Birth Order in the Very Long-Run: Estimating U.S. Firstborn Premiums between 1850 and 1950*

Angela Cools
Davidson College

Jared Grooms
Brigham Young University

Krzysztof Karbownik
Emory University and NBER

Siobhan O'Keefe
Davidson College

Joseph Price
Brigham Young University and NBER

Anthony Wray
University of Southern Denmark

March 26, 2026

The nineteenth- and twentieth-century American family experienced tremendous demographic, economic, and institutional changes. By using birth order effects as a proxy for family environment, and linked census data on both men and women born between 1835 and 1920, we study how the family's role in human capital production evolved over this period. We find firstborn premiums for men for occupational outcomes, marriage, and fertility that are similar across census waves. Marriage and fertility premiums are similar for firstborn women who, however, did not experience the occupational improvements. Irrespective of gender effects of the family environment appear stable over about 100 years of U.S. history.

JEL: J13, J62, N30

Keywords: American family, birth order, parental investments, occupational outcomes, intergenerational mobility, marriage, fertility, women in history

*Cools: Department of Economics, Davidson College (email: ancools@davidson.edu); Grooms: Department of Statistics, Brigham Young University (email: jared.grooms19@gmail.com); Karbownik: Department of Economics, Emory University and NBER (email: krzysztof.karbownik@emory.edu); O'Keefe: Department of Economics, Davidson College (email: siokeefe@davidson.edu); Price: Department of Economics, Brigham Young University and NBER (email: joseph_price@byu.edu); Wray: Department of Economics, University of Southern Denmark (email: wray@sam.sdu.dk). We thank Carver Coleman and J.P. Lefgren for excellent research assistance, Scott Barkowski, Anne Sofie Beck Knudsen, Steven Rivkin, and Marian Vidal-Fernandez for helpful comments, and seminar and conference participants at Emory University, Kennesaw State University, UC Merced, the University of Melbourne, University of New Mexico, the University of New South Wales, and the University of Sydney, the Vanderbilt Economic History Network Conference, the Family and Education Workshop in Uppsala, Sweden, the 2nd Workshop in Economic History in Uppsala, Sweden, the Social Science History Association Annual Meeting in Chicago, IL, the Society for Economics of the Household Annual Meeting in London, England, the Southeastern Micro Labor Conference in Columbia, SC, the Southern Economic Association Annual Meeting in Houston, TX, and the 2024 NBER Summer Institute Children and Families Meeting.

1 Introduction

Economists and policymakers have long focused on understanding the human capital production function and now give more emphasis to the role of the family environment as a key input (Cunha et al. 2010; Almond and Currie 2011; Almond et al. 2018). Since children spend most of their early years interacting with parents and siblings rather than peers or teachers, it seems intuitive that household conditions contribute meaningfully to intermediate and long-run outcomes. Yet, reliable and consistently-reported measures of family environment paired with individual-level economic and social outcomes are rarely available over long time spans. Thus, we know very little about the contribution of the family environment to human capital formation and labor market outcomes in the US (or other countries) over the second half of the nineteenth and first half of the twentieth centuries, or how it may have changed over this long time period. Furthermore, we lack evidence about the effects of childhood circumstances on socioeconomic outcomes of women during this time period, whether we consider birth order or other inputs to human capital.

Such knowledge gaps exist despite the fact that the century-long period from 1850 to 1950 spanned the demographic and health transitions (Greenwood and Seshadri 2002; Costa 2015) and the emergence of modern economic growth (Galor 2005). Compared to the 1835 birth cohort, individuals born in 1920 experienced rising life expectancy (a 79% increase at birth) and education levels (a 50% increase in school enrollment rates and a 20-fold increase in high school graduation rates), and a decline in total fertility by 43% (United States. Bureau of the Census 1949; Snyder 1993). These trends were accompanied by institutional change – the abolition of slavery, the introduction of compulsory schooling laws, and the rise of women's suffrage – and historic macroeconomic shocks in the form of the demographic transition, the second industrial revolution, the Great Depression and the start of the baby boom. These policy changes and macroeconomic shocks may plausibly have interacted with the family environment either as complements, thereby enhancing the firstborn advantage, or as substitutes, thereby diminishing within-family inequality.

In this paper, we use birth order as a within-family quasi experiment to study the effects of family environment on long-run labor market and social outcomes for both men and women born

between 1835 and 1920 in the U.S. We ask the following research questions: How large were birth order effects on labor market outcomes? Were they stable or changing over time? Did birth order affect intergenerational occupational mobility? Were the effects limited to economic outcomes or did birth order affect educational and social outcomes as well? Finally, we consider two types of heterogeneity. First, we ask if the effects differed for men and women. Second, we consider differences by socioeconomic status of the childhood household. These questions allow us to understand the role that family environment played during the critical decades that shaped the modern US, particularly for women who were largely omitted in prior research due to data limitations.

