Changes in Parental Gender Preference in the United States: Evidence from 1850–2019

Todd R. Jones, Matthew Millington, and Joseph Price*

March 9, 2023

Abstract

We examine the degree to which parental gender preferences in the US have changed over time. To quantify levels of parental sex preference, we compare the likelihood that mothers have a third child given the gender makeup of their first two children. We construct a novel dataset of women's fertility histories using full-count censuses from 1850-1880 and 1900-1940 and extend the sample to 2019 using more recent datasets. We find a preference for having a mix of genders with only a small preference for sons. We find that women are about 2 percentage points more likely to have a third child if the sex of her first two children are the same, and this effect was very stable from 1850 to 1940. In contrast, we find that this effect gets much larger after 1940, reaching a high point in 1990-2000 of about 6-7 percentage points.

*Jones: Mississippi State University and IZA. Millington: Arizona State University. Price: Brigham Young University, NBER, and IZA. We would like to thank Angela Cools for helpful feedback.

1 Introduction

Parents exhibit various preferences about the gender composition of their children. In many parts of the world, there has been a strong preference among parents to have a son (Williamson, 1976). In contrast, the majority of the evidence for the US indicates a preference for having at least one son and one daughter. One objective way to measure the strength of gender preference is to examine differences in fertility decisions based on the current gender makeup of a woman's children. Parents who have a son preference will be more likely to have additional children if all of their children have been girls. In the US, parents are more likely to have an additional child if the previous children are of the same gender (Tian and Morgan, 2015).

In this paper, we investigate how this effect of sex composition on fertility has evolved over the past 170 years. There are roughly two opposing forces over this time. The first is that the US became more gender neutral, which Pollard and Morgan (2002) have predicted would result in a decrease in gender preferences of parents. On the other hand, fertility has decreased over time, making it less likely that parents will have both a son and daughter through the natural course of their childbearing.

Our primary data sources are the full-count US Censuses for 1850–1940, provided by IPUMS (Ruggles et al., 2021). We also extend our main analysis to include samples of the US census and the American Community Survey from 1950 to 2019. We find that from 1850 to 1940, mothers whose first two children were the same gender were 2 percentage points more likely to have a third child. This effect size is surprisingly stable over this 90-year time period, varying from 1.3 percentage points in 1880 to 2.7 percentage points in 1940. The effect is smaller but still significant for the decision to have a fourth and fifth child based on the sex composition of all previous children. In contrast, mothers whose first two children are girls are only about 0.3 percentage points more likely to have a third child than mothers whose first two children are boys. Thus, the parental preference for sons, in terms of fertility decisions, is about 7 times smaller than the preference to

have a child of each gender. When we consider the time period from 1950 to 2019, we find consistently higher estimates for the decision to have a third child, with estimated coefficients (on the first two children being boys and the first two children being girls) ranging from 3 to 7 percentage points.

The results in this paper highlight the importance of documenting fertility patterns and other family decisions over a longer historical time period. Analysis using just more recent census data would provide a misleading view about changes in patterns over time, since there was a long period of stability for almost a century, followed by sharp rise after World War II, with a decline only occurring in the last two decades. The strength of the relationship that we document also lends support to research using sex composition of the first two children as an instrumental variable for fertility.

1.1 Related Literature

Gender preference is an active area of research around the world. In the US, researchers have used our same method to analyze fertility data from 1960 to 2010, [1] and each paper finds a preference for both genders (Tian and Morgan, 2015; Pollard and Morgan, 2002; Yamaguchi and Ferguson, 1995; Ben-Porath and Welch, 1976). Pollard and Morgan (2002) hypothesize that gender preference in fertility would disappear as gender norms change, but Tian and Morgan (2015) shows that the effect did not weaken during the 2000s. We extend this research back to 1850 to provide a much longer view of changes in parental preferences over time, which avoids concerns about biases that may arise from any specific time periods, such as the fertility boom immediately following World War II.

One paper which uses a similar method over a similar span is Aaronson et al. (2021), which studies the effect of fertility on female labor supply for many countries, including the US from 1860 to 2010. While not the focus of their paper, in an online appendix, they

[1] An exception is Bohnert et al. (2012), who find a preference for boys sex in the period of 1850 through 1900. They consider only for one state/territory (Utah), while we consider the entire US.

report estimates of the effect on having a third child if the first two children are of the same gender by decade. We consider both the effects of the first two children being boys and the first two children being girls, allowing us to study preference for a mix of genders compared to boy and girl preference. We additionally study higher-order births.

Similar research has been done for other countries. Generally, there appears to be a preference for sons over daughters in low-income countries and Asia (Filmer et al., 2009; Bongaarts, 2013; Basu and Das Gupta, 2001), though some countries may be changing over time. A notable example is South Korea, which has experienced a steep drop in son preference over a relatively short period of time (Choi and Hwang, 2020). Like the US, the preference for mixed genders dominates son preference in Europe (Mills and Begall, 2010; Andersson et al., 2006; Hank and Kohler, 2000) and Australia (Kippen et al., 2007). There is also evidence of a preference for daughters in Sweden (Miranda et al., 2018).

Besides the study of fertility given the gender of children, there is contemporary evidence of son preference in the US on the part of fathers (Dahl and Moretti, 2008; Raley and Bianchi, 2006; Depew and Price, 2018). This son preference affects marriage and divorce decisions such that mothers who have a son are more likely to marry and stay married to the father of their child, and women are more likely to have a second child when their first child is a girl. (Blau et al. (2020) find that such son preference has decreased in recent years.) In this paper, we show that while there has consistently been a small and detectable son preference in the US, it is much smaller than the preference that we observe for having a child of each gender.

As stated above, the gender composition of children is also commonly used as an instrument for fertility. This approach was first used by Angrist and Evans (1998) to study the effect of fertility on female labor supply, and the approach has been applied to different contexts by Aaronson et al. (2021), Cruces and Galiani (2007) and Ebenstein (2009). The same instrument has been used to look at the effect of fertility on other outcomes in Angrist et al. (2010), and Conley and Glauber (2006), Ananat and Michaels (2008) and

Bedard and Deschênes (2005). Thus, one strength of our paper is that it provides a closer study of the first stage in this rich area of research.

The patterns we examine in this paper also have interesting implications for evolutionary models of changes in fertility patterns over time (Colleran, 2016). While most of these models have focused on explanations for the decline in fertility over time, those models could be tested for their relevance towards changes in sex preferences. Colleran (2016) notes that these changes over time are a multi-level phenomenon in which individual decisions interact with smaller social groups and also with larger societies and countries. While there is research about the cultural persistence over time (Falk et al., 2018), one thing that is striking in our setting is that we observe a rather quick change in preference without any evident evolutionary advantage.

2 Empirical strategy

2.1 Data and Sample Selection

Our primary data sources are the full-count US Censuses for 1850 to 1940, provided by IPUMS (Ruggles et al., 2021). We also extend our main analysis to include the available samples of the US census data from 1950 to 2019. We construct a sample of families for which 1) children can be linked to their parents; 2) the mother has stopped having children; and 3) no children have left the household. To achieve these goals, we make the following sample restrictions:[2] We keep households that 1) have one father and one mother; 2) have a male in the first census enumeration slot and a female in the second slot, where these individuals are married to each other; 3) do not include the parents of the head of household, the grandchildren of the head of the household, or institutional inmates; 4) do not have non-children listed before any of the children of the household

[2]See Online Appendix Table A1 for the data cleaning steps we take, with the changes in sample size after each step. The steps included in this table are common across all tables; we make additional sample restrictions in individual tables.

head; and 5) do not have children-in-law. We additionally drop households where the surname of the children does not match the surname of the mother.

We also employ sample restrictions based on the age of the mother and the age of her oldest and youngest child. We restrict our main analysis to mothers whose age is between 30 and 35 years old, a range for which the oldest child of these women is still likely living at home. We restrict the sample to mothers whose oldest observed child is less than 18 years old (to minimize the chance that we include families with children who have left the household). We also restrict our sample to households where the youngest child is at least five years old. This is to minimize the chance that we include mothers who have not completed their fertility. In our analysis, we document the degree to which our estimates change as we vary these restrictions.

