

Adoption and Adult Outcomes in the Early 20th Century

Chiaki Moriguchi* and John M. Parman[†]

August 20, 2014

Abstract

Modern research has found strong links between family structure and children's outcomes. One of the robust findings is that stepchildren and adopted children have worse adult outcomes compared to biological children. However, we know very little about how non-biological children fared historically. In this study, by linking adopted children across U.S. federal censuses in the first half of the 20th century, we create a new dataset that contains rich information on both their childhood households and adult outcomes. To control for household heterogeneity, we also follow (non-adopted) siblings of adopted children into their adulthood. This unique dataset enables us to compare the long-run outcomes of adopted children and biological children controlling for observable and unobservable household characteristics. Our preliminary analysis suggests that educational attainment, income, and marriage patterns of adopted children differed significantly from non-adopted children. Overall, our study brings new historical evidence to the research on family structure.

JEL classifications: I2, J1, N3 Keywords: Adoption, Household Economics, Family Resource Allocation

PRELIMINARY DRAFT

Do Not Cite, Quote, or Circulate Without Permission

*Institute of Economic Research, Hitotsubashi University, e-mail: chiaki@ier.hit-u.ac.jp

[†]Department of Economics, College of William and Mary and NBER, 130 Morton Hall, Williamsburg, VA 23187 e-mail: jmparman@wm.edu

[‡]This project has been funded with the support of a grant from the Japan Society for the Promotion of Science.

1 Introduction

Modern research points to significant consequences of family structure for children's socioeconomic success. A sizable literature exists demonstrating that stepchildren and adopted children fare worse than biological children across a range of outcomes. While these findings are robust across a broad set of modern data sources, there are no comparable data to assess whether living with non-biological parents impacted children's educational and labor market outcomes in a similar way historically. This paper introduces a newly constructed longitudinal dataset of adopted children and their siblings and uses those data to estimate the effects of adoption on socioeconomic outcomes in the first half of the twentieth century.

Studies of adopted children in modern settings have demonstrated that adopted children tend to be more vulnerable to a range of emotional, behavioral and educational problems as children which can translate into worse socioeconomic outcomes as adults. The existing literature points to a variety of reasons that adopted children fare worse than children raised by their biological parents: the households that adopt children may be negatively selected from the general population of households, the psychological effects of being adopted may translate into problems in school and social settings, and parents may treat adopted children differently than their non-adopted siblings.

While the modern literature suggests that adoption has complex effects on socioeconomic outcomes, there is very little evidence suggesting whether adoption had similar effects historically. The early twentieth century United States witnessed adoption rates that, while lower than modern adoption rates, were still of the same order of magnitude yet those adoptions were taking place under very different economic, social and legal environments. Extending analysis of adoption to the first half of the twentieth century would provide an opportunity to explore how those differences in the environment faced by adopted children and their adoptive families altered the effects of adoption.

A lack of appropriate data has limited the study of adoption in the early twentieth century. There are no historical datasets comparable to the modern longitudinal studies and clinical studies looking at outcomes of adopted children. In this paper, we construct a

new dataset linking adopted children and their siblings from the 1910 federal census to the 1940 federal census to create a longitudinal dataset that rivals modern data sources in both size and scope. Our dataset includes detailed information on the characteristics of the adoptive family in 1910 and a range of adult outcomes in 1940 for both the adopted children and their non-adopted siblings. These outcomes include educational and occupational outcomes as well as information on family formation and geographic mobility. These data offer an unprecedented opportunity to examine the historical impacts of adoption on adult outcomes controlling for both the observed and unobserved characteristics of the adoptive households. We complement these data with cross-sectional data from the 1900, 1910, 1920 and 1930 public use micro samples of the federal census to examine the school attendance and labor force participation of adopted individuals and their siblings as children.

Our preliminary results demonstrate that adopted children did fare worse than the general population. As children, adopted individuals had lower school attendance rates compared to the general population and to their non-adopted siblings. We find that these lower school attendance rates translated into lower educational attainments as adults. Additionally, adopted individuals earned substantially lower incomes on average and worked fewer hours than their peers.. Adopted individuals also exhibited different family formation patterns than the general population: they were more likely to be married, tended to have larger households and were more likely to move across states. These effects held not just for the adopted individual but also for that individual's siblings as well. The biological children of adoptive parents also had larger households, lower incomes and exhibited greater geographic mobility than the general population suggesting that the channels through which adoption impacts adult outcomes may operate at a household rather than strictly individual level.

2 Adoption and Adult Outcomes in Modern and Historical Times

There has been considerable interest in the effects of adoption on childhood and adult outcomes in the economics, sociology and psychology literatures. Much of the interest centers around questions of how family structure influences outcomes in terms of the psychological development of children as well as the distribution of household resources and investments in education. Economists have also viewed adoption as an opportunity to disentangle the effects of nature and nurture, looking at how influential the characteristics of an adopted child's biological parents are relative to the characteristics of the adoptive parents. These studies typically find that children from nontraditional family structures including adopted children and stepchildren ultimately have worse outcomes across a range of measures.¹

A variety of clinical and epidemiological studies exist linking adoption to adverse behavioral and emotional problems that can impact adult socioeconomic problems either directly or indirectly through their effects on academic performance. Epidemiological studies have found that adopted children constitute a disproportionately high percentage of the children in residential care facilities, inpatient psychiatric settings and outpatient mental health clinics (Zill, 1985; Brodzinsky, 1987; Piersma, 1987; Rogeness et al., 1988). Researchers have suggested that this a product of both higher incidence of mental health issues among adopted children and of a greater willingness of parents and others to seek out mental health resources for adopted children as compared to biological children (Hersov, 1990; Warren, 1992). These results suggest the possibility of differential treatment of adopted and biological children by parents, something we hope to explore with the unique longitudinal data on both adoptees and siblings at the heart of this study. Complementing these

¹It is important to note that the general empirical approach in the literature compares the outcomes of adopted children or stepchildren to children raised by their biological parents. This is not necessarily the correct counterfactual when asking what the impacts of adoption and adoptive family characteristics are on a child that has already lost his or her biological parents. It is quite likely that adoption in this case leads to better outcomes than any relevant alternatives such as moving from one temporary household to another while in the foster care system. It should be kept in mind throughout this paper than when we refer to better or worse outcomes for adopted children, we are referring to how their outcomes compare to the general population, not to children displaced from their families who are not subsequently adopted.

epidemiological studies are clinical studies investigating the prevalence of personality disorders, substance abuse and learning disabilities among adopted children. The results from these studies are mixed, with some studies showing higher rates of behavioral problems, substance abuse, attention deficit disorder and learning disabilities among adopted children while others find little to no difference between adopted and biological children in clinical settings (see Brodzinsky (1993) for an overview of these studies).

While these epidemiological and clinical studies help identify the particular issues adopted children face that could impact their educational attainments and labor market outcomes, they are limited in the questions they can investigate due to the nature of the data being employed. In nearly all cases, the data come from rather very small samples of adoptees severely limiting the power of any statistical analysis. Consider Warren (1992) who uses the 1981 National Health Survey to study 3,698 adolescents, only 45 of whom were adopted. Clinical studies fare no better in terms of sample sizes. Dickson et al. (1990) offer a typical example. Their psychiatric hospital inpatient data include only 44 adopted children. Furthermore, these types of study looking at the prevalence of adopted children among all children in health surveys or clinical data cannot address the adult outcomes of being adopted as they rely on cross-sectional data observed during childhood.

To assess the long term outcomes of adopted children, researchers have turned to longitudinal datasets that track those children into their teenage years and adulthood. These data sources have revealed that children raised by non-biological parents not only receive lower investments as children, as Case & Paxson (2001) show for health investments, but also ultimately have lower educational attainments and occupational outcomes than children raised by their biological parents and in particular their biological mother (Case et al., 1999, 2000, 2001). While these longitudinal surveys are a major improvement in terms of the amount of data available relative to clinical studies, sample sizes remain an issue. The data used by Case et al. (2001) from the Panel Study of Income Dynamics (PSID) contain only 93 individuals with an adoptive mother and 130 individuals with an adoptive father. In his work using the experience of adopted children to compare the relative importance of

nature and nurture by looking at the effects of biological parental characteristics relative to adoptive parental characteristics, Sacerdote (2000, 2002) reveals that this is a problem common to all of the major longitudinal surveys familiar to economists. Sacerdote relies on 128 adoptees in the National Childhood Development Survey (NCDS), 198 adoptees in the National Longitudinal Survey of Youth 1979 (NLSY79), and 183 adoptees tracked by the Colorado Adoption Project (CAP). These small samples of adoptees severely restrict the insights we can gain from modern data.

The difficulty of quantifying the effects of adoption in modern times with the limited number of adoptees in modern longitudinal surveys or clinical data is magnified when considering the historical effects of adoption as comparable longitudinal or clinical data do not even exist. Most major longitudinal studies did not begin until at least the 1970s offering little hope of individual level data on adoptees in the first half of the twentieth century. Even aggregate data is scarce historically. Collection of national statistics on adoption did not begin until 1944, when the National Center for Social Statistics began compiling adoption level statistics from state-level court records. However, reporting was voluntary with varying numbers of states reporting from one year to the next, requiring extensive assumptions to piece together national trends in historical adoption rates (Zarefsky, 1946; Bonham, 1977; Maza, 1984).