We answer these questions by leveraging a combination of large-scale census data and a sibling fixed effects estimation strategy. Our samples link historical US full-count census data in 30-year intervals: 1850-1880, 1870-1900, 1880-1910, 1900-1930, 1910-1940, and 1920-1950.[1] While holding constant all time-invariant family characteristics, we compare outcomes for the firstborn child in a family, either male or female, with their later-born closest same-sex sibling. We always make within sex comparisons given large level and trend differences in both opportunities and outcomes between men and women over our extended time period. In doing so, we build on a long history of studying birth order as a proxy for family circumstances in social psychology and economics (Adler 1928; Behrman and Taubman 1986; Black et al. 2005a; Roher et al. 2015; Breining et al. 2020).

For men, we find remarkably stable birth order effects over the time frame considered. Our point estimate for the 1835-1850 birth cohorts implies a firstborn occupational premium of 2.6% of a standard deviation (SD) while the point estimate for the 1905-1920 birth cohorts is 3.0% of a SD. The four intermediate estimates range from 2.4% of a SD to 3.8% of a SD. We find similar patterns with alternative measures of occupational attainment such as the probability of working in a white collar occupation. Turning to intergenerational mobility, we find that firstborn sons were more likely to experience upward mobility and less likely to be either occupationally

[1] Since the 1890 US Census records were destroyed in a fire, we cannot create the 1860-1890 or 1890-1920 links. We also conduct robustness analyses with 20-year and 40-year links (Figures E.3 and E.4).

immobile or experience downward mobility, suggesting that family environment played a role in intergenerational mobility in U.S. history. The effects also extend beyond occupational premiums and into social outcomes. Firstborn males were less likely to remain single as adults, with effect sizes ranging from 9.3% to 13.7%, and more likely to have children, with effect sizes ranging from 1.4% to 2.8%.

For women, we find firstborns had similar social advantages over their sisters. They were around 9% less likely to remain single and they had increased fertility on both the intensive and extensive margins, with firstborn sisters 2.4% more likely to have children, having them 2.3% earlier, and having 5.7% more children. However, firstborn sisters did not experience the same occupational status premium as firstborn brothers. Compared to their younger sisters, their husbands are in lower status occupations by 2.7% of a SD while they themselves are less likely to be in the labor force by 4.8% (although this penalty vanishes for the 1905-1920 birth cohorts). Irrespective of the cohort, however, first born sisters are more likely to marry someone in the same occupation micro-class as their father, while their younger sisters marry men with higher ranked occupation.

Taken together evidence for both sexes suggests that the firstborn premia is not a modern phenomena. Male and female firstborns in the United States historically saw strong social advantages. However, while men also experienced a firstborn premia in the labor market, the firstborn sister was unable to translate her advantages into an improvement in economic status, at least until 1950, highlighting the importance of considering outcomes for both genders in historic settings. Our findings are robust to changes in the age ranges of children or mothers that we include in the sample, the choice of control group, adjusting for fertility stopping rules and selection, the use of different time spans for linked census data, or different linkage methods.

This paper contributes to the literature that investigates the role of family environment in human capital production through the lens of birth order. Prior work has documented stark differences between earlier and later born children in health (Jayachandran and Pande 2017; Brenøe and Molitor 2018; Abdellaoui et al. 2022), parental investments (Price 2008; Averett et al. 2011; Lehmann et al. 2018), test scores (Silles 2010; Haan et al. 2014), educational attainment (Black et al. 2005a;

de Haan 2010; Abdellaoui et al. 2022; Clark and Cummins 2024), earnings (Black et al. 2005a), income (Abdellaoui et al. 2022), occupational attainment (Nuevo-Chiquero et al. 2023; Clark and Cummins 2024), farm inheritance (Haws et al. 2023), non-cognitive skills and personality (Roher et al. 2015; Black et al. 2018), and risky behaviors and criminal activity (Argys et al. 2006; Breining et al. 2020).[2] Within this literature, only Black et al. (2005a), Black et al. (2018), Abdellaoui et al. (2022), Nuevo-Chiquero et al. (2023), and Clark and Cummins (2024) document effects on labor market outcomes, while only Nuevo-Chiquero et al. (2023) investigate if birth order matters for intergenerational transmission between parents and their children. Furthermore, only Black et al. (2005a) estimate long-run labor market effects separately for men and women using modern data from Norway. To the best of our knowledge there is no evidence on the effects of birth order on long-run labor market outcome of women in the U.S. or any country in historical setting of second half of the nineteenth and first-half of the twentieth century. We view this as an important gap that this paper intends to fill.

