \gamma_0, \gamma_1)$ in equation (8). It compares these estimates to an OLS version of equation (8) that does not include unobserv-

Table 17: Selection Corrected Model Outcome Estimates

Outcome Index	OLS		Selection Corrected	
	Constant Effect	Interaction with Treatment	Constant Effect	Interaction with Treatment
	(1)	(2)	(3)	(4)
ATE	0.796	-	0.888	-
	(0.065)		(0.187)	
Unobservable selection	-	-	-0.024	-0.037
(γ_0, γ_1)			(0.175)	(0.208)
Male	-0.544	0.243	-0.539	0.258
	(0.092)	(0.128)	(0.100)	(0.159)
Age 15 (ref: age 14)	-0.269	0.147	-0.268	0.150
	(0.170)	(0.183)	(0.164)	(0.197)
Age 16	-0.357	0.216	-0.359	0.220
	(0.172)	(0.171)	(0.179)	(0.184)
Age 17	-0.352	0.094	-0.354	0.104
	(0.258)	(0.278)	(0.278)	(0.343)
Race white (ref: race other)	-0.008	-0.003	-0.012	0.003
	(0.205)	(0.255)	(0.214)	(0.325)
Race black	0.071	-0.199	0.074	-0.199
	(0.211)	(0.268)	(0.233)	(0.345)
Race hispanic	0.138	0.056	0.142	0.045
	(0.222)	(0.259)	(0.243)	(0.327)
Number children	5,113		5,113	

Notes: This table presents estimates of the parameters in (18) using the economic and social outcome index defined in the text. Columns (1) and (2) present OLS estimates of this equation. Column (1) gives the average treatment effect (ATE) estimate of the coefficient estimate on the treatment variable, and estimates of coefficients on all the demographic observables listed. Column (2) gives estimates of interactions between all the demographic observables and the treatment indicator. Columns (3)-(4) estimate selection corrected estimates of (18) using the control function procedure described in the text. The only additional terms are estimates of selection on unobservable levels in column (3) and selection on unobservable gains in column (4). The sample of estimation is the outcome sample used previously. Standard errors for all parameters are computed using a block bootstrap where the blocks are counties with 100 bootstrap replications.

able preference estimates. The table shows average treatment effect estimates in columns (1) and (3) of 0.796 and 0.888 for the OLS and selection corrected models respectively. While the ATE and the implied LATE is not necessarily close to the LATE estimated in Table 8 it is larger than the OLS estimated ATE which mirrors Table 8. Part of this discrepancy is likely due to functional form and parametric assumptions imposed by the model that are necessary for the policy counterfactuals done below.[23]

Second, the point estimates of unobservable selection on levels (column (3)) and gains (column (4)) in row 2 are both small at -0.024 and -0.037. While the standard errors are large and the

[23]For example, if we implement a direct Heckman (1979) probit treatment model of (3), we get an ATE estimate of 1.597, showing that some of this discrepancy is likely unavoidable with standard parametric assumptions. Other tests of the model confirm that this is not due to the fact that the instrument is driving the estimates of $\hat{v}_i$. When implementing the same first stage probit model using the standard normal assumptions, the correlation between the estimated $\hat{v}_i$ with the $\hat{v}_i$ in the model implemented here is 0.702.

estimates cannot reject larger unobservable selection, the model suggests that selection on unobservables is minimal compared to other sources of selection.

The last set of rows show the point estimates for β_0 and β_1. Boys and older children are generally more likely to have worse outcomes, as shown by the estimates in columns (1) and (3). However, they are also more likely to benefit from treatment as seen in columns (2) and (4). Once again, standard errors prevent precise conclusions and the interacted treatment male dummy is only marginally statistically significant, but these patterns generally suggest that children that are more likely to have worse outcomes on observables also benefit the most.

To increase precision with the observables and explore the relationship between family preferences and foster child outcomes more parsimoniously, I explore an observable preference index following Walters (2018). In particular, I consider

$$\mathcal{P}_i^o = X_i\hat{\alpha} \tag{19}$$

which collapses all observables into a one-dimensional index based on family preferences. This preference index can be compared with an unobservable preference index:

$$\mathcal{P}_i^u = v_i. \tag{20}$$

Finally, these two preference indices can be combined into an overall preference index that combines both observables and unobservables:

$$\mathcal{P}_i^f = \mathcal{P}_i^o + \mathcal{P}_i^u. \tag{21}$$

Table 18 provides selection corrected estimates using these indices. The estimates of unobservable selection on gains and levels do not change appreciably. Estimates on the observable preference index show that children more preferred by families on observables tend to have better outcomes overall, but also benefit substantially less from treatment, at a rate of -0.344 over the preference index. When observables and unobservables are collapsed into a single preference index, the same pattern remains.

The results in Table 18 will be important in driving the results in the counterfactuals examined in the next section. In particular, they provide suggestive evidence that re-allocating which children are matched to families, especially on observable characteristics, could improve average outcomes of foster children, and also reduce inequality in outcomes between foster children.

Table 18: Child Outcome Equation Estimates: OLS and Selection Corrected, Preference Index Summaries

Outcome Index	OLS		Selection Corrected Decomposed Model		Selection Corrected Aggregate Model	
	Constant Effect (1)	Interaction with Treatment (2)	Constant Effect (3)	Interaction with Treatment (4)	Constant Effect (5)	Interaction with Treatment (6)
ATE	0.793 (0.066)	-	0.890 (0.182)	-	0.674 (0.175)	-
Unobservable preference index ($\mathcal{P}_i^u$)	-	-	-0.032 (0.169)	-0.032 (0.185)	-	-
Observable preference index ($\mathcal{P}_i^o$)	0.780 (0.121)	-0.323 (0.166)	0.772 (0.245)	-0.344 (0.249)	-	-
Overall preference index ($\mathcal{P}_i^f$)	-	-	-	-	0.299 (0.133)	-0.257 (0.154)
Number children	5,113		5,113		5,113	

Notes: This table presents estimates of the parameters in (18) using the economic and social outcome index defined in the text and the indices defined in (19), (20), and (21). Columns (1) and (2) present OLS estimates of this equation using only the observable preference index as a regressor and the interaction between the observable preference index and treatment as a regressor. Columns (3)-(4) estimate selection corrected estimates of (18) using the control function procedure described in the text for the unobservable preference index, and including the observable preference index as another regressor along with an interaction between the observable preference index and the treatment variable. Columns (5) and (6) estimate selection corrected estimates of (18) using the control function procedure where the only regressor is the overall preference index and an interaction between the overall preference index and the treatment variable. The sample of estimation is the outcome sample used previously. Standard errors for all parameters are computed using a block bootstrap where the blocks are counties with 100 bootstrap replications.

5.4 Model Fit Exercise

To assess model fit I also report on two model fit exercises. The first model fit exercise simply reports the R-squared of the available families model estimated in (6). This equation achieves an R-squared of 0.7868, suggesting that fitting the number of available families on the instrument and county and month-by-year fixed effects provides good in-sample fit.

The second model fit exercise uses the fact that the model parameters are not estimated on all counties and so it is possible to look at out-of-sample fit for markets not included in estimation. Table 19 compares covariances in the raw data and model predicted covariances for this hold-out set of markets. These covariances are used in model estimation, but not for these counties. The model predicted and raw covariances are very close, and the model picks up important patterns across sex, age and race.

Table 19: Model Fits on Holdout Sample of Smaller Counties

Child Observable	Raw Data Covariance with Placement	Model Fit Covariance with Placement
Male	0.146	0.179
Age 15	0.105	0.124
Age 16	0.107	0.127
Age 17	0.0701	0.0836
Race white	0.138	0.153
Race black	0.143	0.180
Race hispanic	0.0846	0.0940
Number of children	16,524	16,524

Notes: This table reports model fit exercises for the structural model. It compares the raw covariances between child placement with a family and child demographics and compares it to the predicted model covariances on the sample of smaller counties that are not used in the model estimation. To get the predicted model covariances I simulate equilibrium according to the model in these other markets and then compute the covariances in the same way.

6 Foster Family Preferences, Supply Increases and Reallocation in Foster Care

This section studies the main counterfactuals motivated by understanding the role of family preferences in determining foster child allocations and outcomes, and evaluating how different foster care placement policies affect foster children's outcomes.

6.1 Foster Family Preferences and the Random Allocation

This subsection explores the role of family preferences in determining foster child aggregate foster child outcomes through the model's allocation mechanism. To quantify this, I compute a counterfactual exercise in which I randomly allocate foster children in each market to the number of available families in that market. Conceptually, when comparing outcomes in this counterfactual to outcomes in the observed equilibrium allocation, I can measure how foster families preferences and equilibrium choices influence aggregate foster child outcomes.

Table 20 shows the results from this counterfactual exercise. The average foster child has better outcomes under the random allocation scheme by 2.1%. This is mainly driven by the increase in outcomes for boys, who represent about half of the foster care market but are placed at substantially lower rates than girls and benefit more from placement.

Overall these results suggest that the accounting for how children are selected by families on observables, family preferences reduce average outcomes of foster children in equilibrium. This is consistent with a scenario reflected in the estimates in Table 17 where families prefer children that are easier to take care of (younger, girls), but these children also benefit less from placement in a

Table 20: Foster Child Outcomes Under Baseline and Random Allocation

	Equilibrium child allocation	Random child allocation
	(1)	(2)
Average foster child outcome	0.963	0.983
Average foster boy outcome	0.666	0.764
Average foster girl outcome	1.31	1.25
Average white foster child outcome	0.849	0.871
Average black foster child outcome	0.849	0.883
Average hispanic foster child outcome	1.17	1.19
Number foster children considered	40,762	40,762

Notes: Column (1) gives mean equilibrium child outcomes predicted using model parameter estimates from (18) and the observed placements of foster children entering between ages 14 and 17. Column (2) gives mean child outcomes predicted using model parameter estimates from (18) and randomly choosing which children are allocated to families in equilibrium, holding fixed the number of available families using the equilibrium condition. It also gives mean outcomes by different demographic groups.

foster family relative to a group home.