The primary goal of these age restrictions in to ensure that we correctly identify the number of children that the mother has had. In Online Appendix Table A2, we investigate the degree to which we identify correctly the number of children the mother has using unique data in the 1900 and 1910 censuses which asks about the number of children ever born and number of children surviving. The table reports the fraction of observations where the number of children in the household is not equal to 1) the number of children ever born to the mother, and (separately) 2) the number of children born to the mother who are surviving. We find that it is more common for the number of children in the household to not equal the number of children ever born than it is for the number of children in the household to not equal the number of children surviving.[3] We interpret this to mean that we do a good job of considering households where the oldest child has not left the household, but that it is common for children to pass away. Later, we use the 1900 and 1910 censuses and restrict to cases where the number of children in the

[3]To see this visually, Online Appendix Figure A1 shows the percentage of households for which the number children in the household equal the number of children ever born. This is declining with the age of the mother. Online Appendix Figure A2 is similar, but focuses on surviving children. Online Appendix Figure A3 shows the mean number of children in the household, children ever born, and surviving children, graphed against mother age. Note that none of these figures to households in which the age of the youngest child is at least five.

household is equal to the number of children ever born and surviving to test sensitivity; our results are qualitatively similar.

Finally, we exclude certain twin/multiple births if this eliminates the choice to have another child. For example, if the outcome is having a third child, we exclude instances where the second and third child are the same age. This is conservative as it excludes those who are not twin births but are born within a year of each other.[4]

2.2 Specification

Our objective is to analyze the likelihood of having an additional child conditional of the gender mix of the prior children. In our main analysis, we consider the likelihood of having a third child conditional on the first two children being either all girls or all boys (with the first two being of opposite genders being the omitted category). We estimate the following regression:

$$HaveThird_i = \beta_0 + \beta_1 TwoGirl_i + \beta_2 TwoBoy + X_i, \tag{1}$$

where the unit of analysis i is a mother. X_i contains, depending on the specification, the mother's age, the mother's birthplace, or mother age/birthplace fixed effects. We cluster standard errors at the mother birthplace level. A positive coefficient of x on $TwoGirl$ ($TwoBoy$) means that mothers whose first two children are girls (boys) are x percentage points more likely to have a third child than mothers whose first two children are a girl and boy or boy and girl.

We estimate several variations on the above. First, we vary the age range of the mother. Second, we restrict to having two or more children as opposed to having 2-3 children.

[4]There could be additional restrictions one could make in order to be even more conservative. For instance, for the 1900 Census, we also limited to households that are coded as married couple family households, dropped households containing children with a probable step/adopted mother and/or father, kept only households considered as such under the 1970 Census definition, kept only households where the number of subfamilies is none or not applicable, and kept only households with one family. Most of these variables are not available for all Census years. We re-ran the main regression and obtained similar results.

Third, we replace the outcome of having a third child with having a second, fourth, fifth, and sixth child. We also estimate the above separately by year to examine changes over time.

Our strategy relies on the gender of children being randomly assigned. While we cannot directly test this, in Table 1, we examine attributes of the mother across mothers whose first two children are 1) girls; 2) boys; and 3) one of each. We find that the attributes across these three types of mothers are nearly identical, with the average value of each (binary) characteristic being within 1 percentage point of each other, the exceptions being urban and farm, which are up to 2 percentage points different. With that said, an F-test rejects equality of means in each case. While not exactly identical, we interpret these findings as at least weak evidence during this time period that gender composition of a mother's first two children was more-or-less random (thus providing the exogenous source of variation for our identification strategy).[5]

3 Results

Table 2 contains the results of our main analysis. Column 1 provides a baseline with no fixed effects where the sample contains mothers with 2 or 3 children in the household. We find evidence of a preference to have children of mixed genders: mothers whose first two children are girls are 2.3 p.p. more likely to have a third child than are mothers whose first two children are of mixed gender, with a similar 2.0 p.p. increase in the probability for mothers with two boys. These represent a 6-7% increase relative to mean for the omitted group which is 0.317. The results are not sensitive to the inclusion of mother age, mother birthplace, or mother age-by-mother birthplace fixed effects (columns 2-4). In column 5, we expand the sample to mothers with 2 or more children and change the outcome to

[5] Another consideration is selection into the main sample based on the sex of the first child. Families with a first born girl are less likely to have a second child, but the estimated coefficient is tiny: -0.004 (0.4p.p.). This is less than 1% of the omitted mean of 0.464, suggesting that there is very little selection along this dimension.

having three or more children. The coefficients in column 5 are larger, which is consistent with the fact that the outcome for the omitted group as a higher mean in this sample.

The main conclusion from Table 2 is that there is strong evidence for a mixed-gender preference but only small evidence of a preference for sons. The gap in the probability of having a third child between have mixed-gender children and same-gender children is over 7 times larger than the gap between having two boys compared to having two girls (0.21 vs 0.03).

Before proceeding to alternate specifications and samples, we first test the sensitivity of our main results in the two census years (1900 and 1910) where we can compare the number of children in the household to the number of surviving children. This helps avoid any mismeasurement that will arise if some of the mother's children have already left the home. We can also compare it to the number of children that have been born to the mother which can help identify cases where children who have died prior to the census (though many of these deaths likely occurred close to the time of birth).

In Table 3, we start by estimating our baseline results without any restrictions. In Column 2, we require that the number of children that the mother has ever had match the number that are currently in the home. This causes our estimates to increase slightly (consistent with removing that source of measurement error). However, this restriction also reduces our sample size by 37% (consistent with the high levels of infant mortality during this time period in the US). In Column 3, we restrict the sample to just those mothers for whom the number of children in the home is the same as the number of surviving children that they have. This produces estimates that are identical to the baseline specification and decreases the sample size by just 7%. The sample size differences between columns 2 and 3 indicate that the main cause of a family not having the same number of children in the household as were ever born is death, not moving out of the home. Column 4 makes both restrictions and the results are similar to what we find when we in Column 2, which is slightly larger in magnitude than the baseline results.

We further test the sensitivity of our main results in Table 4 by varying the restrictions we place on the age of the mother and the ages of her oldest and youngest children. In columns 1-3, we focus on the age of the mother. This is designed to address the concern that women may not be finished having children, or more specifically, that a woman who will later have a third (or more) child is observed as having only two. The older the age of the woman, the less likely she is to have additional children (see Online Appendix Figure A4). Column 1 considers mothers aged 25-29. They are less likely to have finished fertility, but it is very unlikely that their children will have left the house. We find statistically significant coefficients on both variables. The magnitude grows in column 2, which restricts the age to 30-35 and in column 3, which restricts to 36-40, when women are much more likely to have completed fertility. In column 4, we return to mothers aged 30-35, but to make it less likely that we include household that have had children leave, by requiring that the oldest child in the household be no older than 13. Results are similar. In column 5, we instead require the youngest child to be 5 or older, to make it more likely that the women in the sample have finished having children. This is the specification we use throughout the paper (and is the same as column 4 in Table 2. Results are again similar, as they are in column 6 when we make both restrictions. Results for mothers age 25-29 deviate the most from the other columns; in column 7, we consider this age group but also require that the youngest child be 5 or older. Estimates are higher than in column 1, but still lower than in other columns.

In Table 5, we estimate our main results in 10-year intervals from 1850-2019.[6] We find that the effect of sex composition on fertility was very stable from 1850 to 1940 with a gap of about 2 percentage points between families with two children of the same gender and families with both a son and daughter. When we compare families with two girls compared to a family with a mixed-gender pair, we find that the gap varies 1.3 percentage

[6]The surname similarity variables is not available for most of the years after 1940, so we do not restrict on this variable for this analysis. Results for 1850-1940 for which the sample selection uses surname similarity are presented in Online Appendix Table A3; we find similar results.

points in 1870 to 2.7 percentage points in 1940, but nearly all of the estimates are very close to 2 percentage points. We find about the same level of similarity over time when comparing families with two boys to families with a mixed-gender pair. We also find a similar pattern in Online Appendix Table A4 when we include mothers with more than 3 children in the sample.

In Table 5, we also report the mean of the omitted group, which shows that women in our sample were more likely to have a third child in 1850 (38%) compared to 1940 (28%).[7] Thus, if we estimate the gap in fertility relative to the sample mean for the omitted group, we find that the estimated impact of sex composition on fertility decisions did increase during this time period from about 6% in 1850 to 9% in 1940.