Our contribution to this literature is threefold. This is the first paper to document long-run effects of family environment over extended period of U.S. economic and social development. We use full count census data and high quality linkage procedures to study unprecedented number of American families with children born between 1835 and 1920. Importantly we trace the outcomes of these 85 cohorts – involving over 21 million unique individuals – into adulthood and measure their labor market success and social outcomes at ages 30 to 45 likely closely approximating life-cycle effects (see Bhuller et al. (2017) using administrative data from Norway and Figures A.1 and A.2a for approximate life-cycle occupational and fertility profiles for all the censuses we use in this paper) In complementary ongoing work, both Nuevo-Chiquero et al. (2023) and Clark and Cummins (2024) investigate birth order effects in different historical contexts: the Netherlands and England, respectively. We differ from these two papers in several dimensions. Most obviously,

[2]Most of these studies document that later-born children have worse outcomes. However, this result is not universal as Haan et al. (2014) find the opposite pattern in Ecuador where unlike in other countries, parents tend to invest more in later- rather than earlier-born children. This is especially pronounced in poor and low-educated families while in rich and high-educated families the effect reverses to a firstborn advantage. Given massive changes in development and poverty over the time period we study, it is plausible that firstborn premiums might have changed signs in the US over the course of our sample.

we study the US rather than a European country. Given economic and social differences, it's unclear if the results of European studies would translate to the US setting. We also have orders of magnitude larger sample sizes that are representative of the U.S. population e.g., the largest sample size in Clark and Cummins (2024) who uses families with rare surnames is under 55,000 while in Nuevo-Chiquero et al. (2023) it is under 28,000 despite using Dutch census data. Unlike the aforementioned papers, we also focus specifically on the evolution of first-born premiums over an extended period of this that covered important changes in U.S. social and economic development.

Second, to the best of our knowledge this is the first paper estimating the effects of family circumstances in childhood on long-run social and economic outcomes of a large sample of directly linked U.S. women in historical context. Given name changes at marriage almost no prior work was able to link women, at individual level, between their childhood and adulthood census.

Third, our results also speak to the literature on intergenerational inequality. Existing studies document correlations between parents and their children in education (Black et al. 2005b), preferences (Dohmen et al. 2012), health (Bütikofer et al. 2024), non-cognitive skills (Grönqvistt et al. 2017), socio-emotional skills (Attanasio et al. 2021), income (Chetty et al. 2014), and wealth (Black et al. 2020). Furthermore, Song et al. (2019) show that intergenerational occupational mobility in the US declined between 1850 and 2015 – with much of the decline occurring for the early birth cohorts that we consider in this paper. Ward (2023) re-examines these historical trends by accounting for measurement error and race, while Buckles et al. (2023b) additionally use high-recall links and study women. Their results suggest "low but increasing mobility" in contrast to Song et al. (2019) who find "high but decreasing mobility."[3] We contribute to this literature by documenting that firstborn premiums (intragenerational within-family inequality) influence the likelihood of the intergenerational transmission of occupations. To the best of our knowledge, only Nuevo-Chiquero et al. (2023) in parallel work estimated intergenerational effects of birth order on immobility but their sample is limited to men and based on Dutch data. Interestingly, our results for men are directly opposite to what they find. More broadly there is also limited evidence on the

[3]In this paper we use the same Census Tree links as Buckles et al. (2023b).

effects of childhood circumstances on intergenerational transmission and no evidence on this for women.

2 Data

This paper utilizes public-use versions of the 1850-1950 full-count decennial US Censuses of Population from IPUMS (Ruggles et al. 2024). The 1890 Census is omitted since the records were destroyed in a fire while enslaved Black individuals are not included in the 1850 sample. The full-count US decennial censuses include characteristics such as a person's name, birth year, birthplace, gender, race, place of residence, and occupation. The data are organized by household, allowing us to observe family relationships and characteristics of household members. We link individuals between the IPUMS versions of the full-count censuses using publicly available crosswalks from the Census Tree database (Price et al. 2021; Buckles et al. 2023a; Price et al. 2023) for years 1850-1940. We use crosswalk files from the IPUMS Multigenerational Longitudinal Panel (MLP) project in combination with the Census Tree crosswalks to link individuals to the 1950 census (Helgertz et al. 2024).[4]

The starting point for studying the effects of birth order is identifying the set of individuals who were born to the same parents. In our baseline specifications we define siblings as those sharing a common father (or, if missing, a common mother), conditional on the parent(s) and children being enumerated in the same childhood household in the census. In doing so, we exclude stepfathers from the pool of parents due to potential ambiguity in shared parentage. We define birth order among siblings based on relative ages and the order of enumeration within a household (those listed first are considered older when more than one child in a sibship reports the same age).

Next, we impose a series of household-level sample restrictions that address measurement error in birth order.[5] First, we drop households in which the oldest sibling is above the age of 15. As the share of children living with their parents declines during the teenage years, the presence of a

[4] In Figure E.5 we show that our results are robust to alternative linking procedures.