6.2 Adding More Foster Families

This subsection evaluates a policy that increases the supply of foster families using the model. To simulate such a policy, accounting for market conditions, I consider policies that increase the number of available families $|J_t|$ in each market t by some percentage p so that in the counterfactual market $|J_t|_{CF} = (1+p)|J_t|$. If this number is not an integer I round it to the nearest integer. This allows for proportional increases in family supply across markets and counties, and accounts for the fact that these proportional increases are more likely to be equivalently costly for local governments than absolute increases in a number of families across all markets due to population differences between counties.

The top left panel in Figure 4 provides model predictions on the average outcome of foster children, according to the economic and social outcome index, for different values of an increase in the number of families according to the parameter p. As would be naturally predicted by the reduced form results, an increase in families increases average outcomes. The model further incor-

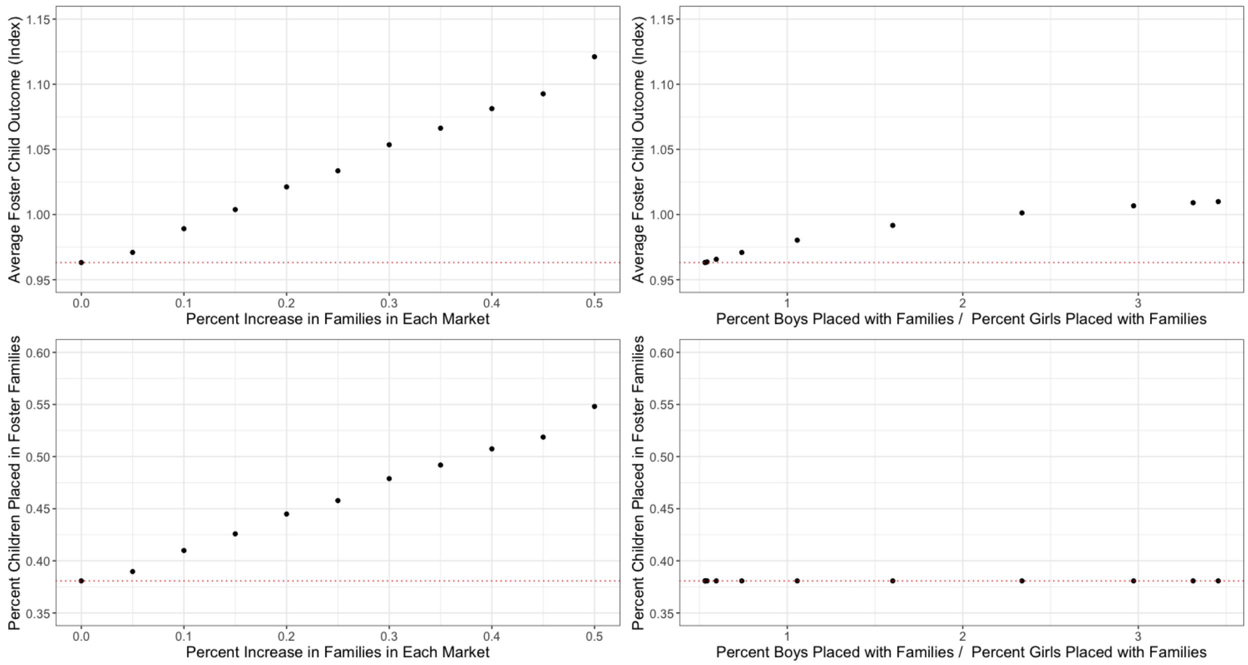

Figure 4: Average Foster Child Outcomes and Family Placement under Family Supply and Allocation Policies

Notes: This figure present counterfactual estimates of average foster child outcomes using parameter estimates and the model in (18), the family availability model assumptions, and the family preference assumptions. The left panel gives the average foster child outcomes as a function of p where for each market considered I add p percent of families to the stock of available families and resimulate equilibrium allocations and foster child outcomes, holding fixed all other parts of the model. The right panel gives average foster child outcomes as a function of the mean placement of boys divided by the mean placement of girls (percent of boys placed / percent of girls placed). These outcomes are simulated using the same model but by shifting α_{boy} in the preference model holding fixed all other parts of the model.

porates that the marginal children selected by families in this policy may be different from average children on both observables and unobservables. With these factors taken into account the model predicts that a 10% increase in the availability of families leads to an approximate 2.6% increase in average foster child outcomes, a 20% increase leads to an approximate 5.7% increase in average foster child outcomes, and a 50% increase in the availability off families leads to an approximate 15.1% increase in average foster child outcomes.

The bottom left panel in Figure 4 shows how the percentage of children placed in foster families increases as the counterfactual parameter p varies. As expected, adding foster families increases the percent of children placed in families.

An important assumption in these counterfactuals is the assumption that an added family finds all foster children acceptable despite their ranking over children. If this is not the case, then this model will overpredict the benefits of family supply. For small p, i.e. marginal policy changes, it appears reasonable to think this is the case.

6.3 Reallocation of Children to Families in Foster Care

The selection corrected outcome estimates and preference estimates suggest another that changing the allocation of children to families, holding constant the number of available families, could also be an effective policy. This may be desirable to policymakers, as recruiting foster families can be expensive. As a concrete example, the results so far found that boys benefit substantially more than girls from placement in foster families. Since boys are placed at a much lower rate than girls, a policy that encourages or subsidizes the placement of boys with families in foster care could increase average outcomes, and also reduce inequality between foster children (since boys also have worse outcomes on average). These types of policy comparisons and counterfactuals are similar to Abaluck et al. (2016), which compares inefficiencies in health care utilization on both the intensity and allocation margins.

To examine the effectiveness of policies that reallocate children to families holding families fixed, I focus on a simply policy in which I change preference parameter $\alpha_{CF}^{boy} = \hat{\alpha}^{boy} + \zeta$ where ζ can be seen as the subsidy parameter for the foster boys in the foster care market. A concrete real world analogy to this policy involves subsidizing placements with boys more highly than girls. This type of payment difference is already present in other demographics, as older children command a higher subsidy than younger children. An important next step that is out of the scope of the model here would be to measure the quantitative implications of the cost of the policy implemented here in terms of dollars.

The right top panel in Figure 4 shows how average outcomes of foster children change as the ζ parameter changes where the x-axis is measured in terms of the ratio of the percent of boys placed to the percent of girls placed. As the figure shows, if ζ is changed sufficiently so that the mean placement rate of boys and the mean placement rate of girls was equalized, the foster care system would be able to achieve an approximately 1.75% increase in average foster children outcomes. This is comparable to the gains achieved by a 5-10% increase in family availability shown in the left panel. The bottom right panel in Figure 4 shows that the overall percentage of children placed in foster families does not change in this counterfactual, since this is only a reallocation exercise.

This policy does not have to be specific to boys and girls, but can be generalized to all observable demographics considered in the model. On all the demographic observables used in the model, the sign of the preference parameters is flipped so that $\alpha_{CF} = -\hat{\alpha}$ there are similar gains to child outcomes, on the order of 3.5%. This is due to the selection patterns and treatment effect interactions on observables found in the previous section in Table 17. Figure A4 shows the results from this exercise.

6.4 Policy Comparison Discussion

These results raise questions about what the policy priorities of foster care policymakers should be in their goals to improve child outcomes. Policies that change the allocation of foster children also appear to have scope to improve foster child outcomes and may be less costly to implement. The subsidies paid by counties for children of different ages already reflect some moves in the right direction: families are generally given more money and allowance for taking care of older children than for taking care of younger children. The results here suggest that policies that continue to move in this direction, subsidizing the placement of older children and perhaps boys, could be as effective as policies that increase family supply, and also not leave behind children more prone to the worse outcomes studied in this paper.

7 Conclusion

Foster care is an important social service in the US affecting hundreds of thousands of abused and neglected children every year. Previous results in the literature find negative effects of foster care on children's later outcomes, which may lead to skepticism of the foster care system's effectiveness, at least for marginal child cases. This paper shows that the distinction between a child placed with a family in foster care and a child placed in a group home is important for understanding these effects. The results here imply that a focus towards increasing the reliance on non-kin families for caring for foster children could substantially improve their later outcomes. But this paper also found that, even if adding non-kin families is too costly, average outcomes for foster children can be improved by considering how children are allocated to foster families.

These results point out that if policymakers are committed to removing older children from their homes, then they should be ready to provide substitute families for these children to remain with instead of group homes. One margin of analysis that this paper does not speak to is whether non-kin families or kin families represent better placements for improving children's outcomes. The focus in this paper was only on non-kin families and group homes. To assess these causal effects require a much broader and stronger array of instruments to shift children between the three different placements (Kirkeboen, Leuven and Mogstad, 2016). The results in the literature so far suggest minor differences between kin and non-kin families on intermediate outcomes such as placement stability (Andersen and Fallesen, 2015) or rely on strong assumptions on either specific institutionally derived instrumental variables or how children are allocated to families based on observables (Font, 2014; Lovett and Xue, 2020). Future work incorporating kin placements in the analysis would allow for a more comprehensive picture of foster care placements and foster child outcomes.

References

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

Abaluck, Jason, Leila Agha, Chris Kabrhel, Ali Raja, and Arjun Venkatesh (2016) "The Determinants of Productivity in Medical Testing: Intensity and Allocation of Care," *American Economic Review*, 106 (12), 3730–3764.

Abdulkadiroğlu, Atila, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters (2020) "Do Parents Value School Effectiveness?" *American Economic Review*, 110 (5), 1502–1539.

Adams, Char (2020) "Foster care crisis: More kids are entering, but fewer families are willing to take them in," NBC News, December.

