Not Without My Sister: The Impact of Same Sex Marriage on the Placement Outcomes of Foster Children[*†]

Chandler Inman
University of Michigan

October 11, 2025

Click here for the most recent version.

Abstract

How did same-sex marriage reform impact the capacity of social workers to find families for children in foster care? I find that extending marriage rights increased the probability children in their first year of care were matched to families by roughly 4 percentage points. In addition, by proposing a new method to impute siblings using administrative data, I find the reform made it easier to match groups of siblings to families, increasing the rate of match by 5-9 percentage points. These gains enabled social workers to keep more siblings together, decreasing the rate of sibling separation by roughly 7 percentage points. I develop a quantitative foster care matching model to determine the relative gain in welfare across child demographics, and parse the portion of gain attributable to decreased sibling separation. I find the highest welfare gains accrued to older children and that aggregate welfare rose by 5 percent.

Keywords: adoption, children, marriage, foster care, same-sex, matching
JEL Codes:

[*] The terms "foster care market" and "matching market" are used throughout this paper. The use of the word "market" here is a purely technical term, defined as the setting in which foster families seek to undertake the care of a foster child and social workers seek to find placements for the foster children under their care. While families have preferences over which children they would like to care for and children have different needs, meaning their placement outcomes will depend on the types of families they are placed with, the term "market" clearly does not adequately convey the deeply human dimensions of significance in the choices made here. The lived experiences of foster children, their families, and the social workers they interact with, must always take precedent over the economic abstractions that, while necessary to answer particular questions, will always miss this irreducible component of reality.

[†] The data used in this publication 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 Adoption and Foster Care Analysis and Reporting System (AFCARS) were originally reported 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 receiver 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.

> Why do these problems seem so intractable, so often only redefined, rather than remedied...The answers, I believe..lie in the child welfare system's place as a political battleground for abiding national conflicts over race, religion, gender, and inequality
>
> Nina Bernstein — *The Lost Children of Wilder: The Epic Struggle to Change Foster Care*

1 Introduction

Every year, roughly 250,000 children are placed in foster care in the US, and of these, 10-20 percent are placed in congregate care or group home settings. While for some children, these placements provide needed access to resources, children in congregate care routinely experience worse outcomes, running away at rates twice as high compared to traditional foster family settings and being three times more likely to engage in drug use (Breland-Noble et al., 2004; Karam and Robert, 2013; Lee et al., 2011). By not being placed with a family, these children are deprived of the stability a traditional familial foster home affords. It is for these reasons it is often considered best practice for social workers to use these placement settings only when necessary.

Sadly, for many foster children, there simply are not enough familial foster homes available (Doyle, 2007). Chronic foster care shortages plague many states and counties, and because some foster children are not as sought after, these shortages tend to be acutely felt by especially vulnerable groups of children: those that are older, non-white, or diagnosed with a disability.

A less often analyzed potential consequence of these shortages is how they may induce increased sibling separation. If social workers are unable to find a home for two children together, they must decide whether to place them apart or place them in a congregate care setting as a unit. Small-scale studies suggest that preserving a sibling group is associated with better mental health outcomes, better socialization, and better educational outcomes (Hegar and Rosenthal, 2011; Tarren-Sweeney and Hazell, 2005). This creates the potential for a heartbreaking tradeoff, where social workers can either place children in traditional familial homes or keep siblings together, but not both. This tradeoff has acute welfare implications that are clearly relevant for crafting policy.

In analyzing these shortages, there is one group of families that is being increasingly recognized as playing a key role due their greater propensity to foster: same-sex couples.

Same-sex couples are six times more likely to adopt, about 40-50 percent of adopted children are from foster care, and twice as likely to foster relative to different-sex couples (American Community Survey). In addition, as their ability to marry has increased, empirical work suggests so has their incentive to invest in relationship specific capital, taking on the form of stronger mutual commitments which encompasses the decision to care for a child. Evidence for this mechanism can be found in Martin and Rodriguez (2022) who find same sex marriage reform increased the rate of foster care adoption by 9-18 percent.

This paper seeks to determine whether same-sex marriage reform improved the capacity of social workers to match children to foster families and whether these initiatives impacted the rate of sibling separation in foster care. By using the heterogenous difference-in-differences design proposed by Callaway and Sant'Anna (2021) and comparing a subset of institutionally similar states, the estimates suggest same sex marriage reform increased the probability a foster child was matched to a family in their first year of care by 4 percentage points. These effects were heterogeneous across demographic characteristics, with older children, those with a disability, those who are male, and those who were non-white tending to benefit the most.

In addition, by using a range of demographic and administrative characteristics, this paper is able to establish plausible sibling sets in national foster care census data. By doing so, the impact of same-sex marriage reform on the rate of match across different child-groups as well as its impact on the rate of sibling separation can be estimated and valuable descriptive information can be learned. The estimates suggest non-homogeneous sibling groups[1] and sibling groups composed of older children benefited the most, experiencing declines on the order of 4 to 8 percentage points respectively. The rate of sibling separation was also found to decline, but this decline was primarily driven by non-homogenous sibling groups, with the probability of separation declining by 4 to 9 percentage points.

In order to quantify the welfare gains of these effects, make theoretically grounded comparisons of these gains across groups, and explore the relationship between sibling separation and the probability of going unmatched, this paper develops, estimates, and tests a quantitative foster care matching model. The model is inspired by the framework presented by Choo and Siow (2006) and extended by Chiappori et al. (2017) in the context of marriage matching. In the context of this paper, heterogeneous child-groups are matched to heterogeneous foster families. The key innovation is that social workers endogenously decide whether or not to separate children in order to maximize their group-specific expected welfare. The decentralized nature of the US foster care system enables the use of variation

[1] Groups classified as non-homogeneous are those that don't share a sex, a race, a gender, or are classified as being in a different age bin. Each of these characteristics are analyzed separately.

across many individual foster care systems, allowing identification of key parameters, such as the type-specific variance in idiosyncratic preferences, which is typically not identified in these models (Chiappori, 2017). The model performs well at matching out of sample moments, and the estimated parameter values make intuitive sense, providing evidence that the model captures important empirical mechanisms.

Finally, this paper calculates the effect of reform on the welfare of foster children by computing counterfactual welfare across child groups using the previously estimated average treatment effects on the probability of being unmatched induced by reform. The model suggests reform induced an aggregate gain in welfare of roughly 4.7%, with groups of children with at least one older child (those aged greater than 10) experiencing the bulk of welfare gains. The model attributes the majority of these gains to reductions in the probability of going unmatched rather than a fall in sibling separation. This is due to the latter being driven by reduced separation of those children already on the margin of being separated. Given that these children had a lower value of remaining together[2], this implies lower aggregate gains for this population. Nonetheless, the model suggests roughly 3 percent of sibling pairs composed of one child age less than 10 and one child aged greater than 10 who were placed together after reform would have been separated from one another in the absence of treatment, which is significant regardless. The model suggests the greatest beneficiaries were older children who entered foster care without a sibling. These children experienced a 15 percent increase in welfare due to reform, the largest increase, despite not experiencing the largest treatment effects. This is attributable to a higher variance in their child specific need for specific foster homes, which tends to imply higher average gains when a match is made, and a high tendency to go unmatched, which leaves more room for gain.

Taken as a whole, this paper documents new benefits of allowing same-sex couples to marry and adopt for some of the most vulnerable children. It creates new linkages using administrative data that provide a plausible means to investigate sibling groups and sibling separation, a crucial avenue through which child welfare is impacted. Finally, this paper provides a policy relevant structural means to think through and estimate the welfare impacts of social policy on foster care at a national level, possessing the benefit of not requiring data on the potential supply of foster homes, which is rarely available.

[2]This value of remaining together is from the perspective of the social worker.

2 Institutional Details

2.1 Foster Care in the United States

Five percent of children in the US will be placed in foster care at some point in their lifetimes (Bald et al., 2022), with this probability increasing to 10 percent for black children and 15 percent for Native American children (Gross and Baron, 2022). Children in foster care have been removed from their homes for a variety of reasons: from neglect to abuse to parental death. For most children, more than 50 percent, the goal is reunification with their original caretakers while for about 30 percent, the primary goal is permanent placement through adoption.[3] Children are the responsibility of their state or county's foster care system and social workers must provide children a place to be cared for. While the use of relatives or kinship care has been increasing over time, constituting roughly a third of placements today, most children are placed in either non-relative foster homes, group homes, or an institution.

Foster care tends to be a bellwether for the already instantiated inequalities in the economy. As mentioned previously, older children, non-white children, and children with a disability are typically harder to place than younger, white, non-disabled children (Bald et al., 2022). Foster care policy is primarily conducted at the state and county level. It is a highly decentralized set of markets, each possessing their own system of financial incentives, licensing requirements, and policies that govern who is able to foster and adopt. While it is possible to place children out of state, this is not often done, as states are still responsible for the children under their care when they are sent out of state,[4] making such placements more difficult to monitor and so, more costly.

Foster care matching in the US tends to be family driven, where social workers announce available sets of children to foster foster families and respond to the ensuing offers. Information on the available sibling sets is typically posted electronically, and regulations are often in place to ensure sufficient care is taken to find the best placement for the child.[5] The alternative to this design, which is also occasionally used, is caseworker driven search, where caseworkers contact families directly given their stated restrictions and preferences (Dierks et al., 2025). States differ on the age at which foster children may make decisions for themselves, with older age groups typically possessing more freedom to accept or decline placements. In any case, social workers must make decisions, such as whether to accept a placement or separate a sibling group, to maximize the welfare of the foster children under

[3] Other case goals include emancipation, transition to relative care, or long term foster care.

[4] The rules regulations governing out of state placement is outlined in the Interstate Compact on the Placement of Children (ICPC)

[5] For example, in Florida, regulations require that each prospective adoptive family be evaluated based on their capacity to meet the needs of the child.(Dierks et al., 2025)

their care. This is emphasized in the social work literature, where best practice is generally described as taking the child's particular circumstances and needs into account when making a placement decision (Pecora, 2000).

2.2 Same Sex Marriage and Adoption

The right of same-sex couples to marry, to foster, and to adopt are interrelated, but depend on the state, even locality, in question. For some states, passing same-sex marriage reform was effectively equivalent to granting same-sex couples the same legal adoption rights as married couples. In other states, the passage of same-sex adoption rights preceded marriage rights. Despite these nuances, it is correct to say that between 2000 and 2021, the rights of same-sex couples to do all of these underwent major shifts across the nation. In the year 2000, only 3 states allowed same-sex couples to jointly adopt a child[6] and none gave them the right to marry. In 2021, same-sex couples in all states possessed both the right to marry and the right to jointly adopt after marriage. Interestingly, the right to foster, at least in legislation, has undergone a recent reversal. Between 2015-2020, nine states adopted legislation explicitly allowing private foster care placement agencies to discriminate against LGBTQ couples, if placing a child with them would violate their religious beliefs.[7] While these pieces of legislation may have simply put into law what was already done in practice, their passage demonstrates the continuing policy relevance of these issues.

It is crucial when using state-level variation of this type to evaluate the institutional consequences of these reforms to ensure that the same type of treatment is being analyzed. In order to ensure that I am properly comparing similar types of reforms, I group states into three separate samples. The first group of states are those that had not yet legalized same-sex joint adoption before legalizing same-sex marriage, but did not have an explicit second-parent adoption ban.[8] Prior to the legalization of same-sex marriage, same-sex couples could theoretically adopt their partner's child in these states, but could not do so simultaneously through joint adoption. This restricted their access to several financial benefits[9] as well as complicated the process of adoption. The second set of states are those that had previously legalized joint adoption for same-sex couples before same-sex marriage

[6]These included Massachusetts, Vermont, and New Jersey

[7]These states include Michigan, Mississippi, Kansas, Montana, Oklahoma, South Carolina, South Dakota, Tennessee, and Texas.

[8]See Levendis and Lowen, 2023. The following localities are included: Arizona, Arkansas, California, DC, Georgia, Idaho, Iowa, Louisiana, Michigan, Minnesota, Montana, New Mexico, New York, North Dakota, Oklahoma, Pennsylvania, South Carolina, South Dakota, Tennessee, Texas, Virginia, Washington, West Virginia, Wyoming

[9]These include the federal adoption tax credit, and, until the second parent adoption could be approved, other legal rights parents have over their children, such as a say in medical treatment.

reform.[10] Although same-sex couples did not have the same rights as married couples before marriage reform, same-sex couples could still adopt children together through similar domestic channels accessible to straight married couples.[11]. The third group of states are those that explicitly banned second-parent adoption or had some other type of restrictive regime keeping same-sex couples from fostering or adopting.[12] These include states like Nevada, where same-sex couples were unable to become foster parents, and Florida, which had an explicit ban preventing LGBT *individuals* from adopting. It is important to emphasize that while each of these groups possessed similar state-level laws, at the local level, discriminatory practices were highly heterogeneous even within states and likely existed in some form in every state ("Legal Recognition of LGBTQ Families", n.d.).

For this analysis, the states analyzed are those that effectively legalized same-sex joint adoption and same-sex marriage simultaneously, but that did not possess a unique ban on same-sex fostering or adoption before reform. This group constitutes 25 states and comprises 57 percent of the foster care population over 2000-2020. This ensures that a similar reform for all states is being analyzed, but has the disadvantage of being unable to disentangle the marriage rights effect from the adoption rights effect.

3 Data

The data used in this paper is the Adoption and Foster Care Analysis Reporting system (AFCARS): the census of all children in foster care reported by states to the federal government. It contains child-level information on their reasons for removal from home, their placement setting, their case goal[13], their location, and other demographic characteristics. In addition, it contains information on whether this is a foster child's first time in foster care and whether or not the child was in foster care in the previous year or later years. Only children placed in foster care once during their first year of care are analyzed. This avoids problems associated with having to decide whether children only partially exposed to a new regime are treated vs in the control. AFCARS has important limitations. It does not contain very extensive information on foster parents, nor does it typically possess outcomes for children after they exit care.[14]. Strikingly, it doesn't directly record which children are

[10] The following states are included: Massachusetts, New Jersey, Connecticut, Indiana, Maine, Oregon, New Hampshire, Nevada, Delaware, Hawaii, Illinois, Rhode Island, Maryland, and Montana

[11] The following states are included: Mississippi, Florida, Nevada, Alabama, Kansas, North Carolina, Utah, Wisconsin, Kentucky, and Ohio

[12] The following states are included:

[13] The goal is typically reunification with their principle care takers. Other goals include adoption, long term care, and emancipation. See Bald et al., 2022

[14] While the National Data Archive on Child Abuse and Neglect does outcome information for children that age out of foster care, outcomes for other children requires detailed administrative state level data

in the same placement setting nor which children are siblings. This paper seeks to use the information that is available to both impute which children are likely siblings and whether or not these siblings are in the same placement setting.

The proposed method avoids making any assumptions concerning the race of siblings, which is essential, as it avoids the issues of racial misclassification that have plagued administrative data in the context of child welfare.[15]

3.1 Imputing Siblings

In order to impute which children are siblings, children are first linked by their location information, a multi-step process is used.

First, children are grouped by location, which includes their state, their county (if there are over 700 cases), and a 13-level rural urban continuum code. Next, children are linked by the available "principal caretaker" characteristics. Principle caretakers are typically biological parents, but include other types of legal guardians as well. These characteristics include their dates of birth and their family structure: married couple, unmarried couple, single male, and single female. Next, children are linked by two "removal reasons" that are highly likely to be shared by siblings. Removal reasons in AFCARS are meant to record the primary reasons for removal of the child from home. Multiple removal reasons can be recorded for a particular child. These removal reasons are "Removal Reason: Parental Death" and "Removal Reason: Parental Incarceration".

Next, the timing of removal and the timing that a child's case is recorded into the foster agency's system are used. AFCARS modifies ALL date variables by changing the birth date of children to the 15th of the month, while keeping the relative distance for all other date variables to this modified birthday date the same.[16] This means there are no identifying dates for children, only identifying intervals of time. Children are next linked by the length of time between the date of removal and *the date the removal is entered into the foster care agency's system*. What this administrative assumption amounts to is that siblings are separated together and then their information is recorded at the same time. A priori, this seems like a reasonable assumption. If both children must be removed, barring other circumstances, it would be more costly to remove them on different days. The assumption that the entry of this information would occur on similar days also seems similarly intuitively plausible.[17]

In order to validate whether this imputation exercise is actually identifying sibling

[15] Akee et al., 2026; Baron et al., 2024b; Baron et al., 2024a

[16] The is not the case for year level information, only month and day level.

[17] Performing all of these steps in the imputation results in assigning 85 percent of all children under 5 siblings and tags 45 percent of all children as not possessing a sibling in care

groups, the removal reasons of children relative to the removal reasons of their identified sibling is a useful indicator. More precisely, an intuitive check that real siblings are being identified is to examine the probability an imputed sibling group all share a removal reason, conditional on at least one of those children also possessing that removal reason. These results are recorded in Table 3. What these calculations show are intuitive correlations between those removal reasons one would expect children to share and what one would expect to be more child dependent. For example, when a child has the removal reason of “Child Drug Abuse”, their siblings only have a 35 percent chance of also having this removal reason. In contrast, while relinquishment is only a removal reason for 1 percent of the foster care population, if one member of a sibling group has this reason, the probability the others also have this reason is 70 percent.

In order to form an even sharper classification, children are further linked by removal reasons that seem most likely to be identifying siblings.[18] The final grouping matches 90 percent of the sample to 5 siblings or less, with 81 percent possessing 3 or fewer, and 50 percent being single. Table 4 produces descriptive statistics for these child groups. While the quantity of singles may appear rather high, given that the majority of children have a sibling, this may be due to the fact that some children were removed without the removal of their sibling, as may be the case if one child was determined to be in danger while the other was not. Nonetheless, it is an imperfect measure that requires trading off between misidentifying children as siblings vs inappropriately classifying children as unrelated.