[5] Prior to the steps discussed below, we drop a small number of observations with duplicate HISTIDs, the IPUMS unique identifier, for which the uniqueness of an individual record is unclear. We also drop a handful of outlier households in which the number of siblings exceeds 15.

sibling older than age 15 increases the likelihood that an even older sibling had formed their own household and is not observed together with their parents. In such cases, the risk of misclassifying birth order is elevated.[6] Next, we drop sibling groups that include foreign-born children since the potential presence of siblings abroad is another source of measurement error in birth order. Finally, we drop mixed-race sibling groups and restrict to individuals who report their race as Black or White in order to avoid conflating birth order differences with unobserved racial differences.[7]

After imposing the sample restrictions, our estimation sample includes firstborn-male (female) sibships in which the firstborn and at least one younger male (female) sibling are linked to the census during adulthood.

Since census enumerators did not ask about biological parents, another potential source of measurement error in sibling status and birth order stems from the fact that we identify siblings based on grouping individuals by common social environment and guardianship (i.e. nurture not nature). For instance, given that familial ties are defined in relation to the household head, who was typically a man, some siblings may not share the same mother. Relatedly, our birth order measure may not align with biological birth order when children did not survive long enough to be enumerated in the census. Since we posit that the effects of birth order have a social or familial origin rather than a biological origin, and are thus interested in the birth order of surviving children, we are unconcerned by the deviation in these measures of birth order due to children who die in infancy or very early childhood.

We consider multiple long-run outcomes – all measured at ages 30-45 in our preferred specification – some of which are identical for men and women and some which differ by sex given historical differences in labor market participation of women. For men, our primary outcome variable is an occupational income score derived from the occupational title reported in the census

[6] Although commonly used in the extant literature the age cutoff of 15 is arbitrary and could be less plausible in historical setting. Therefore, as a robustness check, we also present results when we restrict the sample to families in which the oldest child in the household is age 12 or younger (Figures E.1 and E.2). Our main results barely change.

[7] The granularity of the racial categories varies across censuses. In the 1900-1920 censuses, some mixed-race individuals report their race as "mulatto." We exclude these individuals from the estimation sample to maintain a consistent and comparable classification of race across all censuses.

when an individual is ages 30-45 (Ruggles et al. 2024). We use the OCCSCORE variable from IPUMS and standardize it across the full population by gender within each census year.[8] The use of OCCSCORE facilitates comparisons over time while other socioeconomic measures such as income or education are not consistently reported across censuses. We also construct binary measures of occupational status based on the OCC1950 variable from IPUMS that follows the 1950 Census Bureau classification with 283 occupational categories.[9] For women, given that their labor force participation, and thus valid occupational information, was only 16.7% in 1880 Census and 30.6% in 1950 Census we consider the OCCSCORE of their spouse. We also experiment with alternative formulations for this outcome variable where we use spousal occupation only when a women does not have their own occupational code. We observe the extensive margin of labor supply for both sexes and code this variable as one if an individual is employed and zero otherwise. Between 1880 and 1950, labor force participation was consistently high for men and increased by 14 percentage points for women. When it comes to intergenerational outcomes we consider both immobility defined as one if an individual (or women's spouse) has the same microclass occupation as their father as well as upward mobility defined as one if an individual (or women's spouse) has microclass occupation higher than their father. We observe that immobility declined 22 percentage points for men and 25 percentage points for women between 1880 and 1950. Reflecting increased opportunities in the economy throughout this period, upward mobility increased 23 percentage points for men and 29 percentage points for women over the same time period.

Lastly, we construct social outcomes using information in the census. These outcomes are identical for men and women. We consider two marital and four fertility outcomes. When it comes to the former, we define an indicator for being single based on reported marital status as well as an indicator for being married to first born spouse. The latter variable investigates assortative match-

[8]OCCSCORE "assigns each occupation in all years a value representing the median total income (in hundreds of 1950 dollars) of all persons with that particular occupation in 1950." Zero represents missing (Ruggles et al. 2024).

[9]We use OCC1950 codes to assign an occupation to one of four categories: white collar, skilled, farmers, and unskilled. White-collar occupations include professionals (OCC1950 codes 0-99), managers (200-299), clerical workers (300-399), and sales workers (400-499). Skilled workers are craftsmen (500-599). Farmers include farm owners and tenants (OCC1950 code 100) in addition to farm managers (OCC1950 code 123). Unskilled workers include operatives (600-699), service workers (700-799), farm laborers and general laborers (800-970). Occupations with codes above 970 are excluded from the sample as they signify being out of the labor force.

ing based on birth order and whatever effects it envelopes.[10] For fertility, we define the probability of having any children, probability of having any children by age 30, number of children conditional on having any children, and adult's age at the birth of their first child to investigate shifts in birth timing as well as a proxy for marriage timing.