Agarwal, Nikhil (2015) "An Empirical Model of the Medical Match," *American Economic Review*, 105 (7), 1939–1978.

Agarwal, Nikhil, Charles Hodgson, and Paulo Somaini (2020) "Choices and Outcomes in Assignment Mechanisms: The Allocation of Deceased Donor Kidneys," Working Paper 28064, National Bureau of Economic Research, Series: Working Paper Series.

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.

Andersen, Signe Hald and Peter Fallesen (2015) "Family matters? The effect of kinship care on foster care disruption rates - ScienceDirect," *Child Abuse & Neglect*, 48, 68–79.

Angrist, Joshua D. and Jörn-Steffen Pischke (2008) *Mostly Harmless Econometrics*: Princeton University Press, Publication Title: Mostly Harmless Econometrics.

Bailey, Martha J., Hilary W. Hoynes, Maya Rossin-Slater, and Reed Walker (2020) "Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence from the Food Stamps Program," Working Paper 26942, National Bureau of Economic Research, Series: Working Paper Series.

Bald, Anthony, Eric Chyn, Justine S. Hastings, and Margarita Machelett (2019) "The Causal Impact of Removing Children from Abusive and Neglectful Homes," January, NBER Working Paper 25419.

Barth, Richard P. (2002) "Institutions vs. Foster Homes: The Empirical Base for a Century of Action,"Technical report, Jordan Institute for Families, School of Social Work, UNC.

Becker, Marion A., Neil Jordan, and Rebecca Larsen (2007) "Predictors of successful permanency planning and length of stay in foster care: The role of race, diagnosis and place of residence," *Children and Youth Services Review*, 29 (8), 1102–1113.

Berry, Steven and Ariel Pakes (2007) "The Pure Characteristics Demand Model*," *International Economic Review*, 48 (4), 1193–1225, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2354.2007.00459.x.

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

Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Løken, and Magne Mogstad (2020) "Incarceration, Recidivism, and Employment," *Journal of Political Economy*, 128 (4), 1269–1324, Publisher: The University of Chicago Press.

Biehal, Nina (2014) "A Sense of Belonging: Meanings of Family and Home in Long-Term Foster Care," *The British Journal of Social Work*, 44 (4), 955–971.

Brinch, Christian N., Magne Mogstad, and Matthew Wiswall (2017) "Beyond LATE with a Discrete Instrument," *Journal of Political Economy*, 125 (4), 985–1039, Publisher: The University of Chicago Press.

California Department of Social Services (2021) "Continuum of Care Reform."

Card, David (1999) "Chapter 30 - The Causal Effect of Education on Earnings," in Ashenfelter, Orley C. and David Card eds. *Handbook of Labor Economics*, 3, 1801–1863: Elsevier.

Cherry, Donna J. and John G. Orme (2013) "The vital few foster mothers," *Children and Youth Services Review*, 35 (9), 1625–1633.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. 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.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014) "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States *," *The Quarterly Journal of Economics*, 129 (4), 1553–1623.

Child Welfare Monitor (2020) "Around the country, states face shortage of placements for youth with greater needs," Online Article, January.

Children's Bureau (2016) "Child Maltreatment 2016."

——— (2020) "AFCARS Report #27."

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

Cohen, Elior (2021) "Housing the Homeless: The Effect of Housing Assistance on Recidivism to Homelessness, Economic, and Social Outcomes," June.

Collinson, Robert and Davin Reed (2018) "The Effects of Evictions on Low-Income Households," Working Paper.

County, Los Angeles (2020) "0900-511.10, Rates for Placement and Related Services Los Angeles County."

Dahl, Gordon B., Andres Ravndal Kostol, and Magne Mogstad (2014) "Family Welfare Cultures," *The Quarterly Journal of Economics*, 129 (4), 1711–1752.

Danielson, Caroline and Helen Lee (2010) "Foster Care in California Achievements and Challenges,"Technical report, Public Policy Institute of California.

DeGarmo, Dr John (2019) "The Foster Care Crisis: The Shortage Of Foster Parents In America," July.

van Dijk, Winnie (2019) "The socio-economic consequences of housing assistance," January, Working Paper.

Dippel, Christian, Robert Gold, Stephan Heblich, and Rodrigo Pinto (2020) "Mediation Analysis in IV Settings With a Single Instrument," July.

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), 201–240.

Doyle, Joseph J. (2007a) "Can't buy me love? Subsidizing the care of related children," *Journal of Public Economics*, 91 (1), 281–304.

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

——— (2008) "Child Protection and Adult Crime: Using Investigator Assignment to Estimate Causal Effects of Foster Care," *Journal of Political Economy*, 116 (4), 746–770.

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

Fagereng, Andreas, Magne Mogstad, and Marte Rønning (2021) “Why Do Wealthy Parents Have Wealthy Children?” *Journal of Political Economy*, 129 (3), 703–756, Publisher: The University of Chicago Press.

Font, Sarah A. (2014) “Kinship and Nonrelative Foster Care: The Effect of Placement Type on Child Well-Being,” *Child Development*, 85 (5), 2074–2090.

Font, Sarah, Lawrence M. Berger, Jessie Slepicka, and Maria Cancian (2021) “Foster Care, Permanency, and Risk of Prison Entry,” *Journal of Research in Crime and Delinquency*, 00224278211001566, Publisher: SAGE Publications Inc.

Freundlich, Madelyn and Rosemary J. Avery (2006) “Transitioning from congregate care: Preparation and outcomes,” *Journal of Child and Family Studies*, 15 (4), 503–514.

Gross, Max and E. Jason Baron (Forthcoming) “Temporary Stays and Persistent Gains: The Causal Effects of Foster Care,” *American Economic Journal: Applied Economics*.

Guryan, Jonathan, Jens Ludwig, Monica P. Bhatt et al. (2021) “Not Too Late: Improving Academic Outcomes Among Adolescents,” March, NBER Working Paper 28531.

Heckman, James J. (1979) “Sample Selection Bias as a Specification Error,” *Econometrica*, 47 (1), 153–161.

Heckman, James J. and Edward Vytlacil (2005) “Structural Equations, Treatment Effects, and Econometric Policy Evaluation1,” *Econometrica*, 73 (3), 669–738, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0262.2005.00594.x.

Herman, D B, E S Susser, E L Struening, and B L Link (1997) “Adverse childhood experiences: are they risk factors for adult homelessness?” *American Journal of Public Health*, 87 (2), 249–255, Publisher: American Public Health Association.

Hitsch, Gunter J., Ali Hortacsu, and Dan Ariely (2010) “Matching and Sorting in Online Dating,” *American Economic Review*, 100 (1), 130–163.

Imbens, Guido W. and Joshua D. Angrist (1994) “Identification and Estimation of Local Average Treatment Effects,” *Econometrica*, 62 (2), 467–475, Publisher: [Wiley, Econometric Society].

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman (2001) "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *The Quarterly Journal of Economics*, 116 (2), 607–654.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad (2016) "Field of Study, Earnings, and Self-Selection," *The Quarterly Journal of Economics*, 131 (3), 1057–1111.

Kline, Patrick and Christopher R. Walters (2016) "Evaluating Public Programs with Close Substitutes: The Case of Head Start*," *The Quarterly Journal of Economics*, 131 (4), 1795–1848.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz (2007) "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75 (1), 83–119.

Koegel, P, E Melamid, and m A Burnam (1995) "Childhood risk factors for homelessness among homeless adults.," *American Journal of Public Health*, 85 (12), 1642–1649, Publisher: American Public Health Association.

Koh, Eun and Mark F. Testa (2008) "Propensity Score Matching of Children in Kinship and Nonkinship Foster Care: Do Permanency Outcomes Differ?" *Social Work Research*, 32 (2), 105–116.

Lancaster, Kelvin J. (1966) "A New Approach to Consumer Theory," *Journal of Political Economy*, 74 (2), 132–157, Publisher: University of Chicago Press.

Lee, David S. (2009) "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *The Review of Economic Studies*, 76 (3), 1071–1102.

Lockwood, Katie K., Susan Friedman, and Cindy W. Christian (2015) "Permanency and the Foster Care System," *Current Problems in Pediatric and Adolescent Health Care*, 45 (10), 306–315.

Lovett, Nicholas and Yuhan Xue (2020) "Family first or the kindness of strangers? Foster care placements and adult outcomes - ScienceDirect," *Labour Economics*.

Ludwig, Jens, Jeffrey B. Liebman, Jeffrey R. Kling, Greg J. Duncan, Lawrence F. Katz, Ronald C. Kessler, and Lisa Sanbonmatsu (2008) "What Can We Learn about Neighborhood Effects from the Moving to Opportunity Experiment?" *American Journal of Sociology*, 114 (1), 144–188, Publisher: The University of Chicago Press.

McFadden, Daniel (1989) "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57 (5), 995–1026, Publisher: [Wiley, Econometric Society].

Mersky, J. P., J. Topitzes, and A. J. Reynolds (2013) "Impacts of adverse childhood experiences on health, mental health, and substance use in early adulthood: A cohort study of an urban, minority sample in the U.S.," *Child Abuse & Neglect*, 37 (11), 917–925.

National Institute on Drug Abuse (2021) "Is drug addiction treatment worth its cost?."

Nelson III, Charles A, Charles H. Zeanah, Nathan A. Fox, Peter J. Marshall, Smyke Anna T., and Donald Guthrie (2007) "Cognitive Recovery in Socially Deprived Young Children: The Bucharest Early Intervention Project," *Science*.

Olea, José Luis Montiel and Carolin Pflueger (2013) "A Robust Test for Weak Instruments," *Journal of Business & Economic Statistics*, 31 (3), 358–368.

Ours, Jan C. van and Jenny Williams (2015) "Cannabis Use and Its Effects on Health, Education and Labor Market Success," *Journal of Economic Surveys*, 29 (5), 993–1010, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/joes.12070.