In the empirical application, sibling groups that are size 4 or greater are dropped, leaving 81 percent of the sample. This is done for two reasons. First, some of these siblings groups are uncharacteristically large,[19] indicating that some children are being misclassified as siblings when they are not. Second, larger sibling groups are less likely to be properly characterized by the type-level distinctions analyzed below. For example, homogeneous vs heterogeneous child groups relative to particular child-level demographic characteristics (e.g. all old vs mixed age, all white vs mixed race, etc.) are examined. Analyzing groups along this dimension seems less informative as the size of sibling groups increases. Table 4 provides descriptives on a number of characteristics across imputed sibling group size, and reveals that dropping these children will cause some selection, as those imputed sibling groups of size greater than 3 tend to be less white and more rural than those imputed as being of size less than 4. This is likely because county level information is only available for those counties with more than 700 cases, which tends to occur most often in rural counties, and

[18] These removal reasons are: Parental Alcohol Abuse, Parental Drug Abuse, Inability to Cope, Inadequate Housing, Abandonment, and Relinquishment

[19] This is on the order of hundreds for particular groups

non-white families tend to have more children relative to white families.

3.2 Imputing Shared Placements, Matches, and Sibling Separation

For children identified as having at least one sibling in foster care, this paper imputes which subsets of each sibling group are matched to the same placement. This provides a measure of whether a sibling group was matched together or separately. Whether a child is matched to the same placement as each of their siblings is inferred based on each child's current placement setting and on the characteristics of the foster caretakers.[20]

These imputations allow for the measurement of the two primary outcomes of interest. The first is the probability a sibling group is separated. If a child is placed in a different setting than their siblings, that sibling group is called "separated". The next outcome of interest is the probability different types of child-groups go unmatched. When a sibling group is split-apart, one reason for this may be because it would enable some of the children to gain access to a foster family that would be unwilling or unable to care for all the children together. It is for this reason that the probability of going unmatched, measured as the probability of being placed in a group home or an institution, is calculated at the child-group level, rather than the sibling level. For example, the probability a "single" young-child is unmatched is the probability that a young child, placed in a setting with zero siblings, is placed in a group home or an institution. The probability that an all young sibling group, placed together, is unmatched is the probability that they, given they are an all young group, are placed in a group home or institution.

3.3 Descriptive Statistics on Outcomes

Table 5 provides descriptive details on the probability of being unmatched across groups of children and the probability of being separated across different sibling groups.[21] As mentioned previously, older children, disabled children, and male children all exhibit a greater likelihood of being placed in a congregate care setting. The disparities are striking. A single old child has a 64% chance of being placed in a congregate care setting, possessing an average age of 15. Compare this to a single young child, with an average age of 2,

[20] These characteristics include the foster family structure (married couple, unmarried couple, single female, or single male), and the year of birth, race, and hispanic origin of each foster caretaker. If all members of a sibling group are placed in a group home or they are all placed in an institution, it is assumed they have been placed in the same group home or the same institution

[21] It is important to mention that kinship care is abstracted from in this analysis, meaning only children in institutions, children in group homes, children in non-relative foster care, and children in pre-adoptive homes are considered. If it is assumed that if possible, social workers place children with kin, the population analyzed may better be understood as those children for whom kinship care is simply not possible or not an option. While this makes the descriptive information less totally generalizable, the disparities they reveal are no less striking and the population they pertain all the more vulnerable and important to understand.

that has a 4-8% chance. Interestingly, while single non-white and white children have a similar probability of being unmatched, this equivalence fades when examining all-white vs all-nonwhite vs multiple race sibling groups. Part of the reason for this may be that errors in racial classification are more likely for individual children, but at a group level, are less likely as more information about race can be gained from the children themselves.

The differences in separation rates are also quite informative. Groups of young children are very unlikely to be separated, at roughly 23%, while groups of old children are roughly 46% more likely to be separated and mixed age groups are more likely than not to be separated, at 57%. When considering the possibility of potentially problematic sibling dynamics, these numbers make sense. Alongside young children, the sibling group least likely to split apart are those composed of all girls. Groups of sisters have a separation rates of 22% to 25%, compared to 42% when the group is all male, and 44% when they are mixed sex.

4 Empirical Methodology

Despite the similar legal regimes in selected states, the diverse local practices and populations across cohorts implies there likely was substantial treatment-effect heterogeneity induced by reform. In order to estimate aggregated average treatment effects that are robust to this concern and properly account for the staggard roll-out over time, the methodology proposed by Callaway and Sant'Anna (2021) is used. The control group is composed of "not-yet-treated" states, meaning that the outcomes of child groups in treated states are compared to the outcomes of child groups in all states in the sample that had not yet experienced treatment, i.e. had not yet adopted same-sex marriage reform. Figure 5 plots this roll-out across states over time, illustrating which states are treated, and which states are control in each year-level treatment effect.

The identifying assumption to estimate these treatment effects is that in the absence of treatment, the treated and control states would have evolved in parallel. Treatment effects are aggregated by length of time since reform. Care must be taken when interpreting these estimates, as different states are included in each treatment effect by length of exposure estimate, again, highlighted in Figure 5. It's for this reason the treatment effects 0-3 years after reform are analyzed, as these treatment groups comprise the largest shares of the foster care population.

5 Reduced Form Results

Four major child-level demographic categories are analyzed: white vs non-white, young vs old, male vs female, and disabled vs not diagnosed as disabled.[22] These demographic characteristics are further split at the child group level as either homogeneous, meaning all children in that group possess the same demographic trait, or heterogeneous, meaning at least one child in the group does not share that trait. For example, a "mixed aged" sibling group, is a sibling group where at least one child is categorized as young and one child is categorized as old.

5.1 How Reform Impacted the Rate of Match for Individual Foster Children

The first two columns of Tables 6 through 8, show the treatment effects for individual children matched alone across demographic characteristics. Figures 6 and 7 plot these treatment effects along with the placebos.

Column (1) from each table shows upon immediate adoption of reform, children diagnosed with a disability, older children, non-white children and male children all experienced a 3 to 4 percentage point drop in the probability of being placed in a congregate care setting. For the former three groups, the estimates suggest these gains were persistent and growing, up to 7 to 9 percentage points 1 to 3 years after reform. In contrast, the analyses presented in column (2) mostly do not detect statistically significant gains. On average, there was a fall in the use of congregate care setting for non-disabled children, but the treatment effects are not distinguishable from zero several years after reform.

These results together suggest that those children with the highest propensities to be placed in a group home or institution ex-ante benefited from reform, and that these gains were relatively persistent or grew over time.

5.2 How Reform Impacted the Rate of Match for Groups of Foster Children

For groups of sibling sets placed together of size 2-3, the sample sizes are lower artificially because a sibling group is treated as a single observation, which avoids any issues of double counting. These results are presented in Tables 6 through 8, in columns (3)-(5) and Figures 8-10

Column (3)-(4) and Figures 8-9 report the average treatment effects for sibling groups that share some demographic characteristics. The estimates suggest groups composed of all

[22] A child is categorized as non-white if they are marked as having any other race or recorded as being Hispanic. A child is categorized as young if the are aged 10 or under, and old otherwise. A child is categorized as male if their sex is labeled as such. Finally, a child is categorized as disabled if they have been diagnosed with a physical or mental disability, or have been categorized as suffering from emotionally disturbed, implying a greater need for psychiatric help.

old children tended to experience statistically significant declines in the probability of being in a group home or institution. The binned three year treatment effect estimate for older groups of children suggests a 8.1 percentage point drop in the probability of being placed in a congregate care setting. Statistically significant declines could not be detected for other groups, but this may be a product of the lower power induced by the smaller sample sizes.

Column (5) in Tables 6 through 8 and Figure 10 reports results across sibling groups that are non-homogeneous across the examined demographic characteristic. Here, persistent declines are detected along mixed disability status, age, and race. These declines are on the order of 5 to 10 percentage point drops in the probability of being placed in a congregate care setting. These declines are significant and persistent.

5.3 How Reform Impacted the Rate of Sibling Separation

Tables 10 and 11 report the impact of reform on sibling separation across the same demographic characteristics. The corresponding treatment effects are plotted in Figures 11-13.

Generally, these results show evidence of non-parallel trends in the pre-period, which limits the interpretability of the point estimates. Regardless, the estimates generally point to a reduction in the rate of sibling separation among homogeneous sibling groups.

In contract, columns (3) and (5) corresponding to Figure 13, suggest large statistically significant declines in the rate of sibling separation for heterogeneous sibling groups along the dimension of age, race, and sex. These declines tend to be on the order of 3-5 percentage points immediately after reform, with evidence that these effects tended to be persistent or grow.

5.4 Discussion

These results suggest that same-sex marriage reform induced increases in the rate of match for many children, that these gains were especially pronounced for disadvantaged children, and that heterogeneous sibling groups tended to experience the greatest decline in sibling separation. In order to quantify the welfare gains associated with these effects, make theoretically grounded welfare comparisons across child types, and trace the relationship between sibling separation and the probability of finding a match, the next sections build a quantitative foster care matching model. The model is first presented in its most general form, but then empirical assumptions are made to ensure the mechanisms it suggests are testable and estimation is tractable.

6 Theoretical Framework

6.1 An Intuitive Exposition of the Model and Its Properties

The model is an adaptation and extension of several celebrated structural matching models used most often in the marriage literature (Choo and Siow, 2006; Chiappori et al., 2017). The model has several core components that allow for equilibrium analysis of a quite complicated matching environment, allowing for computations of expected welfare and counterfactual simulations that would ordinarily not be tractable.

The first component assumes social workers act according to the best interests of their assigned sibling group, making decisions to maximize their expected welfare. Social workers must make decisions entirely for, in the case of young children, or in collaboration with the foster children under their care. To the extent that social workers have other incentives or preferences not directly aligned with those of the children, the expected welfare measures presented here would have to be re-interpreted to accommodate this.[23]

Second, the model assumes *transferable utility*.[24] In the context of the model, when families and child groups match together, a total surplus or total level of welfare is generated. This welfare is then divided, without costs, between agents, allowing the stable (equilibrium) matching to be produced through a process of transfers. Again, this is an abstraction, but it can be rationalized partly by the incentives of foster families and social workers. For example, suppose a particular type of child, say, a young child, is especially sought after by foster families. In order to be placed with this type of child, families must convince social workers that they are an especially good match. They can do this in a number of ways, such as making improvements to their home to create an especially effective home study, ensuring their schedules allow for sufficient time to care for children, or generally signaling through their profile or application that they would provide the best environment for the type of child or group of children they are seeking to care for.[25] Alternatively, consider a child that may be particularly difficult to place, such as an older child. The social worker's outside option, if they are unable to find a family for the child, is to place them in an institution

[23]One interpretation could be the "utility" the social worker obtains given a particular outcome is obtained for the children under their care.

[24]See Chiappori, 2017 for an excellent primer on these types of models. See Chade et al., 2017 and Agarwal, 2015 for other varieties of matching models that do not rely on this assumption. See Galichon and Salanié, 2022 for a generalization of perfectly TU models.

[25]While regular check-ins by social workers help ensure that these were not mere signals, a complete analysis would take into account this informational asymmetry between foster care takers and social workers. It is also important to note that social workers are empowered to end particular matches. A complete analysis of welfare would also take match break up into account (Chiappori and Salanié, 2016). Each of these are currently outside the scope of this paper, but could be highly important mechanisms.

or a group home. To the extent the family is expected to provide more support than these other environments, the social worker will act in the best interests of the child and place them there.[26] In the language of the model, to the extent that a foster family only needs to provide a marginally better welfare outcome than the institution, the social worker will be willing to accept a smaller transfer of welfare to the child.[27]

Finally, the model assumes a specific decision process for social workers. Social workers are assigned a child or a sibling group and observe the childrens' idiosyncratic need to remain together. In addition, the social worker observes the expected welfare of particular *types* of child groups in their particular foster care system. The social worker then decides which children to place together, and only when they have made this decision, matches each child group according to that group's specific needs. What this decision process is meant to capture is how the separation decision on the part of social workers may be influenced by conditions in the foster care system. Taking another concrete example, suppose a sibling group enters foster care composed of one old child and one young child. The social worker has two options: match them together or match them each alone. If they are kept together, there may simply not be a home willing to care for both children together. However, this particular sibling group may have an especially high need to remain together, implying they may be better off being placed in a group home as a unit. The social worker must make this decision on behalf of the children.

7 Model Details

7.1 Families, Children, and Child Groups

Consider a set of foster care markets indexed by m. Within each market there are N_f^m families and N_g^m child groups.[28] These child groups are those that have been endogenously chosen to be kept together or split apart by social workers. Each family is indexed by j and is a discrete type, $J = 1, ..., K_J$. These discrete types are common across markets and include any observable characteristics that may influence the welfare generated by a match. Each child group is indexed by g. Child groups are composed of individual children,indexed by i_g. Children can be be one of a discrete type, $I = 1, ..., K_I$. These types can include age, race, whether or not the child has a disability, or any other discretizable characteristic

[26]With the caveats regarding the principle agent problem discussed previously

[27]An interesting direction for more research is the potentially endogenous response of institutions themselves, in response to changes in increased numbers of children entering foster care. To the extent that institutions provide better care in response to an increase, the welfare measures reported in this paper will understate the gains from reform.

[28]In this framework, a market will be a state$\times$year pair.

that may influence the welfare they receive when matched to a family in foster care. These are also common across markets. Group types are determined by the number of and type of children they are composed of. Specifically, the group type of an individual group g is determined by an ordered pair (I_g, z_g), where I_g are the types of children within group g, and z_g is a group specific function that encodes the number of children of each type, i.e. $z_g : I_g \to \mathbb{N}$, where $\mathbb{N}$ is the set of positive integers. In other words, the group type is only dependent on the types of children composing the group and the number of each type. Assuming a maximum number of children within a group is finite, the number of group types will also be finite

7.2 Matching Environment For Child Groups and Families

Within a foster care market m, families and and child groups are matched in a frictionless setting [29]. The total surplus generated by matching child group g of type G with family j of type J is given by

$$S_m(g, j) = Z_m(G, J) + \epsilon_{g,j} \tag{1}$$

where $Z_m(G, J)$ is the surplus generated purely through the child group type and family type. In other words, this is the part of the surplus that is generated from matching a type G child group to a type J family, that is common to all children and families within the market.[30] $\epsilon_{g,j}$ is the surplus generated that is match specific. Meaning, this is the surplus generated from matching this particular family j to this particular child group g.

In addition, families and child groups may go unmatched, the surplus generated from which will constitute the outside option. In the data, a child group is defined as going unmatched if the group is placed in a group home or an institution. A family going unmatched is conceptualized as the family possessing the means of receiving an additional placement, but not receiving one or choosing to leave that placement unfilled.[31] The surplus generated when child group g goes unmatched is given by

[29] Including search frictions is a dimension along which I would like to generalize this model. Importantly, however, it is assumed that children are indifferent between families of the same type, this assumption does not imply that all children observe all families and choose the best one.

[30] If data on the quantity of unmatched foster families was available, $Z_m(G, J)$ could be estimated. Looking across foster care markets, this would be an interesting measure for total relative efficiency of a market. While policy makers are likely more interested in the expected welfare of children, which can be identified, the total surplus generated, including the surplus flowing to families, likely matters a great deal as this will influence the decision of foster parents to enter the market.

[31] It is possible for foster families to have multiple "slots" for placements, meaning they can be matched to multiple child groups, so long as it is assumed there are no complementarities between child groups. In addition, because markets are large, once a child group is split up, the probability that a particular child group is matched with their sibling is probability zero.

$$S_m(g, \emptyset) = Z_m(G, \emptyset) + \epsilon_{g,\emptyset} \tag{2}$$

The $\epsilon_{g,\emptyset}$ can be thought of as the idiosyncratic group specific surplus generated by going unmatched, while $Z_m(G, \emptyset)$ is normalized to zero. This normalization implies that the actual value of being matched to a family for a child group will always be calculated relative to the outside option of remaining unmatched.

Similarly, the surplus generated from family j going unmatched is given by

$$S_m(\emptyset, j) = Z_m(\emptyset, J) + \epsilon_{\emptyset,j} \tag{3}$$

where $Z_m(\emptyset, J)$ is also normalized to zero.

Assumption 1: Separability *Assume the error term* $\epsilon_{g,j}$ *is of the form*

$$\epsilon_{g,j} = \tilde{\alpha}_g^J + \beta_j^G \tag{4}$$

where the random variables $\tilde{\alpha}_g^J,\ g \in G,\ J = 1, ..., K_J;$ *and* $\beta_j^G,\ j \in J,\ G = 1, ..., K_I^G$ *are independent*

This assumption is taken directly from Chiappori, Salanié, and Weiss (2017). What it amounts to is that the idiosyncratic contribution to the surplus generated by child group g, depends only on family j's type, J, and that the idiosyncratic contribution to the surplus generated by family j depends only on child group g's type, G. In other words, while families and child groups can have different preferences over types, families and child groups are indifferent between partners of the same type.

There are several good justifications for why separability may be a reasonable assumption in this context. The urgent nature of the foster cases involved often requires social welfare workers to make placement decisions based on limited information quite quickly. This implies that the key drivers of the matching process are likely quite closely linked to observables. While different children may have different needs, which implies different preferences over family types, there is likely limited capacity to make match specific decisions. In other words, and realistically, foster parents do not optimize over every available child but rather the types of children available[32] In addition, the set of family types is allowed to be incredibly flexible in the model's empirical application. Hence, child group indifference between family types is not an unreasonable assumption, given the only restriction is that the set of types if finite.

[32] An important exception to this is kinship placement. Social workers more and more frequently are seeking to find eligible kin with whom to place children, which changes the problem and the incentives involved enormously. Kinship placements are not considered in this analysis.

7.3 Assume Transferable Utility

In addition to the separability assumption, for this context, perfectly transferable utility is assumed, meaning the utility of a foster family and the utility of a child group may be costlessly exchanged. In one-to-one matching models with perfectly transferable utility, determining a stable match is equivalent to determining which set of matches maximize total surplus, which makes determining the equilibrium stable match much more tractable. While the theoretical properties of these models have been well articulated and analyzed, such as in Chiappori and Salanié (2016), it is worth thinking through the practical implications of this assumption in this context.