3 Methods

The main estimating equation compares a firstborn child of a given sex to their brother or sister who is closest in birth order and linked across censuses:

$$Y^{G}_{ijcms} = \beta_0 + \beta_1 \text{firstborn}_i + \mathbf{\Theta X_i} + \gamma_c + \delta_s + + \alpha_j + \varepsilon_{ijcs} \tag{1}$$

where i denotes individuals, j denotes families, c denotes birth cohort, and s denotes state of birth. We index Y_{ijcs} by G to highlight the fact that we always make same-sex comparisons either male-male or female-female. A vector of control variables, X_i, that is used in select specifications includes a race indicator as well as fixed effects for family size and father's occupational status. We use the same categorization procedure for fathers occupations as for occupational outcomes of sons. We do not control for mother's occupational status given that most mothers, especially early in our sample, did not work. All regressions control for child's year of birth (γ_c) and child's state of birth (δ_s) fixed effects. Our preferred specification further includes a family fixed effect, α_j, which effectively forces the comparison of outcomes to be within a family. The family fixed effects are appealing because they control for factors that are invariant within families and might be unobserved or observed with measurement error in historical records. For example, results in Clark and Cummins (2024) are often sensitive to the inclusion of family fixed effects highlighting the value of this more credible identification strategy in historical setting. Thus, in this equation, β_1 compares outcomes of the firstborn son (daughter) in the family with his (her) younger brother (sister) closest in age and linked to the census in adulthood. We cluster standard errors, ε_{ijcms}, at the family level.

There are two issues related to estimating birth order effects that need to be addressed. First,

[10] We code an individual as single if their reported marital status is "Never married/single." Other response categories are: married (with or without spouse present in the household), separated, divorced, or widowed.

family size can have its own effects on children other than directly through birth order. For example, Bagger et al. (2021) note that birth order and family size are jointly determined, i.e., one cannot manipulate family size while holding the within-family distribution of birth order constant. Second, families with same-sex children among their firstborns are more likely to have additional births — a feature that is often used as an instrument for family size (Angrist et al. 2010; Conley and Glauber 2006; Jones et al. 2023).[11] To address these issues, we include family fixed effects which control for family size and its potential effects on the outcomes. We also show estimates separately for smaller and larger families (Figures D.2 and D.3) which for most outcomes and periods are similar.

Another issue is that if a child requires costly parental investments, for example due to health problems (Ejrnaes and Pörtner 2004), parents may choose not to continue having children and thus last-born children may be negatively selected, thereby biasing the firstborn premium upward. We address this issue by excluding families in which the firstborn child is compared to the youngest child enumerated in the household and the results are virtually unchanged (Figures E.1 and E.2).

Finally, a more nuanced point relates to the interpretation of our estimates in the context of nineteenth and early-twentieth century inheritance laws and norms. Specifically, if primogeniture – a historical practice in which the eldest son inherited a family's land and titles – was common during our study period, then the firstborn advantage in male-male sample could solely reflect prevailing laws and customs rather than effects of family environment. Shammas et al. (1987) highlights (in their Table 3.1) that by 1790 most colonies had equal rights to inheritance and that "by 1800 in most states, sons and daughters received equal shares in real and personal property." This suggest that although primogeniture was initially in place in the US, it was no longer a common practice by the time the earliest cohorts in our sample were born in 1835. Furthermore, Haws et al. (2023) find a limited role of birth order in farm inheritances in the US. While their 1900-1920 estimates imply a small birth order penalty, it vanishes in census links over 30 to 40 years. Based

[11] Black et al. (2005a) also use this instrument in their analysis of family size and note that same-sex siblings could have an independent effect on child outcomes and that the birth order effects are both stable when instrumenting for completed family size and orders of magnitude larger than the family size effects. See Tan (2019) for a discussion of family size effects in a historical context.

on these facts it is unlikely that inheritance customs played a meaningful role in our setting and thus we prefer to interpret the estimates as the effects of family environment.

4 Results

4.1 Birth Order Effects for Men

We begin by presenting birth order effects on long-run labor market outcomes for men for six 30-years census links, beginning with 1850-1880 and ending with 1920-1950. Figure 1a displays these results from specification that includes both family fixed effects and controls. First-born males have higher occupational income score (panel A) and labor force participation (panel D) than their next born younger brother, with the exception of the first census pairs where labor force participation estimates are imprecise. The extensive margin effects sizes are small: ranging from 0.1 to 0.3%. The primary reason for this is very high levels of male labor force participation in these samples – between 95.4% in 1940 and 97.9% in 1910. On the other hand, for occupational income we find persistent, always statistically significant, and similar over time first-born premia. The coefficients range from 2.4% of a SD in 1940 target census to 3.8% of a SD in 1930 target census and there is no clear trend in our estimates. The pooled coefficient for individuals born between 1835 and 1920 implies 3.0% of a SD first-born occupational advantage. We view these magnitudes as economically meaningful. For example, at the beginning of our sample period, this occupational upgrading could represent moving from a general laborer to an apprentice in a trade, and at the end this would be consistent with moving from a sales clerk to a general clerical worker.[12]

In Figure B.1 we further investigate whether the occupational effects of birth order are con-