Despite the paucity of historical data on adoptions, extending studies of the consequences of adoption to the first half of the twentieth century is a particularly worthwhile albeit difficult endeavor. Historical data would offer an opportunity to gain a much richer understanding of how adoption affects household resource allocation decisions and socioeconomic conditions. The first half of the twentieth century witnessed transitions in the structure of the economy, changes in attitudes toward the family as an economic unit, and shifts in the legal framework governing adoption. All of these shifts offer promising opportunities to empirically identify the motivations behind adoption and the channels through which adoption influences adult outcomes. America entered the twentieth century with the majority of its residents living in rural areas where farm life accommodated large family

sizes. By the middle of the century, the majority of the population had shifted to urban areas, changing the nature of work and the nature of the family. The High School Movement took hold in the first decades of the twentieth century, offering children greater access to secondary education and the jobs that increasingly demanded that education (Goldin & Katz, 2009). The Progressive Movement in the early twentieth century also ushered in a new emphasis on protecting the welfare of children. While the major results of this emphasis on child welfare were the introduction of compulsory schooling and child labor laws, the movement also changed the nature of adoption laws. One key piece of legislation was the Children's Code of Minnesota of 1917 which became a model followed by other states in the following years. With the Children's Code, Minnesota became the first state to require an investigation to determine the suitability of a prospective adoptive home. The code also set up other means of providing for the welfare of adopted children such as establishing a six-month probationary period in the adoptive home and having the state's Board of Control review adoption petitions. The next two decades saw similar legislation passed in other states, with 44 states enacting or revising stricter adoption laws by 1937 (see Carp (2009); Katz (1996); Gordon (1994); Heisterman (1935) for extensive discussion of adoption legislation during the Progressive Era).

All of these changes in American society taking place during the early twentieth century suggest promising sources of variation in the legal and social institutions that could influence the practice of adoption. As America shifted from the orphan trains of the late nineteenth century to modern adoption practices by the middle of the century, it is quite possible that there were equally dramatic changes in the outcomes of adopted children.² Researchers have begun to piece together evidence on the magnitude of these changes as new data sources become available. In particular, the recent digitization of the complete census returns through 1940 have offered an opportunity to assess how adoptive households have changed from the late nineteenth century through the first half of the twentieth

²See Askeland (2006) for discussion of these historical transitions in the nature of adoption, from the informal adoption of colonial times to the orphan trains of the late 1800s to modern sealed-document adoption.

century.³ Moriguchi (2013) examines households with adopted children or stepchildren from 1880 through 1930 in the federal census and compares them to households in 2000, finding that the conditions of adoptive households have improved substantially relative to biological households over the past century. This evidence that adoptive households have changed over time suggests that there may also be changes in the impact of adoption on adult outcomes over time. The goal of this study is to build on the work of Moriguchi (2013) and exploit the ability to link individuals across historical censuses to estimate the effects of adoption on educational and occupational outcomes as well as family formation in the early twentieth century.

3 Using Cross-sectional Data to Explore the Impacts of Adoption

Before turning to the construction of a longitudinal data set to examine adult outcomes of adoptees, it is instructive to exploit existing cross-sectional data on adoptees as children to assess the scope of adoption in the early twentieth century and the characteristics of adopted children and their adoptive families that may have implications for adult outcomes. The Integrated Public Use Microdata Series (IPUMS) samples of the federal census provide an excellent opportunity to observe adopted children while living in their adoptive households and compare their school attendance and labor force participation to both their siblings and the general population.

The IPUMS samples for the 1900, 1910, 1920 and 1930 provide a reasonably large number of adopted children. These samples range from 1% samples of the US population for the 1920 and 1930 censuses to 1.4% and 2.5% samples for the 1910 and 1900 censuses, respectively. Table 1 provides the estimated number of two-parent households by child type and census year for the entire US population based on the IPUMS samples. The number

³These data sources are only now becoming available both because of the massive amount of time required to digitize the records and because of the 72-year-rule which says that census records only become public 72 years after the census took place.

of households with adopted children ranges from roughly 35,000 in 1920 to over 70,000 in 1930. Given the sampling frequency of the IPUMS samples, these numbers lead to several hundred households with adopted children observed in each IPUMS data set, relatively large samples sizes compared to modern adoption studies. The figures in Table 1 help underscore both the feasibility and importance of a study of historical adoption. Adoption was a widespread phenomenon in the early twentieth century and historical federal census data offer a promising means of studying the conditions and outcomes of adopted children during this period.

It should be noted that these numbers are based on people self-identifying as adopted in the federal census. The historical federal censuses did not specifically ask about adoption. Instead, the status of adopted individuals is inferred from the response to a question asking for the “relationship of this person to the head of the family”. The instructions to the census enumerator listed several examples of relationships but made no mention of adoption.⁴ Consequently, all instances of adopted sons, adopted daughters or any other adopted members of the family in the census were self reported . The number of adopted individuals in any of these historical censuses may exclude adopted individuals who are listed simply as sons or daughters in the census and may include individuals that are informally adopted by families but may not meet the stricter definitions of adopted used in more modern surveys.

The advantage of using a cross-section of adopted individuals as children is that they can be observed in their childhood households. This allows for comparing the adopted individuals not only to all children but also to their non-adopted siblings. The size of the IPUMS historical census samples makes such comparisons possible: hundreds of households each census sample from 1900 to 1930 contain both adopted and biological children. Comparing adopted children to their non-adopted siblings provides an opportunity to control for household characteristics that are both observed and unobserved, affording an oppor-

⁴In the 1900 federal census, the first with a large number of adopted individuals reported, the instructions to the enumerator list the following examples for relationships to the head of the family: “head, wife, mother, father, son, daughter, grandson, daughter-in-law, aunt, uncle, nephew, niece, boarder, lodger, servant, etc.”.

tunity to determine whether any observed differences between adopted children and the general population are driven by the characteristics of households that tend to adopt rather than by how the adopted children fare within those households.

Table 2 and Table 3 present the results of OLS regressions of various outcomes of interest on household characteristics, individual characteristics and adoption status. The limited nature of the information reported for children in the federal census restricts the outcomes we can examine to whether the child is currently attending school, whether the child is in the labor force, and whether the child is literate. The set of individual, parental, and household characteristics that we can control for is far richer including the age, race, employment status, socioeconomic status and nativity of both parents, the location of the household in terms of the state, urban/rural status, and whether it is in a metropolitan area, whether the house is owned and whether the family lives on a farm. Columns (1), (3) and (5) of Table 2 and (1), (3), (5) and (7) of Table 3 include these full sets of controls for observable characteristics and an indicator for being adopted in order to compare adopted children to the general population. Columns (2), (4) and (6) of Table 2 and (2), (4), (6) and (8) of Table 3 include household fixed effects, allowing us to control for both observed and unobserved household characteristics. For these regressions, the coefficient on the adopted indicator variable is capturing the marginal effect of being adopted relative to other children raised in the same household environment.

The results suggest that adopted children fared worse than the general population in terms of schooling. Younger adopted children ages 6 to 11 were roughly four percent less likely to attend primary school than non-adopted children while older children ages 12 to 17 were over five percent less likely to attend secondary school than non-adopted children. In terms of educational outcomes, we find that adopted children are significantly less likely to be literate than non-adopted children. While the signs of these marginal effects remain the same when including household fixed effects, the coefficients are no longer statistically significant. However, we do find statistically significant differences within families when looking at labor force participation. Adopted children were less likely to be in school and

out of the labor force than their non-adopted siblings. These results raise the possibility that the childhood schooling and labor force experiences of adopted children may translate into disadvantages as adults in terms of overall educational attainment and associated labor market outcomes. The next sections explore the extent to which adoptees differ from the general population and their siblings as adults through the construction of a longitudinal data set.

4 Constructing a Historical Longitudinal Sample of Adoptees

As discussed in the previous sections, a primary challenge of studying historical adoption practices and their consequences for children's adult outcomes is the scarcity of data on adoptees. As noted earlier, the types of longitudinal surveys best suited to investigating the adult outcomes of adopted children were not undertaken until recent decades. However, the lack of privacy restrictions on historical census data offers a means of creating a longitudinal dataset that can actually exceed the scale of these modern studies. After 72 years, federal census records become public. This means that researchers have access to information such as name and birth year that can be used to identify the same individual across multiple federal censuses, offering an opportunity to observe the same individual both in his childhood household and as an adult. The original federal manuscript pages, and in particular the 1940 federal census, offer a wealth of information about individuals and their households that provides detail on family composition, geographic mobility, and socioeconomic status that rivals the variables available in modern data sources. Furthermore, because the complete census manuscripts are available, the entire US population can be observed leading to large numbers of adoptees available for analysis.