First, suppose a particular child group is especially in need of placement, due to this group having an especially low value for the outside option. The contribution to the surplus this child group generates from its potential matches may be used to attract families, as they may transfer this surplus to the family they are matched with directly. The same is true for families. As mentioned previously, there are many ways to conceptualize these transfers, but the important point is that the value ultimately gained by the child group will be determined by an equilibrium process, where emergent transfers will guarantee match stability and will be set such that no child group family pair would prefer to defect to a match outside their equilibrium match (given the transfers), and no child or family matched, would prefer to go unmatched.[33]

7.4 Matching Equilibrium and Equilibrium Welfare

Under Assumption 1 and under transferable utility, Choo and Siow (2006) and Chiappori et al. (2017) prove that under the stable matching the following holds:

Proposition CSW *Assume the surplus takes the form of (1) and the error terms satisfy Assumption 1, separability. Then, for each market m, there exist values U_m^{GJ} and V_m^{GJ}, $G = 1, ..., K_I^G$, $J = 1, ..., K_J$ such that*

1. *For all J and G*

$$U_m^{GJ} + V_m^{GJ} = Z_m(G, J) \tag{5}$$

2. *If g is matched with j at the stable matching, then payoffs are given by*

[33] This is clearly a simplification of reality, but appears to be no more unreasonable that the contexts in which it is traditionally used. The question is not whether this assumption is literally true, but rather whether the mechanisms it helps illuminate are empirically important and its implications, empirically useful.

$$u_g = U_m^{GJ} + \tilde{\alpha}_g^J, \quad v_j = V_m^{GJ} + \beta_j^G \tag{6}$$

3. *The expected payoffs for a child group type and family type will be given by*

$$\mathbb{E}[u_g|G, m] = \mathbb{E}[\max_J U_m^{GJ} + \tilde{\alpha}_g^J|G], \quad \mathbb{E}[v_j|J, m] = \mathbb{E}[\max_G V_m^{GJ} + \beta_j^G|J] \tag{7}$$

In other words, the matching problem reduces to a set of discrete choice problems. With the appropriate distributional assumptions, the matching decision of the child-groups may be analyzed independently from the matching decisions of the foster families. This is an incredibly useful property in this context, given the limited availability of foster family level information. It allows the analysis of market specific, expected, child-group level welfare, without requiring information on the supply of unmatched foster homes nor requiring the specification of the actual foster family types. The expected equilibrium payoffs,$\mathbb{E}[u_g|G, m]$, will be taken as inputs that will be used by social workers to determine how to group together siblings when they enter the foster care market.

7.5 Endogenizing the Child Groups

Sibling groups enter each foster care market exogenously, and are each assigned a social worker. Consider an individual sibling group s in market m. Each child in the sibling group possesses a type I and each pair of children possess an idiosyncratic gain if they are kept together, η_{ij}^s. The sibling group as a whole possesses a group type $G_s = \{I_s, z_s\}$ which would be the group type on the matching market if the social worker chooses to keep them all together. The social worker's goal is to maximize the expected welfare of the sibling group, given the conditions in their particular matching market.

The social worker observes their own market specific conditions, observing $\mathbb{E}[u_g|G, m]$ for all G, as well as a collection of group specific child terms, η_{ij}^s, that constitute the additional welfare generated when child i and child j are placed in the same group together. What remains unobserved are the idiosyncratic preferences of possible child groups, over the range of family types, the $\tilde{\alpha}_g^J$s.

The social worker chooses which children to group together and which to split apart. Formally, the social worker chooses a set of child groupings, C_s, that specifies which individual siblings in sibling group s will be together and the group types associated with each of those groupings. The set of groupings ultimately chosen by the social worker, C_s^*, must satisfy two conditions. The first is feasibility.

Feasibility *A set of child groupings C_k is a feasible set of groupings for sibling group s of group type $G_s = \{I_s, z_s\}$ if, for the groups G_k in C_k*

1. $I_{g_k} \subset I_s \; \forall \; k$

2. $\sum_k z_{g_k} = z_s$

These conditions simply state that the groupings must only be composed of child types that are also contained in G_s and that the number of children of each type, within all groups, must be exactly equal to the number of children of each type in G_s.

The second condition that must be satisfied is optimality. The total expected welfare of an arbitrary C_k is the sum of expected utilities of each child group, given the conditions in market m, and the sum of pair specific surplus terms η^s_{ij} which are determined by the individual pairs of children in each group.

Optimality *Let $C_1, ..., C_n$ be the list of all feasible child groupings for sibling group s. Without loss of generality, let C_1 be the set of child groupings composed of only one group, the group where all children are kept together. In addition, let C_n be the set of child groupings where no child is kept together. Let the group type of each individual child, when they are kept alone be given by I_i. Then, C^*_s must satisfy*

$$C^*_s = \text{argmax}_{C_1,...,C_n}\{\mathbb{E}[u_g|G^*, m] + \sum_j \sum_{i\neq j} \eta^s_{ij}, ..., \sum_i \mathbb{E}[u_i|I_i, m]\} \tag{8}$$

In other words, the social worker will group children together to maximize the total expected welfare of the sibling group, which includes the pair specific gains. Assuming that each η^s_{ij} is independently drawn from an atomless distribution, the solution to this discrete choice problem will be unique with probability 1.

8 Quantitative Model

8.1 Assumptions on Preference Heterogeneity

Assume that $\tilde{\alpha}^J_g = \sigma^G \alpha^J_g$, where α^J_g is distributed according to the the type 1 extreme value distribution. The parameter σ^G measures the degree of idiosyncratic preference variation across group types G, which is assumed to be constant across markets. A key reason for using a structural model in this context is to make consistent comparisons in relative welfare gains across child group types. As the level of heteroskedasticity shows up directly in the expected welfare calculation[34] and different child groups will likely have different levels of variation in their idiosyncratic needs for different household types, allowing heteroskedasticity is crucial in order to assess which children benefited most, at least along the dimension of expected utility.

[34]For child group G in market m, expected welfare for that group will be given by $\mathbb{E}[u_g|G, m] = \sigma^G \ln\left(\frac{1}{\text{Pr(Group is Unmatched}|G,m}\right)$

Finally, assume that $\eta_{ij}^G \sim N(\mu_G, 1)$ and are independently drawn. In other words, η's are drawn from distributions that possess type specific means but the same variance. This distributional assumption allows for the average value of keeping two children together to vary across group types.

It is worth noting what assumptions have not been necessary that typically must be assumed in the literature. First, almost no structure has been placed on β_j^G, save for the requirement of separability and independence from α_j^G.

8.2 Assumptions on Group Types

Next, assume that there are only five child group types $\{Y, O, (YY), (OY), (OO)\}$, which corresponds to a single young child, a single old child, a pair of young children, an old and a young child grouped together, and a pair of older children.[35] While other groupings are possible, there are two reasons for choosing this specific set. First, there is enough data for each of these sets of groupings across the analyzed markets, which will be constituted by a state×year pair. It is necessary for estimation to have well-measured moments, and so reasonably large sample sizes for each type in each market are necessary. Second, the child's age is one of the primary determinants of the social worker's capacity to find them a placement, and hence, the assumption of substitutability within these child group types is more reasonable. Other unmodeled characteristics will impact estimation to the extent that the demographic characteristics of each group type differs. For example, if older children also tend to be more likely to be diagnosed with a disability, familial capacity to deal with a disability will be embodied in the estimates relevant to age.

8.3 Choosing Markets

The key to identification in this approach is the existence of many, independent, foster care markets. It is rare for foster children to be placed across state lines, as this makes monitoring, which is the responsibility of the state under whose supervision the child has been placed, much more difficult.[36] In fact, foster care systems are often administered at the county level, with different counties even providing different benefits to foster families within the same state. The highly localized nature of the US foster care system enables the use of this many markets approach.[37] The next dimension of a market used in this analysis is time. A market is defined as a state×year pair. To make this demarcation plausible, only children in their first year of care who have only entered foster care for the first time are

[35] A young child is a child aged less than 10. An old child is a child aged 10 or greated

[36] AFCARS suggests that only about 2% of foster children are placed across state lines

[37] While this is useful for estimation, it is likely not optimal, given a larger market provides the opportunity for better matches.

used. This means social workers are looking at contemporaneous conditions, for children who have just entered their foster care system.

Only the states that were used in the reduced form analysis are used. In total, 404 markets are utilized across 25 states over approximately 15 years of data, with some state year pairs being dropped to ensure appropriate sample sizes.

For estimation, variation in the relative probabilities of being unmatched across types and variation in market level separation rates must exist to identify the type specific parameters. Figure 2 and Figure 3 plot the variation across markets on the probability of being unmatched across the five types. As mentioned, while single young and groups of young children tend to have similar separation rates across markets, what is truly needed is variation in relative separation rates, which is induced by the variation across the other types. Figure 4 shows that there is also sufficient variation across markets in the share of sibling sets that are split up.

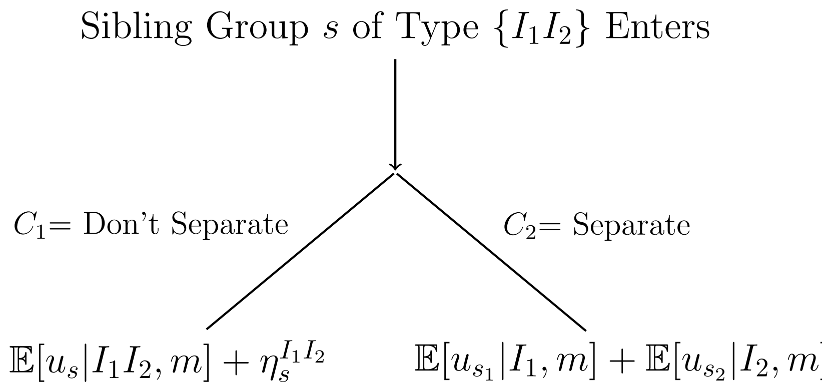
8.4 Market Level Equations to Estimate Parameters

Within a market, m, the social worker's estimate for the expected welfare received by a group type G must be calculated. These quantities are identified by the following empirical relationship (Chiappori, 2017):

$$\mathbb{E}[u_m^G|G,m] = \sigma_G \ln\left(\frac{1}{\Pr(\text{Unmatched}|G,m)}\right) = \sigma_G z_{Gm} \tag{9}$$

$\Pr(\text{Unmatched}|G,m)$ is computed directly at the market level, by calculating the share of child groups of group type G placed in a group home or an institutional setting. Importantly, groups can be composed of a single child. There are five of these shares calculated for each market, one for each group. The parameters σ^G will be estimated.

Next, three moments for each market are generated using the social worker's separation decision. The tree below outlines the decision process for an individual social worker, making decisions for a sibling group s composed for two children of type $I_1 I_2 \in \{YY, OO, OY\}$, where the expected generated welfare composes the bottom of the decision tree:



The social worker's discrete choice problem and the distributional assumptions imposed on η, where $\eta_s^{I_1 I_2} \sim N(\mu_{I_1, I_2}, 1)$, imply the following conditions ought to hold in market m:

$$\begin{aligned} \Pr(\text{Sibling Group s Split}|I_1 I_2, m) &= \\ \Pr(\eta_s^{I_1 I_2} \leq \mathbb{E}[u_{s_1}|I_1, m] + \mathbb{E}[u_{s_2}|I_2, m] - \mathbb{E}[u_s|I_1 I_2, m]) &= \\ \Phi(\sigma_{I_1} z_{mI_1} + \sigma_{I_2} z_{mI_2} - \sigma_{I_1 I_2} z_{mI_1 I_2} - \mu_{I_1 I_2}) \end{aligned} \tag{10}$$

where it Φ is the normal cdf function. Both the far left hand side of this relationship and the z_{mG}, which is defined in (9), are directly measured in the data. As there are three groups with more than 1 sibling, $\{YY, OO, OY\}$, this relationship will generate three moment conditions per market, with eight parameters that must be estimated: $\{\sigma_Y, \sigma_O, \sigma_{YY}, \sigma_{OY}, \sigma_{OO}, \mu_{YY}, \mu_{OY}, \mu_{OO}\}$. The variation across markets used here is variation in separation rates combined with the variation across markets in relative z's, which are pinned downed by market level probabilities in being unmatched.

8.5 Estimation and Moment Conditions

Let y_{sm} be a binary outcome for whether or not sibling group s in market m is separated. Let I_1 and I_2 be the individual type of each child in the group. Let X be a matrix indicating the market in which sibling group s resides and the group type of sibling group s, which can be $\{YY, OY, OO\}$. Given the market level equations, 1212 moment conditions can be defined to estimate the 8 parameters: 3 for each market across 404 markets. The moment conditions are given by

$$\mathbb{E}[X(y_s - \Phi(\sigma_{I_1} z_{mI_1} + \sigma_{I_2} z_{mI_2} - \sigma_{I_1 I_2} z_{mI_1 I_2} - \mu_{I_1 I_2}))] = 0 \tag{11}$$

The exogeneity assumption being invoked is that the prediction error of the model, which estimates the probability that child group s is separated given their market and type, is independent of their group type and market location. This assumption would be violated if the model systematically over or under estimated particular groups in particular markets. The model is highly over-identified, possessing 1208 degrees of freedom.[38]

8.6 Parameter Estimates

Table 1: Parameter Estimates

Parameter	Estimate	Std Error	p-value
σ_O	0.179	0.0124	0.000
σ_Y	0.0146	0.00388	0.000
σ_{OY}	0.0912	0.00566	0.000
σ_{OO}	0.0801	0.0100	0.000
σ_{YY}	0.0583	0.00469	0.000
μ_{YY}	0.787	0.0135	0.000
μ_{OY}	-0.228	0.0156	0.000
μ_{OO}	-0.0236	0.0137	0.0845

Note: The parameters are estimated across 404 markets using 1212 moment conditions utilizing one step GMM, the moment conditions being defined in (11). Standard errors are calculated using the the asymptotic variance formula that emerges when using the identity weighting matrix.

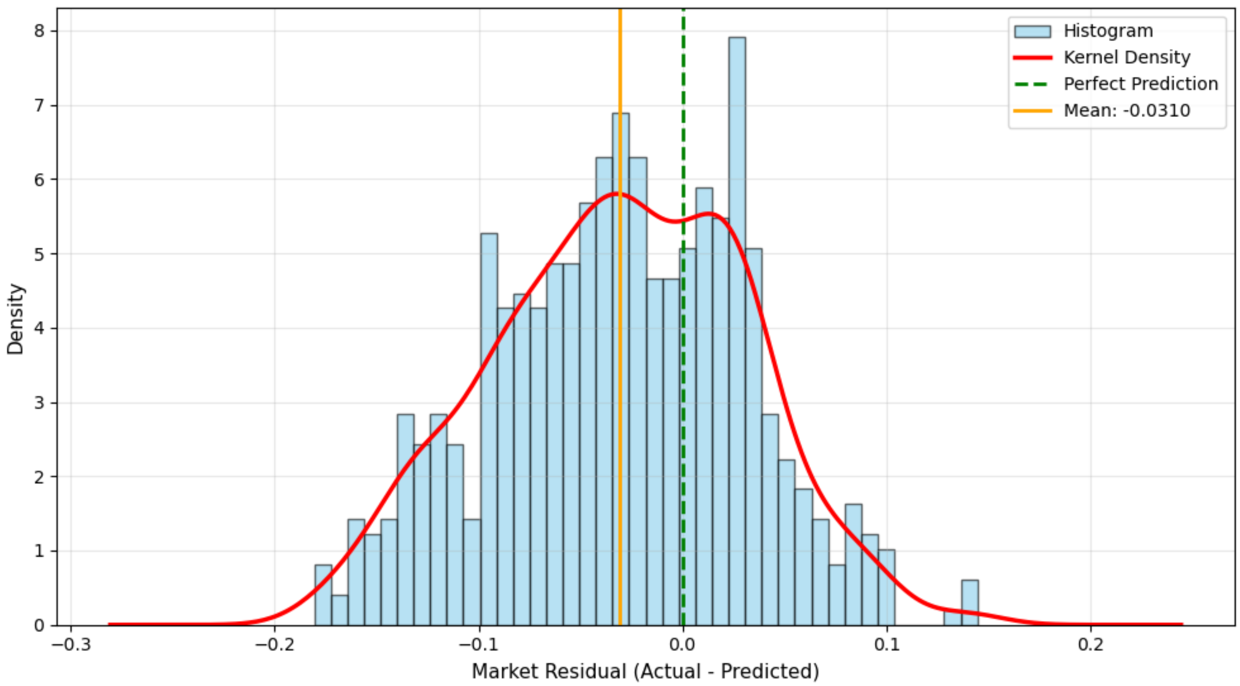
The parameter estimates are both revealing and make intuitive sense in this context. First, all σ terms are positive and statistically different than zero. Given that these measure the variance of group level idiosyncratic preference terms, this is comforting. In addition, the magnitude across these variances make sense. The group type with the smallest variance in idiosyncratic preferences is the young child grouped alone, while the highest is the single old child grouped alone. It makes sense that the needs across different household types for a single young child would tend to be more similar, compared to the needs of a single old child, who is also actually able to make their specific needs known to the social worker. The ordering of the μ's, which measures the average group specific preference for remaining together, also makes intuitive sense. Given that two young children are composed of two siblings both aged less ten, the probability of a problematic sibling dynamic is lower, and hence, the average η, at least as perceived by the social worker, tends to be higher among these children.

[38]Given the sparsity of matrix X, one-step GMM is performed, as the numerical issues associated with inversion made calculating the optimal weighting matrix impractical. The standard errors are adjusted accordingly

8.7 Validating the Model

In order to assess the performance of the model, an out of sample prediction exercise is performed to evaluate its capacity to match un-targeted markets: i.e. un-targeted moments. Figure 1 plots these results. The procedure is presented in Appendix B.

Figure 1: Out of Sample Prediction Errors on Untargeted Markets



Note: This demonstrates model performance at out of sample prediction, by plotting the residual prediction error of markets not used in used estimation. The outcome being predicted is the market level rate at which sibling pairs are separated within a foster care market. The moments here may also be called un-targeted because the aggregate rate of sibling separation is not explicitly used in estimation. The results show a mean of -.03, close to one, and a fairly tight distribution of standard errors of .06. There are no outliers and the shape of the distribution is what would be expected of unbiased estimation with random noise.