After 1940, we cannot use the full-count censuses for our analysis, but we can extend the sample to 2019 using samples of the decennial censuses for 1950–2000 and the American Community Survey for 2010 and 2019.[8] We first point out that ultrasound technology and legalized abortion both became available during this later time frame. Both coefficients fluctuate over these later years, but in all cases, both coefficients are higher for each year between 1950–2019 than they are for any year between 1850–1940. With the caveat that the sample sizes are very low for some of the years (and keeping in mind that the baseline propensity to have a third child—and to have between two and three children and thus be in the sample—differs over time), the estimates indicate an increased preference for families of mixed gender over time.

The results so far have focused on the decision to have a third child. In Table 6, we consider the decisions to have higher order births. For the decision to have a Xth child,

[7] Online Appendix Figure A5 shows that the number of children in the household for mothers aged 30-35 is declining by census year.

[8] For 1950, we use the 1950 1% sample. For 1960, we use the 5% sample. For 1970, we combine the 1% state form 1, the 1% state form 2, the 1% metro form 1, the 1% metro form 2, the 1% neighborhood form 1, and the 1% neighborhood form 2 samples together. For 1980, we combine the 5% state, the 1% metro, the 1% urban, the 1% labor market areas, and the 1% metro/nonmetro samples together. For 1990, we combine the 5% state and 1% metro samples together. For 2000, we combine the 5% and 1% weighted samples together. For 2010, we use the ACS. For 2019, we also use the ACS. All samples are obtained from IPUMS USA (Ruggles et al., 2022). We do not use weights; results should be interpreted accordingly.

we limit the sample to mothers with either $X-1$ or X children. The first column reproduces column 4 of Table 2. The coefficients decrease for higher order births, but there continues to be a statistically significant effect of sex composition on the decision to have a fourth or fifth child. In Online Appendix Table A5, we present these same results but expand the mothers who had at least X children.

Finally, in Online Appendix Table A6, we present results for two subsamples: mothers with Asian ancestry and African American mothers. In more recent years, recent immigrants to Canada from South and East Asian countries have been shown to exhibit a son preference (Almond et al., 2013). In Panel A, we consider mothers whose race is coded as "Chinese", "Japanese," or "Other Asian or Pacific Islander" and present results using our main specification (Column 4 of Table 2). The first column considers the years 1850–1940 together. The sample size is very small, with no observations before 1870, and fewer than 200 per year before 1920. We fail to detect an effect, though the coefficient on two girls is of the same magnitude as our main results in Table 2 albeit with a much larger standard error. The second column consider the years between 1950–2019 (specifically 1950, 1960, 1970, 1980, 1990, 2000, and 2019). We detect positive and significant coefficients for mothers with two boys as well as two girls, evidence for a preference for mixed gender children. The former coefficient is roughly twice as large as the latter, indicating a relative preference for boys compared to girls for mothers with Asian ancestry since 1950. However, one problem with grouping 1950–2019 together is that the sample size varies widely by year.[9] Therefore, the remaining columns present the results in progressively smaller windows. If anything, the preference for a son increases over time.

Online Appendix Table A6 Panel B presents the results for African American mothers. In the United States, African American mothers are more likely to have girls at birth than are White mothers (Anderson and Ray, 2010). If this is known among African American mothers, it may have implications for gender preferences. In the early period (1850–1940),

[9]There are 9 observations in 1950, 298 in 1960, 619 in 1970, 3,069 in 1980, 2,507 in 1990, 2,142 in 2000, 317 in 2010, and 237 in 2019.

we find evidence that women are more likely to have a third child if the first two children are boys, but not if the first two children are girls, though the coefficient is relatively small.[10] In the latter years, we find a preference for mixed gender families, with the effect staying roughly equal over time.[11]

4 Conclusion

The impact of the sex composition of children on fertility decisions is an important area of research for two reasons. First, it provides an objective measure of parental preferences that can be consistently measured over time and across different cultures. We extend this research to cover nearly a century of data for the US with our main analysis and expand it to cover nearly 160 years using more recent samples of the census data. We find that from 1850-1940, the preference for mixed gender children is surprisingly consistent even as various aspects of families and the economy changed. Our results are also robust to a variety of specifications and indicate that mothers are about 2 percentage points more likely to have a third child if their first two children are the same gender. It is only after 1940 that we see both a dramatic increase in this estimate and variation over time.

Second, the impact of sex composition on fertility has been used as an important instrumental variable for studies that examine the impact of fertility on female labor supply, including studies by Angrist and Evans (1998), Aaronson et al. (2021), Cruces and Galiani (2007), and Ebenstein (2009). Our results show that this first stage relationship is strong and consistent over time, providing support for the use of this instrumental variable in past and future research on the impact of fertility. The availability of the full-count US census data will allow future research on the impact of fertility to focus on specific sub-

[10] Before 1870, there are very few African Americans in our sample: less than 1,000 in each of 1850 and 1860, compared to over 11,000 in 1870 and over 35,000 in 1940.

[11] We again note the issue with the number of observations varying greatly across years: 559 in 1950, 3,010 in 1960, 5,348 in 1970, 12,603 in 1980, 6,940 in 1990, 5,405 in 2000, 342 in 2010, and 194 in 2019.

groups of women and the ability to link these historical census records across time will also make it possible to examine the impact of fertility on very long-run outcomes.

References

Aaronson, D., R. Dehejia, A. Jordan, C. Pop-Eleches, C. Samii, and K. Schulze (2021). The Effect of Fertility on Mothers' Labor Supply over the Last Two Centuries. *The Economic Journal 131*(633), 1–32.

Almond, D., L. Edlund, and K. Milligan (2013). Son preference and the persistence of culture: evidence from south and east asian immigrants to canada. *Population and Development Review 39*(1), 75–95.

Ananat, E. O. and G. Michaels (2008). The Effect of Marital Breakup on the Income Distribution of Women with Children. *The Journal of Human Resources 43*(3), 611–629.

Anderson, S. and D. Ray (2010). Missing women: age and disease. *The Review of Economic Studies 77*(4), 1262–1300.

Andersson, G., K. Hank, M. Rønsen, and A. Vikat (2006). Gendering family composition: sex preferences for children and childbearing behavior in the Nordic countries. *Demography 43*(2), 255–267.

Angrist, J., V. Lavy, and A. Schlosser (2010). Multiple Experiments for the Causal Link between the Quantity and Quality of Children. *Journal of Labor Economics 28*(4), 773–824.

Angrist, J. D. and W. N. Evans (1998). Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *The American Economic Review 88*(3), 450–477.

Basu, A. and M. Das Gupta (2001). Family Systems and the Preferred Sex of Children. In N. J. Smelser and P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences*, pp. 5350–5357. Oxford: Pergamon.

Bedard, K. and O. Deschênes (2005). Sex Preferences, Marital Dissolution, and the Economic Status of Women. *Journal of Human Resources XL*(2), 411–434.

Ben-Porath, Y. and F. Welch (1976). Do Sex Preferences Really Matter? *The Quarterly Journal of Economics 90*(2), 285–307.

Blau, F. D., L. M. Kahn, P. Brummund, J. Cook, and M. Larson-Koester (2020). Is there still son preference in the united states? *Journal of Population Economics 33*(3), 709–750.

Bohnert, N., H. L. Jåstad, J. Vechbanyongratana, and E. Walhout (2012). Offspring sex preference in frontier america. *Journal of Interdisciplinary History 42*(4), 519–541.

Bongaarts, J. (2013). The Implementation of Preferences for Male Offspring. *Population and Development Review 39*(2), 185–208.

Choi, E. J. and J. Hwang (2020). Transition of Son Preference: Evidence From South Korea. *Demography 57*(2), 627–652.

Colleran, H. (2016). The cultural evolution of fertility decline. *Philosophical Transactions of the Royal Society B: Biological Sciences 371*(1692), 20150152.

Conley, D. and R. Glauber (2006). Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility. *The Journal of Human Resources 41*(4), 722–737.

Cruces, G. and S. Galiani (2007). Fertility and female labor supply in Latin America: New causal evidence. *Labour Economics 14*(3), 565–573.