Pakes, Ariel and David Pollard (1989) "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57 (5), 1027–1057, Publisher: [Wiley, Econometric Society].

Paxson, Christina and Jane Waldfogel (2003) "Welfare reforms, family resources, and child maltreatment," *Journal of Policy Analysis and Management*, 22 (1), 85–113, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pam.10097.

Reiersøl, Olav (1941) "Confluence Analysis by Means of Lag Moments and Other Methods of Confluence Analysis," *Econometrica*, 9 (1), 1–24, Publisher: [Wiley, Econometric Society].

Robinson-Cortes, Alejandro (2019) "Who Gets Placed Where and Why? An Empirical Framework for Foster Care Placement," Working Paper.

Roos, Leslie E., Natalie Mota, Tracie O. Afifi, Laurence Y. Katz, Jino Distasio, and Jitender Sareen (2013) "Relationship Between Adverse Childhood Experiences and Homelessness and the Impact of Axis I and II Disorders," *American Journal of Public Health*, 103 (Suppl 2), S275–S281.

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.

Sacerdote, Bruce (2007) "How Large are the Effects from Changes in Family Environment? A Study of Korean American Adoptees*," *The Quarterly Journal of Economics*, 122 (1), 119–157.

Stock, James and Motohiro Yogo (2005) *Identification and Inference for Econometric Models*, Chap. Testing for Weak Instruments in Linear IV Regression, 80–108, New York: Cambridge University Press, Pages: 80-108.

Trends, Child (2018) "Child Welfare Agency Spending SFY 2018: California,"Technical report, Child Trends.

Unrau, Yvonne A., Sarah A. Font, and Glinda Rawls (2012) "Readiness for college engagement among students who have aged out of foster care," *Children and Youth Services Review*, 34 (1), 76–83.

Van Ours, Jan C. (2007) "The Effects of Cannabis Use on Wages of Primeage Males," *Oxford Bulletin of Economics and Statistics*, 69 (5), 619–634.

Van Ours, Jan C. and Jenny Williams (2015) "CANNABIS USE AND ITS EFFECTS ON HEALTH, EDUCATION AND LABOR MARKET SUCCESS," *Journal of Economic Surveys*, 29 (5), 993–1010.

Wager, Stefan and Susan Athey (2018) "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests," *Journal of the American Statistical Association*, 113 (523), 1228–1242.

Waldfogel, Jane (2000) "Child welfare research: How adequate are the data?" *Children and Youth Services Review*, 22 (9), 705–741.

Walters, Christopher R. (2018) "The Demand for Effective Charter Schools," *Journal of Political Economy*, 126 (6), 2179–2223.

WeHaveKids.com (2020) "Getting Paid to Be a Foster Parent: State-by-State Monthly Guide," July.

Wong, Yin-Ling Irene, Jung Min Park, and Howard Nemon (2006) "Homeless Service Delivery in the Context of Continuum of Care," *Administration in Social Work*, 30 (1), 67–94.

Wooldridge, Jeffrey M. (1999) "Asymptotic Properties of Weighted M-estimators for variable probability samples," *Econometrica*, 67 (6), 1384–1406.

Appendix

Table A1: First Stage

	Outcome Sample (1)	Eligible Sample (2)	Eligible Weighted (3)	Old Children Sample (4)	Old Children Weighted (5)
Panel A:					
Instrument: Non-kin exits					
First stage coefficient and s.e.	0.00206 (0.00032)	0.00128 (0.00019)	0.00091 (0.00018)	0.00083 (0.00025)	0.00088 (0.00015)
Cluster robust F-statistic	41.7	43.5	23.7	10.7	33.2
Panel B:					
Instrument: Non-kin exits / log(county pop)					
First stage coefficient and s.e.	0.0319 (0.0049)	0.0195 (0.0032)	0.0146 (0.0029)	0.0125 (0.0042)	0.0137 (0.0026)
Cluster robust F-statistic	43.0	36.8	25.1	9.0	27.8
Panel C:					
Instrument: Non-kin exits / log(county pop) w/ total entry control					
First stage coefficient and s.e.	0.0315 (0.0048)	0.0209 (0.0031)	0.0161 (0.0031)	0.0154 (0.0035)	0.0153 (0.0026)
Cluster robust F-statistic	42.3	45.0	26.7	19.6	33.6
Panel D:					
Instrument: log(1+ non-kin exits)					
First stage coefficient and s.e.	0.0234 (0.0142)	0.0064 (0.0084)	0.0465 (0.0171)	0.0202 (0.0035)	0.0314 (0.0119)
Cluster robust F-statistic	2.7	0.6	7.4	33.2	7.0
County, month x year fixed effects	Y	Y	Y	Y	Y
Child demographic, entry controls	Y	Y	Y	Y	Y
Weighted by county representation in outcome sample	N	N	Y	N	Y
Number observations (children)	5,113	18,461	18,461	209,075	209,075

Notes: This table reports OLS first stage coefficients and cluster robust F-statistics where standard errors and F-statistics are computed with county-clustered robust standard errors. Column (1) shows results in the outcome sample, column (2) shows results in the eligible sample, column (3) shows results in the eligible sample where observations are weighted by county representation in the outcome sample (observation weight = percent of observations in outcome sample with same county as observation), column (3) shows results in the old children sample and column (4) shows results in the old children sample where observations are weighted by county representation in the outcome sample. Panel A presents specifications with the raw instrument and no county normalization. These coefficients can be interpreted as the probability increase in placement with a family for one more exit of a child from a non-kin family in the same county-month-year in which the child exits through reunification or emancipation. Panel B presents specifications with instrument divided by log county population. Panel C presents specifications where the instrument is divided by log county population with an additional covariate of total entries in that same county-month-year. Panel D presents specifications where the instrument is log(1+exits) where exits is defined as in Panel A.

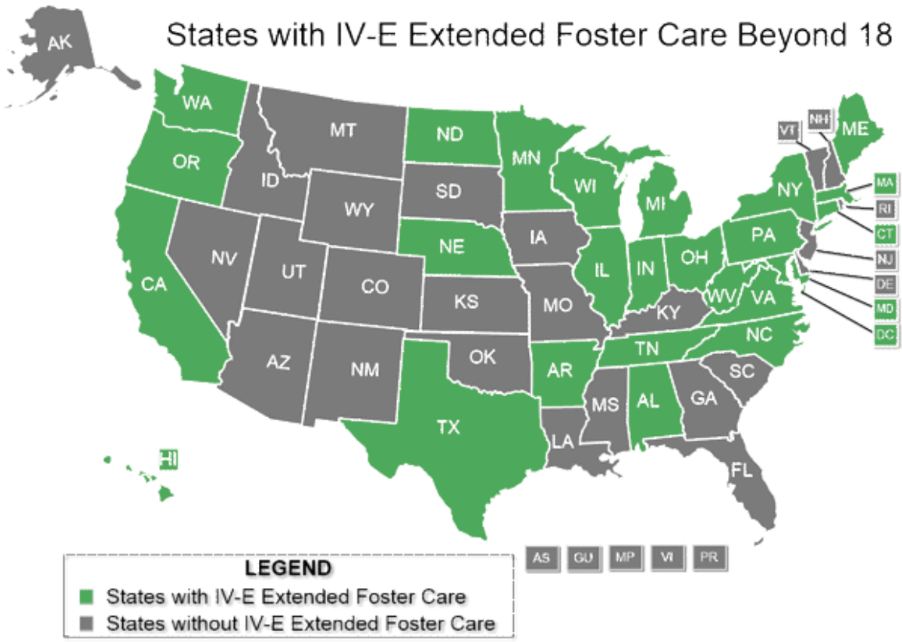

Figure A1: State Foster Care Maximum Age

Notes: This figure shows the states in the US that extend foster care support beyong the age of 18. Source: NCSL https://www.ncsl.org/research/human-services/extending-foster-care-to-18.aspx.

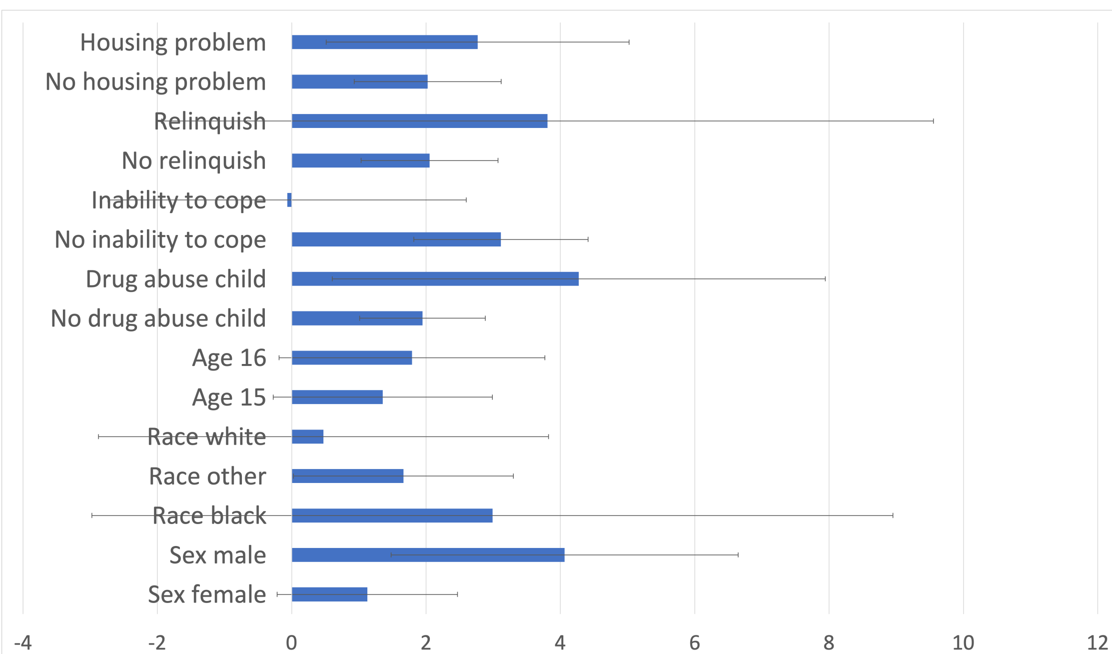

Figure A2: Heterogeneous Treatment Effect Summary

Notes: This figure plots heterogeneous treatment effects of initial non-kin foster family placement vs. initial group home placement on the outcome index considered in Table 8 along with 95% confidence intervals. These heterogeneous treatment effects are estimated by restricting to different subsamples. County and month by year fixed effects are included in these regressions. Standard errors are clustered at the county level.