Figure 1 plots how well the model predicts the aggregate rate of sibling separation on markets not used in estimation. The average residual is -.03, which is quite close to zero, and indicates roughly unbiased predictions are being made. The standard deviation of .06 indicates predictions hover around the mean, meaning prediction errors generated by the model do not vary wildly. The shape of the distribution is also quite encouraging as it appears roughly normal, which is the distribution one would expected for an unbiased estimator in the presence of random noise. Taken together, the figure provides good evidence that the model can produce useful counterfactuals for estimation and welfare analysis.

9 Combining the Treatment Effects and Structural Model

While the supply of foster homes is not observable, the treatment effect of same sex marriage and adoption reform on the probability each sibling group type,$\{Y, O, YY, OY, YO\}$, is unmatched can be estimated. Combining these treatment effects with the structural model allows for the estimation of counterfactual welfare before and after reform across child groups. In addition, because sibling separation is an endogenous process and the distribution of η is estimated, expected welfare can be estimated pre-separation decision. This is the policy relevant measure of welfare, as it captures the welfare implication that siblings may be separated upon entering foster care. This procedure is inspired by the intuition of Choo and Siow (2006).[39]

9.1 Using Estimated Treatment Effects to Generate Counterfactuals

The group level treatment effects estimated using the difference in differences approach are applied in order to calculate counterfactual outcomes: $\hat{z}_{mG}$. Treatment effects are binned across three years of exposure and the same treatment effect is imposed across all markets. Naturally, only those markets exposed to treatment are used. This treatment effect homogeneity assumption is appropriate given the goal is to measure broad average changes in welfare across all relevant markets, for each child group.[40] It would be inappropriate to use this measure to look at distributional changes across markets, as this would require more knowledge of market level treatment effects.[41]

9.2 Computing Changes in Expected Welfare Due to Reform

With the estimated $\hat{z}_{mG}$ and z_{mG}, along with the parameter estimates, the expected change in welfare for children that enter foster care alone can be estimated. Let W_R be average expected welfare after reform and W_{NR} be expected welfare given no reform had taken place. Let n_{mG} be the number of groups of type G in market m, pre-separation decision and let M be the number of markets. Then the change in welfare caused by reform is given by:

[39] In their paper, they estimate the impact of abortion policies on the matching patterns of couples, to determine how the return to marriage changed for woman before and after reform

[40] Ideally, of course, a market specific treatment effect would be used.

[41] Given that dynamic effects are estimated, it is possible to estimate these effects across time, decomposing some of this heterogeneity.

$$
\begin{aligned}
\mathbb{E}[\Delta W|Y] &= \frac{\hat{\sigma}_Y}{M}\sum_{m=1}^{M} n_{mY}(z_{mY} - \hat{z}_{mY}) \\
\mathbb{E}[\Delta W|O] &= \frac{\hat{\sigma}_O}{M}\sum_{m=1}^{M} n_{mO}(z_{mO} - \hat{z}_{mO})
\end{aligned} \tag{12}
$$

Hence, the percent change in expected welfare induced by reform for young and old children that enter with no sibling is given by:

$$\frac{\mathbb{E}[W_R|Y] - \mathbb{E}[W_{NR}|Y]}{\mathbb{E}[W_{NR}|Y]} \tag{13}$$

$$\frac{\mathbb{E}[W_R|O] - \mathbb{E}[W_{NR}|O]}{\mathbb{E}[W_{NR}|O]} \tag{14}$$

Once these treatment effects are calculated, the rates of sibling separation implied by the parametric assumptions of the model, given $\hat{z}_{mG}$ and z_{mG}, are calculated.[42] This requires incorporating the fact that siblings actually kept together are those with $\eta' s$ that were large enough. A useful quantity to define is the market level expected welfare effect of sibling separation, $\rho_{mI_1I_2}$:

$$\rho_{mI_1I_2} = \hat{\sigma}_{I_1} z_{mI_1} + \hat{\sigma}_{I_2} z_{mI_2} - \hat{\sigma}_{I_1I_2} z_{mI_1I_2} - \hat{\mu}_{I_1I_2} \tag{15}$$

When counterfactual z's are used, define $\hat{\rho}_{mI_1I_2}$ to be:

$$\hat{\rho}_{mI_1I_2} = \hat{\sigma}_{I_1} \hat{z}_{mI_1} + \hat{\sigma}_{I_2} \hat{z}_{mI_2} - \hat{\sigma}_{I_1I_2} \hat{z}_{mI_1I_2} - \hat{\mu}_{I_1I_2} \tag{16}$$

Let W_R be average expected welfare after reform and W_{NR} be expected welfare given no reform had taken place. The welfare calculation for a sibling group I_1I_2, where $I_i \in \{Y, O\}$, is given by:

$$
\begin{aligned}
\mathbb{E}[W_R|I_1I_2] = {} & \frac{1}{M}\sum_{m=1}^{M} n_{I_1I_2m}\left[(1-\Phi(\rho_{mI_1I_2}))\times\left(\underbrace{\hat{\sigma}_{I_1I_2} z_{mI_1I_2} + \mathbb{E}\left[\eta^{I_1I_2}|\eta^{I_1I_2} > \rho_{mI_1I_2} + \hat{\mu}_{I_1I_2}\right]}_{\text{Expected Welfare of staying together given not separated}}\right)\right] \\
& + \frac{1}{M}\sum_{M} n_{I_1I_2m}\left[\Phi(\rho_{mI_1I_2})\times\underbrace{(\hat{\sigma}_{I_1} z_{mI_1} + \hat{\sigma}_{I_2} z_{mI_2})}_{\text{Expected welfare if separated}}\right]
\end{aligned}
$$

[42] Interestingly, this provides a further means to evaluate the mechanisms of the model, as the the treatment effect of same sex marriage reform on sibling separation can be calculated in the data. The computed treatment effects are not statistically different from the counterfactual change in separation rates implied by the model

The expected value of welfare given no reform, i.e. the expected value given the $\hat{z}$'s, can be calculated similarly.[43] Hence, the percent change in welfare under the counterfactual of no reform for a child group of type $I_1 I_2$ is given by:

$$\%\Delta W = \frac{\mathbb{E}[W_R | I_1 I_2] - \mathbb{E}[W_{NR} | I_1 I_2]}{\mathbb{E}[W_{NR} | I_1 I_2]} \tag{17}$$

Finally, in order to compute the fraction of gain due to reduced sibling separation, the expected gain for those who would have been separated in the absence of reform must be calculated. Of course, there is nothing in the model that guarantees reduced child separation. What determines this is the relationship between $\rho_{mI_1I_2}$ and $\hat{\rho}_{mI_1I_2}$. If, in a market $\hat{\rho}_{mI_1I_2} - \rho_{mI_1I_2} > 0$, the expected welfare effect of separating two children of type $I_1 I_2$ has declined, and so the rate of separation ought to decline. If $\hat{\rho}_{mI_1I_2} - \rho_{mI_1I_2} \leq 0$, the rate of separation is projected to not decline or even increase. The total expected increase in welfare attributable to decreased sibling separation, $\mathbb{E}[\Delta W_{SS} | I_1 I_2]$, is given by:

$$\mathbb{E}[\Delta W_{SS} | I_1 I_2] = \frac{1}{M} \sum_{m=1}^{M} n_{mI_1I_2} \left(\underbrace{\Phi(\hat{\rho}_{mI_1I_2}) - \Phi(\rho_{mI_1I_2})}_{\text{Change in Separation Rate}} \right) \times$$
$$\left[\underbrace{(\hat{\sigma}_{I_1I_2} z_{mI_1I_2} + \mathbb{E}\left[\eta^{I_1I_2} | \rho_{mI_1I_2} + \mu_{I_1I_2} < \eta^{I_1I_2} < \hat{\rho}_{mI_1I_2} + \mu_{I_1I_2}\right]}_{\text{Welfare Given Not Separated After Reform}} - \underbrace{(\sigma_{I_1} \hat{z}_{mI_1} + \sigma_{I_2} \hat{z}_{mI_2})}_{\text{Welfare of Split in Counterfactual}} \right]$$

These gains are composed of two terms. The first is the change in probability a sibling group $I_1 I_2$ are separated, given reform, in a particular market. The next term is the expected change in welfare induced by reform, given that these siblings are no longer separated, but were separated in the counterfactual.

43

$$\mathbb{E}[W_{NR} | I_1 I_2] = \frac{1}{M} \sum_{m=1}^{M} n_{I_1I_2m} \left[(1 - \Phi(\hat{\rho}_{mI_1I_2})) \times \left(\hat{\sigma}_{I_1I_2} \hat{z}_{mI_1I_2} + \mathbb{E}\left[\eta^{I_1I_2} | \eta^{I_1I_2} > \hat{\rho}_{mI_1I_2} + \hat{\mu}_{I_1I_2}\right] \right) \right]$$
$$+ \frac{1}{M} \sum_{m=1}^{M} n_{I_1I_2m} \left[\Phi(\hat{\rho}_{mI_1I_2}) \times (\hat{\sigma}_{I_1} \hat{z}_{mI_1} + \hat{\sigma}_{I_2} \hat{z}_{mI_2}) \right]$$

The share of welfare gains due to reduced child separation is given by:

$$W_{SS}^{Share} = \frac{\mathbb{E}[\Delta W_{SS}|I_1 I_2]}{\mathbb{E}[W_R|I_1 I_2] - \mathbb{E}[W_{NR}|I_1 I_2]} \tag{18}$$

9.3 Estimated Welfare Impact of Reform

Table 2 shows the estimated impact of same sex marriage reform, given the treatment effect on the probability of being unmatched induced by reform.

Table 2: The Impact of Treatment on Welfare

Type	ATET	$\%\Delta W$	W_{SS}^{Share}
Old-Old	-0.0745**	4.83%	0.03%
	(0.0407)		
Mixed Age	-0.0462***	6.96%	2.63%
	(0.0121)		
Young-Young	0.0024	-1.14%	0.00%
	(0.0065)		
Single Young	0.0006	-0.69%	
	(0.0057)		
Single Old	-0.0451**	14.99%	
	(0.0196)		
All		4.76%	

Note: The ATETs are the estimated treatment effects on the probability of being unmatched, binned across years 0-3 post treatment. The sample is restricted to the 25 treated states. The $\%\Delta W$ are the percent changes in expected welfare induced by treatment for each group estimated using equations (13), (14), and (17). Finally, $W_S^{Share}S$ is the fraction of gain due to reduced sibling separation, and is computed at the group level using equation (18).

A number of interesting results from Table 2 are worth highlighting. First, treatment resulted in an aggregate increase in welfare of 4.7%. The magnitude of this rise is significant, and comprises gains from increased match probabilities and changes in sibling separation. Next, the variation in gains is quite striking. While gains could not be detected for single young children or groups of two young children, groups composed of at least one child aged greater than 10 saw significant increases in welfare. Strikingly, despite a larger reduction in the probability of being unmatched for groups of children containing an older sibling, older children entering with no sibling had the largest percentage increase in expected welfare. This is driven partially by their higher σ term, which measures the level of child group specific heterogeneity in need across family types: the value a specific child group, g, has for matching with a particular family type, J. The fact that more of these children found a match and more of these children have highly specific needs, at least from the perspective of the social worker, implies that the same increase in match probability for that child group type results in a greater increase in expected welfare. The other factor that may be driving this larger relative increase is the lower initial level of welfare experienced by old children entering alone, possessing the lowest unconditional probability of finding a match.

Finally, the share of gain due to reduced child separation is, on the surface, quite small, being largest for sibling groups of mixed age, at 2.63%. The reason for this is actually quite intuitive. The children who gained were those already on the margin of being split up by the social worker, as those with high η's would always be kept together and those with low η's would always be split apart. Nonetheless, the estimates suggest approximately 3% of sibling pairs composed of one young and one old child who were matched together would have been separated without reform, which is significant regardless of the imputed change in welfare.

10 Conclusion

This paper has provided evidence that same-sex marriage reform in a set of institutionally similar states improved the capacity of social workers to match children to families, as well as keep more children together. The gains from this reform flowed disproportionately to the children and groups of siblings most difficult to match, and resulted in sizable welfare gains. There are a number of additional considerations that would be vital for a total analysis, but nonetheless had to be abstracted from. First, foster care placement is a dynamic process with match breakups, re-entry, and even exit through emancipation. Incorporating and considering each of these processes into a search and matching model would provide an even fuller picture of how social policy may impact the lives of foster children. Second, this analysis abstracted from the variety of different requirements states have in order to become a foster family. This understudied policy lever may be critical, as it effectively bars certain families from being foster parents, while also potentially protecting children by providing a minimum quality of care.[44] Finally, while this framework did not require information on the supply of foster homes or foster families, in order to appropriately ascertain the precise mechanisms for the documented gains, more work would be needed to document the dynamic response of foster home supply and its general determinants.

Nonetheless, the most straightforward explanation, that same-sex couples were induced

[44] Informational asymmetries between foster families and social workers were also not considered, which could be very important.

to enter the foster care market, seems to be the most plausible. First, this could explain the distribution of the gains. Queer children are roughly 2.5 times more likely to be placed in foster care (Fish et al., 2019), and to the extent that older children are more likely to know their own sexual orientation, and to the extent that same sex couples would be more prone to want to help queer children compared to straight couples, this could manifest as an increase in placements for older children relative to younger children. Second, as mentioned previously, there is evidence that same sex families in response to marriage reform, married more, adopted more, and that adoptions increased from foster care with its heterogeneous roll out (Martin and Rodriguez, 2022; *Twenty Years of Legal Marriage for Same-Sex Couples in the United States*, 2024). All of this suggests that marriage reform led to an increase in fostering among same-sex couples, affording the opportunity to place more children in need of homes with a family wanting to care for them.

References

Agarwal, N. (2015). An empirical model of the medical match. *American Economic Review*, *105*(7), 1939–1978. https://doi.org/10.1257/aer.20131006

Akee, R., Katz, L. F., & Loewenstein, M. A. (2026). *Race, ethnicity, and economic statistics for the 21st century* (tech. rep.). National Bureau of Economic Research.

Baccara, M., Collard-Wexler, A., Felli, L., & Yariv, L. (2014). Child-adoption matching: Preferences for gender and race. *American Economic Journal: Applied Economics*, *6*(3), 133–158. https://doi.org/10.1257/app.6.3.133

Bald, A., Doyle, J. J., Gross, M., & Jacob, B. A. (2022). Economics of foster care. *Journal of Economic Perspectives*, *36*(2), 223–246. https://doi.org/10.1257/jep.36.2.223

Baron, E. J., Doyle Jr, J. J., Emanuel, N., Hull, P., & Ryan, J. (2024a). Discrimination in multiphase systems: Evidence from child protection. *The Quarterly Journal of Economics*, *139*(3), 1611–1664.

Baron, E. J., Doyle Jr, J. J., Emanuel, N., Hull, P., & Ryan, J. P. (2024b). *Unwarranted disparity in high-stakes decisions: Race measurement and policy responses* (tech. rep.). National Bureau of Economic Research.

Baron, E. J., & Gross, M. (2022, April). *Is there a foster care-to-prison pipeline? evidence from quasi-randomly assigned investigators* (w29922). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w29922

Baron, E. J., Lombardo, R., Ryan, J., Suh, J., & Valenzuela-Stookey, Q. (2024, April). *Mechanism reform: An application to child welfare* (w32369). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w32369

Bernstein, N. (2011). *The lost children of wilder: The epic struggle to change foster care*. Knopf Doubleday Publishing Group.

Breland-Noble, A. M., Elbogen, E. B., Farmer, E. M. Z., Dubs, M. S., Wagner, H. R., & Burns, B. J. (2004). Use of Psychotropic Medications by Youths in Therapeutic Foster Care and Group Homes. *Psychiatric Services*, *55*(6), 706–708. https://doi.org/10.1176/appi.ps.55.6.706

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Chade, H., Eeckhout, J., & Smith, L. (2017). Sorting through search and matching models in economics. *Journal of Economic Literature*, *55*(2), 493–544. https://doi.org/10.1257/jel.20150777

Chiappori, P.-A. (2017). *Matching with transfers: The economics of love and marriage* [OCLC: ocn962355514]. Princeton University Press.

Chiappori, P.-A., & Salanié, B. (2016). The econometrics of matching models. *Journal of Economic Literature*, *54*(3), 832–861. Retrieved April 16, 2025, from https://www.jstor.org/stable/43932477

Chiappori, P.-A., Salanié, B., & Weiss, Y. (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review*, *107*(8), 2109–2167. https://doi.org/10.1257/aer.20150154

Choo, E., & Siow, A. (2006). Who marries whom and why. *Journal of Political Economy*, *114*(1), 175–201. https://doi.org/10.1086/498585

Dierks, L., Olberg, N., Seuken, S., Slaugh, V. W., & Ãœnver, M. U. (2025, July). *Search and matching for adoption from foster care* (Boston College Working Papers in Economics No. 1093). Boston College Department of Economics. https://doi.org/None

Donald, M., & E, D. (2023). *Foster care: A dynamic matching approach* (Working Paper No. 2023-01). Working Papers. Retrieved April 16, 2025, from https://www.econstor.eu/handle/10419/296957

Doyle, J. J. (2007). Child protection and child outcomes: Measuring the effects of foster care. *American Economic Review*, *97*(5), 1583–1610. https://doi.org/10.1257/aer.97.5.1583

Fish, J. N., Baams, L., Wojciak, A. S., & Russell, S. T. (2019). Are sexual minority youth overrepresented in foster care, child welfare, and out-of-home placement? findings from nationally representative data. *Child abuse & neglect*, *89*, 203–211.