Dahl, G. B. and E. Moretti (2008). The Demand for Sons. *The Review of Economic Studies 75*(4), 1085–1120.

Depew, B. and J. Price (2018). Marriage and the economic status of women with children. *Review of Economics of the Household 16*(4), 1049–1061.

Ebenstein, A. (2009). When Is the Local Average Treatment Close to the Average? Evidence from Fertility and Labor Supply. *The Journal of Human Resources 44*(4), 955–975.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics 133*(4), 1645–1692.

Filmer, D., J. Friedman, and N. Schady (2009). Development, Modernization, and Childbearing: The Role of Family Sex Composition. *The World Bank Economic Review 23*(3), 371–398.

Hank, K. and H.-P. Kohler (2000). Gender Preferences for Children in Europe: Empirical Results from 17 FFS Countries. *Demographic Research 2*(1).

Kippen, R., A. Evans, and E. Gray (2007). Parental preference for sons and daughters in a Western industrial setting: Evidence and implications. *Journal of Biosocial Science 39*(4), 583–597.

Mills, M. and K. Begall (2010). Preferences for the sex-composition of children in Europe: A multilevel examination of its effect on progression to a third child. *Population Studies 64*(1), 77–95.

Miranda, V., J. Dahlberg, and G. Andersson (2018). Parents' Preferences for Sex of Children in Sweden: Attitudes and Outcomes. *Population Research and Policy Review 37*(3), 443–459.

Pollard, M. S. and S. P. Morgan (2002). Emerging Parental Gender Indifference? Sex Composition of Children and the Third Birth. *American Sociological Review 67*(4), 600–613.

Raley, S. and S. Bianchi (2006). Sons, Daughters, and Family Processes: Does Gender of Children Matter? *Annual Review of Sociology 32*(1), 401–421.

Ruggles, S., S. Flood, S. Foster, R. Goeken, J. Pacas, M. Schouweiler, and M. Sobek (2021). Ipums usa: Version 11.0 [dataset]. *Minneapolis, MN: IPUMS 10*, https://doi.org/10.18128/D010.V12.0.

Ruggles, S., S. Flood, R. Goeken, M. Schouweiler, and M. Sober (2022). Ipums usa: Version 12.0 [dataset]. *Minneapolis, MN: IPUMS*, https://doi.org/10.18128/D010.V11.0.

Tian, F. F. and S. P. Morgan (2015). Gender Composition of Children and the Third Birth in the United States. *Journal of Marriage and Family 77*(5), 1157–1165.

Williamson, N. E. (1976). *Sons or daughters: a cross-cultural survey of parental preferences*. Sage library of social research ; vol.31. Beverly Hills ; London: Sage Publications for the National Council on Family Relations.

Yamaguchi, K. and L. R. Ferguson (1995). The Stopping and Spacing of Childbirths and Their Birth-History Predictors: Rational-Choice Theory and Event-History Analysis. *American Sociological Review 60*(2), 272–298.

5 Tables

Table 1: Summary Statistics

	First Two Girl	First Two Boy	First Two Mixed	F-Test p-value
Born in US	0.836	0.835	0.834	0.014*
	(0.371)	(0.371)	(0.372)	
White	0.931	0.935	0.936	0.000**
	(0.253)	(0.246)	(0.244)	
Urban	0.539	0.523	0.530	0.000**
	(0.499)	(0.499)	(0.499)	
Metro	0.457	0.452	0.454	0.000**
	(0.498)	(0.498)	(0.498)	
Farm	0.224	0.242	0.235	0.000**
	(0.417)	(0.428)	(0.424)	
Literate	0.687	0.684	0.687	0.000**
	(0.464)	(0.465)	(0.464)	
Age of Oldest Child in HH	11.654	11.679	11.653	0.000**
	(2.603)	(2.624)	(2.607)	
Observations	628,359	675,826	1,350,716	

This table reports summary statistics, split by mothers whose first two children were girls (column 1), boys (column 2), and of mixed gender (column 3). Means are reported with standard deviations in parentheses. Column 4 reports the p-value from an F-test that the means are the same. Sample is limited to mothers aged 30-35 with 2 or 3 children where the 2nd and 3rd children are not of the same age, and where the youngest child is at least five years old. The variables refer to the mother. Born in the US includes being born in American Samoa, Guam, Puerto Rico, and the US Virgin Islands. Urban, metro, and farm refer to the mother living in an urban area, a metro area, and on a farm. Literate means the mother can both read and write. The sample size is smaller for being literate due to this question not being available for the 1940 census. ** 0.01, * 0.05.

Table 2: Effect of Gender Composition of Children on the Likelihood of Having a Third Child

	(1)	(2)	(3)	(4)	(5)
First Two Girl	0.023**	0.023**	0.022**	0.022**	0.026**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
First Two Boy	0.020**	0.020**	0.020**	0.020**	0.023**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Omitted Y Mean	0.317	0.317	0.317	0.317	0.427
Observations	2,654,901	2,654,901	2,654,895	2,654,857	3,181,960
Sample (# Births)	2-3	2-3	2-3	2-3	2+
FEs	N/A	Age	Bplc	Age-Bplc	Age-Bplc

This table reports the results from a regression of an indicator for having three children—or more than three children in column 5— on indicators for the first two children being girls the first two children being boys. The sample is limited to mothers with 2 or 3 children in all columns except column 5, which limits to mothers with 2 or more children. The sample is limited to mothers aged 30-35 whose youngest child is five years or older. Cases where the second and third child are the same age are excluded. No fixed effects are included in the first column; column 2 includes mother age fixed effects; column 3 includes mother birthplace fixed effects; and columns 4-5 includes mother age -by- mother birthplace fixed effects. Standard errors are included in parentheses and are clustered at the level of mother birthplace. ** 0.01, * 0.05.

Table 3: Had 3 Children; Sensitivity Check, 1900, 1910

	(1)	(2)	(3)	(4)
Panel A: 1900				
First Two Girl	0.019**	0.022**	0.020**	0.022**
	(0.002)	(0.003)	(0.002)	(0.003)
First Two Boy	0.018**	0.019**	0.018**	0.019**
	(0.002)	(0.002)	(0.002)	(0.002)
Omitted Y Mean	0.343	0.319	0.339	0.318
Observations	242,941	152,791	224,681	151,734
Kids HH==Kids Ever		Yes		Yes
Kids HH==Kids Survive			Yes	Yes
Panel B: 1910				
First Two Girl	0.018**	0.022**	0.019**	0.022**
	(0.002)	(0.003)	(0.002)	(0.003)
First Two Boy	0.019**	0.021**	0.018**	0.020**
	(0.002)	(0.002)	(0.002)	(0.002)
Omitted Y Mean	0.328	0.303	0.324	0.302
Observations	305,375	203,343	283,471	201,822
Kids HH==Kids Ever		Yes		Yes
Kids HH==Kids Survive			Yes	Yes

This table reports the results from a regression of an indicator for having three children on indicators for the first two children being girls the first two children being boys. Panel A is for 1900 and Panel B is for 1910. Columns 2 and 4 restrict the sample to mothers who have the same number of children in the household as the number ever born; column 5 restricts to when these variables do not match. Columns 3 and 4 restrict to the number of children in the household being the same as the number of a mother's children ever surviving; column 6 restricts to when these variables do not match. Column 7 limits to both the same number of children in the household not matching the number ever born -and- the number of children in the household not being the same as the number of a mother's children ever surviving. The sample is limited to mothers aged 30-35 with 2 or 3 children where the youngest child is five years or older. Cases where the second and third child are the same age are excluded. We include mother age -by- mother birthplace fixed effects in all columns. Standard errors are included in parentheses and are clustered at the level of mother birthplace. ** 0.01, * 0.05.