The process of creating a longitudinal sample of adoptees begins by identifying the set of all adopted children present in the 1910 federal census. The 1910 federal census is chosen for two primary reasons. First, of the federal censuses that are public, only the 1900 through

1940 censuses contain a sizable number of adoptees, as shown in Table 4. These counts of adoptees are not official statistics. Rather, they are the number of records returned when searching the complete census records on the genealogy website Ancestry.com for the keyword “adopted” appearing in the relationship to the head of household variable. As noted in the previous section, these historical censuses did not specifically prompt individuals to report whether or not they were adopted. The set of adopted individuals identified on Ancestry.com may differ from the set of individuals who would meet modern definitions of adopted. Regardless of whether they are undercounts or overcounts of the number of adopted individuals, the electronically indexed census records do not provide a sizable number of adopted individuals until 1900.

The second reason for focusing on the 1910 federal census is that we want to match adopted children from the chosen census to a later census in order to observe them as adults. The ideal census for observing the individuals as adults is the 1940 census. This is the first census in which individuals were asked to report their annual income and their educational attainment. Earlier censuses contain occupation but not income and literacy but not years of educational attainment. The 1940 census also offers a rich set of variables asking for information about the individual in 1935, offering additional longitudinal data not available in other censuses. Given that the 1940 census will be used for adult outcomes, the 1910 census offers the largest sample of children who will be old enough in 1940 to be moved out of their parents’ household and young enough to still be active participants in the labor market.

Once the set of all adopted individuals in the 1910 federal census is obtained from the electronic index of census records available on Ancestry.com, the digitized information on each individual is used to restrict the sample to males with another household member born within two years the adopted individual.⁵ The restriction to males is necessary because individuals are matched to the 1940 federal census on the basis of first and last name. This

⁵The digitized information includes the individual’s first name, last name, birth year, birth state, father’s birth state, mother’s birth state, township, county and state of residence, gender, race, relationship to the head of the household and the names and ages of other household members.

requires that the last name stay the same into adulthood which will not be true of females who marry and take their husband's name. The restriction to individuals with another household member close in age enables us to have a sibling that can also be matched to 1940 federal census. Matching multiple siblings from the same family to the 1940 census will allow us to include family fixed effects in our analysis adopted children, controlling for both observed and unobserved childhood household characteristics common to all children in the family. This will be particularly useful to account for the negative selective of households with nonbiological parents identified by Moriguchi (2013).

The resulting set of roughly 15,000 adopted individuals is then searched for one at a time in the electronic index of the 1940 federal census by first and last name, birth year and birth state.⁶ Information on the best and second best matches in the search results are recorded including the first and last name, birth year and birth state. Of the roughly 15,000 adopted individuals, 34 percent have a good match in the 1940 federal census defined as a match for which the birth year is within two years, the birth state matches, the last name is identical and the first name is identical or a common nickname for the first name in the 1910 census.⁷ For roughly half of these individuals with a good match, the second best match result also meets this definition of a good match. We drop these individuals from the sample of matched adopted individuals, keeping only those who have a unique good match. Finally, we transcribe the complete information for the individual from the 1940 federal census.

For those individuals who were successfully matched to the 1940 federal census, the three siblings from their childhood household closest in age to the adopted individual are also

⁶Note that race, while reported, is not used as a primary search criteria. Race is recorded in the census as one or two handwritten characters: often with 'W' for white, 'B' for black or 'N' for negro, 'Mu' for mulatto, and 'Ch' for Chinese. Without the complete words recorded, poor handwriting and poor scans of the manuscript pages can often lead to incorrectly recorded races in the digitized records throwing off search results. Race is used to distinguish between otherwise identical search results.

⁷Allowing for the birth year given in the 1940 census to be within two years of the birth year in the 1910 census is necessary because of the high levels of age misreporting in historical records. In a comparison of ages on death certificates matched to federal census records in 1960, only 73 percent of white males and 65 percent of white females had ages that were in agreement between the two sources(Hambright, 1969). Rates of agreement were even worse for minorities, with only 43 percent of nonwhite males and 34 percent of nonwhite females having the same ages reported in both sources.

searched for in the 1940 federal census. These searches are done on the basis of first name, last name, birth year and birth state and evaluated in the same manner as the searches for the adopted individual. If a unique match is found in the 1940 federal census, the complete information for the individual from the 1940 federal census is transcribed. The linking of the siblings suffers the same problems as the linking of the adopted individuals: roughly one third have a good match in the 1940 census and only half of those are unique matches. For those adopted individuals who are successfully linked to the 1940 federal census and have at least one sibling successfully linked to the 1940 federal census, the final data set contains information on the childhood household characteristics of the adopted individual, the childhood schooling experience of both the adopted individual and his siblings, the adult household of the adopted individual, the adult occupational, income and educational attainment outcomes for the adopted individual, and the same set of adult characteristics and outcomes for his siblings. The relevant variables in this constructed longitudinal dataset are detailed in Table 5 and Table 6. The variables available for the unlinked adopted individuals and unlinked siblings are given in the first two columns. Variables for the linked adoptees and linked siblings are given in the third and fourth columns, respectively.

One concern with this approach to creating a longitudinal dataset by linking individuals across multiple censuses is that those individuals successfully linked will not be representative of the general population. Given the relatively low rates of finding a unique, good match to the 1940 census, this could introduce a potentially large bias. In order to assess the extent of the problem, Table 7 presents summary statistics of the adopted individuals by the outcome of the linking process. Summary statistics for all children under the age of 20 in the Integrated Public Use Microsample (IPUMS) one percent sample of the 1910 federal census are also provided to compare both the linked and unlinked adopted individuals to the general population.

Several differences between adopted individuals and the general population stand out from Table 7. A far greater percentage of adopted children are black as compared to the general population. This is also the one dimension along which the linking procedure

leads to noticeable selection problems, with black individuals far less likely to be successfully linked to the 1940 census than white individuals. This is consistent with the literature on census enumeration in the early twentieth centuries that cites issues with age misreporting being particularly prevalent in the census returns and other official documents of black individuals (Hambricht, 1969; Elo et al., 1996; Rosenwaike & Hill, 1996; Elo & Preston, 1994). Adopted individuals are more likely to come from rural areas than the general population and are more geographically mobile. Neither of these dimensions has a significant effect on the likelihood of an adopted individual being linked to the 1940 census. Overall, aside from racial composition, the set of linked adopted individuals appears quite similar to the set of unlinked adoptees helping reduce concerns of selection bias arising from the linking procedure.

5 Empirical Approach and Preliminary Results Using the Longitudinal Data

With the longitudinal dataset constructed, it is possible to directly compare the adult outcomes of adopted individuals in terms of educational attainment, occupational status, annual income and family structure to the outcomes of their siblings. Regressing these outcomes on individual characteristics and including a family fixed effect offers a way to control for household characteristics, both observed and unobserved, identifying the impact of being adopted on adult outcomes relative to individuals raised in the exact same environment but who are the biological children of the parents.

While this approach offers an unprecedented opportunity to investigate the outcomes of adopted children relative to their siblings in a historical setting, it has two drawbacks. The first is the sample size problem faced by modern adoption studies. While the use of the complete federal census returns provides an initial sample of adopted individuals that dwarfs the samples used in modern studies, the pairs of adopted individuals and their siblings that can be linked to the 1940 census number only in the hundreds. When including

family fixed effects, it is difficult to obtain precise estimates of the effect of being adopted, a problem exacerbated by typically only observing two individuals per family. The second drawback is that the use of family fixed effects prevents looking at systematic differences between the outcomes of individuals from families containing an adopted children, whether those individuals are themselves adopted or the biological children of the parents, and children from traditional families.

To address this issue of how adopted children and their siblings differ from the population, we also consider a second empirical approach. We combine the data on the adopted individuals with data on males between the ages of 30 and 59 from the IPUMS one percent sample of the 1940 federal census. Combining the adoptee sample with the IPUMS sample provides a large comparison group of children raised in traditional families but sacrifices the ability to control for childhood characteristics; these variables are not available for the IPUMS sample.⁸ When using these combined samples, we include dummy variables for being adopted or being the sibling of an adopted child to capture the effects of adoption on adult outcomes.

Table 8 presents summary statistics for the adopted individuals who were linked to the 1940 federal census, the successfully linked siblings, and the sample of 30 to 59 year old males from the IPUMS one percent sample of the 1940 census. The summary statistics reveal that adopted children and their siblings look fairly similar as adults while both groups differ from the general population along several important dimensions. Adopted children and their siblings tend to be more geographically mobile, have higher rates of marriage and have larger households than the general population. When looking at the differences in means between the adopted individuals and their siblings, adopted children have fewer years of schooling on average but higher incomes and a greater number of weeks worked in the previous year.

To assess whether these differences between adopted individuals, their siblings and the

⁸In ongoing data collection efforts, we are linking a random sample of children living with their biological parents from the 1910 federal census to the 1940 federal census to provide a comparison group that does have childhood household information.

general population persist after controlling for individual characteristics, we run linear regressions for a variety of outcomes on a dummy variable for being adopted, a dummy variable for being the sibling of an adopted individual, and a set of individual characteristics including whether the individual lives in an urban area, whether the individual moved across states, race, and a quadratic in age. All regressions also include state fixed effects, an important set of controls given that the geographic distribution of adoptees, as shown in Table 7 and Table 8, differs from the general population with adopted individuals being more likely to reside in the South Atlantic states and less likely to live in the Middle Atlantic states than biological children.