Font, S. A., & Mills, C. (2020). Safe from harm? peer effects and criminal capital formation in foster care. Retrieved April 16, 2025, from https://appam.confex.com/appam/2020/meetingapp.cgi/Paper/37529

Galichon, A., & Salanié, B. (2022). Cupid's invisible hand: Social surplus and identification in matching models. *The Review of Economic Studies*, *89*(5), 2600–2629. https://doi.org/10.1093/restud/rdab090

Gates, G., Badgett, L., Macomber, J., & Chambers, K. (2007, March 27). *Adoption and foster care by lesbian and gay parents in the united states*. Urban Institute. https://www.urban.org/research/publication/adoption-and-foster-care-lesbian-and-gay-parents-united-states

Gross, M., & Baron, E. J. (2022). Temporary stays and persistent gains: The causal effects of foster care. *American Economic Journal: Applied Economics*, *14*(2), 170–199. https://doi.org/10.1257/app.20200204

Hansen, M. E. (2007). Using subsidies to promote the adoption of children from foster care. *Journal of Family and Economic Issues*, *28*(3), 377–393. https://doi.org/10.1007/s10834-007-9067-6

Hegar, R. L., & Rosenthal, J. A. (2011). Foster children placed with or separated from siblings: Outcomes based on a national sample. *Children and Youth Services Review*, *33*(7), 1245–1253. https://doi.org/10.1016/j.childyouth.2011.02.020

Karam, R., & Robert, M. (2013). Understanding runaway behaviour in group homes: What are runaways trying to tell us? *Journal of Community Positive Practices*, *13*(2).

Kojima, F., Pathak, P. A., & Roth, A. E. (2010, May). Matching with couples: Stability and incentives in large markets. https://doi.org/10.3386/w16028

Lee, B. R., Bright, C. L., Svoboda, D. V., Fakunmoju, S., & Barth, R. P. (2011). Outcomes of Group Care for Youth: A Review of Comparative Studies. *Research on Social Work Practice*, *21*(2), 177–189. https://doi.org/10.1177/1049731510386243

Legal recognition of LGBTQ families [National center for lesbian rights]. (n.d.). Retrieved April 18, 2025, from https://www.nclrights.org/get-help/resource/legal-recognition-of-lgbt-families/

Levendis, J., & Lowen, A. (2023). What same-sex adoption laws can tell us about the gender wage gap in the united states. *Journal of Family and Economic Issues*, *44*(2), 473–489. https://doi.org/10.1007/s10834-022-09835-2

Levitt, H. M., Schuyler, S. W., Chickerella, R., Elber, A., White, L., Troeger, R. L., Karter, J. M., Preston, J. M., & Collins, K. M. (2020). How discrimination in adoptive, foster, and medical systems harms LGBTQ+ families: Research on the experiences of prospective parents. *Journal of Gay & Lesbian Social Services*, *32*(3), 261–282. https://doi.org/10.1080/10538720.2020.1728461

Martin, J., & Rodriguez, Z. (2022, December 19). The effect of same-sex marriage legalization on adoptions and family formation. https://doi.org/10.2139/ssrn.4307175

Movement advancement project — foster care laws & regulations. (n.d.). Retrieved April 18, 2025, from https://www.lgbtmap.org/equality-maps/foster_care_laws

Pecora, P. J. (Ed.). (2000). *The child welfare challenge: Policy, practice, and research* (2nd ed). Aldine de Gruyter.

Tarren-Sweeney, M., & Hazell, P. (2005). The mental health and socialization of siblings in care. *Children and Youth Services Review*, *27*(7), 821–843. https://doi.org/10.1016/j.childyouth.2004.12.014

Taylor, C. (2024). Why do families foster children? a beckerian approach. *Review of Economics of the Household*, *22*(1), 261–293. https://doi.org/10.1007/s11150-023-09654-w

Twenty years of legal marriage for same-sex couples in the united states: Evidence review and new analyses. (2024). RAND Corporation. https://doi.org/10.7249/RRA2912-1

A Tables and Figures

Table 3

Removal Reason	**Conditional Probability of Shared Removal Reason**	**Unconditional Population Average**
Physical Abuse	73%	15%
Sexual Abuse	67%	5%
Neglect	90%	60%
Parental Alcohol Abuse	81%	7%
Parental Drug Abuse	87%	28%
Child Drug Abuse	35%	1%
Child Disability	27%	3%
Child Behavioral Problem	53%	11%
Inability to Cope	78%	17%
Abandonment	75%	5%
Relinquishment	70%	1%
Inadequate Housing	82%	11%

Note: The conditional probability of shared removal reason is the probability that if one sibling has the given removal reason, all their siblings share that same reason. Shares don't sum to one, since each child may have multiple removal reasons.

Table 4: Pairwise Difference in Means Tests

	Sibling Group Size			**Differences**		
	1	*2-3*	*4+*	*1 vs 2-3*	*1 vs 4+*	*2-3 vs 4+*
Share Age > 10	0.480	0.308	0.334	0.172***	0.146***	-0.027***
Mean Age (among > 10)	14.729	13.498	13.683	1.231***	1.046***	-0.185***
Mean Age (among ≤ 10)	2.411	3.931	3.831	-1.520***	-1.420***	0.100***
Share Disabled	0.383	0.306	0.365	0.078***	0.018***	-0.059***
Share Non-white	0.474	0.454	0.626	0.019***	-0.152***	-0.172***
Share Male	0.541	0.514	0.515	0.028***	0.026***	-0.002
Share in Metro Area	0.825	0.817	0.915	0.007***	-0.090***	-0.097***
Share in Rural Area	0.018	0.018	0.009	0.000	0.009***	0.010***
Observations	810,320	456,698	253,407	1,267,018	1,063,727	710,105

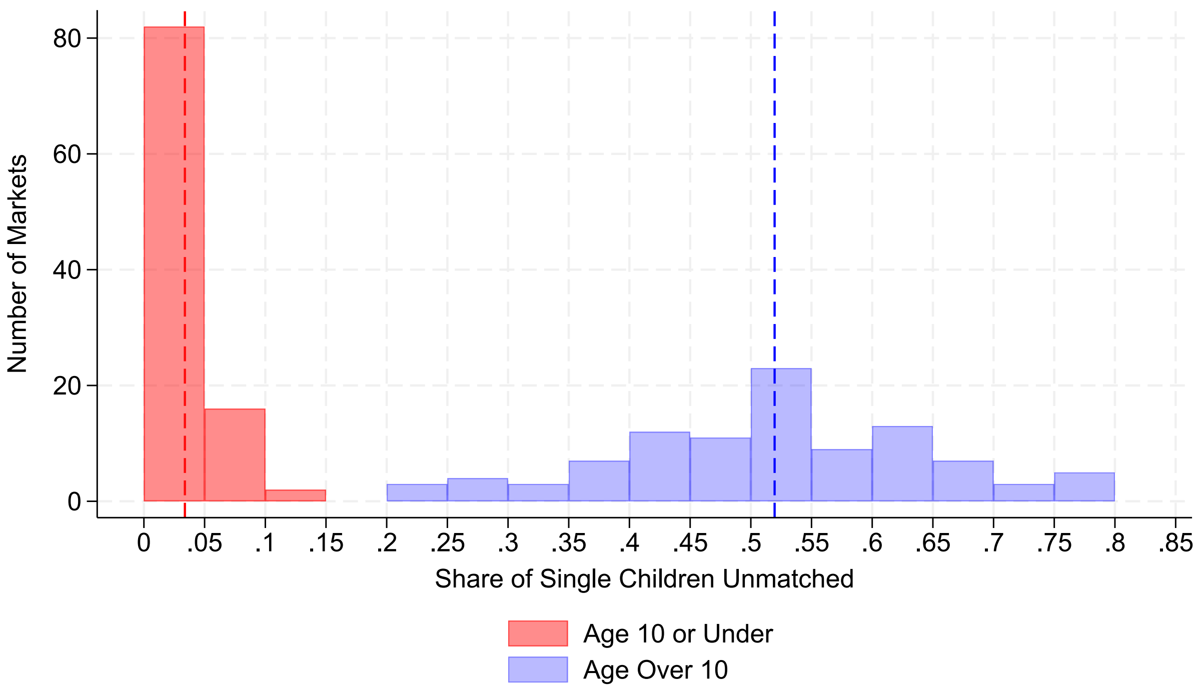
Note: The sample includes all treated states observed between 2008 and 2015. It includes all children in their first year of care, who are placed in a pre-adoptive home, a non-relative foster home, a group home, or another institution. Being in a metro area is defined as residing in a county with a 2013 USDA Rural-Urban Continuum Code of 1-3, and being in a rural area is defined as residing in a county with a RUCC of 8-9.
* p<0.10, ** p<0.05, *** p<0.01

Table 5: Probability a Child or Sibling Group is Unmatched or Split Up by Characteristics

	Probability Unmatched		**Probability Separated**	
Sibling Groups	*Pre-Treatment*	*Post-Treatment*	*Pre-Treatment*	*Post-Treatment*
All Age > 10	54.76%	39.23%	53.19%	47.24%
All Age ≤ 10	8.85%	4.04%	23.86%	22.18%
Multiple Age Groups	18.95%	11.43%	57.72%	56.76%
All Disabled	22.68%	15.73%	44.16%	37.72%
All Non-Disabled	17.5%	8.23%	28.59%	24.4%
Multiple Disability Statuses	24.24%	10.98%	62.45%	61.28%
All White	16.51%	8.88%	28.35%	24.15%
All Non-White	19.08%	9.85%	39.75%	34.61%
Multiple Races	41.6%	14.42%	71.83%	69.96%
All Male	25.11%	11.75%	42.72%	38.94%
All Female	15.51%	8.67%	27.56%	22.16%
Multiple Sexes	16.78%	8.77%	44.24%	40.23%
	Probability Unmatched			
Children Matched as Singles	*Pre-Treatment*	*Post-Treatment*		
Age > 10	61.24%	62.09%		
Age ≤ 10	8.19%	3.86%		
Disabled	39.15%	38.7%		
Non-Disabled	32.49%	22.1%		
White	34.9%	25.79%		
Non-White	33.77%	27.74%		
Male	38.86%	29.97%		
Female	29.29%	22.93%		

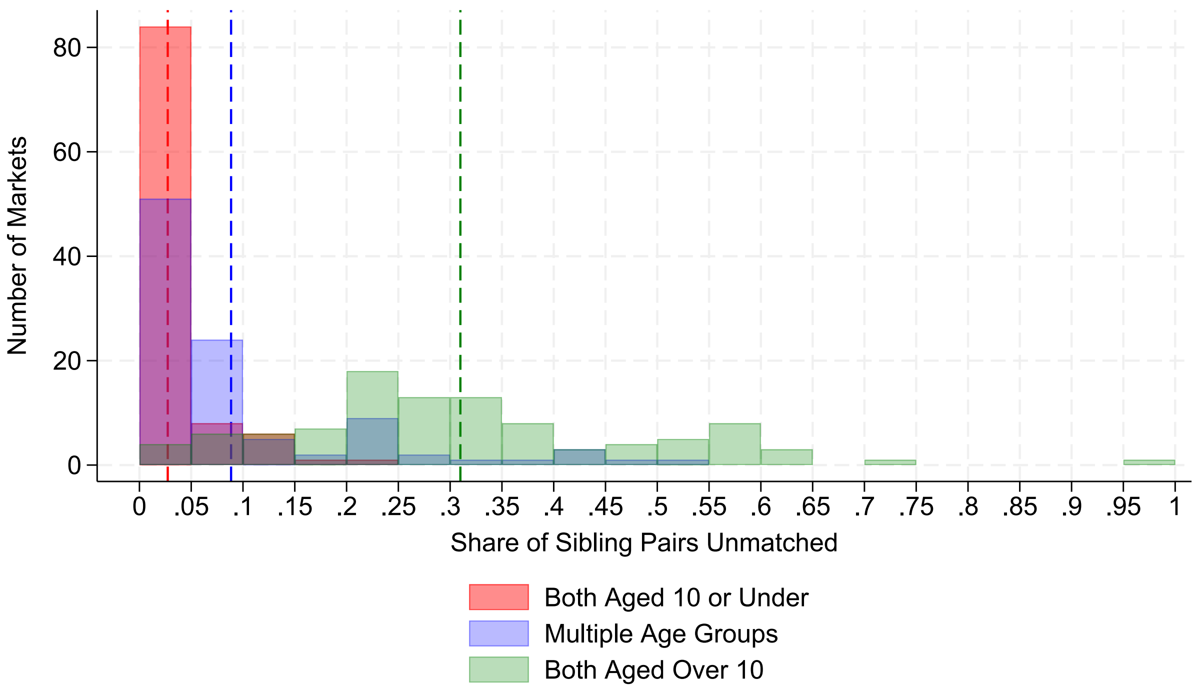
Note: The sample includes treated states observed between 2000 and 2021. It includes all children identified as having two or fewer siblings at the time of entry to foster care, who are placed in a pre-adoptive home, a non-relative foster home, a group home, or another institution. Children matched as singles are those who either entered without any siblings, or were not placed with any of their identified siblings. Sibling groups matched together include sibling groups comprised of 2-3 children who are placed in the same setting. The probability of being unmatched is the probability of being placed in a group home or institution, and the probability of being separated is the probability that at least one sibling was placed in a different setting.

Figure 2: Distribution of Share of Single Children Unmatched Across Markets



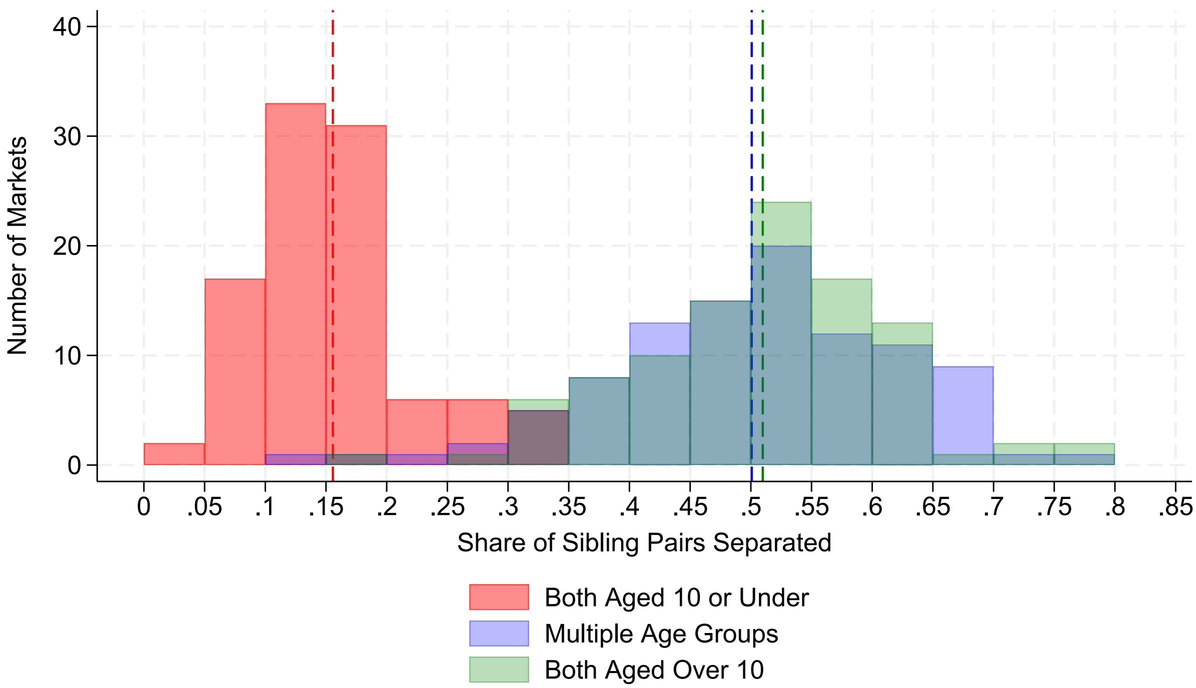
Note: This figure plots the variation in the market-level probability of being unmatched for children who are matched as singles, by child age group. Markets are defined as state-year pairs. There are 404 markets across the 25 treated states (listed in Figure 5). The unweighted means across markets are represented by the vertical dashed lines.

Figure 3: Distribution of Share of Sibling Pairs Unmatched Across Markets



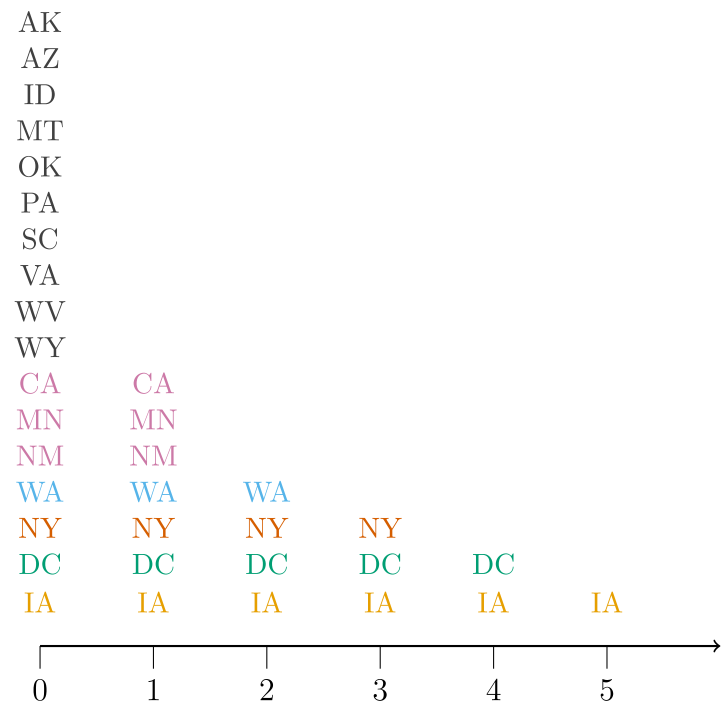
Note: This figure plots the variation in the market-level probability of being unmatched for sibling pairs matched together, by sibling pair age type. Markets are defined as state-year pairs. There are 404 markets across the 25 treated states (listed in Figure 5). The unweighted means across markets are represented by the vertical dashed lines.

Figure 4: Distribution of Share of Sibling Pairs Separated Across Markets



Note: This figure plots the variation in the market-level probability of being separated for children entering as a sibling pair, by age type. Markets are defined as state-year pairs. There are 404 markets across the 25 treated states (listed in Figure 5). The unweighted means across markets are represented by the vertical dashed lines.