Table 4: Had Three Children; Vary Age of Mother and Oldest/Youngest Child

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
First Two Girl	0.014**	0.020**	0.021**	0.020**	0.022**	0.023**	0.016**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
First Two Boy	0.010**	0.017**	0.023**	0.017**	0.020**	0.018**	0.012**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Omitted Y Mean	0.382	0.418	0.422	0.405	0.317	0.276	0.227
Observations	6,027,186	6,614,140	3,765,542	5,744,646	2,654,857	1,970,774	834,330
Mother Age	25-29	30-35	36-40	30-35	30-35	30-35	25-29
Oldest Kid Age $\leq$	17	17	17	13	17	13	17
Youngest Kid Age $\geq$	0	0	0	0	5	5	5

This table reports the results from a regression of an indicator for having three children—or more than three children in column 2— on indicators for the first two children being girls the first two children being boys. The sample is limited to mothers with 2 or 3 children. The sample is limited to mothers aged 25-29 in column 1; 30-35 in columns 2 and 4-6; and 36-40 in column 3. Cases where the second and third child are the same age are excluded. In columns 1-3 and 5, the age of the oldest child is limited to be 17 and younger, while in columns 4 and 6 it is 13 and younger. In columns 1-4, the age of the youngest child is unrestricted, while in columns 5 and 6 the age of the youngest child is 5 and older. All columns include mother age -by-mother birthplace fixed effects. Standard errors are included in parentheses and are clustered at the level of mother birthplace. ** 0.01, * 0.05.

Table 5: Had Three Children; by Census Year

	1850	1860	1870	1880	1900	1910	1920	1930	1940
First Two Girl	0.023**	0.021**	0.013**	0.019**	0.019**	0.017**	0.021**	0.022**	0.027**
	(0.006)	(0.005)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
First Two Boy	0.018**	0.025**	0.028**	0.019**	0.017**	0.018**	0.017**	0.020**	0.022**
	(0.006)	(0.005)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Omitted Y Mean	0.384	0.366	0.372	0.362	0.343	0.328	0.324	0.312	0.282
Observations	39,407	62,343	101,302	133,920	254,622	322,043	459,031	669,414	738,851
	1950	1960	1970	1980	1990	2000	2010	2019	
First Two Girl	0.033**	0.044**	0.047**	0.060**	0.070**	0.072**	0.045**	0.043*	
	(0.011)	(0.005)	(0.004)	(0.002)	(0.003)	(0.004)	(0.014)	(0.018)	
First Two Boy	0.038**	0.046**	0.056**	0.054**	0.062**	0.060**	0.035*	0.065**	
	(0.011)	(0.004)	(0.004)	(0.002)	(0.003)	(0.004)	(0.014)	(0.018)	
Omitted Y Mean	0.309	0.307	0.399	0.254	0.230	0.243	0.263	0.252	
Observations	11,797	66,587	98,120	207,933	122,356	84,326	6,529	4,073	

This table reports the results from a regression of an indicator for having three children on indicators for the first two children being girls the first two children being boys. Each column corresponds to a different census year. The sample is limited to mothers aged 30-35 with 2 or 3 children where the youngest child is five years or older. Years 1850-1940 are based on full count censuses. The remaining years are based on samples from censuses, or the ACS, depending on the year. See the text for more details. Sample selection does not include surname similarity. Cases where the second and third child are the same age are excluded. Mother age -by- mother birthplace fixed effects are included. Standard errors are included in parentheses and are clustered at the level of mother birthplace. ** 0.01, * 0.05.

Table 6: Had 3 through 6 Children

Had X Births	3	4	5	6
All Girls	0.022**	0.009**	0.009**	0.005
	(0.001)	(0.001)	(0.003)	(0.005)
All Boys	0.020**	0.005**	0.010**	0.009
	(0.001)	(0.001)	(0.002)	(0.005)
Omitted Y Mean	0.317	0.288	0.243	0.201
Observations	2,654,857	1,248,359	491,123	155,637
Sample (# of Births)	2-3	3-4	4-5	5-6

This table reports the results from a regression of an indicator for having a given number of children on indicators for the first number of children being girls the first number of children being boys. Each column refers to a different sample, which is indicated at the bottom of the table. "All Girls" and "All Boys" are column-specific; for example, for the column labeled "3," "All Girls" means that the first two children are girls and "All Boys" means that the first two children are boys. For the column labeled "4," "All Girls" ("All Boys") means that the first three children are girls (boys). The sample is limited to mothers aged 30-35 whose youngest child is five years or older. Cases where the second and third are the same age are excluded in the column labeled "3," cases where the third and fourth are the same age are excluded in the column labeled "4," and so on. Mother age -by- mother birthplace fixed effects are included. Standard errors are included in parentheses and are clustered at the level of mother birthplace. ** 0.01, * 0.05.

6 Online Appendix Tables and Figures

Table A1: Dataset Cleaning

	1850	1860	1870	1880	1900	1910	1920	1930	1940
Step 1	3,706,255	5,341,529	7,767,847	10,423,752	16,874,876	21,178,389	25,736,893	31,658,676	37,780,475
Step 2	2,308,635	3,298,431	4,546,439	6,161,044	9,184,248	11,182,608	13,411,298	15,615,878	16,885,140
Step 3	2,282,383	3,268,353	4,496,573	6,110,829	9,103,009	11,069,559	13,284,602	15,495,714	16,688,909
Step 4	2,264,872	3,244,252	4,455,194	6,079,803	9,060,622	11,017,904	13,227,966	15,430,854	16,620,787
Step 5	2,257,405	3,233,751	4,438,391	6,070,559	9,038,967	10,967,518	13,186,455	15,382,344	16,563,836
Step 6	2,235,007	3,203,971	4,401,621	5,989,620	8,936,262	10,865,272	13,081,844	15,265,632	16,456,943
Step 7	2,197,002	3,159,088	4,339,982	5,797,005	8,848,928	10,753,441	12,925,706	15,082,333	16,259,574
Step 8	2,190,694	3,151,105	4,329,617	5,763,908	8,777,274	10,682,703	12,816,096	14,961,527	16,098,221
Step 9	1,228,464	1,790,457	2,409,056	3,105,347	4,651,319	5,681,684	6,932,550	7,927,379	8,114,921
Step 10	1,171,363	1,715,847	2,321,008	2,804,462	4,449,909	5,418,438	6,643,837	7,600,357	7,789,016
Step 11	1,059,881	1,556,828	2,089,796	2,542,750	4,022,143	4,933,626	6,110,180	6,931,974	7,050,171

This table reports how the number of households (HHs) changes when performing various cleaning steps:
(Step 0: For 1860 and 1870, remove a tiny fraction of HHs that include an individual with non-unique identifier)
Step 1: Full Sample
Step 2: Keep HHs with exactly one mother and one father
Step 3: Keep HHs where a male is in census position 1 and female in census position 2
Step 4: Keep HHs where first two individuals are spouses
Step 5: Keep HHs where the mother is in census position 2
Step 6: Drop HHs with grandchildren, parents, or institutional inmates (in relation to head of HH)
Step 7: Drop HHs with non-children listed before any of head of HH's children
Step 8: Drop HHs with children-in-law
Step 9: Keep HHs with mothers between age 25 and 40
Step 10: Drop HHs where surnames of children do not match surname of mother
Step 11: Drop if a child is older than 17 or if age of youngest child is coded as more than that of oldest
Note that we make further restrictions in individual tables (such as further restrictions to the age of the mother and age of youngest child); these are not included here.

Table A2: Misclassification

	(1)	(2)	(3)	(4)	(5)	(6)
N in HH!=N Ever Born	0.296	0.362	0.444	0.350	0.350	0.327
N in HH!=N Survive	0.048	0.060	0.116	0.048	0.073	0.052
Observations	305,375	203,343	283,471	201,822	530,982	163,644
Mother Age	25-29	30-35	36-40	30-35	30-35	30-35
Oldest Kid Age $\leq$	17	17	17	13	17	13
Youngest Kid Age $\geq$	0	0	0	0	5	5

This table reports the results of a misclassification exercise using the 1900 and 1910 censuses, which include the number of children ever born and number of children surviving variables. It reports the fraction of observations that did not have 1) the number of children in the household equal the number of children ever born and (separately) 2) the number of children in the household equal the number of children surviving. The sample is limited to mothers with 2 or 3 children. The sample is limited to mothers aged 25-29 in column 1; 30-35 in columns 2 and 4-6; and 36-40 in column 3. In columns 1-3 and 5, the age of the oldest child is limited to be 17 and younger, while in columns 4 and 6 it is 13 and younger. In columns 1-4, the age of the youngest child is unrestricted, while in columns 5 and 6 the age of the youngest child is 5 and older. Cases where the second and third child are the same age are excluded. Mother age -by- mother birthplace fixed effects are included.