The results of these regressions are presented in Table 14 for educational and labor market outcomes and Table 15 for outcomes related to household characteristics and family structure. When focusing on the educational and labor market outcomes, the various individual characteristics are typically highly statistically significant and have the expected signs. Urban individuals have higher educational attainments and higher incomes than rural individuals, white individuals have higher incomes and levels of schooling than black individuals. More geographically mobile individuals tend to have better outcomes than less geographically mobile individuals as demonstrated by the large, positive coefficients on the dummy variable for moving across states. What are of main interest, though, are the coefficients on the indicator variables for being adopted or being the sibling of an adopted individual. We find that adoption does have a negative impact on adult outcomes. Adopted individuals earn seven percent less in annual income relative to general population and work over an hour less per week. While not statistically significant, the negative coefficients on the adopted variable for years of schooling and weeks worked are also consistent with adopted children having worse outcomes than their biological counterparts.

While these negative impacts of adoption on adult outcomes are consistent with the modern literature and with Moriguchi's analysis of school attendance rates for adopted children in the early twentieth century, it is still possible that the effects are being driven by negative selection of adoptive households. If the type of household that tends to adopt

children differs systematically from households with traditional family structures then the negative consequences of adoptive status on adult outcomes may be a product of these differences in childhood household characteristics rather than being specific to the experience of being adopted. One way to assess this possibility is to include the linked siblings in the regressions in Table 14 with an indicator variable for being the sibling of an adopted child. If it is the childhood household characteristics driving the result for adopted children, similar effects should be present for their siblings. We find that this is not the case. While the coefficients on the sibling indicator are negative for income, weeks worked and hours worked, suggesting that siblings of adoptees also fare worse than the general population, the coefficients are statistically insignificant and, in the case of the weeks worked and hours worked variables, far smaller in magnitude than the coefficients for the adopted indicator variable. More telling is the coefficient on for the sibling indicator in the years of schooling regression. The coefficient on the sibling indicator is statistically significant and has the opposite sign of the adopted indicator. This suggests that in terms of schooling, it is not simply that adopted children were disadvantaged relative to the general population, they were disadvantaged relative to their siblings as well.

When turning to the household characteristics and family structure regressions in Table 15, we find once again that adopted children have adult outcomes that differ substantially from the general population.⁹ Adopted children tended to have larger households as adults, were more likely to be married and were more likely move across states. Once again, these effects could be the product of the adoptive household's characteristics rather than the marginal effect of being adopted. Including the adopted children's siblings helps disentangle these two effects. We find that the siblings of adopted children also tended to have larger families and were more geographically mobile than the general population but were less likely to be married, not only relative to their adopted siblings but also to the

⁹Note that while it may seem sensible to use a model more suited to binary incomes for the marriage variable and the moved across states variable, we use a linear probability model so that the inclusion of family fixed effects will be feasible. Nonlinear models would suffer from the incidental parameters problem when including the family fixed effects. The use of the linear probability model for the binary outcome variables allows us to keep estimates comparable between the non-family fixed effects and family fixed effects regressions.

general population.

In an attempt to better control for the characteristics of adoptive households and further explore these differences between adopted children and their siblings, we re-estimated the models with the inclusion of family fixed effects. The results are presented in Table 16. Unfortunately, given the small number of families for which both the adopted individual and at least one sibling could be linked to the 1940 federal census, the estimates suffer from large standard errors making it difficult to interpret the results. We are hopeful that the next wave of data collection for this project will provide sufficient sample sizes to obtain precise, meaningful results from the family fixed effects regressions.

6 Direction of Future Data Collection

The preliminary results presented in the previous section offer hope that our approach of constructing longitudinal data from historical censuses will shed new light on the relationship between adoption and adult outcomes in the early twentieth century. Already, the preliminary sample has demonstrated that adopted individuals do have significantly worse labor market outcomes than the general population and make different decisions about marriage, household size and household location. However, with the expansion of our dataset, we hope to provide an even more detailed picture of adoption and its consequences in the early twentieth century.

There are three areas in which we are expanding our data collection efforts: (1) increasing the number of linked adoptee siblings, (2) creating a linked sample of children from traditional families, and (3) transcribing additional data from images of the original 1910 census manuscript pages. In this section, we will briefly describe each of these phases of the project and discuss the ways in which the additional data will sharpen our estimates and allow us to ask more nuanced questions about the effects of adoption.

Expanding the number of linked siblings of adopted children will allow us to obtain better estimates of how biological children from adoptive families differ from the general

population. The preliminary sample of linked siblings was created by linking the siblings of those adopted individuals who had been successfully linked to the 1940 federal census. We are extending our data collection efforts to also link siblings for the adopted individuals who could not be linked to the 1940 federal census (the information on siblings needed to link them is taken from the 1910 census and therefore available regardless of whether the adopted individual can be linked to 1940). If the rates of successfully linking these siblings are comparable to the rates experienced in the preliminary sample, we should be able to add roughly 2,500 linked siblings to the the sample even if only looking at the sibling closest in age to the adopted individual. Even larger sample sizes will be possible if we expand the data collection efforts to link the second and third closest siblings. When thinking about how large a sample of several thousand siblings of adopted children observed both in childhood and as adults is, it is worth remembering that the best modern longitudinal studies have adoptive families numbering in the hundreds, with many far below even that. This approach of linking historical censuses yields sample sizes that are at least an order of magnitude greater than modern longitudinal surveys.

The second ongoing data collection effort is the construction of a linked sample of biological children from the 1910 federal census to the 1940 federal census. This linking procedure will begin with a random sample of children from the 1910 federal census and will follow the same steps as the linking of the adopted children. This sample of linked biological children will serve two purposes. The first is to better address concerns of sample selection arising from the linking procedure. Having a linked sample of biological children will allow us run regressions comparable to those in Table 14 and Table 15 using the sample of linked biological children rather than the IPUMS sample of the general population. This will provide a general population comparison group for the adopted children that suffers from the same sample selection biases as the adoptee sample, helping ensure that the coefficient on the adopted indicator variable is picking up the effects of being adopted rather than the influence of characteristics correlated with the probability of being successfully linked to the 1940 census. The second reason for creating a sample of linked biological children is that it

will allow for us to control for childhood household characteristics when including biological children in regressions. These childhood household characteristics are not available from the 1940 IPUMS sample but are crucial for allowing us to properly control for the influence of systematic differences between adoptive and traditional households.

This desire to better control for childhood household characteristics leads to the final current data collection effort, the transcription of additional information from the 1910 census manuscript pages. While the complete information from the 1940 federal census has been digitized, only the variables relevant to genealogical research have been digitized for the complete 1910 federal census. However, digital images of the original census manuscript pages are available and can be used to obtain additional details about the individual and household characteristics. In particular, we will obtain information on literacy and school attendance of each child in the 1910 household, the occupations of all household members, years married for the parents, and the number of children ever born and children surviving for the mother. Transcribing these details will give us a far richer set of controls for all of the regressions.

As an example of the usefulness of these additional controls, consider one household characteristic that is available from the digitized version of the 1910 census, the household location. We have used the IPUMS sample of the 1910 census to determine the percentage of residents living in urban areas in each county and used this to construct an indicator variable for whether an individual lives in an urban county in 1910, defined as a county in which greater than 25 percent of its residents live in an urban area. One could imagine that the causes and consequences of adoption will differ substantially between rural and urban areas. An additional household member may provide a useful source of labor on a farm, the marginal costs of increasing family size will be substantially lower in rural rather than urban areas, and so on. To assess whether the effects of adoption on adult outcomes differ by whether the adoptive household was located in an urban or rural area, we modified the basic regressions to include an interaction term between the adopted indicator variable and the indicator for growing up in an urban household.

The results when including these interactions are presented in Table 17 and offer evidence that the consequences of being adopted do vary with the conditions under which the adopted child grew up. We find that the schooling advantage enjoyed by adopted individuals siblings' is present for rural adoptive households but disappears for urban households. Additionally, the significantly lower earnings associated with being adopted appear to be specific to adopted individuals growing up in rural households. These preliminary results suggest that the effects of being adopted on adult outcomes and perhaps the motivations behind adoption and resource allocation decisions of households may vary across rural and urban households. Transcribing the additional information available in the 1910 federal census will provide potentially compelling results about how adoptees fared across different household types, locations and labor markets, information that is crucial to explaining why families chose to adopt and the channels through which adoption impacted adult outcomes.

7 Conclusion

Growing up in a non-traditional family structure has significant consequences for the outcomes of children. A large literature points to worse outcomes in terms behavioral problems, schooling outcomes, and labor market outcomes for children not raised by their biological parents. However, these non-traditional family structures are not a modern phenomenon. Historically, there have been a large number of children raised by stepparents or adoptive parents yet we have little quantitative evidence of how these children fared. A lack of historical data on adoptive families and the outcomes of adopted children has left many open questions regarding the motivations behind adoption, the ways in which adoptive parents allocated resources between adopted and biological children, and the adult outcomes of those children. In this study, we have constructed a historical longitudinal dataset of adopted children and their siblings to shed light on these questions. Given the availability of complete historical federal census records, we are able to track thousands of adopted children and their siblings from their childhood households in 1910 to their adult households

in 1940, leading to a dataset that rivals modern longitudinal studies in terms of sample size and the scope of its variables.