Figure 5: States in Treatment Group by Number of Years Since Treatment



Note: This figure lists the states in the treated group by the number of years since treatment that they remain in the regression sample. Colors denote the following legalization years: 2014, 2013, 2012, 2011, 2010, 2009. Control states included: AR, GA, LA, MI, ND, SD, TN, and TX.

Table 6: Treatment Effect on the Probability of Being Unmatched by Disability Status

	Child Matched Alone		**2-3 Siblings Matched Together**		
Years Since Treatment	(1) *Disabled*	(2) *Non-Disabled*	(3) *All Disabled*	(4) *All Non-Disabled*	(5) *Mixed Disability*
-5	-0.0146	-0.0220	-0.0738*	-0.0677*	-0.0178
	(0.0219)	(0.0257)	(0.0425)	(0.0359)	(0.0302)
-4	-0.0210	0.0082	-0.0821	0.0134	-0.0204
	(0.0204)	(0.0122)	(0.0675)	(0.0159)	(0.0183)
-3	0.0333	-0.0286*	0.0288	-0.0644***	-0.0192
	(0.0289)	(0.0173)	(0.0580)	(0.0236)	(0.0130)
-2	-0.0230	-0.0233***	-0.0147	-0.0041	-0.0119
	(0.0242)	(0.0083)	(0.0315)	(0.0150)	(0.0231)
-1	-0.0015	0.0049	0.0434**	0.0121	-0.0163
	(0.0155)	(0.0112)	(0.0195)	(0.0080)	(0.0182)
0	-0.0394**	-0.0270***	-0.0340	-0.0157	-0.0510*
	(0.0193)	(0.0100)	(0.0310)	(0.0116)	(0.0296)
1	-0.0785***	-0.0196	-0.0493	-0.0066	-0.0870***
	(0.0209)	(0.0216)	(0.0319)	(0.0155)	(0.0321)
2	-0.0664***	0.0305	-0.0377	0.0247	-0.1216***
	(0.0222)	(0.0427)	(0.0644)	(0.0171)	(0.0262)
3	-0.0630	-0.0422	-0.0049	0.0168	-0.1728***
	(0.0400)	(0.0792)	(0.0693)	(0.0216)	(0.0320)
Aggregated ATE	-0.0577***	-0.0198	-0.0350	-0.0078	-0.0777***
	(0.0177)	(0.0168)	(0.0348)	(0.0126)	(0.0262)
Observations	327,964	826,949	28,172	139,600	27,715
Mean Pre-Reform	0.45	0.278	0.196	0.106	0.165

Note: The regression samples for columns 3 - 5 include sibling groups comprising 2 or 3 children who were matched to the same placement setting. The outcome is measured at the sibling group level, and is equal to one if all siblings who were placed together were unmatched (i.e. placed in a group home or institution). Pre-reform means are measured one year before reform. The aggregated ATE is calculated by aggregating across years 0 to 3 post-reform.

Table 7: Treatment Effect on the Probability of Being Unmatched by Age

Years Since Treatment	**Child Matched Alone** (1) *Age > 10*	(2) *Age ≤ 10*	**2-3 Siblings Matched Together** (3) *All Age > 10*	(4) *All Age ≤ 10*	(5) *Multiple Ages*
-5	-0.0589	0.0022	-0.1426***	-0.0061	-0.0409
	(0.0399)	(0.0089)	(0.0476)	(0.0126)	(0.0276)
-4	0.0060	-0.0057	0.0018	-0.0011	-0.0409
	(0.0241)	(0.0058)	(0.0611)	(0.0091)	(0.0379)
-3	-0.0138	-0.0034	-0.0744	-0.0067	-0.0015
	(0.0184)	(0.0075)	(0.0518)	(0.0075)	(0.0359)
-2	-0.0183	0.0035	-0.0554	-0.0002	-0.0273
	(0.0160)	(0.0077)	(0.0434)	(0.0078)	(0.0299)
-1	0.0162*	0.0012	0.0460	0.0044	0.0417**
	(0.0095)	(0.0052)	(0.0383)	(0.0067)	(0.0177)
0	-0.0353***	-0.0018	-0.0503	-0.0090*	-0.0608***
	(0.0125)	(0.0062)	(0.0515)	(0.0053)	(0.0104)
1	-0.0446**	-0.0047	-0.0993***	-0.0014	-0.0520***
	(0.0188)	(0.0081)	(0.0371)	(0.0102)	(0.0144)
2	-0.0451	0.0197***	-0.0609	0.0101	-0.0509**
	(0.0433)	(0.0073)	(0.0637)	(0.0090)	(0.0250)
3	-0.0934**	0.0115	-0.2769***	0.0264	-0.0573*
	(0.0434)	(0.0091)	(0.0503)	(0.0192)	(0.0298)
Aggregated ATE	-0.0451**	0.0006	-0.0816**	-0.0021	-0.0570***
	(0.0196)	(0.0057)	(0.0386)	(0.0066)	(0.0110)
Observations	535,550	619,363	33,173	130,274	32,040
Mean Pre-Reform	0.670	0.052	0.461	0.059	0.163

Note: The regression samples for columns 3 - 5 include sibling groups comprising 2 or 3 children who were matched to the same placement setting. The outcome is measured at the sibling group level, and is equal to one if all siblings who were placed together were unmatched (i.e. placed in a group home or institution). Pre-reform means are measured one year before reform. The aggregated ATE is calculated by aggregating across years 0 to 3 post-reform.

Table 8: Treatment Effect on the Probability of Being Unmatched by Race

	Child Matched Alone		**2-3 Siblings Matched Together**		
Years Since Treatment	(1) *Non-White*	(2) *White*	(3) *All Non-White*	(4) *All White*	(5) *Multiple Races*
-5	-0.0163	-0.0171	-0.0606	-0.0730**	-0.0551
	(0.0247)	(0.0220)	(0.0416)	(0.0301)	(0.0471)
-4	-0.0122	-0.0007	-0.0224	-0.0044	-0.0242
	(0.0181)	(0.0125)	(0.0363)	(0.0177)	(0.0482)
-3	0.0039	-0.0314**	-0.0502	-0.0343	-0.0158
	(0.0153)	(0.0146)	(0.0374)	(0.0271)	(0.0305)
-2	-0.0228**	-0.0194*	0.0071	-0.0045	-0.0538
	(0.0105)	(0.0105)	(0.0247)	(0.0139)	(0.0553)
-1	0.0144	0.0134	0.0006	0.0226**	0.0577**
	(0.0089)	(0.0096)	(0.0100)	(0.0096)	(0.0291)
0	-0.0408***	-0.0112	-0.0101	-0.0195	-0.1088**
	(0.0106)	(0.0101)	(0.0140)	(0.0125)	(0.0539)
1	-0.0492***	-0.0049	-0.0190	-0.0087	-0.1413***
	(0.0121)	(0.0166)	(0.0153)	(0.0131)	(0.0409)
2	-0.0410	0.0449	0.0448*	0.0055	-0.3283***
	(0.0336)	(0.0310)	(0.0269)	(0.0180)	(0.1044)
3	-0.0822**	-0.0122	-0.0197	0.0073	-0.3434***
	(0.0351)	(0.0555)	(0.0177)	(0.0196)	(0.0189)
Aggregated ATE	-0.0475***	0.0032	-0.0087	-0.0124	-0.1553***
	(0.0143)	(0.0135)	(0.0158)	(0.0118)	(0.0469)
Observations	509,346	645,567	76,562	105,776	13,149
Mean Pre-Reform	0.342	0.331	0.124	0.116	0.298

Note: The regression samples for columns 3 - 5 include sibling groups comprising 2 or 3 children who were matched to the same placement setting. The outcome is measured at the sibling group level, and is equal to one if all siblings who were placed together were unmatched (i.e. placed in a group home or institution). Pre-reform means are measured one year before reform. The aggregated ATE is calculated by aggregating across years 0 to 3 post-reform.

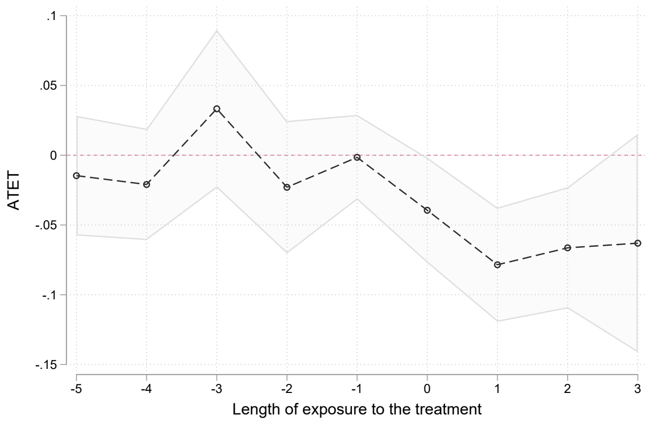
Table 9: Treatment Effect on the Probability of Being Unmatched by Sex

	Child Matched Alone		2-3 Siblings Matched Together		
Years Since Treatment	(1) *Male*	(2) *Female*	(3) *All Male*	(4) *All Female*	(5) *Multiple Sexes*
-5	-0.0149	-0.0186	-0.0932**	-0.0632*	-0.0443**
	(0.0292)	(0.0169)	(0.0375)	(0.0325)	(0.0205)
-4	0.0056	-0.0147	-0.0058	0.0028	-0.0307
	(0.0176)	(0.0124)	(0.0396)	(0.0154)	(0.0249)
-3	-0.0223	-0.0109	-0.0734	-0.0591***	-0.0096
	(0.0152)	(0.0164)	(0.0496)	(0.0214)	(0.0138)
-2	-0.0165*	-0.0257***	-0.0178	0.0103	-0.0052
	(0.0093)	(0.0081)	(0.0282)	(0.0198)	(0.0080)
-1	0.0174**	0.0101	0.0281*	0.0406**	-0.0049
	(0.0087)	(0.0075)	(0.0150)	(0.0202)	(0.0085)
0	-0.0321***	-0.0153	-0.0370*	-0.0262	-0.0129
	(0.0081)	(0.0113)	(0.0217)	(0.0205)	(0.0137)
1	-0.0432***	-0.0109	-0.0194	-0.0233	-0.0246**
	(0.0139)	(0.0124)	(0.0313)	(0.0153)	(0.0105)
2	-0.0076	0.0117	0.0210	-0.0037	-0.0143
	(0.0459)	(0.0270)	(0.0428)	(0.0176)	(0.0151)
3	-0.0721	-0.0283	-0.0220	-0.0373**	-0.0248*
	(0.0630)	(0.0239)	(0.0206)	(0.0160)	(0.0137)
Aggregated ATE	-0.0363**	-0.0120	-0.0249	-0.0240*	-0.0174*
	(0.0163)	(0.0125)	(0.0244)	(0.0143)	(0.0102)
Observations	615,850	539,063	54,471	48,672	92,344
Mean Pre-Reform	0.389	0.274	0.172	0.122	0.113

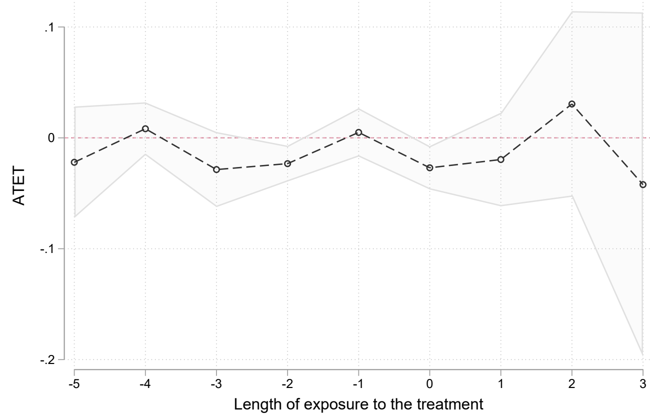
Note: The regression samples for columns 3 - 5 include sibling groups comprising 2 or 3 children who were matched to the same placement setting. The outcome is measured at the sibling group level, and is equal to one if all siblings who were placed together were unmatched (i.e. placed in a group home or institution). Pre-reform means are measured one year before reform. The aggregated ATE is calculated by aggregating across years 0 to 3 post-reform.

Figure 6: Aggregated ATET on Probability Unmatched - Single Children by Disability and Age

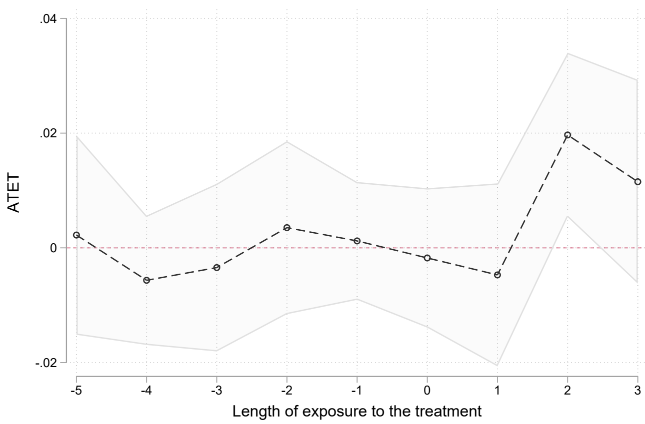
(a) Single Disabled Children



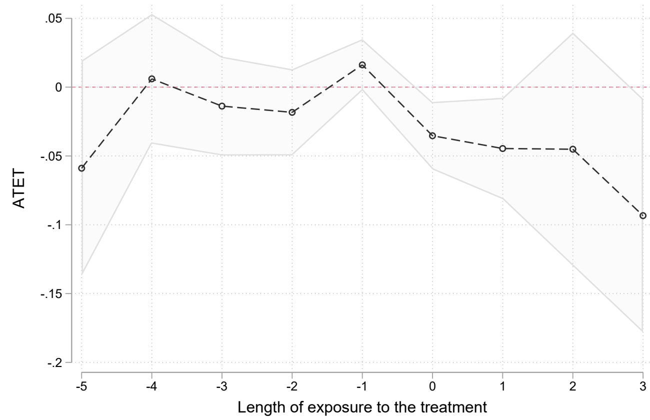
(b) Single Non-Disabled Children



(c) Single Children 10 or Under



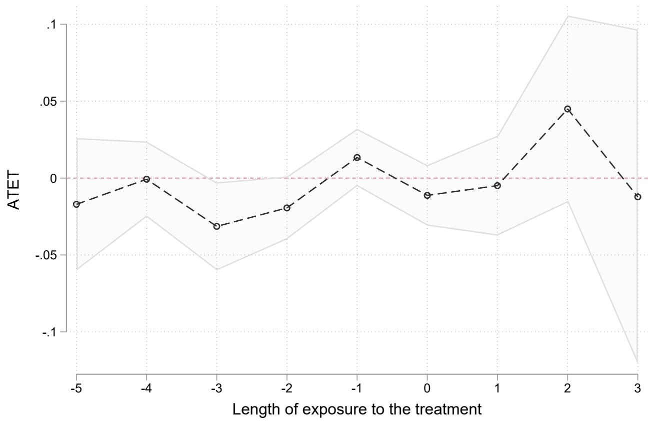
(d) Single Children Over 10



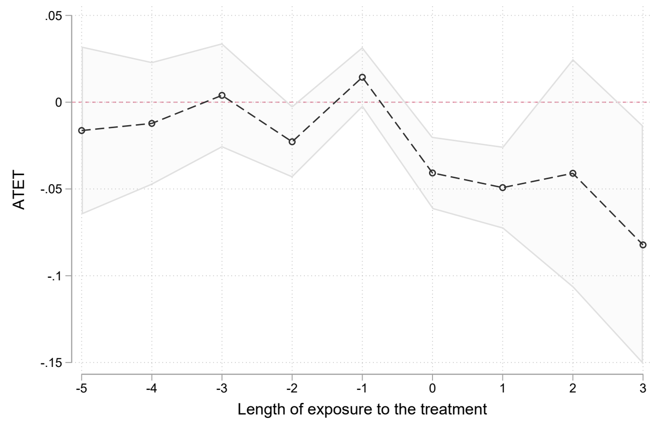
Note: This figure reports the coefficients from the Callaway and Sant'Anna (2021) difference in differences model on the probability of being unmatched (i.e. placed in an institution or group home) among children placed as a single. The coefficients are aggregated across treatment cohorts–sets of states that experienced reform at the same time–by the length of time since treatment occurred. The coefficients in periods $T = \{0, \ldots, 3\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to the year immediately before reform, while the coefficients in periods $T = \{-5, \ldots, -1\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to period $T-1$. The regression sample includes children who enter without any siblings, as well as those who enter with siblings, but are placed in a setting alone. Subfigures (a) and (b) report the results by child disability status, and subfigures (c) and (d) report the results by age.