Table A3: Had Three Children; by Census Year; Sample Selection Uses Surname Similarity

	1850	1860	1870	1880	1900	1910	1920	1930	1940
First Two Girl	0.025**	0.020**	0.014**	0.017**	0.019**	0.018**	0.022**	0.023**	0.028**
	(0.006)	(0.005)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
First Two Boy	0.015*	0.025**	0.026**	0.020**	0.018**	0.019**	0.018**	0.021**	0.022**
	(0.006)	(0.005)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Omitted Y Mean	0.385	0.366	0.372	0.361	0.343	0.328	0.324	0.312	0.281
Observations	37,186	59,336	96,795	123,894	242,941	305,375	438,929	641,518	708,532

This table reports the results from a regression of an indicator for having three children on indicators for the first two children being girls the first two children being boys. The sample selection uses surname similarity. Each column corresponds to a different census year. The sample is limited to mothers aged 30-35 with 2 or 3 children where the youngest child is five years or older. Years 1850-1940 are based on full count censuses. The remaining years are based on samples from censuses, or the ACS, depending on the year. See the text for more details. Cases where the second and third child are the same age are excluded. Mother age -by- mother birthplace fixed effects are included. Standard errors are included in parentheses and are clustered at the level of mother birthplace. ** 0.01, * 0.05.

Table A4: Had Three+ Children; by Census Year

	1850	1860	1870	1880	1900	1910	1920	1930	1940
First Two Girl	0.025**	0.021**	0.019**	0.021**	0.021**	0.018**	0.026**	0.026**	0.032**
	(0.005)	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
First Two Boy	0.021**	0.027**	0.031**	0.022**	0.018**	0.022**	0.021**	0.022**	0.025**
	(0.005)	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Omitted Y Mean	0.529	0.498	0.502	0.482	0.461	0.440	0.435	0.420	0.380
Observations	52,050	79,163	129,044	165,979	311,879	388,174	552,437	798,502	861,644

	1950	1960	1970	1980	1990	2000	2010	2019
First Two Girl	0.048**	0.047**	0.051**	0.068**	0.080**	0.080**	0.057**	0.055**
	(0.010)	(0.004)	(0.003)	(0.002)	(0.003)	(0.004)	(0.014)	(0.018)
First Two Boy	0.047**	0.048**	0.055**	0.058**	0.067**	0.067**	0.035*	0.070**
	(0.010)	(0.004)	(0.003)	(0.002)	(0.003)	(0.004)	(0.014)	(0.018)
Omitted Y Mean	0.422	0.399	0.541	0.320	0.276	0.294	0.312	0.309
Observations	14,335	77,373	130,875	230,019	131,280	91,306	7,043	4,464

This table reports the results from a regression of an indicator for having three children on indicators for the first two *or more* children being girls the first two children being boys. Each column corresponds to a different census year. The sample is limited to mothers aged 30-35 with 2 or more children where the youngest child is five years or older. Years 1850-1940 are based on full count censuses. The remaining years are based on samples from censuses, or the ACS, depending on the year. See the text for more details. Sample selection does not include surname similarity. Cases where the second and third child are the same age are excluded. Mother age -by- mother birthplace fixed effects are included. Standard errors are included in parentheses and are clustered at the level of mother birthplace. ** 0.01, * 0.05.

Table A5: Had 3 through 6 Children, Unconditional

Had X Births	3	4	5	6
All Girls	0.026**	0.012**	0.011**	0.008
	(0.001)	(0.002)	(0.003)	(0.005)
All Boys	0.023**	0.009**	0.014**	0.014**
	(0.001)	(0.001)	(0.002)	(0.005)
Omitted Y Mean	0.427	0.368	0.299	0.240
Observations	3,181,960	1,409,728	530,982	163,644
Sample (# of Births)	2+	3+	4+	5+

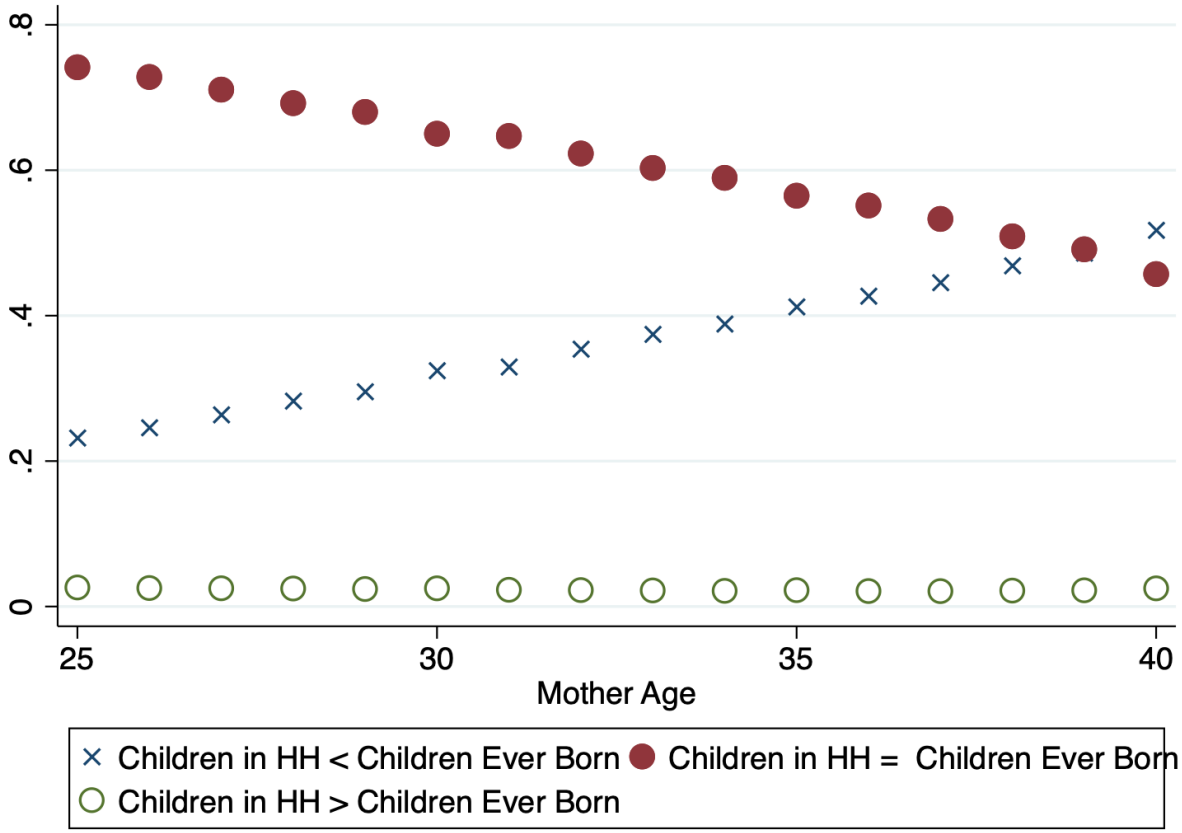
This table reports the results from a regression of an indicator for having a given number *or more* children on indicators for the first number of children being girls the first number of children being boys. Each column refers to a different sample, which is indicated at the bottom of the table. "All Girls" and "All Boys" are column-specific; for example, for the column labeled "3," "All Girls" means that the first two children are girls and "All Boys" means that the first two children are boys. For the column labeled "4," "All Girls" ("All Boys") means that the first three children are girls (boys). In contrast to Table 6, this table does not set an upper limit on the number of births. The sample is limited to mothers aged 30-35 whose youngest child is five years or older. Cases where the second and third are the same age are excluded in the column labeled "3," cases where the third and fourth are the same age are excluded in the column labeled "4," and so on. Mother age by mother birthplace fixed effects are included. Standard errors are included in parentheses and are clustered at the level of mother birthplace. ** 0.01, * 0.05.