Table A2: Heterogeneous Effects: Gender and Race

	First Stage	Reduced Form	IV	OLS	First Stage Eligible Sample
	(1)	(2)	(3)	(4)	(5)
Subgroup: Female					
Coefficient and s.e.	0.044	0.0498	1.125	0.682	0.0269
	(0.0091)	(0.0308)	(0.671)	(0.079)	(0.0067)
Cluster robust F-statistic	23.5	-	-	-	15.88
Number of children		2,967			
Instrument		Non-kin exits / log(population)			
Subgroup: Male					
Coefficient and s.e.	0.038	0.153	4.064	0.943	0.0271
	(0.010)	(0.036)	(1.292)	(0.124)	(0.0050)
Cluster robust F-statistic	14.0	-	-	-	28.50
Number of children		2,146			
Instrument		Non-kin exits / log(population)			
Subgroup: Black					
Coefficient and s.e.	0.0238	0.071	2.46	0.731	0.020
	(0.0122)	(0.070)	(3.12)	(0.121)	(0.0072)
Cluster robust F-statistic	3.8	-	-	-	7.8
Number of children		1,532			
Instrument		Non-kin exits / log(population)			
Subgroup: Hispanic					
Coefficient and s.e.	0.0541	0.163	3.092	1.127	0.043
	(0.0129)	(0.043)	(0.872)	(0.151)	(0.0047)
Cluster robust F-statistic	17.6	-	-	-	85.9
Number of children		1,051			
Instrument		Non-kin exits / log(population)			
Subgroup: White					
Coefficient and s.e.	0.0467	0.022	0.473	0.979	0.0236
	(0.015)	(0.079)	(1.68)	(0.12)	(0.011)
Cluster robust F-statistic	10.13	-	-	-	4.25
Number of children		2,265			
Instrument		Non-kin exits / log(population)			

Notes: This table shows OLS results for the first stage (initial placement with family indicator on instrument), reduced form (outcome index on instrument), IV treatment effects and OLS treatment effects in the outcome sample, and the first stage in the eligible sample in different subgroups of the data based on child observables. All models do not include child demographic and entry reason controls. All models include county and month-by-year fixed effects, and standard errors are clustered at the county level. Blank entries in the table indicate that the sample size is too small to provide informative estimates.

Table A3: Heterogeneous Effects: Age

	First Stage	Reduced Form	IV	OLS	First Stage Eligible Sample
	(1)	(2)	(3)	(4)	(5)
Subgroup: Age 14					
Coefficient and s.e.	-0.007	0.0264	-	0.773	-0.023
	(0.0023)	(0.0867)		(0.178)	(0.025)
Cluster robust F-statistic	0.09	-	-	-	0.86
Number of children		615			1,809
Instrument		Non-kin exits / log(population)			
Subgroup: Age 15					
Coefficient and s.e.	0.061	0.083	1.357	0.959	0.0435
	(0.015)	(0.055)	(0.815)	(0.131)	(0.0070)
Cluster robust F-statistic	17.0	-	-	-	38.7
Number of children		1,454			4,560
Instrument		Non-kin exits / log(population)			
Subgroup: Age 16					
Coefficient and s.e.	0.034	0.061	1.792	0.949	0.0276
	(0.0078)	(0.030)	(0.989)	(0.098)	(0.00345)
Cluster robust F-statistic	19.3	-	-	-	64.3
Number of children		2,722			10,272
Instrument		Non-kin exits / log(population)			
Subgroup: Age 17					
Coefficient and s.e.	0.0623	0.875	-	1.140	0.020
	(0.0500)	(0.206)	-	(0.592)	(0.020)
Cluster robust F-statistic	1.6	-	-	-	1.0
Number of children		322			1,820
Instrument		Non-kin exits / log(population)			

Notes: This table shows OLS results for the first stage (initial placement with family indicator on instrument), reduced form (outcome index on instrument), IV treatment effects and OLS treatment effects in the outcome sample, and the first stage in the eligible sample in different subgroups of the data based on child observables. All models do not include child demographic and entry reason controls. All models include county and month-by-year fixed effects, and standard errors are clustered at the county level. Blank entries in the table indicate that the sample size is too small to provide informative estimates.

Table A4: Heterogeneous Effects: Physical and Sexual Abuse

	First Stage (1)	Reduced Form (2)	IV (3)	OLS (4)	First Stage Eligible Sample (5)
Subgroup: No Physical Abuse					
Coefficient and s.e.	0.049	0.0806	1.648	0.930	0.0288
	(0.0052)	(0.0236)	(0.510)	(0.071)	(0.0042)
Cluster robust F-statistic	88.5	-	-	-	46.8
Number of children		4,550			16,666
Instrument		Non-kin exits / log(population)			
Subgroup: Physical Abuse					
Coefficient and s.e.	0.0277	0.143	5.185	0.664	0.0177
	(0.0191)	(0.061)	(3.261)	(0.231)	(0.00863)
Cluster robust F-statistic	2.1	-	-	-	4.2
Number of children		563			1,795
Instrument		Non-kin exits / log(population)			
Subgroup: No Sexual Abuse					
Coefficient and s.e.	0.0459	0.0888	1.935	0.921	0.0290
	(0.0057)	(0.0223)	(0.514)	(0.070)	(0.00379)
Cluster robust F-statistic	65.4	-	-	-	58.7
Number of children		4,670			17,241
Instrument		Non-kin exits / log(population)			
Subgroup: Sexual Abuse					
Coefficient and s.e.	0.0254	-0.151	-	0.156	0.0391
	(0.0421)	(0.098)		(0.251)	(0.0150)
Cluster robust F-statistic	0.4	-	-	-	6.8
Number of children		443			1,220
Instrument		Non-kin exits / log(population)			

Notes: This table shows OLS results for the first stage (initial placement with family indicator on instrument), reduced form (outcome index on instrument), IV treatment effects and OLS treatment effects in the outcome sample, and the first stage in the eligible sample in different subgroups of the data based on child observables. All models do not include child demographic and entry reason controls. All models include county and month-by-year fixed effects, and standard errors are clustered at the county level. Blank entries in the table indicate that the sample size is too small to provide informative estimates.

Table A5: Heterogeneous Effects: Neglect and Inability to Cope

	First Stage (1)	Reduced Form (2)	IV (3)	OLS (4)	First Stage Eligible Sample (5)
Subgroup: No Neglect					
Coefficient and s.e.	0.0504	0.147	2.917	0.891	0.0334
	(0.0076)	(0.035)	(0.754)	(0.095)	(0.00498)
Cluster robust F-statistic	44.4	-	-	-	44.8
Number of children		3,109			11,688
Instrument		Non-kin exits / log(population)			
Subgroup: Neglect					
Coefficient and s.e.	0.0248	-0.0036	-0.145	0.736	0.0058
	(0.0072)	(0.0341)	(1.382)	(0.113)	(0.00495)
Cluster robust F-statistic	11.78	-	-	-	1.36
Number of children		2,004			6,773
Instrument		Non-kin exits / log(population)			
Subgroup: No Inability to Cope					
Coefficient and s.e.	0.0458	0.143	3.116	0.840	0.0309
	(0.0062)	(0.024)	(0.648)	(0.088)	(0.0061)
Cluster robust F-statistic	55.32	-	-	-	25.5
Number of children		3,964			14,681
Instrument		Non-kin exits / log(population)			
Subgroup: Inability to Cope					
Coefficient and s.e.	0.0360	-0.0024	-0.067	0.931	0.0293
	(0.010)	(0.0478)	(1.334)	(0.145)	(0.0068)
Cluster robust F-statistic	12.4	-	-	-	18.7
Number of children		1,149			
Instrument		Non-kin exits / log(population)			

Notes: This table shows OLS results for the first stage (initial placement with family indicator on instrument), reduced form (outcome index on instrument), IV treatment effects and OLS treatment effects in the outcome sample, and the first stage in the eligible sample in different subgroups of the data based on child observables. All models do not include child demographic and entry reason controls. All models include county and month-by-year fixed effects, and standard errors are clustered at the county level. Blank entries in the table indicate that the sample size is too small to provide informative estimates.