Our preliminary analysis of this new dataset reveals that adopted individuals do differ from their peers as adults. They worked less, earned lower incomes and had lower levels of educational attainment. They also differed in terms of their household characteristics. Adopted individuals tended to have higher rates of marriage, larger families and greater geographic mobility as adults. A unique feature of our new dataset is the ability to also track these adopted individual's siblings into adulthood. We find that the siblings of adopted individuals tended to have greater educational attainment, lower rates of marriage and higher geographic mobility. The differences in the educational attainments and marriage patterns of adopted individuals and the siblings of adopted individuals suggest that the observed effects of being adopted are not simply the product of the characteristics of the households that tended to adopt children. Instead, they are product of forces that differentially affected the adopted individual relative to his siblings whether through the direct impact of adoption or the resource allocation decisions of the adoptive household.

As we continue to expand the linked dataset, we will be able to address more nuanced questions related to historical adoption. We will be able to control for a broader range of individual and childhood household characteristics and test whether the effects of adoption differed depending on the type of household and community in which the adopted child grew up. This will provide critical insights into why families chose to adopt and how families allocated household resources in the first half of the twentieth century.

References

- Askeland, L. (2006). *Children and youth in adoption, orphanages, and foster care: A historical handbook and guide*. Greenwood Publishing Group.
- Bonham, G. S. (1977). Who adopts: The relationship of adoption and social-demographic characteristics of women. *Journal of Marriage and the Family*, (pp. 295–306).
- Brodzinsky, D. M. (1987). Adjustment to adoption: A psychosocial perspective. *Clinical*

- Psychology Review*, 7(1), 25–47.
- Brodzinsky, D. M. (1993). Long-term outcomes in adoption. *Adoption*, 3(1), 153–166.
- Carp, E. W. (2009). *Adoption in America: Historical Perspectives*. University of Michigan Press.
- Case, A., Lin, I.-F., & McLanahan, S. (1999). Household resource allocation in stepfamilies: Darwin reflects on the plight of cinderella. *American Economic Review*, (pp. 234–238).
- Case, A., Lin, I.-F., & McLanahan, S. (2000). How hungry is the selfish gene? *The Economic Journal*, 110(466), 781–804.
- Case, A., Lin, I.-F., & McLanahan, S. (2001). Educational attainment of siblings in stepfamilies. *Evolution and human behavior*, 22(4), 269–289.
- Case, A., & Paxson, C. (2001). Mothers and others: who invests in childrens health? *Journal of health economics*, 20(3), 301–328.
- Dickson, L. R., Heffron, W. M., & Parker, C. (1990). Children from disrupted and adoptive homes on an inpatient unit. *American Journal of Orthopsychiatry*, 60(4), 594–602.
- Elo, I. T., & Preston, S. H. (1994). Estimating african-american mortality from inaccurate data. *Demography*, 31(3), 427–458.
- Elo, I. T., Preston, S. H., Rosenwaike, I., Hill, M., & Cheney, T. P. (1996). Consistency of age reporting on death certificates and social security records among elderly african americans. *Social Science Research*, 25(3), 292–307.
- Goldin, C. D., & Katz, L. F. (2009). *The race between education and technology*. Harvard University Press.
- Gordon, L. (1994). *Pitied but not entitled: Single mothers and the history of welfare, 1890-1935*. Free Press New York.
- Hambright, T. Z. (1969). Comparison of information on death certificates and matching 1960 census records: age, marital status, race, nativity and country of origin. *Demography*, 6(4), 413–423.
- Heisterman, C. A. (1935). A summary of legislation on adoption. *The Social Service Review*, 9(2), 269–293.
- Hersov, L. (1990). The seventh jack tizard memorial lecture* aspects of adoption. *Journal of child Psychology and Psychiatry*, 31(4), 493–510.
- Katz, M. B. (1996). *In the shadow of the poorhouse: A social history of welfare in America*. Basic Books.
- Maza, P. (1984). Adoption trends: 1944-1975. *Child welfare research notes*, 9, 1–11.

- Moriguchi, C. (2013). Adopted children and stepchildren in twentieth century america: Evidence from federal census microdata, 1880-1930 and 2000. *Working Paper*.
- Piersma, H. L. (1987). Adopted children and inpatient psychiatric treatment: A retrospective study. *Psychiatric Hospital*.
- Rogeness, G. A., Hoppe, S. K., Macedo, C. A., Fischer, C., & Harris, W. R. (1988). Psychopathology in hospitalized, adopted children. *Journal of the American Academy of Child & Adolescent Psychiatry*, 27(5), 628–631.
- Rosenwaike, I., & Hill, M. E. (1996). The accuracy of age reporting among elderly african americans evidence of a birth registration effect. *Research on Aging*, 18(3), 310–324.
- Sacerdote, B. (2000). The nature and nurture of economic outcomes. *NBER Working Paper*, (w7949).
- Sacerdote, B. (2002). The nature and nurture of economic outcomes. *American Economic Review*, 92(2), 344–348.
- Warren, S. B. (1992). Lower threshold for referral for psychiatric treatment for adopted adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, 31(3), 512–517.
- Zarefsky, J. (1946). Children acquire new parents. *The Child*, 10, 142–4.
- Zill, N. (1985). Behavior and learning problems among adopted children: Findings from a us national survey of child health. *Child Trends, Inc., Washington, DC*.

8 Figures and Tables

Table 1: Distribution of two-parent households by child type, 1900-1930.

Year:	1900	1910	1920	1930
	<u>White households</u>			
Total number of households	7,848,520	9,624,013	11,534,399	13,352,190
Number of households by type of children:				
Biological only	7,644,560	9,352,992	11,297,632	12,954,765
Adopted only	36,480	47,997	23,210	53,934
Biological and adopted	15,120	14,611	11,911	18,685
Stepchildren only	76,440	112,854	104,961	175,235
Biological and stepchildren	75,440	94,746	96,180	147,258
Adopted and stepchildren	200	597	404	707
Biological, adopted and stepchildren	280	216	101	606
	<u>Black households</u>			
Total number of households	838,240	973,781	1,031,611	13,351,190
Number of households by type of children:				
Biological only	778,360	894,298	968,742	980,205
Adopted only	8,360	13,063	9,085	17,574
Biological and adopted	3,600	4,991	3,735	6,060
Stepchildren only	26,440	36,515	30,672	43,531
Biological and stepchildren	21,280	24,341	18,975	20,200
Adopted and stepchildren	80	489	202	404
Biological, adopted and stepchildren	120	84	200	505

Source: IPUMS 1900 2.5% sample, 1910 1.4% sample, 1920 1% sample, and 1930 1% sample.

Notes: Numbers given are estimates for the total population based on the IPUMS samples. The race of a household is defined by the race of a household head. Children are defined as any person under age 18 residing in a household whose relationship with household head is recorded as 'child,' including biological, adopted and stepchildren. Only white and black children are included. Only households with both married parents present are included. Children with not well-identified mother or father are excluded. Households in Alaska and Hawaii are excluded.

Table 2: OLS analysis of childhood educational status, 1900-1930, white male children.

Dependent Variable	In primary school (1=yes)		In secondary school (1=yes)		Literate (1=yes)	
	White males, age 6-11		White males, age 12-17		White males, age 10-17	
Sample:						
Including family fixed effects:	no	yes	no	yes	no	yes
	(1)	(2)	(3)	(4)	(5)	(6)
Adopted (1=yes)	-0.0391*** (0.0133)	-0.0957 (0.0634)	-0.0542*** (0.0146)	-0.0688 (0.0826)	-0.0151*** (0.0058)	-0.0100 (0.0200)
Age	0.3419*** (0.0047)	0.2987*** (0.0068)	0.3973*** (0.0090)	0.4203*** (0.0150)	0.0277*** (0.0016)	0.0251*** (0.0020)
Age Squared	-0.0166*** (0.0003)	-0.0149*** (0.0004)	-0.0167*** (0.0003)	-0.0170*** (0.0005)	-0.0009*** (0.0001)	-0.0009*** (0.0001)
Birth Order	0.0015* (0.0009)	-0.0265*** (0.0034)	0.0193*** (0.0013)	0.0537*** (0.0046)	-0.0001 (0.0005)	-0.0039*** (0.0010)
No. of Siblings	0.0014** (0.0006)		-0.0111*** (0.0005)		-0.0005** (0.0002)	
Father literate (1=yes)	0.0446*** (0.0038)		0.0634*** (0.0044)		0.0705*** (0.0029)	
Mother literate (1=yes)	0.0559*** (0.0037)		0.0615*** (0.0042)		0.1051*** (0.0028)	
Father's SEI	0.0040*** (0.0004)		0.0168*** (0.0004)		0.0003*** (0.0001)	
Mother's SEI	0.0059*** (0.0017)		0.0106*** (0.0019)		0.0038*** (0.0007)	
Father working (1=yes)	0.0196*** (0.0069)		-0.0344*** (0.0065)		0.0006 (0.0023)	
Mother working (1=yes)	-0.0227*** (0.0063)		-0.0424*** (0.0071)		-0.0198*** (0.0034)	
House owned (1=yes)	0.0222*** (0.0015)		0.0443*** (0.0018)		0.0135*** (0.0007)	
Mean Probability	0.8354		0.7918		0.9256	
No. of Observations	255,797	256,038	206,656	206,909	286,511	286,848
Adjusted R-squared	0.1942	0.1804	0.2188	0.2763	0.1526	0.0077

The sample consists of white male children residing in married two-parent households in IPUMS 1900 2.5%, 1910 1.4%, 1920 1%, and 1930 1% samples. In (1), (3), and (5), independent variables include a constant and fixed effects for year, state, and their interactions. We also control for the number of relatives and non-relatives in HH, the age, race, nativity and employment status of father and mother, a female-headed indicator, and urban, metropolitan and farm indicators. Standard errors are clustered at HH level. In (2), (4), and (6), household fixed effects are included. Standard errors are reported in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 3: OLS analysis of school attendance and labor force status, 1900-1930, white male children.