Table A6: Effect of Gender Composition of Children on the Likelihood of Having a Third Child; Asian Origins and Black/African American Subsamples

	1850-1940	1950-2019	1960-2019	1970-2019	1980-2019	1990-2019
Panel A: Asian Ancestry Subsample						
First Two Girl	0.024	0.069**	0.070**	0.071**	0.078**	0.091**
	(0.017)	(0.011)	(0.011)	(0.011)	(0.010)	(0.011)
First Two Boy	-0.009	0.036*	0.036*	0.039**	0.031*	0.038*
	(0.009)	(0.014)	(0.014)	(0.013)	(0.013)	(0.017)
Omitted Y Mean	0.407	0.216	0.216	0.210	0.199	0.175
Observations	2,155	9,198	9,189	8,879	8,257	5,153
	1850-1940	1950-2019	1960-2019	1970-2019	1980-2019	1990-2019
Panel B: African American Subsample						
First Two Girl	0.003	0.039**	0.039**	0.039**	0.047**	0.039**
	(0.002)	(0.007)	(0.007)	(0.007)	(0.007)	(0.011)
First Two Boy	0.010**	0.043**	0.044**	0.046**	0.046**	0.036**
	(0.003)	(0.005)	(0.005)	(0.006)	(0.006)	(0.008)
Omitted Y Mean	0.378	0.339	0.338	0.335	0.317	0.304
Observations	167,392	34,401	33,841	30,830	25,477	12,852

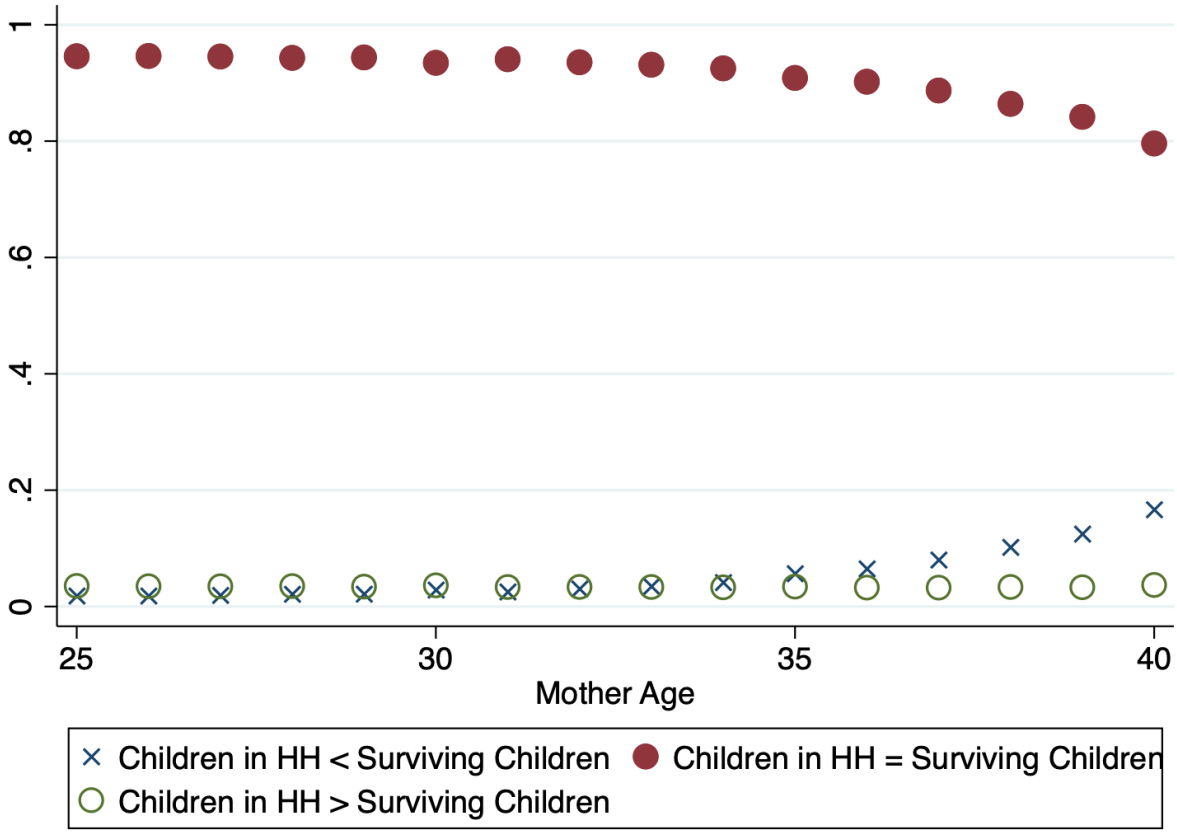
This table reports the results from a regression of an indicator for having three children on indicators for the first two children being girls the first two children being boys. The sample is limited to those whose race is coded as "Chinese", "Japanese," or "Other Asian or Pacific Islander" in Panel A and to "Black/African American/Negro" in Panel B. Data from 1850–1940 is used in column 1; no weights or year controls are used. Data from 1950–2019 is used in column 2, with the remaining columns being limited to 1960–2019, 1970–2019, 1980–2019, and 1990-2019, respectively; no weights or year controls are used—sample sizes differ largely by year due (in part) to the different samples used. The sample is limited to mothers with 2 or 3 children in all columns, and to mothers aged 30-35 whose youngest child is five years or older. Cases where the second and third child are the same age are excluded. All columns include mother age -by-mother birthplace fixed effects. Standard errors are included in parentheses and are clustered at the level of mother birthplace. ** 0.01, * 0.05.

Figure A1: Relationship between Number of Children in Household and Number of Children Ever Born, by Mother Age



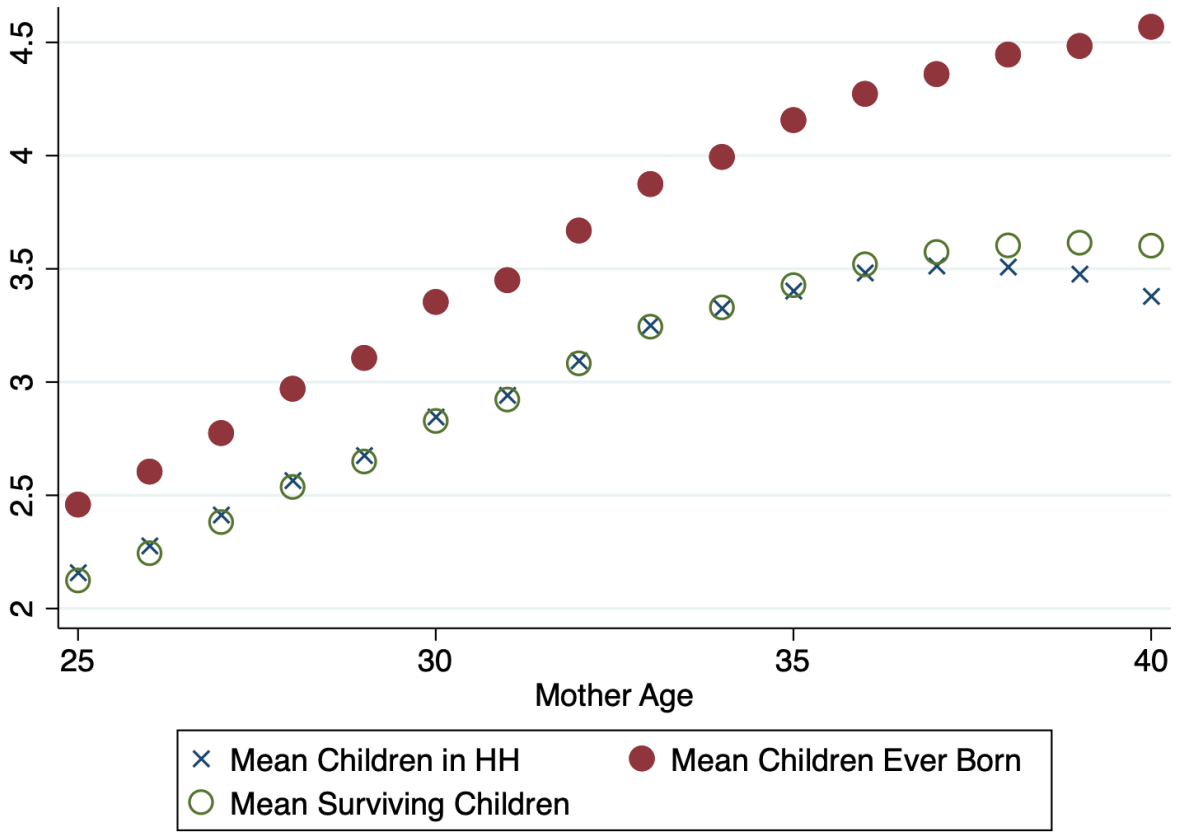
Notes: This figure shows the mean of the relationship (less than, equal to, and greater than) between the number of children in the household and the number of children ever born. Sample is limited to 1900 and 1910. Sample is not limited by age of youngest child.