[12] In Figure E.3 we verify that this pattern is not an artifact of using the 30-years links. We find a very similar pattern of estimates for the firstborn premium when using 20-year links, which include additional census waves, with slightly larger effects for the earliest (1850-1870) and latest (1920-1940) census waves. We do not include the results based on the 20-year links among our preferred estimates because they measure occupational outcomes at ages 20-35 (rather than ages 30-45) which might be an imperfect proxy for life-cycle occupational income and labor force participation – especially for the more recent cohorts for whom educational attainment increased substantially. This concern is particularly relevant for birth order as the later-born child is more likely to be on the steeper part of the curve for life-cycle labor market outcomes. Therefore, it makes sense that our estimates are larger for the 20-year links. Figure E.3 also reports estimates using 40-years links in which occupational outcomes are measured at ages 40-55. Again, we find reasonably stable firstborn premiums starting with an estimate of 3.9% of a SD for the 1860-1900 link and ending with an estimate of 3.1% of a SD for the 1900-1940 link.

centrated in specific parts of the distribution. We group occupations into four categories based on OCC1950 codes (see Section 3 for details). Firstborns are less likely to work as farmers or unskilled laborers, with larger effect sizes for farming in earlier census waves and increasing magnitudes over time for unskilled occupations. This pattern is consistent with the changing occupational structure of the economy as farming was dominant in the nineteenth century while unskilled employment grew in other sectors in the early twentieth century. Furthermore, the figure shows that firstborns were more likely to enter skilled or white collar occupations and do so at a consistent rate across the time span of our analysis (5.1-6.4% for white collar occupations and 3.1-4.4% for white collar or skilled occupations). Thus, it appears that firstborn premiums affect the entire distribution of occupational outcomes by keeping individuals out of relatively lower-skilled jobs and enabling them to enter higher-status occupations.

Figures 1b and 1c investigate if birth order moderates intergenerational occupational mobility. We follow the approach used by Song et al. (2019) and define immobility as the likelihood that a son stayed in the same micro-class occupation as their father.[13] We find remarkably stable immobility estimates that are always negative and almost always statistically significant at conventional levels. One reason for such stability is the fact that as immobility rates decline over time (from 42.5% for 1850-1880 linked sample to 20.8% for the 1920-1950 linked sample) so do the estimated treatment effects (Table 1). The descriptive finding of declining microclass mobility over time is consistent with Song et al. (2019). The effect sizes range from 2.8% for the 1880-1910 linked sample to 0.6% for the 1870-1900 linked sample. In contrast to our results, Nuevo-Chiquero et al. (2023) find that firstborns in the Netherlands are more likely to end up in the same occupation as their father suggesting that different processes could have been shaping occupational mobility in the U.S. and Europe during this time period.

We then compare the occupational income score of fathers and sons to examine effects on absolute upward (Panel C) mobility. We find statistically significant increases in the probability of upward mobility ranging from 1.3% to 3.4% which again are remarkably stable over time and do

[13]The micro-class occupational scheme aggregates the 283 occupational categories of the IPUMS OCC1950 variable to 70 groups (Song et al. 2019).

not exhibit any particular trend. The pooled estimate of 1.0 pp (or 2% effect size) is consistent with a firstborn advantage that nevertheless appears to be smaller than the approximately 10% effect of lower health capital in childhood on absolute mobility that has been estimated for a similar time period in the UK (Karbownik and Wray 2024). At the same time, hospitalizations investigated by (Karbownik and Wray 2024) are clearly a larger shock compared with birth order differences which estimate, under specific assumptions (Becker and Tomes 1976), marginal effects of family resources on children.

Consistent with improved labor market outcomes, we also find evidence of a firstborn premia for human capital although these are measured differently in different census years. Panel A of Figure C.1 shows that firstborns are 0.1 to 0.4 pp more likely to be literate. Given the high reported rates of literacy in the census, the effects sizes are small (0.1 to 0.4%), but consistent with increased parent investment leading to improved human capital. Panel B focuses on recent censuses where years of educational attainment can be observed. Here we find stronger evidence of the firstborn advantage: a 3.1% to 6.5% increase in the probability of completing a high school degree. Using the Song et al. (2019) adjusted educational score to compare estimates across the entire time period, we see a consistent educational premia for firstborns (0.7 to 2.3%). Higher levels of human capital helps explain why firstborn men have higher occupational income scores and especially why they are more likely to be employed as a skilled craftsman or white collar worker, occupations that require specialized skills. On the other hand, considering the Song et al. (2019) score we observe increasing birth order effects for the first four census pairs that then stabilize for the last two census pairs. Given that our occupational outcomes findings appeared much more stable, the human capital channel is unlikely to be the only driver behind the labor market findings.

Finally, we examine effects of birth order on marriage and fertility outcomes. Figure 3 presents these results for four outcomes probability of being single (panel A), probability of having any children by age 30 (panel B), probability of being married to a firstborn spouse (panel C), and number of children conditional on having any children (panel D). When it comes to probability of being single and assortative matching along the birth order dimension we find very stable results for those

born between 1835 and 1910. First-born males are less likely to remain single in adulthood (effect sizes of -9.3 to -13.7%) and they are more likely to marry first-born spouse (effect sizes of 2.3 to 6.4%) compared to their closest in age younger brother. At the same time, for those linked between 1920 childhood and 1950 adulthood censuses the marriage premium declines modestly while the assortative matching effect goes to zero. In this last set set of matches, the probability of remaining single decreases by around 25 percent, implying the firstborn advantage is less meaningful in the high-marriage environment of the post-World War II baby boom.