Table A6: Heterogeneous Effects: Alcohol Abuse, Drug Abuse Parent

	First Stage (1)	Reduced Form (2)	IV (3)	OLS (4)	First Stage Eligible Sample (5)
Subgroup: No Alcohol Abuse Parent					
Coefficient and s.e.	0.0460 (0.0058)	0.0953 (0.0224)	2.073 (0.529)	0.889 (0.067)	0.0294 (0.0041)
Cluster robust F-statistic	63.2	-	-	-	51.8
Number of children		4,919			17,776
Instrument		Non-kin exits / log(population)			
Subgroup: Alcohol Abuse Parent					
Coefficient and s.e.	-	-	-	-	-
Cluster robust F-statistic	-	-	-	-	-
Number of children		-			-
Instrument		Non-kin exits / log(population)			
Subgroup: No Drug Abuse Parent					
Coefficient and s.e.	0.0459 (0.0056)	0.0896 (0.0241)	1.950 (0.577)	0.937 (0.068)	0.0312 (0.00364)
Cluster robust F-statistic	67.7	-	-	-	73.3
Number of children		4,672			16,607
Instrument		Non-kin exits / log(population)			
Subgroup: Drug Abuse Parent					
Coefficient and s.e.	0.0601 (0.0460)	-0.138 (0.289)	-	0.389 (0.411)	0.0112 (0.0158)
Cluster robust F-statistic	1.7	-	-	-	0.5
Number of children		441			1,854
Instrument		Non-kin exits / log(population)			

Notes: This table shows OLS results for the first stage (initial placement with family indicator on instrument), reduced form (outcome index on instrument), IV treatment effects and OLS treatment effects in the outcome sample, and the first stage in the eligible sample in different subgroups of the data based on child observables. All models do not include child demographic and entry reason controls. All models include county and month-by-year fixed effects, and standard errors are clustered at the county level. Blank entries in the table indicate that the sample size is too small to provide informative estimates.

Table A7: Heterogeneous Effects: Alcohol Abuse, Drug Abuse Child

	First Stage (1)	Reduced Form (2)	IV (3)	OLS (4)	First Stage Eligible Sample (5)
Subgroup: No Alcohol Abuse Child					
Coefficient and s.e.	0.0461 (0.0056)	0.0936 (0.0218)	2.028 (0.513)	0.875 (0.068)	0.0299 (0.0042)
Cluster robust F-statistic	68.3	-	-	-	50.0
Number of children		5,016			18,045
Instrument		Non-kin exits / log(population)			
Subgroup: Alcohol Abuse Child	-	-	-	-	-
Coefficient and s.e.	-	-	-	-	-
Cluster robust F-statistic	-	-	-	-	-
Number of children		-			
Instrument		Non-kin exits / log(population)			
Subgroup: No Drug Abuse Child					
Coefficient and s.e.	0.0471 (0.0057)	0.0917 (0.0205)	1.947 (0.469)	0.853 (0.071)	0.0294 (0.0044)
Cluster robust F-statistic	68.3	-	-	-	44.7
Number of children		4,899			17,358
Instrument		Non-kin exits / log(population)			
Subgroup: Drug Abuse Chld					
Coefficient and s.e.	0.256 (0.109)	1.093 (0.799)	4.274 (1.834)	0.916 (1.011)	0.0554 (0.0177)
Cluster robust F-statistic	5.5	-	-	-	9.8
Number of children		214			1,103
Instrument		Non-kin exits / log(population)			

Notes: This table shows OLS results for the first stage (initial placement with family indicator on instrument), reduced form (outcome index on instrument), IV treatment effects and OLS treatment effects in the outcome sample, and the first stage in the eligible sample in different subgroups of the data based on child observables. All models do not include child demographic and entry reason controls. All models include county and month-by-year fixed effects, and standard errors are clustered at the county level. Blank entries in the table indicate that the sample size is too small to provide informative estimates.

Table A8: Heterogeneous Effects: Child Disability, Behavioral Problem

	First Stage	Reduced Form	IV	OLS	First Stage Eligible Sample
	(1)	(2)	(3)	(4)	(5)
Subgroup: No Child Disability					
Coefficient and s.e.	0.0458	0.103	2.247	0.925	0.0290
	(0.0057)	(0.022)	(0.543)	(0.068)	(0.0040)
Cluster robust F-statistic	63.7	-	-	-	51.3
Number of children		4,905			17,660
Instrument		Non-kin exits / log(population)			
Subgroup: Child Disability					
Coefficient and s.e.	-0.235	0.456	-	0.133	-0.0841
	(0.216)	(1.165)		(0.486)	(0.0365)
Cluster robust F-statistic	1.2	-	-	-	5.3
Number of children		208			801
Instrument		Non-kin exits / log(population)			
Subgroup: No Child Behavior Problem					
Coefficient and s.e.	0.0346	0.0484	1.401	0.795	0.0208
	(0.0057)	(0.0272)	(0.779)	(0.082)	(0.0048)
Cluster robust F-statistic	35.9	-	-	-	18.6
Number of children		3,039			9,886
Instrument		Non-kin exits / log(population)			
Subgroup: Child Behavior Problem					
Coefficient and s.e.	0.0193	0.0755	-	0.628	0.0074
	(0.0124)	(0.0485)		(0.133)	(0.0047)
Cluster robust F-statistic	2.4	-	-	-	2.5
Number of children		2,074			8,575
Instrument		Non-kin exits / log(population)			

Notes: This table shows OLS results for the first stage (initial placement with family indicator on instrument), reduced form (outcome index on instrument), IV treatment effects and OLS treatment effects in the outcome sample, and the first stage in the eligible sample in different subgroups of the data based on child observables. All models do not include child demographic and entry reason controls. All models include county and month-by-year fixed effects, and standard errors are clustered at the county level. Blank entries in the table indicate that the sample size is too small to provide informative estimates.

Table A9: Heterogeneous Effects: Relinquishment, Abandonment, Housing Problems

	First Stage (1)	Reduced Form (2)	IV (3)	OLS (4)	First Stage Eligible Sample (5)
Subgroup: No Abandonment					
Coefficient and s.e.	0.0492	0.0957	1.945	0.887	0.0287
	(0.0063)	(0.0215)	(0.483)	(0.073)	(0.0042)
Cluster robust F-statistic	60.4	-	-	-	47.0
Number of children		4,579			16,556
Instrument		Non-kin exits / log(population)			
Subgroup: Abandonment					
Coefficient and s.e.	0.0345	0.250	-	0.508	0.0656
	(0.0593)	(0.212)		(0.243)	(0.0263)
Cluster robust F-statistic	0.3	-	-	-	6.2
Number of children		534			1,905
Instrument		Non-kin exits / log(population)			
Subgroup: No Relinquishment					
Coefficient and s.e.	0.0452	0.0927	2.052	0.888	0.0286
	(0.0058)	(0.0214)	(0.509)	(0.069)	(0.0041)
Cluster robust F-statistic	61.3	-	-	-	47.9
Number of children		4,987			18,049
Instrument		Non-kin exits / log(population)			
Subgroup: Relinquishment					
Coefficient and s.e.					
Cluster robust F-statistic		-	-	-	
Number of children		126			
Instrument		Non-kin exits / log(population)			
Subgroup: No Housing Problems					
Coefficient and s.e.	0.0434	0.0880	2.025	0.889	0.0318
	(0.0060)	(0.0206)	(0.547)	(0.070)	(0.0040)
Cluster robust F-statistic	52.8	-	-	-	64.2
Number of children		4,861			17,482
Instrument		Non-kin exits / log(population)			
Subgroup: Housing Problems					
Coefficient and s.e.	0.215	0.594	2.768	2.234	-0.0477
	(0.0381)	(0.301)	(1.127)	(0.923)	(0.0162)
Cluster robust F-statistic	31.7	-	-	-	8.6
Number of children		252			979
Instrument		Non-kin exits / log(population)			

Notes: This table shows OLS results for the first stage (initial placement with family indicator on instrument), reduced form (outcome index on instrument), IV treatment effects and OLS treatment effects in the outcome sample, and the first stage in the eligible sample in different subgroups of the data based on child observables. All models do not include child demographic and entry reason controls. All models include county and month-by-year fixed effects, and standard errors are clustered at the county level. Blank entries in the table indicate that the sample size is too small to provide informative estimates.

Table A10: OLS Models on Exit Time

	In Foster Care Age 18		Age 19		Age 20	
	Max Age in FC	Max Exit Age	Max Age in FC	Max Exit Age	Max Age in FC	Max Exit Age
	(1)	(2)	(3)	(4)	(5)	(6)
Initial non-kin	0.0770	0.0606	0.0832	0.0608	0.0250	0.0237
family placement	(0.0165)	(0.0178)	(0.0117)	(0.0122)	(0.0133)	(0.0126)
County, month x year fes	Y	Y	Y	Y	Y	Y
Child demographic, entry reason controls	Y	Y	Y	Y	Y	Y
NYTD 2011 only	N	N	N	N	Y	Y
Mean outcome	0.732	0.718	0.224	0.191	0.0431	0.0266
Number observations (children)	5,226	4,727	5,226	4,727	1,857	1,727

Notes: This table reports results from OLS regressions of different outcomes of exit time for foster children on initial placement in a non-kin family. Exit age is measured in two ways. First I measure exit age by looking at the maximum age that I observe a foster child in our data. Second, I measure exit age by looking at the maximum age that I observe for a verified exit of the foster child in our data. Columns (1)-(2) use an outcome variable that is an indicator for a child still being in foster care at age 18. Columns (3)-(4) use an outcome variable that is an indicator for a child still being in foster care at age 19. Column (5)-(6) use an outcome variable that is an indicator for a child still being in foster care at age 20. For all columns I look at subsamples of children in the outcome data. For the regression looking at age 20 I only include children that were surveyed at age 17 in 2011, since I do not observe children surveyed at age 17 when they are 20 because I use placement data that ends in the fall of 2016. Throughout county and month by year fixed effects are included, child demographic and entry reason controls are included, and standard errors are clustered at the county level.

Table A11: Treatment Effects on Time in Foster Care

	Months in Foster Care	Months in Non-kin Placements	Percent Time in Non-kin Placements
	(1)	(2)	(3)
Initial non-kin	3.979	33.655	0.710
family placement	(4.028)	(4.630)	(0.137)
Instrument		non-kin exits / log(pop)	
County, month x year fes	Y	Y	Y
Child demographic, entry reason controls	Y	Y	Y
Mean outcome	24.527	13.98	0.482
Mean outcome control	22.014	4.128	0.129
Mean outcome treatment	27.052	23.88	0.838
Number observations (children)	4,646	5,227	5,227

Notes: This table provides IV estimates of regression models that measure months in foster care, months in a non-kin family placement while in foster care, and percent of time in foster care in a non-kin placement for the outcome sample. Months in foster care is a numeric variable that looks at the difference in days between the child's reported entry date and their exit date. If there is no exit date, this is NA. Months in non-kin placements is measured by counting the number of reporting period placements in which a child is in a non-kin placement and multiplying by 6 (reporting periods are 6 months apart). Percent time in non-kin placements looks at the percentage of placements reported for the child at and after entry that are non-kin placements. Standard errors are clustered at the county-level.