Figure A2: Relationship between Number of Children in Household and Number of Surviving Children, by Mother Age



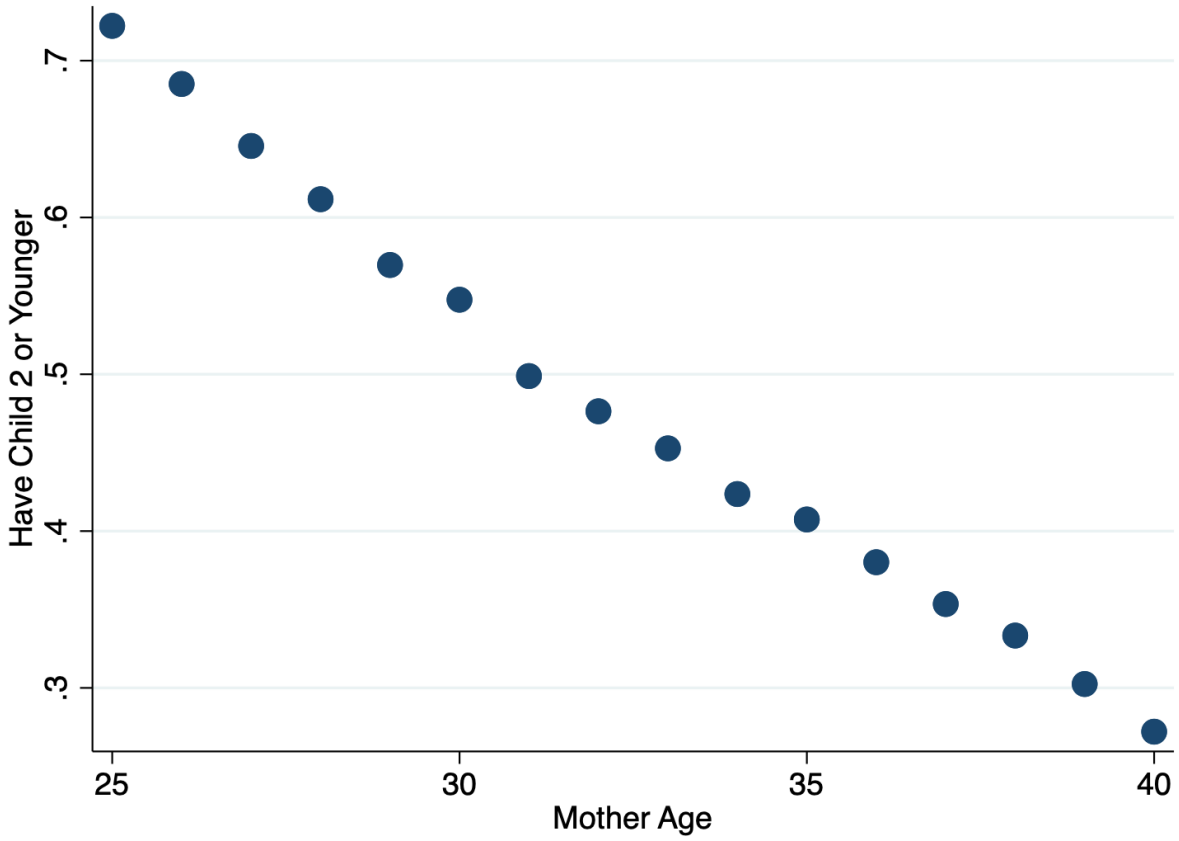
Notes: This figure shows the mean of the relationship (less than, equal to, and greater than) between the number of children in the household and the number of surviving children. Sample is limited to 1900 and 1910. Sample is not limited by age of youngest child.

Figure A3: Number of Children in Household, Children Ever Born, and Surviving Children, by Mother Age



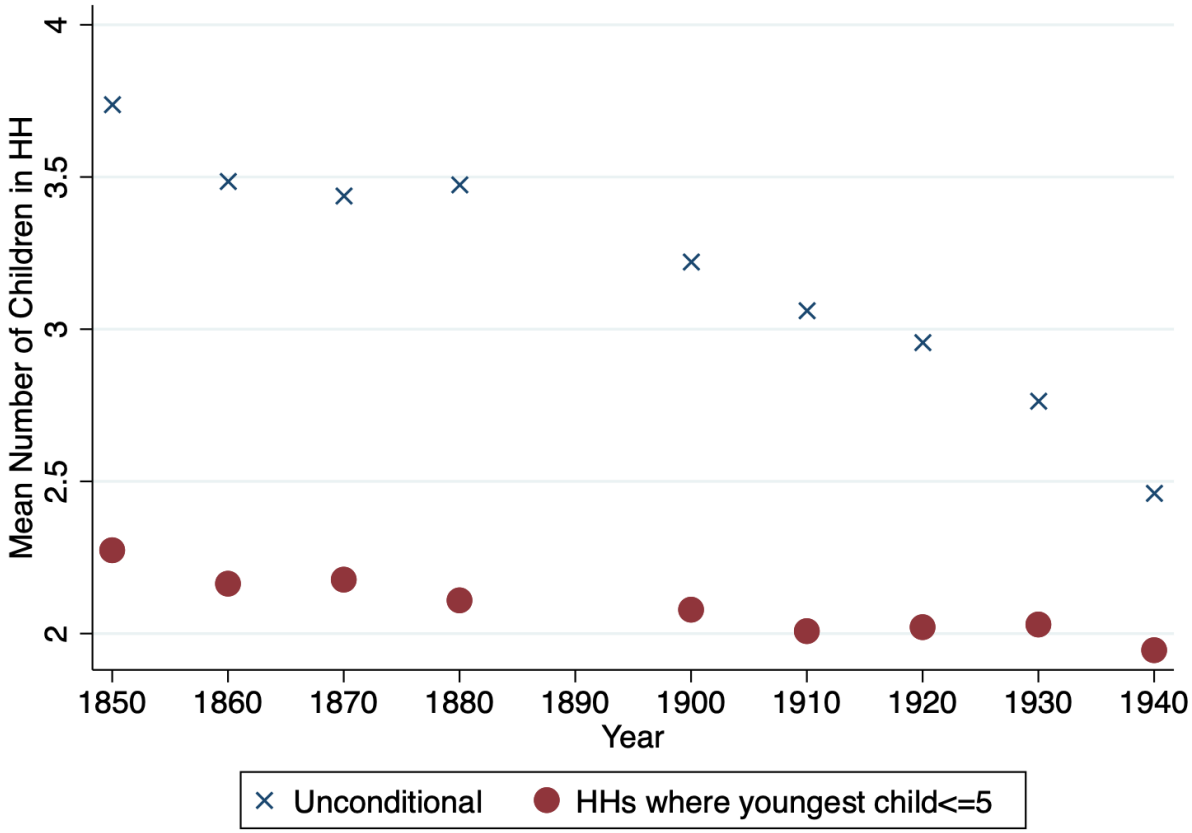
Notes: This figure shows the mean number of children in the household, the mean number of children ever born, and the mean number of surviving children. Sample is limited to 1900 and 1910. Sample is not limited by age of youngest child.

Figure A4: Fraction of Mothers with Child 2 or Younger, by Mother Age



Notes: This figure shows the fraction of mothers with a child 2 or younger, by mother age. Sample is not limited by age of youngest child.

Figure A5: Number of Children in Household, Mothers 30-35, by Year



Notes: This figure shows the mean number of children in the household separately by year. The red dots are not conditional on age of youngest child and the blue X's restrict to households in which the youngest child is five years or older. Sample is limited to mothers aged 30-35.