Dependent variable: Including family fixed effects:	In school and not in labor		In school and in labor force		Not in school and in labor		Not in school and not in labor force (1=yes)	
	no (1)	yes (2)	no (3)	yes (4)	no (5)	yes (6)	no (7)	yes (8)
Adopted (1=yes)	-0.0548*** (0.0123)	-0.1075** (0.0547)	-0.0015 (0.0039)	0.0155 (0.0228)	0.0132** (0.0064)	-0.0480 (0.0402)	0.0431*** (0.0114)	0.1400*** (0.0459)
Age	0.3306*** (0.0038)	0.3564*** (0.0055)	0.0030* (0.0016)	0.0020 (0.0023)	-0.1694*** (0.0025)	-0.1928*** (0.0041)	-0.1643*** (0.0033)	-0.1657*** (0.0046)
Age Squared	-0.0146*** (0.0001)	-0.0153*** (0.0002)	0.0000 (0.0001)	0.0001 (0.0001)	0.0074*** (0.0001)	0.0081*** (0.0001)	0.0072*** (0.0001)	0.0071*** (0.0002)
Birth Order	0.0167*** (0.0010)	0.0358*** (0.0027)	-0.0015*** (0.0005)	-0.0018 (0.0011)	-0.0048*** (0.0005)	-0.0176*** (0.0020)	-0.0104*** (0.0008)	-0.0164*** (0.0023)
No. of Siblings	-0.0111*** (0.0005)	(0.0027)	0.0027*** (0.0002)	(0.0011)	0.0065*** (0.0003)	(0.0020)	0.0018*** (0.0004)	(0.0023)
Father literate (1=yes)	0.0670*** (0.0039)	(0.0027)	-0.0112*** (0.0021)	(0.0011)	-0.0163*** (0.0023)	(0.0020)	-0.0396*** (0.0034)	(0.0023)
Mother literate (1=yes)	0.0644*** (0.0036)	(0.0027)	-0.0062*** (0.0019)	(0.0011)	-0.0233*** (0.0023)	(0.0020)	-0.0350*** (0.0032)	(0.0023)
Father's SEI	0.0138*** (0.0004)	(0.0027)	-0.0012*** (0.0001)	(0.0011)	-0.0081*** (0.0002)	(0.0020)	-0.0045*** (0.0003)	(0.0023)
Mother's SEI	0.0287*** (0.0017)	(0.0027)	-0.0198*** (0.0011)	(0.0011)	-0.0090*** (0.0012)	(0.0020)	0.0000 (0.0012)	0.0000 (0.0012)
Mean Probability	0.8083		0.0261		0.058		0.1076	
No. of Observations	286,511	286,848	286,511	286,848	286,511	286,848	286,511	286,848
Adjusted R-squared	0.2269	0.2862	0.0746	0.0107	0.1569	0.1469	0.0964	0.1064

The sample consists of white male children residing in married two-parent households in IPUMS 1900 2.5%, 1910 1.4%, 1920 1%, and 1930 1% samples. In (1), (3), (5), and (7), independent variables include a constant and fixed effects for year, state, and their interactions. We also control for the number of relatives and non-relatives in HH, the age, race, employment status and nativity of father and mother, a female-headed indicator, and urban, metropolitan and farm indicators. Standard errors are clustered at HH level. In (2), (4), (6), and (8), household fixed effects are included. Standard errors are reported in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 4: Number of adopted individuals in the federal census returns by census year.

Census year	Number of adopted individuals	Total Population
1850	10	23,191,876
1860	10	31,443,321
1870	27	38,588,371
1880	5,774	50,155,783
1890	13	62,947,714
1900	101,764	75,994,575
1910	128,755	91,972,266
1920	88,416	105,710,620
1930	173,485	122,775,046
1940	55,220	131,669,275

Notes: Number of adopted individuals is given by the number of census records returned when searching for the word "adopted" in Ancestry.com's index of the complete census returns. Many of the 1890 census records have been destroyed so the 1890 number is based on only a fragment of the census returns. All of the other years are based on the complete census returns. The total population figures are based on the original counts for census dates and are taken from Table Aa1-5 of the Historical Statistics of the United States.

Table 5: Variables available by variable type and sample.

	Unlinked adopted individuals	Unlinked siblings of adopted individuals	Linked adopted individuals	Linked siblings of adopted individuals	1910 general population (IPUMS sample)	1940 general population (IPUMS sample)
<u>General Characteristics</u>						
Year of birth	X	X	X	X	X	X
Gender	X	X	X	X	X	X
Race	X	X	X	X	X	X
Mother's birth state	X	X	X	X	X	X
Father's birth state	X	X	X	X	X	X
Mother's birth state	X	X	X	X	X	X
<u>Childhood individual characteristics</u>						
Able to read	X	X	X	X	X	
Able to write	X	X	X	X	X	
Attended school in past year	X	X	X	X	X	
<u>Childhood household characteristics</u>						
Township, county and state of residence	X	X	X	X	X	
Urban/rural status	X	X	X	X	X	
Farm status	X	X	X	X	X	
Home owned, mortgaged or rented	X	X	X	X	X	
Years of marriage for parents	X	X	X	X	X	
Number of children ever born for mother	X	X	X	X	X	
Number of children surviving for mother	X	X	X	X	X	
Ages, genders and relationship to head of household for all household members						
Parents' occupations	X	X	X	X	X	
Parents' industries	X	X	X	X	X	
Parents' employment status	X	X	X	X	X	
Literacy of all household members	X	X	X	X	X	

Table 6: Variables available by variable type and sample, continued.

	Unlinked adopted individuals	Unlinked siblings of adopted individuals	Linked adopted individuals	Linked siblings of adopted individuals	1910 general population (IPUMS sample)	1940 general population (IPUMS sample)
<u>Adult individual characteristics</u>						
Marital status			X	X	X	X
Attended school in the past year			X	X	X	X
Highest grade of school completed			X	X	X	X
Occupation			X	X	X	X
Industry			X	X	X	X
Employment status			X	X	X	X
Duration of unemployment			X	X	X	X
Weeks worked in past year			X	X	X	X
Hours worked in last week			X	X	X	X
Income earned in past year			X	X	X	X
Veteran status			X	X	X	X
<u>Adult household characteristics</u>						
Town, county and state of residence			X	X	X	X
Home owned or rented			X	X	X	X
Value of home if owned, monthly rent if rented			X	X	X	X
Urban/rural status			X	X	X	X
Farm status			X	X	X	X
Ages, genders and relationship to head of household for all household members			X	X	X	X

Table 7: Characteristics of children in 1910 by adoption status and linking outcome.

Variable	Adopted children successfully linked to 1940 census	Adopted children who could not be linked	IPUMS 1% sample of the 1910 census
Age	12.38 (6.23)	12.08 (6.43)	9.14 (5.81)
White (1=yes)	0.71 (0.46)	0.59 (0.49)	0.87 (0.33)
Number of household members	7.00 (2.80)	7.22 (2.94)	6.97 (19.09)
Percentage of county that is urban	0.33 (0.32)	0.30 (0.32)	0.41 (0.36)
Child has moved across states (1=yes)	0.19 (0.39)	0.18 (0.39)	0.11 (0.31)
Mother has moved across states (1=yes)	0.23 (0.42)	0.21 (0.41)	0.28 (0.45)
Father has moved across states (1=yes)	0.24 (0.43)	0.22 (0.42)	0.31 (0.46)
<u>Percentage living in:</u>			
New England	6.22	4.77	6.08
Middle Atlantic	11.79	11.13	19.00
East North Central	15.83	12.18	18.35
West North Central	15.27	11.81	12.78
South Atlantic	21.08	25.62	15.10
East South Central	10.26	13.13	10.56
West South Central	11.59	15.01	11.42
Mountain	4.08	3.26	2.73
Pacific	3.88	3.08	3.93
Number of observations	2,511	12,518	194,987

Notes: Standard deviations given in parentheses. Urban percentage is defined as the percentage of individuals in a county designated as living in an urban area in the IPUMS 1% sample. All samples are restricted to males. The IPUMS 1% sample is restricted to children under the age of 20. Individuals are defined as moving across states if the state of residence in 1910 is different than the birth state given in the census.