Table A12: Treatment Effects with Time in Foster Care as Endogenous Variables

	Economic and Social Outcome Index			
	IV		OLS	
	(1)	(2)	(3)	(4)
Months in non-kin family placement	0.0618 (0.0193)		0.0230 (0.002)	
Percent time in non-kin family placement		2.964 (1.107)		0.830 (0.0750)
County, month x year fes	Y	Y	Y	Y
Child demographics, entry reason controls	Y	Y	Y	Y
Sd endogenous variable	15.0	0.437	15.0	0.437
Mean endogenous variable control	4.20	0.131	4.20	0.131
Mean endogenous variable treatment	24.0	0.838	24.0	0.838
Number observations (children)	5,113	5,113	5,113	5,113

Notes: This table reports treatment effects estimated by IV and OLS on two alternative endogenous variables for the economic and social outcome index. The models are identical to those in Table 8 except for the endogenous variables. Months in non-kin family placement is a numeric variable that counts the number of placements recorded at and after entry that are non-kin family placements and multiplies by 6 months (the length between reporting periods). Percent time in non-kin placements looks at the percentage of placements reported for the child at and after entry that are non-kin placements. Standard errors are clustered at the county-level.

Table A13: NYTD Longitudinal Analysis I

	Outcomes at Age 17 or before				Outcomes at Age 19			
	14, 15 Year Old Entry		Children with Age 21 Outcomes		14,15,16,17 Year Old Entry		Children with Age 21 Outcomes	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Initial non-kin family placement	0.6049 (0.0694)	1.6766 (0.8975)	0.5862 (0.0771)	1.0350 (0.7138)	0.6693 (0.0644)	0.9689 (0.9623)	0.5939 (0.0749)	0.7421 (0.9784)
County, month x year fes	Y	Y	Y	Y	Y	Y	Y	Y
Child demographics, entry reason controls	Y	Y	Y	Y	Y	Y	Y	Y
Instrument	non-kin exits / log(pop)							
Number observations (children)	4,584	4,584	5,046	5,046	5,716	5,716	4,031	4,031

Notes: This table provides OLS and IV estimates of regressions that look at the economic and social outcome index constructed at age 17 and before (so that outcomes occur either at age 17 or any time before a child's life, corresponding to the same relative time horizon as the age 21 outcomes), and at age 19 (time horizons correspond to the same as the age 21 outcomes). These are estimated on 14 and 15 year olds entering (columns (1)-(2)), children that have age 21 outcomes (columns (3)-(4) and (7)-(8); these are a subset of the outcome sample used throughout the rest of the paper), and children entering at age 14, 15, 16 and 17. All models include child demographic and entry reason controls, and county and month-by-year fixed effects. Standard errors are clustered at the county level.

Table A14: NYTD Longitudinal Analysis II

Panel A: Correlations between outcomes at different ages	Outcomes Age 21	
	(1)	(2)
Outcomes age 17 or before	0.2944 (0.0151)	
Outcomes age 19		0.4674 (0.0152)
R-squared	0.0756	0.203
Number observations (children)	4,652	3,721

Panel B: Empirical Strategy Controlling for Outcomes at 17	Full Sample	Older Sample (16, 17 year old entry)
	(1)	(2)
Initial non-kin family placement	0.4840 (0.0708)	0.5054 (0.1029)
County, month x year fes	Y	Y
Child demographic, entry reason controls	Y	Y
Number observations (children)	4,652	2,756

Panel C: Validity Test with 17 Year Old Outcomes	Outcome Index Age 17		Outcome Index Age 21	
Non-kin exits month / log(pop)	0.0130 (0.0316)	-0.0326 (0.0287)	0.0914 (0.0427)	0.0712 (0.0503)
County, month x year fes	Y	Y	Y	Y
Child demographic, entry reason controls	N	Y	N	Y
Number observations (children)	2,756	2,756	2,756	2,756

Panel D: Including 19 Year Old Outcomes in Strategy	Outcome Index	Incarceration	Homeless	Substance Abuse	Enrollment or employment
	(1)	(2)	(3)	(4)	(5)
Initial non-kin family placement	2.296	-0.103	-0.348	-0.295	0.394
IV Coefficient	(0.757)	(0.125)	(0.180)	(0.155)	(0.169)
County, month x year fes	Y	Y	Y	Y	Y
Child demographic, entry reason controls	Y	Y	Y	Y	Y
Number observations (children)	6,789	6,724	6,727	6,721	6,789

Notes: This table presents four panels of results. Panel A estimates OLS regressions of outcomes at age 21 on outcomes at age 17 or before and on outcomes at age 19. All these variables are outcome indices constructed the same as the outcome index in age 21. These are univariate regressions. Panel B implements an OLS regression on outcomes at age 21 that include the outcome index measured at age 17 as a control, along with demographic and entry reason controls, and county and month-by-year fixed effects. Standard errors are clustered at the county level. Panel C implements OLS regressions of outcomes at age 17 and before, and outcomes at age 21 on the instrument (the second of these regressions is the reduced form). To make comparisons similar across these outcomes we include only children that have outcomes measured at both age 17 and age 21. We include specifications with and without demographic and entry reason controls. All specifications have county and month-by-year fixed effects. Standard errors are clustered at the county level. Panel D supplements IV results in Table 8 by constructing outcome variables that are indicators for outcomes at age 21, but if outcomes at age 21 do not exist and outcomes at age 19 exist, then measure outcomes at age 19. Thus they are a superset of the outcome sample used in Table 8. All specifications include demographic and entry reason controls, county and month-by-year fixed effects. Standard errors are clustered at the county level.

Table A15: Causal Forest Treatment Effect Summary Statistics

	Causal forest summary of treatment effects
Mean predicted CATE	0.931
SD predicted CATE	2.200
1st quartile predicted CATE	-0.556
Median predicted CATE	0.882
3rd quartile predicted CATE	2.292
Number children estimated on (test set)	511

Notes: This table shows summary statistics from the estimated conditional average treatment effects using child demographics and entry reasons following the method of Wager and Athey (2018).

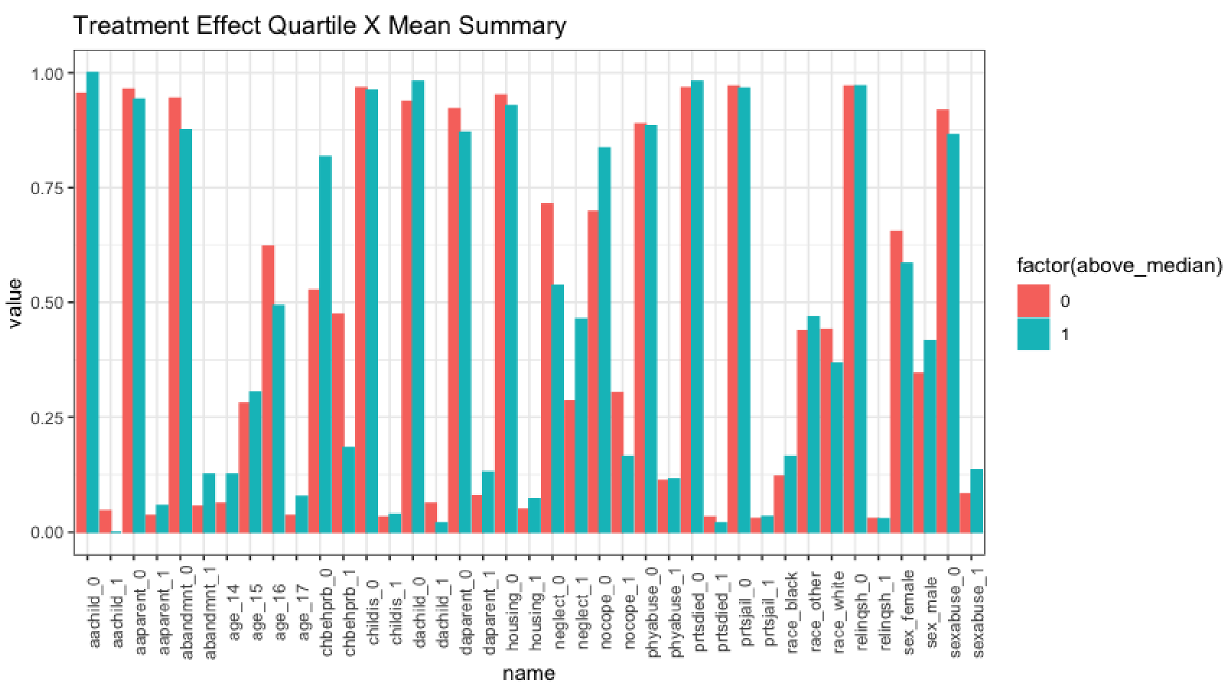

Figure A3: Causal Forest Heterogeneity by Child Demographics and Entry Reasons

Notes: This figure shows the percentage of children in the above and below median conditional average treatment effect group by different observable characteristics following the method of Wager and Athey (2018).