4.2 Birth Order Effects for Women

Our unique set of matches allow us to extend this analysis to include women, which other work on historical birth order is unable to do. During this time period, women's economic status was largely influenced by whether and who they married. Labor force participation, especially for married white women, was often reserved for cases of extreme necessity (Goldin 1990). Therefore, we start by presenting the effect of birth order on women's social outcomes.

Figures 4a and 4b show the birth order effect on the probability a woman is never married and the probability she has children in her household by age 30. Similar to the results for men, the first born premium for marriage is strong and consistent throughout the study period but with notably weaker effect during the post World War II baby boom. For cohorts of sisters born between 1835 through 1920, the first-born sister is 1.4 percentage points less likely to be never-married when observed as an adult. Relative to the sample mean, the pooled estimate represents close to a 10% decrease in the probability of being single as an adult. Given the lack of economic opportunity for women in these cohorts, especially women from lower socioeconomic status backgrounds (Olivetti et al. 2024), this could have represented increased economic security for the older sisters. Consistent with an increased probability of marriage, firstborn sisters are also 2.8 percentage points (or 4.1%) more likely to have had children before 30. Note that this estimate cannot be driven by age difference between first and nextborn sisters given that we observe all women at or after the age of 30 in the adulthood census.

While the birth order effect for social outcomes is similar for males and females, the effects di-

verge when looking at household socioeconomic status. Unlike first-born men, whose occupation score is higher than their brothers, Figure 2a shows that firstborn sisters are generally married to men with lower occupational scores than their younger sisters' husbands. This presents a puzzle: why do firstborn sisters not see an economic status premium, like their male counterparts? Figures 2b and 2c shed light on the dynamics that cause the observed drop in household status. The firstborn sister is significantly more likely to marry someone in the same micro-class of occupations as their father. This contrasts with the overall pattern of marriages in the US over this period. Bailey and Lin (2024) show that this type of intergenerational persistence falls over our sample period, but oldest daughters seem to resist this trend. In turn, their younger sisters are able to take advantage of increased opportunities for upward mobility and the changing economy.

While husband's occupation is an important marker of status for these cohorts of women, it is not the only quality of life measure we observe. Figure 2d shows that firstborn sisters are less likely to be in the labor force (except for the 1920-1950 census pair). Given the types of work available to women as well as the social and legal restrictions on married women's participation in the workforce (Goldin 1990), this is evidence first-born sisters are doing better than their younger sisters on at least one margin. Similarly, Figure C.1 shows older sisters had higher human capital than their younger sisters. While the point estimates for literacy are imprecise they are positive and generally similar to those for men (reported baseline literacy rates were similar for both sexes). After the Census starts recording years of education in 1940, the older sister is 2.1 to 3.0% percent more likely to have completed high school. Similar to modern birth order estimates (Black et al. 2005a) and our own estimates for men, this implies that parental investments in the firstborn child lead to higher human capital, regardless of gender.

An alternative hypothesis is that older daughters, potentially facing increased parental marriage expectations or the desire to escape in-home caretaking responsibilities, marry more quickly.[14] Their decreased search time, while increasing the likelihood of marriage, may decrease the potential gains matching with a higher-earning spouse. While we cannot observe age at marriage

[14]Risk aversion, if higher among firstborn girls, could also lead to earlier marriage. (Schmidt 2008).

directly, we can back out the correlated measure of age at first birth. Panel C of Table 4 shows that older sisters have their first birth about 0.55 (in 1880) to 0.30 (in 1950) years before their younger sisters (or 2.5% and 1.3%, respectively); in a time period before highly reliable birth control, this likely implies that they also married younger.

5 Robustness Checks

This section presents a series of analyses to gauge the robustness of our main results. We have already documented in Figures E.3 and E.4 that our results are largely unaffected by changing the gap between the baseline and follow-up years when linking census waves. Therefore, for brevity, we focus on the 30-year links and family fixed effects models.

We first address the possibility that birth order is assigned incorrectly when older siblings have formed their own households, and therefore are not enumerated in the census with their parents and younger siblings. The green star in Figures E.1 and E.2 restricts the maximum age of a child in the household to be 12 years in the baseline census (rather than 15 years) while the yellow "x" further restricts the sample to younger mothers who were between the ages 18-30 when any of their children were born. Despite a reduction in the sample size by half, the magnitude of our estimates and their pattern over time remain very similar to our preferred specification.

Next we consider different comparison groups of siblings. Recall that our primary results compares the firstborn child to the younger same-sex sibling closest in birth order who can be linked to outcomes in adulthood, ignoring siblings born in between. Our conclusions remain unchanged when we compare the firstborn child to all later-born same-sex siblings linked to outcomes (orange square in Figures E.1 and E.2).