Table 8: Characteristics of adult adoptees and adoptee siblings in 1940.

Variable	Adoptees	Siblings of adoptees	IPUMS 1% sample of the 1940 census
Age	42.26 (6.30)	41.47 (7.55)	37.33 (11.24)
White (1=yes)	0.71 (0.46)	0.76 (0.43)	0.90 (0.29)
Number of household members	4.85 (3.00)	4.84 (2.95)	4.27 (2.34)
Percentage of county that is urban	0.52 (0.33)	0.51 (0.33)	0.56 (0.33)
Moved across states (1=yes)	0.39 (0.49)	0.42 (0.49)	0.29 (0.46)
Years of schooling	7.97 (3.35)	8.71 (3.23)	8.77 (3.65)
Annual income (1940 dollars)	842.58 (949.17)	838.54 (1003.76)	830.88 (928.29)
Hours worked in past week	44.97 (16.42)	45.10 (15.69)	36.01 (22.47)
Weeks worked in past year	41.66 (16.40)	39.77 (18.07)	39.70 (17.71)
Married (1=yes)	0.85 (0.36)	0.72 (0.45)	0.66 (0.47)
<u>Percentage living in:</u>			
New England	5.48	6.11	6.43
Middle Atlantic	15.06	15.84	22.20
East North Central	18.37	17.96	21.00
West North Central	12.15	16.08	10.10
South Atlantic	17.84	15.46	12.82
East South Central	8.18	7.48	7.41
West South Central	9.94	8.60	9.27
Mountain	3.81	3.49	2.99
Pacific	9.17	8.98	7.78
Number of observations	2,511	818	367,425

Notes: Standard deviations given in parentheses. Urban percentage is defined as the percentage of individuals in the county designated as living in an urban area in the IPUMS 1% sample. All samples are restricted to males. The IPUMS 1% sample is restricted to men between the ages of 20 and 59. Individuals are defined as moving across states if the state of residence in 1910 is different than the birth state given in the census.

Table 9: Geographic mobility patterns of adopted children who moved across states between 1910 and 1940.

<u>Five states with the most adoptees exiting the state</u>	
State	Percentage of adoptees who moved living in state in 1910
Pennsylvania	5.13
Virginia	4.93
South Carolina	4.34
Georgia	4.04
Illinois	3.94
<u>Five states with the most adoptees entering the state</u>	
State	Percentage of adoptees who moved living in the state in 1940
California	12.13
New York	8.09
Illinois	7.79
Pennsylvania	6.21
Ohio	4.64

Table 10: OLS analysis of education and labor market outcomes.

Dependent variable:	Years of schooling		Log of annual income		Weeks worked in past year		Hours worked in past week	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adopted (1=yes)	-0.09 (0.11)	-0.09 (0.11)	-0.07*** (0.02)	-0.07*** (0.02)	-0.30 (0.24)	-0.30 (0.24)	-1.19*** (0.37)	-1.19*** (0.37)
Adoptee sibling (1=yes)		0.44*** (0.13)		-0.05 (0.05)		-0.21 (0.48)		-0.94 (0.61)
Urban county in 1940 (1=yes)	0.95*** (0.07)	0.95*** (0.07)	0.47*** (0.03)	0.47*** (0.03)	0.35 (0.23)	0.34 (0.23)	-1.88*** (0.60)	-1.89*** (0.60)
Black (1=yes)	-3.16*** (0.15)	-3.15*** (0.15)	-0.68*** (0.02)	-0.68*** (0.03)	-2.40*** (0.24)	-2.39*** (0.25)	-1.04*** (0.30)	-1.04*** (0.28)
Age in 1940	-0.08*** (0.01)	-0.08*** (0.01)	0.142*** (0.004)	0.142*** (0.004)	1.09*** (0.05)	1.09*** (0.05)	0.20*** (0.02)	0.20*** (0.02)
Age in 1940 squared	0.0001** (0.0001)	0.0001** (0.0001)	-0.002*** (0.000)	-0.002*** (0.000)	-0.01*** (0.001)	-0.01*** (0.001)	-0.002*** (0.0003)	-0.002*** (0.0003)
Moved across states (1=yes)	0.60*** (0.09)	0.60*** (0.09)	0.16*** (0.03)	0.16*** (0.02)	-0.09 (0.15)	-0.09 (0.15)	-0.61*** (0.18)	-0.61*** (0.18)
Number of observations	322,406	323,155	230,230	230,723	291,459	292,097	251,583	252,144
Adjusted R-squared	0.21	0.21	0.23	0.23	0.04	0.04	0.05	0.05

Notes: Standard errors clustered by residence state given in parentheses. All regressions include residence state fixed effects. Adoptee siblings are excluded from the regression sample in regressions (1), (3), (5) and (7). Samples for the weeks worked and hours worked regressions are restricted to individuals reporting positive values. * significant at 10%, ** significant at 5%, *** significant at 1%

Table 11: OLS analysis of adult household characteristics.

Dependent variable:	Number of household members			Moved across states (1=yes)		
	(1)	(2)	(3)	(4)	(5)	(6)
Adopted (1=yes)	0.66*** (0.08)	0.66*** (0.08)	0.06*** (0.01)	0.06*** (0.009)	0.03*** (0.01)	0.03*** (0.01)
Adoptee sibling (1=yes)		0.80*** (0.18)		-0.06*** (0.02)		0.08*** (0.03)
Urban county in 1940 (1=yes)	-0.40*** (0.03)	-0.40*** (0.03)	-0.013*** (0.005)	-0.013*** (0.005)	0.09*** (0.01)	0.09*** (0.01)
Black (1=yes)	0.23*** (0.06)	0.23*** (0.06)	-0.08*** (0.01)	-0.08*** (0.01)	0.14** (0.06)	0.14** (0.06)
Age in 1940	-0.06*** (0.01)	-0.06*** (0.01)	0.10*** (0.002)	0.10*** (0.002)	0.015*** (0.001)	0.015*** (0.001)
Age in 1940 squared	0.0006*** (0.0001)	0.0006*** (0.0001)	-0.001*** (0.00002)	-0.001*** (0.00002)	-0.0001*** (0.00002)	-0.0001*** (0.00002)
Moved across states (1=yes)	-0.37*** (0.04)	-0.37*** (0.04)	0.008* (0.004)	0.007* (0.004)		
Number of observations	322,474	322,835	322,474	322,835	322,474	320,563
Adjusted R-squared	0.04	0.04	0.18	0.18	0.16	0.16

Notes: Standard errors clustered by residence state given in parentheses. All regressions include residence state fixed effects. Adoptee siblings are excluded from the regression sample in regressions (1), (3) and (5). Moved across states equals one if the residence state in 1940 is different than the individual's birth state. * significant at 10%, ** significant at 5%, *** significant at 1%

Table 12: OLS analysis of adult outcomes with family fixed effects.

Dependent variable:	Years of	Log of annual	Weeks worked	Hours worked	Number of	Married	Moved across
	schooling	income	in past year	in past week	household	(1=yes)	states (1=yes)
	(1)	(2)	(3)	(4)	(5)	(6)	(6)
Adopted (1=yes)	-0.57* (0.30)	0.01 (0.13)	2.78 (2.30)	0.19 (2.00)	0.01 (0.32)	0.10** (0.04)	-0.02 (0.03)
Urban county in 1940 (1=yes)	1.12** (0.42)	0.39 (0.28)	-0.34 (3.09)	-5.34 (4.66)	-0.02 (0.43)	-0.02 (0.08)	0.29*** (0.08)
Age in 1940	-0.08 (0.29)	0.14 (0.13)	1.25 (1.75)	0.98 (2.22)	0.32 (0.23)	0.07 (0.05)	-0.04 (0.03)
Age in 1940 squared	-0.0001 (0.003)	-0.002 (0.001)	-0.02 (0.02)	-0.01 (0.03)	-0.004 (0.003)	-0.001 (0.001)	0.0005 (0.0003)
Moved across states (1=yes)	0.12 (0.36)	0.23 (0.25)	1.99 (2.65)	3.53 (3.34)	0.19 (0.48)	-0.04 (0.06)	
Number of observations	2,953	2,010	2,875	2,313	3,043	3,043	3,043
Adjusted R-squared	0.38	0.25	0.05	0.13	0.001	-0.14	0.34

Notes: Standard errors clustered by residence state given in parentheses. Regression sample includes all adoptees and adoptee siblings linked to the 1940 federal census. Moved across states equals one if the residence state in 1940 is different than the individual's birth state. * significant at 10%, ** significant at 5%, *** significant at 1%

Table 13: OLS analysis of adult outcomes with urban childhood household interactions.