Table A16: OLS and Instrument Reduced Form Attrition

Panel A: Correcting for Non-Response Bias with Observables	Economic and Social Outcome Index	
	Non Weighted	Weighted
	(1)	(2)
Initial non-kin family placement	2.056	2.686
	(0.726)	(0.972)
Instrument	Non-kin exits / log(pop)	
Inverse propensity score weighted	N	Y
County, month x year fes	Y	Y
Child demographic, entry reason controls	Y	Y
Number observations (children)	5,112	5,112

Panel B: Lee (2009) Attrition Bounds	OLS	Reduced Form (ITT)	OLS Conditioned on initial sample	Reduced Form (ITT) Conditioned on initial sample	OLS Conditioned on non-sampling states	Reduced Form (ITT) Conditioned on non-sampling states
	(1)	(2)	(3)	(4)	(5)	(6)
Initial non-kin family placement	0.6459		0.6459		0.7090	
	(0.0667)		(0.0667)		(0.0747)	
Non-kin exits / log(pop)		0.1288		0.1288		0.1511
		(0.0710)		(0.0710)		(0.0825)
Lee (2009) upper bound	1.4472	0.1288	1.2410	0.1288	1.2734	0.1511
Lee (2009) lower bound	0.6459	-0.1465	0.6459	0.0146	0.7089	0.0382
Response rate treatment	0.321	0.265	0.621	0.556	0.630	0.560
Response rate control	0.243	0.288	0.521	0.575	0.516	0.575
p-value response rates differ	<0.001	0.0327	<0.001	0.308	<0.001	0.616
County, month x year fes	Y	Y	Y	Y	Y	Y
Child demographic, entry reason controls	Y	Y	Y	Y	Y	Y
Number observations (children)	5,112	5,112	5,112	5,112	3,877	3,877

Notes: This table contains two panels of results. Panel A undertakes the exercise in Sacerdote (2007) suggested by Wooldridge (1999) and corrects for non-response bias on observables by creating a propensity score for response to the survey at age 21 using a logistic regression model, and weighting observations according to $1/$fitted prob response. All demographics and entry reason variables are used to create the weights. Panel B computes Lee (2009) bounds for OLS treatment effects and intent-to-treat effects from the reduced form. The outcome variable is the outcome index used throughout the paper. Columns (1) and (2) use the outcome sample used throughout the paper and the response rate is computed with respect to the entire eligible sample. Columns (3) and (4) use the outcome sample used throughout the paper and the response rate is computed with respect to a smaller subset of children eligible to take the survey: children that responded to the survey at age 17 and that were sampled by states that randomly sample eligible children. Column (5) and (6) use the subset of the outcome sample in states that do not randomly sample eligible children, and compute response rates in those samples, too. Throughout standard errors are clustered at the county level.

Table A17: Robustness to Age Cutoff for Children Included in Sample

Panel A: IV	Children Last Entry 12 Years or Older	Children Last Entry 13 Years or Older	Children Last Entry 15 Years or Older
	(1)	(2)	(3)
Initial non-kin family placement	1.207	1.388	1.723
	(0.529)	(0.521)	(0.915)
Instrument		non-kin exits month / log(pop)	
First stage F-statistic	57.3	68.5	39.2
County, month x year fes	Y	Y	Y
Child demographic, entry reason controls	Y	Y	Y
Number observations (children)	5,699	5,545	4,498
Panel B: OLS			
	(1)	(2)	(3)
Initial non-kin family placement	0.627	0.629	0.658
	(0.065)	(0.066)	(0.071)
County, month x year fes	Y	Y	Y
Child demographic, entry reason controls	Y	Y	Y
Number observations (children)	5,699	5,545	4,498

Notes: This table includes OLS and IV estimates for regressions of the outcome index at age 21 used in Table 8 on an indicator for a child's initial placement in a non-kin family with various samples of children that vary by the age cutoff. Column (1) provides IV (Panel A) and OLS (Panel B) estimates for the sample of foster children that enter between ages 12 and 17. Column (2) provides IV and OLS estimates for the sample of foster children that enter between ages 13 and 17. Column (3) provides IV and OLS estimates for the sample of foster children that enter between ages 15 and 17. All models include demographic and child entry controls, and county and month-by-year fixed effects. Standard errors are clustered at the county level.

Table A18: IV Specification and Index Robustness

Panel A: Specification Tests	Old Child Exits	Drop Outlier County x Month x Years	Drop Very Small Counties	Dropping Endpoints of Data
	(1)	(2)	(3)	(4)
First stage coefficient on	0.0489	0.0449	0.0312	0.0316
instrument	(0.0108)	(0.0120)	(0.00487)	(0.00489)
IV coefficient on	3.545	2.613	1.755	1.987
economic and social outcome index	(0.927)	(0.757)	(0.745)	(0.744)
Instrument	Non-kin exits month 14 years+ / log(population)		Non-kin exits month / log(population)	
County, month x year fes	Y	Y	Y	Y
Child demographic, entry controls	Y	Y	Y	Y
Number observations (children)	5,113	4,277	3,923	5,037

Panel B: Outcome Indices	Incarceration, Homelessness, Substance Abuse Index	Employment, Enrollment Alternate Index	Incarceration, Homelessness, Substance Abuse, Employment, Enrollment Alternate Index with High School Education	Economic and Social Outcome Index with High School Education
	(1)	(2)	(3)	(4)
IV coefficient on	1.949	1.308	3.683	2.482
specified outcome	(0.646)	(0.677)	(1.100)	(0.878)
Instrument		Non-kin exits month / log(population)		
Mean outcome	0.323	0.217	0.661	1.13
Sd outcome	1.94	1.63	3.09	2.44
County, month x year fes	Y	Y	Y	Y
Child demographic, entry controls	Y	Y	Y	Y
Number observations (children)	5,113	5,113	5,113	5,113

Notes: Panel A provides first stage and IV regressions on different subsamples and with different instruments. Column (1) of panel A uses 14 year old non-kin exits as the instrument; column (2) drops county-month-year level observations where the instrument value falls outside the 5th and 95th percentile of the county-specific instrument distribution; column (3) drops all counties with 4 or less children in the sample; column (4) drops children with observed entries in the same month as the first reporting period. Panel B provides IV regressions on different outcome indices. Column (1) uses an index that adds incarceration, homelessness and substance abuse; column (2) uses an index that adds part-time employment, full-time employment and enrollment status; column (3) uses an index that adds the indices in columns (1) and (2) and also adds in high school education; column (4) uses the original index used in the main results and adds high school education. In all regressions standard errors are clustered at the county-level.

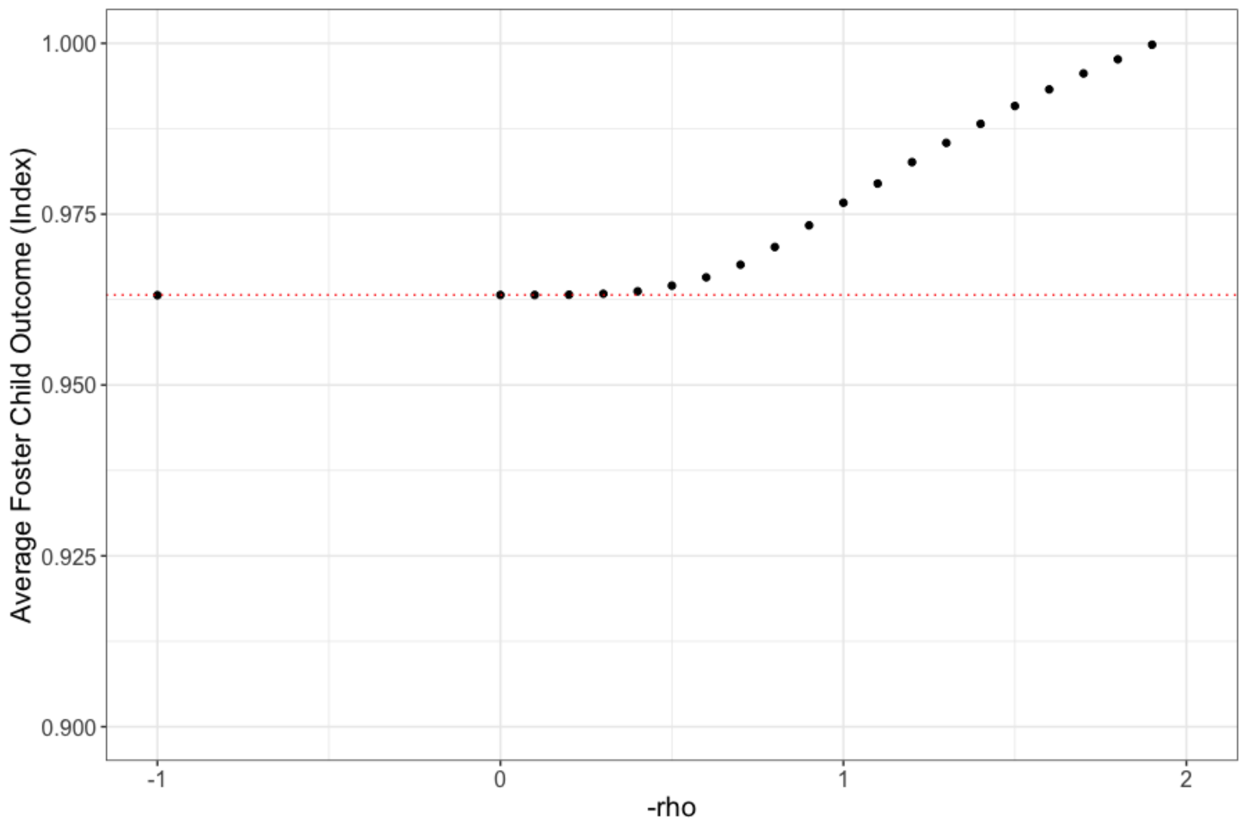

Figure A4: $E[Y_i]$ for Counterfactuals that Adjust All Observables

Notes: This figure shows average foster child outcomes on the outcome index simulated using the model in which family preferences are $u_i^{CF} = X_i(\rho \cdot \alpha) + v_i$ where $-\rho$ is plotted on the x-axis going from left to right, and the number of available families is held fixed. Outcomes are estimated using the estimated outcome model (18).