Figure 7: Aggregated ATET on Probability Unmatched - Single Children by Race and Sex

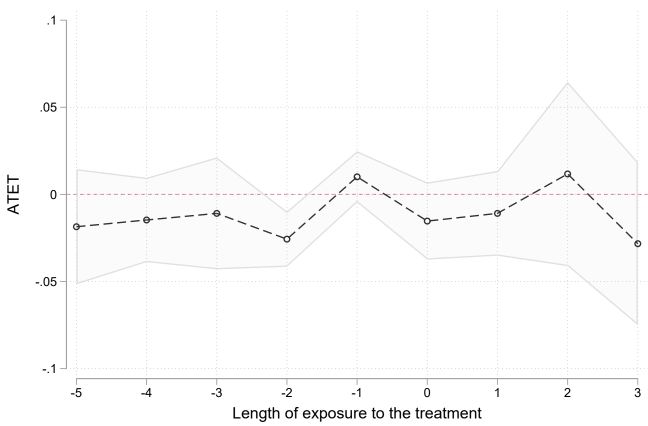
(a) Single White Children



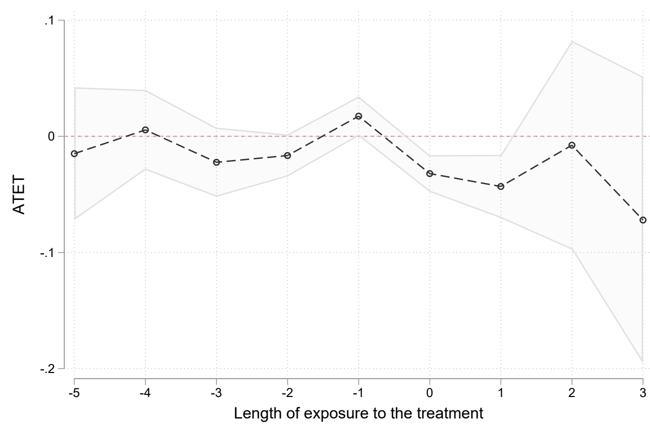
(b) Single Non-White Children



(c) Single Female Children



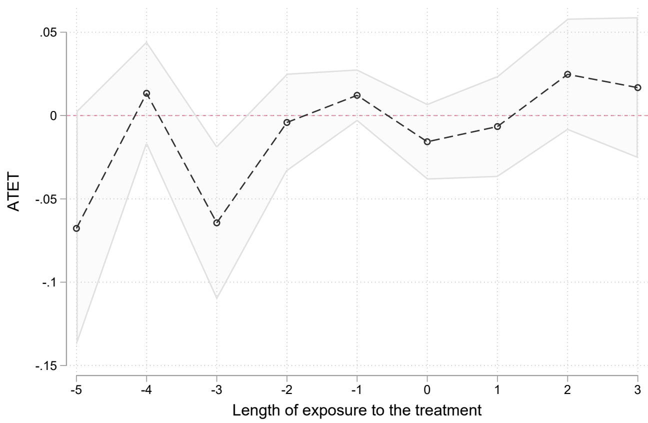
(d) Single Male Children



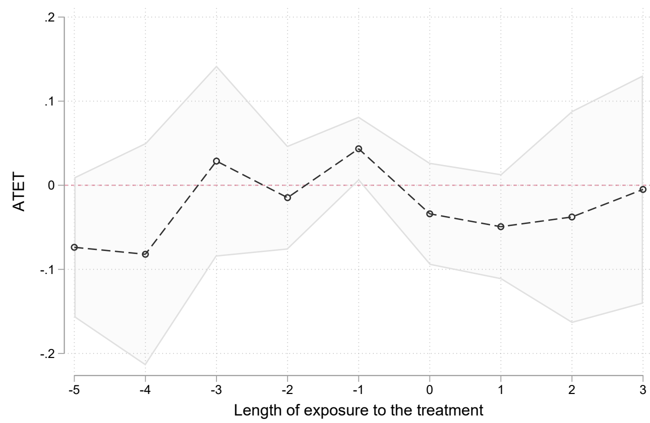
Note: This figure reports the coefficients from the Callaway and Sant'Anna (2021) difference in differences model on the probability of being unmatched (i.e. placed in an institution or group home) among children placed as a single. The coefficients are aggregated across treatment cohorts–sets of states that experienced reform at the same time–by the length of time since treatment occurred. The coefficients in periods $T = \{0, \ldots, 3\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to the year immediately before reform, while the coefficients in periods $T = \{-5, \ldots, -1\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to period $T - 1$. The regression sample includes children who enter without any siblings, as well as those who enter with siblings, but are placed in a setting alone. Subfigures (a) and (b) report the results by child race, and subfigures (c) and (d) report the results by sex.

Figure 8: Aggregated ATET on Probability Unmatched - Homogeneous Siblings by Disability and Age

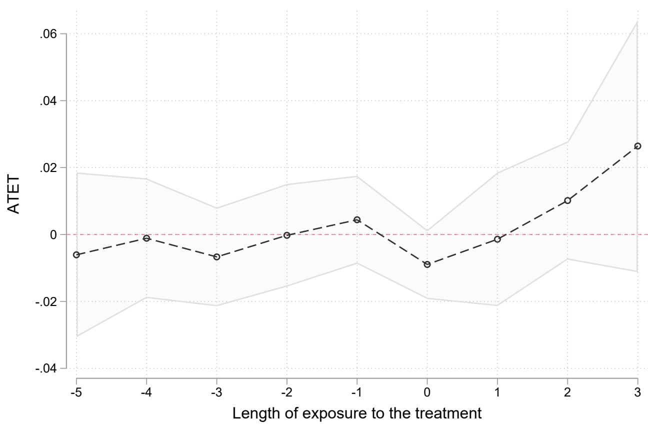
(a) Sibling Groups of Size 2-3 - All Non-Disabled



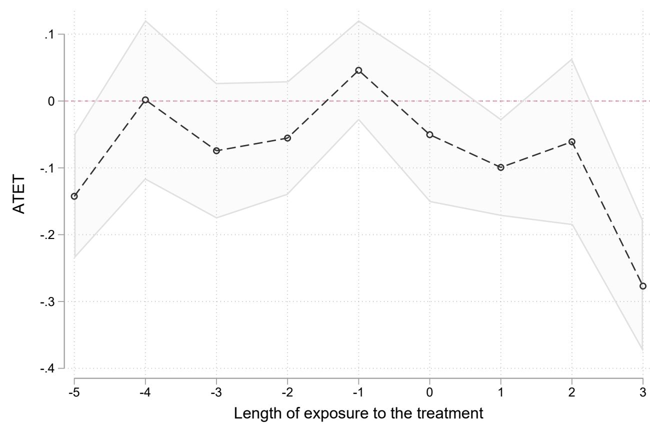
(b) Sibling Groups of Size 2-3 - All Disabled



(c) Sibling Groups of Size 2-3 - All 10 or Under



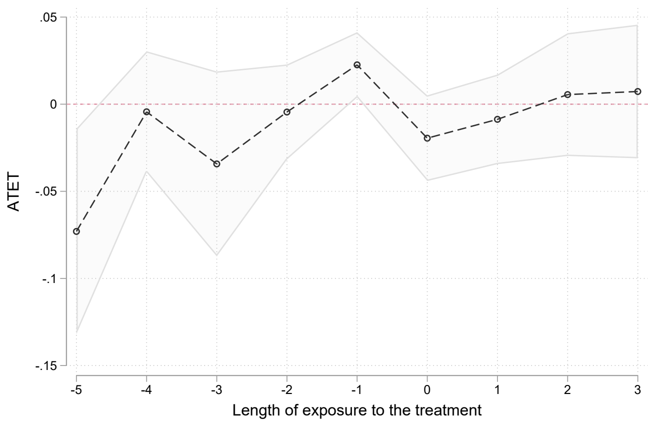
(d) Sibling Groups of Size 2-3 - All Over 10



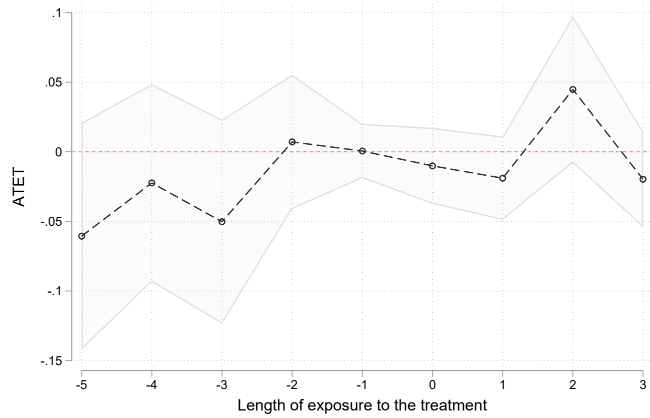
Note: This figure reports the coefficients from the Callaway and Sant'Anna (2021) difference in differences model on the probability of being unmatched (i.e. placed in a group home or institution) among child groups consisting of 2-3 siblings placed together. The coefficients are aggregated across treatment cohorts–sets of states that experienced reform at the same time–by the length of time since treatment occurred. The coefficients in periods $T = \{0, \ldots, 3\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to the year immediately before reform, while the coefficients in periods $T = \{-5, \ldots, -1\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to period $T - 1$. Subfigures (a) and (b) report the results by sibling-group disability status, and subfigures (c) and (d) report the results by sibling-group age.

Figure 9: Aggregated ATET on Probability Unmatched - Homogeneous Siblings by Race and Sex

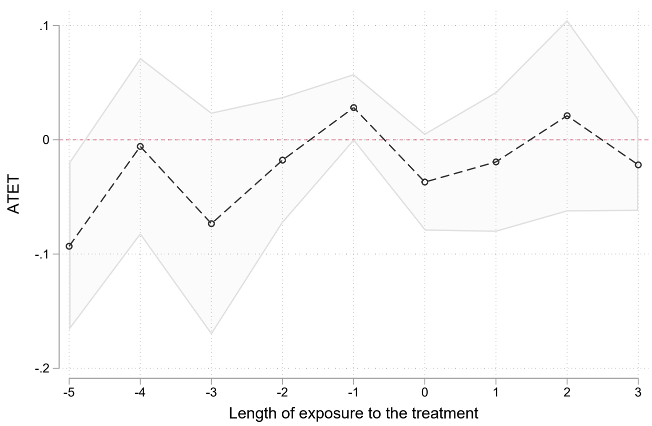
(a) Sibling Groups of Size 2-3 - All White



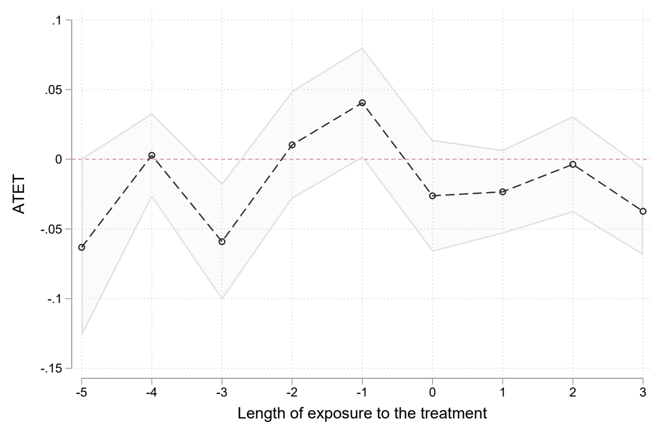
(b) Sibling Groups of Size 2-3 - All Non-White



(c) Sibling Groups of Size 2-3 - All Male



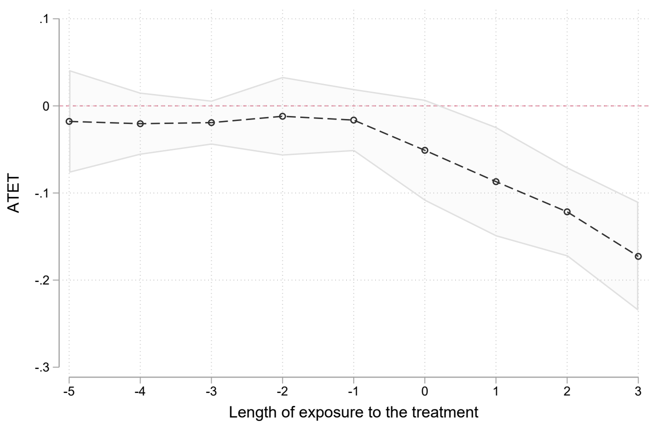
(d) Sibling Groups of Size 2-3 - All Female



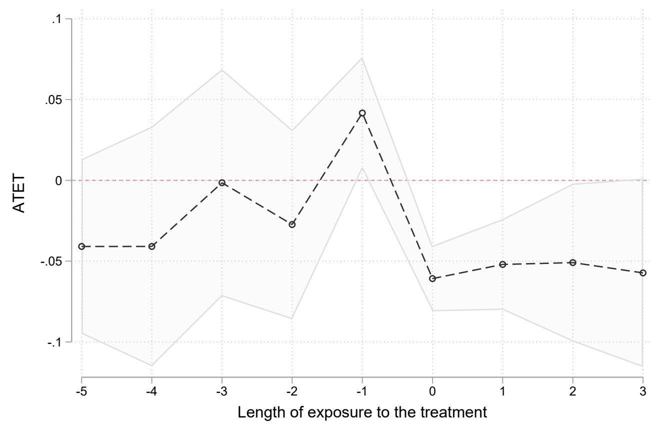
Note: This figure reports the coefficients from the Callaway and Sant'Anna (2021) difference in differences model on the probability of being unmatched (i.e. placed in a group home or institution) among child groups consisting of 2-3 siblings placed together. The coefficients are aggregated across treatment cohorts–sets of states that experienced reform at the same time–by the length of time since treatment occurred. The coefficients in periods $T = \{0, \ldots, 3\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to the year immediately before reform, while the coefficients in periods $T = \{-5, \ldots, -1\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to period $T - 1$. Subfigures (a) and (b) report the results by sibling-group race, and subfigures (c) and (d) report the results by sibling-group sex.

Figure 10: Aggregated ATET on Probability Unmatched - Heterogeneous Siblings

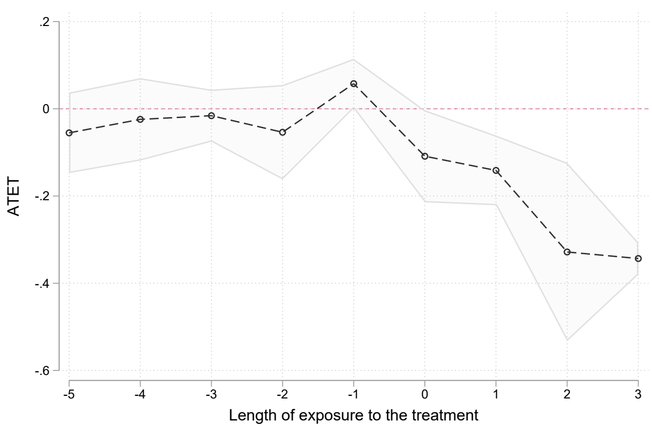
(a) Sibling Groups of Size 2-3 - Multiple Disability Statuses



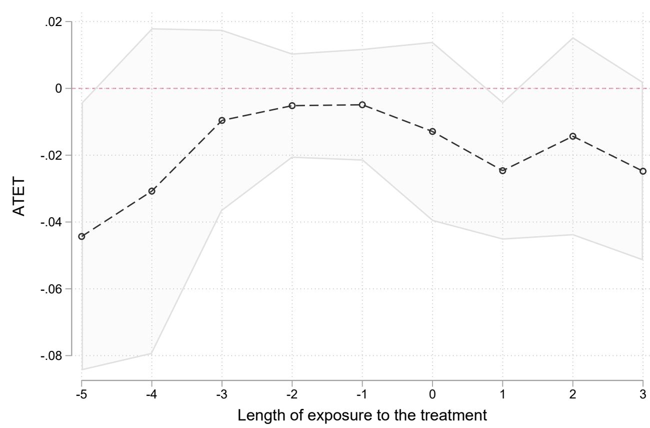
(b) Sibling Groups of Size 2-3 - Multiple Age Groups



(c) Sibling Groups of Size 2-3 - Multiple Races



(d) Sibling Groups of Size 2-3 - Multiple Sexes



Note: This figure reports the coefficients from the Callaway and Sant'Anna (2021) difference in differences model on the probability of being unmatched (i.e. placed in a group home or institution) among child groups consisting of 2-3 siblings placed together. The coefficients are aggregated across treatment cohorts–sets of states that experienced reform at the same time–by the length of time since treatment occurred. The coefficients in periods $T = \{0, \ldots, 3\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to the year immediately before reform, while the coefficients in periods $T = \{-5, \ldots, -1\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to period $T - 1$.

Table 10: Treatment Effect on the Probability a Sibling Group is Split Up by Disability Status and Age

Years Since Treatment	(1) *All Disabled*	(2) *All Non-Disabled*	(3) *Mixed Disability*	(4) *All Age > 10*	(5) *All Age ≤ 10*	(6) *Mixed Age*
-5	0.0170	0.0270	0.0166	0.1006**	0.0073	-0.0193
	(0.0430)	(0.0191)	(0.0263)	(0.0479)	(0.0175)	(0.0185)
-4	0.0135	-0.0131	-0.0023	-0.0062	0.0079	-0.0108
	(0.0348)	(0.0189)	(0.0251)	(0.0226)	(0.0183)	(0.0272)
-3	0.0321	0.0085	-0.0430	0.0662*	0.0009	0.0172
	(0.0258)	(0.0154)	(0.0349)	(0.0395)	(0.0170)	(0.0272)
-2	0.0164	0.0393***	0.0843***	0.0113	0.0387***	0.0540**
	(0.0240)	(0.0145)	(0.0245)	(0.0364)	(0.0090)	(0.0215)
-1	-0.0090	-0.0224*	-0.0720***	-0.0179	-0.0176**	-0.0288
	(0.0347)	(0.0129)	(0.0245)	(0.0187)	(0.0089)	(0.0250)
0	-0.0540*	-0.0287	-0.0599	-0.0286	-0.0225	-0.0531*
	(0.0292)	(0.0197)	(0.0462)	(0.0320)	(0.0139)	(0.0287)
1	-0.0681	-0.0671*	-0.1390*	-0.0496	-0.0573	-0.0995**
	(0.0480)	(0.0371)	(0.0769)	(0.0356)	(0.0358)	(0.0484)
2	-0.0940*	-0.0657*	-0.0999**	-0.0911	-0.0326*	-0.1613***
	(0.0515)	(0.0373)	(0.0436)	(0.1013)	(0.0192)	(0.0225)
3	-0.0589	-0.0529	-0.1278***	-0.1395	-0.0157	-0.0509**
	(0.0523)	(0.0354)	(0.0416)	(0.1082)	(0.0185)	(0.0239)
Aggregated ATE	-0.0642**	-0.0442**	-0.0947*	-0.0496*	-0.0333*	-0.0773***
	(0.0318)	(0.0223)	(0.0505)	(0.0308)	(0.0193)	(0.0276)
Observations	40,117	165,139	58,163	54,588	146,710	62,121
Mean Pre-Reform	0.387	0.233	0.58	0.511	0.207	0.576

Note: The regression samples include sibling groups at time of entry that included 2 or 3 children. The outcome is measured at the sibling group level, and indicates whether at least one child was placed in a different home than at least one of their siblings. For example, if a child entered in a group of three siblings, and two of the siblings were placed together, but one was placed elsewhere, then the outcome variable would be equal to one for that sibling group. Pre-reform means are measured one year before reform. The aggregated ATE is calculated by aggregating across years 0 to 3 post-reform.