Dependent variable: Including family fixed effects:	Years of schooling		Log of annual income	
	no (1)	yes (2)	no (3)	yes (4)
Adopted (1=yes)	-0.04 (0.14)	-0.78** (0.35)	-0.09*** (0.03)	0.03 (0.20)
Adopted x Urban household in 1910	-0.12 (0.16)	0.41 (0.45)	0.05 (0.04)	-0.03 (0.24)
Adoptee sibling (1=yes)	0.65*** (0.19)		-0.06 (0.06)	
Adoptee sibling x Urban household in 1910	-0.43* (0.26)		0.03 (0.08)	
Urban county in 1940 (1=yes)	0.95*** (0.07)	1.14*** (0.42)	0.47*** (0.03)	0.14 (0.13)
Black (1=yes)	-3.16*** (0.15)		-0.68*** (0.03)	
Age in 1940	-0.08*** (0.01)	-0.08 (0.29)	0.142*** (0.004)	0.14 (0.13)
Age in 1940 squared	0.0001** (0.0001)	-0.0001 (0.0033)	-0.002*** (0.000)	-0.002 (0.001)
Moved across states (1=yes)	0.60*** (0.09)	0.11 (0.36)	0.16*** (0.02)	0.24 (0.25)
Number of observations	323,155	2,953	230,723	2,010
Adjusted R-squared	0.21	0.38	0.23	0.24

Notes: Standard errors clustered by residence state given in parentheses. For columns (1) and (3), the regression sample includes all adoptees and adoptee siblings linked to the 1940 federal census and all males between the ages of 20 and 59 from the 1% IPUMS sample of the 1940 census. For columns (2) and (4), the regression sample includes all adoptees and adoptee siblings linked to the 1940 census. Moved across states equals one if the residence state in 1940 is different than the individual's birth state. Black is dropped from the family fixed effects regressions because there is no variation in the variable within families in the regression sample. * significant at 10%, ** significant at 5%, *** significant at 1%

Table 14: OLS analysis of education and labor market outcomes, sample restricted to white males.

Dependent variable:	Years of schooling		Log of annual income		Weeks worked in past year		Hours worked in past week	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adopted (1=yes)	-0.48*** (0.09)	-0.48*** (0.09)	-0.14*** (0.02)	-0.14*** (0.02)	-0.36 (0.22)	-0.36 (0.22)	-0.70* (0.41)	-0.70* (0.41)
Adoptee sibling (1=yes)		0.08 (0.14)	-0.08 (0.05)	-0.08 (0.05)	-0.62 (0.56)	-0.62 (0.56)	-1.20* (0.71)	-1.20* (0.71)
Urban county in 1940 (1=yes)	0.94*** (0.07)	0.94*** (0.07)	0.48*** (0.03)	0.48*** (0.03)	0.56** (0.23)	0.57** (0.23)	-2.20*** (0.62)	-2.21*** (0.62)
Age in 1940	-0.07*** (0.01)	-0.07*** (0.01)	0.15*** (0.003)	0.15*** (0.003)	1.14*** (0.05)	1.14*** (0.05)	0.21*** (0.02)	0.21*** (0.02)
Age in 1940 squared	0.0001 (0.0001)	0.0001 (0.0001)	-0.002*** (0.0000)	-0.002*** (0.0000)	-0.01*** (0.001)	-0.01*** (0.001)	-0.002*** (0.0003)	-0.002*** (0.0003)
Moved across states (1=yes)	0.62*** (0.10)	0.62*** (0.10)	0.15*** (0.02)	0.15*** (0.02)	-0.01 (0.16)	-0.01 (0.16)	-0.59*** (0.16)	-0.59*** (0.16)
Number of observations	289,199	289,769	207,510	207,888	262,036	262,519	226,689	227,296
Adjusted R-squared	0.11	0.11	0.18	0.18	0.04	0.04	0.06	0.06

Notes: Standard errors clustered by residence state given in parentheses. All regressions include residence state fixed effects. Adoptee siblings are excluded from the regression sample in regressions (1), (3), (5) and (7). Samples for the weeks worked and hours worked regressions are restricted to individuals reporting positive values. * significant at 10%, ** significant at 5%, *** significant at 1%

Table 15: OLS analysis of adult household characteristics, sample restricted to white males.

Dependent variable:	Number of household members			Moved across states (1=yes)		
	(1)	(2)	(3)	(4)	(5)	(6)
Adopted (1=yes)	0.76*** (0.08)	0.76*** (0.08)	0.05*** (0.01)	0.05*** (0.01)	0.05*** (0.02)	0.05*** (0.02)
Adoptee sibling (1=yes)		0.63*** (0.12)		-0.06*** (0.02)		0.09*** (0.02)
Urban county in 1940 (1=yes)	-0.38*** (0.03)	-0.38*** (0.03)	-0.01* (0.005)	-0.01* (0.005)	0.09*** (0.01)	0.09*** (0.01)
Age in 1940	-0.05*** (0.01)	-0.05*** (0.01)	0.10*** (0.002)	0.10*** (0.002)	0.01*** (0.002)	0.01*** (0.002)
Age in 1940 squared	0.001*** (0.0001)	0.0005*** (0.0001)	-0.001*** (0.00002)	-0.001*** (0.00002)	-0.0001*** (0.00002)	-0.0001*** (0.00002)
Moved across states (1=yes)	-0.36*** (0.04)	-0.36*** (0.04)	0.01*** (0.004)	0.01*** (0.004)		
Number of observations	289,233	289,821	289,233	289,821	289,233	289,821
Adjusted R-squared	0.04	0.04	0.19	0.19	0.16	0.16

Notes: Standard errors clustered by residence state given in parentheses. All regressions include residence state fixed effects. Adoptee siblings are excluded from the regression sample in regressions (1), (3) and (5). Moved across states equals one if the residence state in 1940 is different than the individual's birth state. * significant at 10%, ** significant at 5%, *** significant at 1%

Table 16: OLS analysis of adult outcomes with family fixed effects, sample restricted to white males.

Dependent variable:	Years of schooling	Log of annual income	Weeks worked in past year	Hours worked in past week	Number of household members		Married (1=yes)	Moved across states (1=yes)
	(1)	(2)	(3)	(4)	(5)	(6)	(6)	(6)
Adopted (1=yes)	-0.56 (0.35)	0.01 (0.13)	3.66* (2.17)	1.09 (2.34)	0.11 (0.36)	0.10***	-0.03 (0.03)	
Urban county in 1940 (1=yes)	0.91** (0.40)	0.48 (0.29)	-1.85 (4.17)	-6.70 (4.92)	0.71 (0.57)	0.001 (0.09)	0.24*** (0.08)	
Age in 1940	-0.18 (0.23)	0.14 (0.11)	0.72 (1.54)	1.33 (2.51)	0.26 (0.23)	0.07 (0.05)	-0.03 (0.04)	
Age in 1940 squared	0.001 (0.003)	-0.002 (0.001)	-0.01 (0.02)	-0.02 (0.03)	-0.003 (0.002)	-0.001 (0.001)	0.0004 (0.0004)	
Moved across states (1=yes)	0.35 (0.38)	0.05 (0.23)	1.60 (2.77)	3.17 (3.43)	0.09 (0.47)	-0.02 (0.06)		
Number of observations	2,150	1,479	2,085	1,680	2,202	2,202	2,202	
Adjusted R-squared	0.35	0.25	-0.002	0.16	-0.01	-0.14	0.40	

Notes: Standard errors clustered by residence state given in parentheses. Regression sample includes all adoptees and adoptee siblings linked to the 1940 federal census. Moved across states equals one if the residence state in 1940 is different than the individual's birth state. * significant at 10%, ** significant at 5%, *** significant at 1%

Table 17: OLS analysis of adult outcomes with urban childhood household interactions, sample restricted to white males.

Dependent variable: Including family fixed effects:	<u>Years of schooling</u>		<u>Log of annual income</u>	
	no (1)	yes (2)	no (3)	yes (4)
Adopted (1=yes)	-0.62*** (0.14)	-0.79 (0.48)	-0.20*** (0.04)	-0.01 (0.22)
Adopted x Urban household in 1910	0.25 (0.20)	0.39 (0.50)	0.11** (0.05)	0.03 (0.25)
Adoptee sibling (1=yes)	0.14 (0.18)		-0.09 (0.08)	
Adoptee sibling x Urban household in 1910	-0.11 (0.25)		0.02 (0.10)	
Urban county in 1940 (1=yes)	0.94*** (0.07)	0.95** (0.40)	0.48*** (0.03)	0.14 (0.11)
Age in 1940	-0.07*** (0.01)	-0.17 (0.23)	0.15*** (0.00)	0.14 (0.11)
Age in 1940 squared	0.0001 (0.0001)	0.001 (0.003)	-0.002*** (0.00004)	-0.002 (0.001)
Moved across states (1=yes)	0.62*** (0.10)	0.35 (0.38)	0.15*** (0.02)	0.04 (0.23)
Number of observations	289,769	2,150	207,888	1,479
Adjusted R-squared	0.11	0.35	0.18	0.24

Notes: Standard errors clustered by residence state given in parentheses. For columns (1) and (3), the regression sample includes all adoptees and adoptee siblings linked to the 1940 federal census and all males between the ages of 20 and 59 from the 1% IPUMS sample of the 1940 census. For columns (2) and (4), the regression sample includes all adoptees and adoptee siblings linked to the 1940 census. Moved across states equals one if the residence state in 1940 is different than the individual's birth state. Black is dropped from the family fixed effects regressions because there is no variation in the variable within families in the regression sample. * significant at 10%, ** significant at 5%, *** significant at 1%