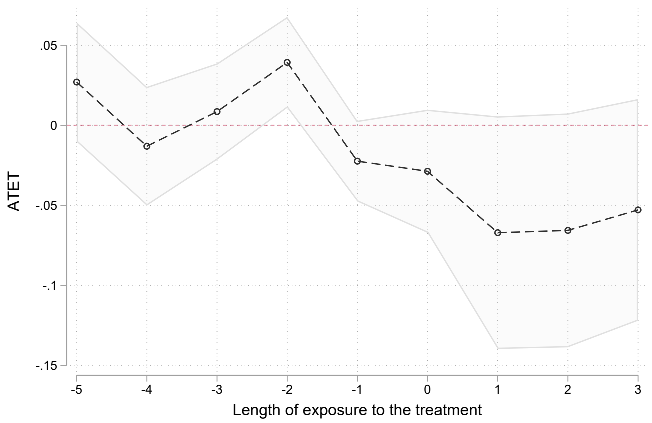
Table 11: Treatment Effect on the Probability a Sibling Group is Split Up by Race and Sex

Years Since Treatment	(1) *All Non-White*	(2) *All White*	(3) *Multiple Races*	(4) *All Male*	(5) *All Female*	(6) *Multiple Sexes*
-5	0.0239	0.0204	0.0182	0.0115	0.0311	0.0059
	(0.0292)	(0.0295)	(0.0260)	(0.0212)	(0.0310)	(0.0198)
-4	-0.0076	0.0018	0.0022	0.0011	0.0017	-0.0128
	(0.0194)	(0.0187)	(0.0307)	(0.0373)	(0.0180)	(0.0177)
-3	0.0089	0.0061	0.0222	0.0453**	-0.0028	-0.0069
	(0.0166)	(0.0159)	(0.0239)	(0.0191)	(0.0220)	(0.0239)
-2	0.0622***	0.0181	0.0169	0.0447**	0.0087	0.0573***
	(0.0159)	(0.0123)	(0.0198)	(0.0193)	(0.0162)	(0.0191)
-1	-0.0410***	-0.0029	-0.0016	-0.0402**	0.0130	-0.0228
	(0.0142)	(0.0132)	(0.0211)	(0.0164)	(0.0167)	(0.0170)
0	-0.0392	-0.0284	-0.0242	-0.0295	-0.0334	-0.0392*
	(0.0279)	(0.0194)	(0.0329)	(0.0230)	(0.0283)	(0.0217)
1	-0.0954**	-0.0446	-0.0801***	-0.1008*	-0.0408	-0.0758**
	(0.0477)	(0.0545)	(0.0251)	(0.0556)	(0.0577)	(0.0315)
2	-0.0534	-0.0571**	-0.1867***	-0.1210**	-0.0175	-0.0992***
	(0.0391)	(0.0249)	(0.0363)	(0.0484)	(0.0430)	(0.0365)
3	-0.0604	-0.0016	-0.2396***	-0.0599	-0.0142	-0.1169***
	(0.0388)	(0.0174)	(0.0340)	(0.0637)	(0.0202)	(0.0390)
Aggregated ATE	-0.0613*	-0.0336	-0.0735**	-0.0627**	-0.0326	-0.0617***
	(0.0330)	(0.0230)	(0.0311)	(0.0305)	(0.0329)	(0.0200)
Observations	102,664	125,809	34,946	72,789	54,020	136,610
Mean Pre-Reform	0.363	0.234	0.705	0.384	0.218	0.376

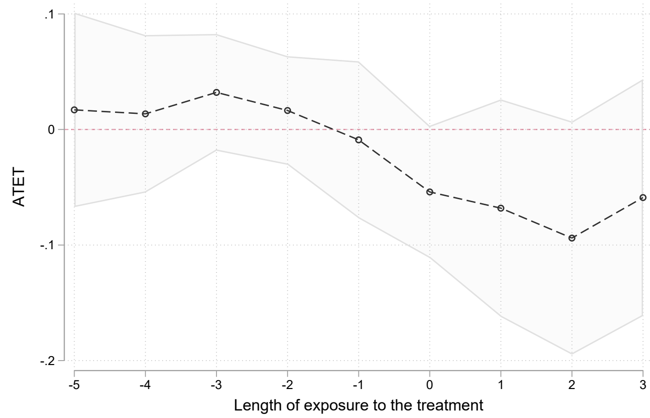
Note: The regression samples include sibling groups at time of entry that included 2 or 3 children. The outcome is measured at the sibling group level, and indicates whether at least one child was placed in a different home than at least one of their siblings. For example, if a child entered in a group of three siblings, and two of the siblings were placed together, but one was placed elsewhere, then the outcome variable would be equal to one for that sibling group. Pre-reform means are measured one year before reform. The aggregated ATE is calculated by aggregating across years 0 to 3 post-reform.

Figure 11: Aggregated ATET on Probability Split Up - Homogeneous Siblings by Disability and Age

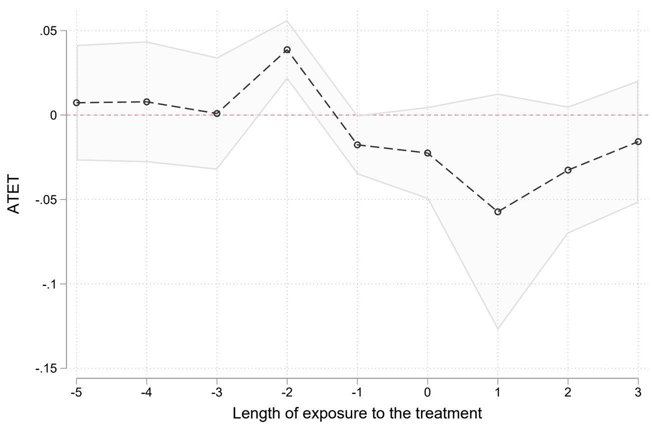
(a) Sibling Groups of Size 2-3 - All Non-Disabled



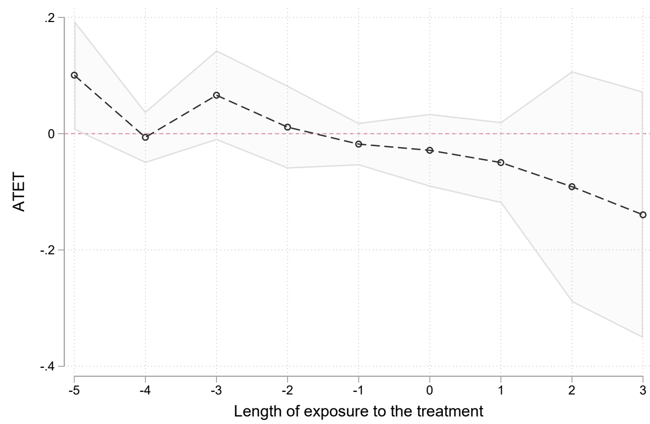
(b) Sibling Groups of Size 2-3 - All Disabled



(c) Sibling Groups of Size 2-3 - All 10 or Under



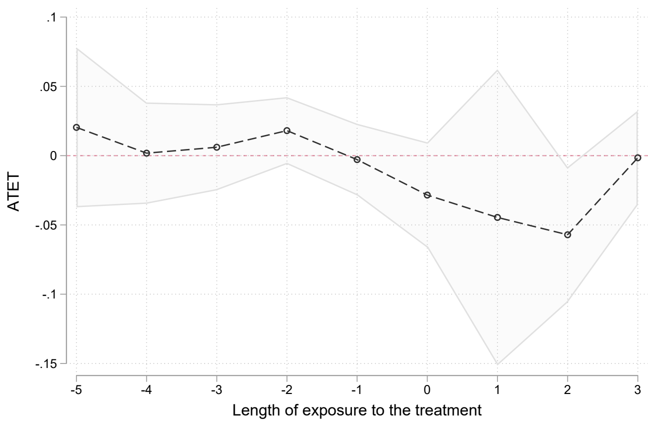
(d) Sibling Groups of Size 2-3 - All Over 10



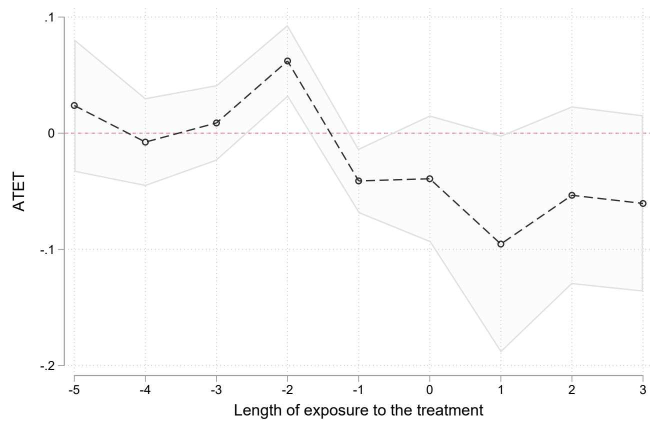
Note: This figure reports the coefficients from the Callaway and Sant'Anna (2021) difference in differences model on the probability of being separated (i.e. not having all siblings placed together) among child groups consisting of 2-3 siblings who entered together. The coefficients are aggregated across treatment cohorts– sets of states that experienced reform at the same time–by the length of time since treatment occurred. The coefficients in periods $T = \{0, \ldots, 3\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to the year immediately before reform, while the coefficients in periods $T = \{-5, \ldots, -1\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to period $T - 1$. Subfigures (a) and (b) report the results by sibling-group disability status, and subfigures (c) and (d) report the results by sibling-group age.

Figure 12: Aggregated ATET on Probability Split Up - Homogeneous Siblings by Race and Sex

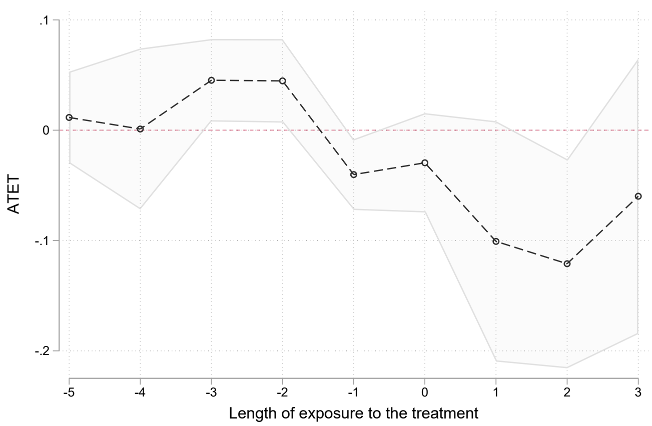
(a) Sibling Groups of Size 2-3 - All White



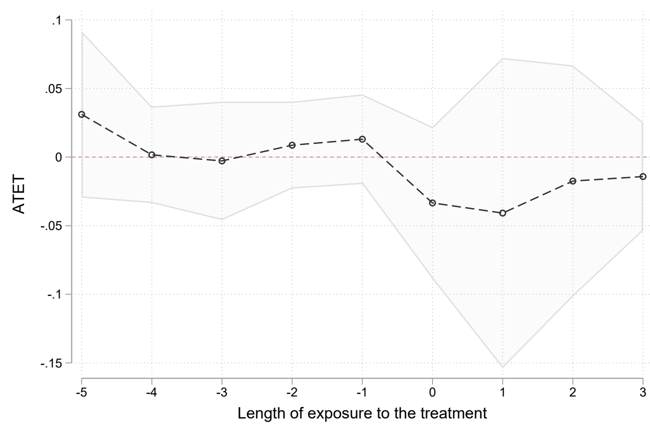
(b) Sibling Groups of Size 2-3 - All Non-White



(c) Sibling Groups of Size 2-3 - All Male



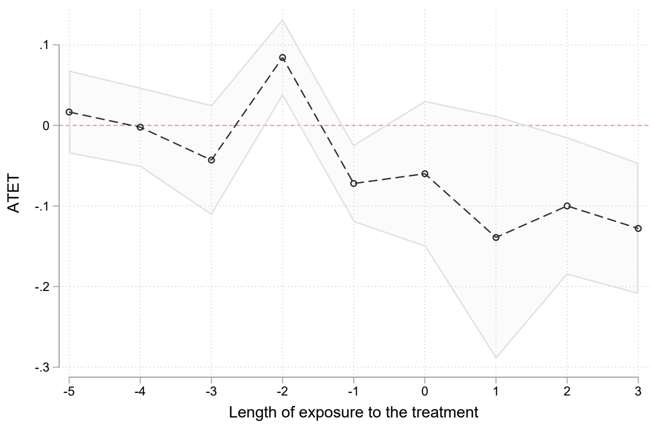
(d) Sibling Groups of Size 2-3 - All Female



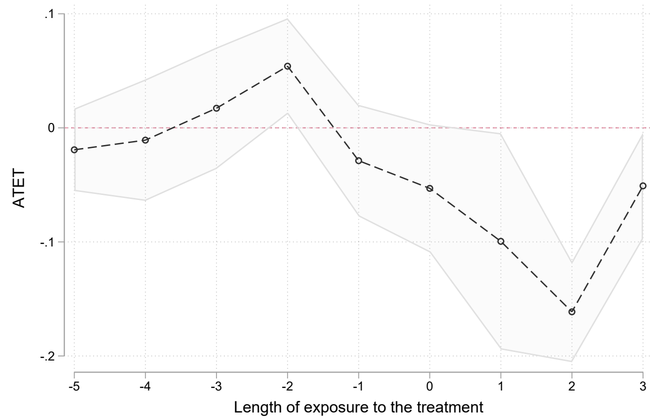
Note: This figure reports the coefficients from the Callaway and Sant'Anna (2021) difference in differences model on the probability of being separated (i.e. not having all siblings placed together) among child groups consisting of 2-3 siblings who entered together. The coefficients are aggregated across treatment cohorts– sets of states that experienced reform at the same time–by the length of time since treatment occurred. The coefficients in periods $T = \{0, \ldots, 3\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to the year immediately before reform, while the coefficients in periods $T = \{-5, \ldots, -1\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to period $T - 1$. Subfigures (a) and (b) report the results by sibling-group race, and subfigures (c) and (d) report the results by sibling-group sex.

Figure 13: Aggregated ATET on Probability Split Up - Heterogeneous Siblings

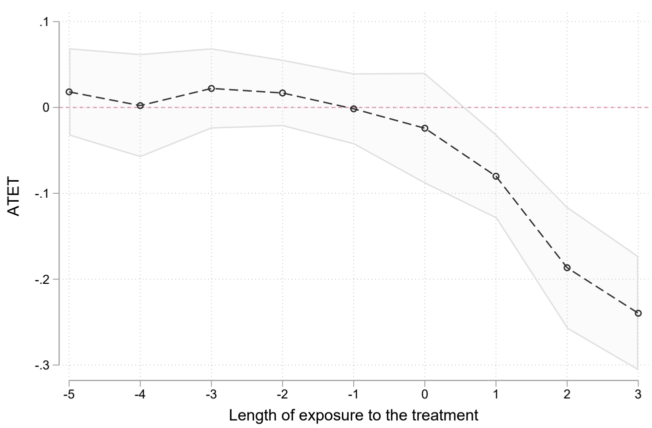
(a) Sibling Groups of Size 2-3 - Multiple Disability Statuses



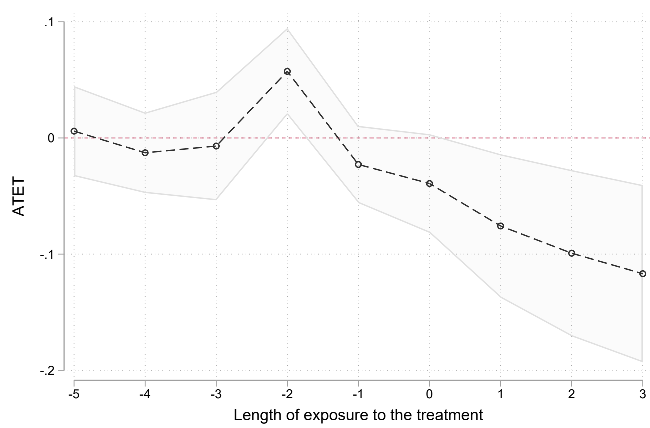
(b) Sibling Groups of Size 2-3 - Multiple Age Groups



(c) Sibling Groups of Size 2-3 - Multiple Races



(d) Sibling Groups of Size 2-3 - Multiple Sexes



Note: This figure reports the coefficients from the Callaway and Sant'Anna (2021) difference in differences model on the probability of being separated (i.e. not having all siblings placed together) among child groups consisting of 2-3 siblings who entered together. The coefficients are aggregated across treatment cohorts–sets of states that experienced reform at the same time–by the length of time since treatment occurred. The coefficients in periods $T = \{0, \ldots, 3\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to the year immediately before reform, while the coefficients in periods $T = \{-5, \ldots, -1\}$ represent the change in outcomes for treated states relative to not-yet-treated states, compared to period $T - 1$.

B Model Validation Procedure

The model validation procedure works as follows:

1. Randomly select half of the 202 markets. This will be the estimation sample. The unselected markets will compose the validation sample.
2. Estimate $\{\sigma_Y, \sigma_O, \sigma_{YY}, \sigma_{OY}, \sigma_{OO}, \mu_{YY}, \mu_{OY}, \mu_{OO}\}$ on the estimation sample
3. Using $\{\hat{\sigma}_Y, \hat{\sigma}_O, \hat{\sigma}_{YY}, \hat{\sigma}_{OY}, \hat{\sigma}_{OO}, \hat{\mu}_{YY}, \hat{\mu}_{OY}, \hat{\mu}_{OO}\}$, for each market in the validation sample, estimate at the individual level the probability each sibling group of size two is split up. For each market in the validation sample, aggregate these estimates to get the predicted market-level rate of sibling separation.
4. Compute the difference between the actual rate of sibling separation within a market and predicted rate of sibling separation, for sibling groups of size 2.
5. Call this collection of residual R_1
6. Repeat steps 1-4, n times, collecting $R_1, ..., R_n$. For this analyis, $n = 3$
7. Plot $R_1, ..., R_n$