We also address the issue of endogenous fertility stopping. We exclude families with exactly two children (turquoise triangle) or families for which the younger sibling included in the estimation sample is the youngest in the household (pink diamond). These additional sample restrictions do not affect our results.

Next, we also explore the sensitivity of our results to the choice of linking method and concerns about the representativeness of our linked samples. Figure E.5 shows results remain unchanged

when we restrict to the subset of links in the Census Tree dataset that are included in the Census Linking Project crosswalks (Abramitzky et al. 2020), are generated by the XGBoost machine learning algorithm, or appear in the Family Tree. Overall, we conclude that our results are robust to a reasonable set of alternative sample and estimation choices.

6 Conclusions

Birth order affects a multitude of short- and long-run outcomes in contemporary data spanning developed and developing countries. Despite its importance for economic outcomes and human capital development, there is limited evidence if and to what extent birth order mattered historically, when fertility and mortality rates were much higher and the quality of parental inputs and knowledge about returns to investment in children were much lower. We contribute to this literature by estimating firstborn premiums for US-born individuals – both men and women – between 1835 and 1920, paying particular attention to the dynamics of the estimates over time.

We find evidence that the firstborn premium was present among families as early as those raising their children in the 1830s. Remarkably, the effects identified for these cohorts are very similar to those who grew up in the first decades of the twentieth century. They extend beyond occupational outcomes and into the social domain including the likelihood of remaining single as well as fertility at both the extensive and intensive margins. The effects on social outcomes are very similar for males and females in our sample. Finally, we find that birth order affected the intergenerational transmission of occupations. Irrespective of the outcome and the period in US history, we find that firstborn males were always better off compared to their younger brothers. The same cannot be true for firstborn females who in terms of occupational outcomes of their spouses appear to be worse off.

Our results have important implications for understanding the role that family resources play in shaping human capital and labor market outcomes of children. To the extent that birth order credibly estimates the marginal returns to family environment, our findings suggest that these returns changed relatively little over an extended period of time during which the American family experienced dramatic cultural, economic, and social changes.

References

Abdellaoui, Abdel, Oana Borcan, Pierre-Ander Chiappori, and David Hugh-Jones (2022) "Trading social status for genetics in marriage markets: Evidence from UK Biobank."

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, Santiago Pérez, and Myera Rashid (2020) "Census Linking Project: Version 2.0 [dataset].," https://censuslinkingproject.org.

Adler, Alfred (1928) "Characteristics of the first, second, and third children," *Children*, Vol. 3, No. 5, p. 14.

Almond, Douglas and Janet Currie (2011) "Human capital development before age five," in David Card and Orley Ashenfelter eds. *Handbook of Labor Economics*, Vol. 4B, pp. 1315–1486.

Almond, Douglas, Janet Currie, and Valentina Duque (2018) "Childhood circumstances and adult outcomes: Act II," *Journal of Economic Literature*, Vol. 56, No. 4, pp. 1360–1446.

Angrist, Joshua, Victor Lavy, and Analia Schlosser (2010) "Multiple experiments for the causal link between the quantity and quality of children," *Journal of Labor Economics*, Vol. 28, No. 4, pp. 773–824.

Argys, Laura, Daniel Rees, Susan Averett, and Benjama Witoonchart (2006) "Birth order and risky adolescent behavior," *Economic Inquiry*, Vol. 44, No. 2, pp. 215–233.

Attanasio, Orazio, Aureo de Paula, and Alessandro Toppeta (2021) "The persistence of socio-emotional skills: Life cycle and intergenerational evidence," Working Paper 27823, National Bureau of Economic Research.

Averett, Susan, Laura Argys, and Daniel Rees (2011) "Older siblings and adolescent risky behavior: Does parenting play a role?," *Journal of Population Economics*, Vol. 24, No. 3, pp. 957–978.

Bagger, Jesper, Javier Birchenall, Hani Mansour, and Sergio Urzu'a (2021) "Education, birth order and family size," *Economic Journal*, Vol. 131, No. 633, pp. 33–69.

Bailey, Martha J and Peter Z Lin (2024) "Marital matching and women's intergenerational mobility in the late 19th and early 20th century US," *The Economic History of American Inequality: New Evidence and Perspectives*, pp. 165–196.

Becker, Gary and Nigel Tomes (1976) "Child endowments and the quantity and quality of children," *Journal of Political Economy*, Vol. 84, No. 4, pp. S143–S162.

Behrman, Jere and Paul Taubman (1986) "Birth order, schooling, and earnings," *Journal of Labor Economics*, Vol. 4, No. 3, pp. 121–145.

Bhuller, Manudeep, Magne Mogstad, and Kjell G Salvanes (2017) "Life-cycle earnings, education premiums, and internal rates of return," *Journal of Labor Economics*, Vol. 35, No. 4, pp. 993–1030.

Black, Sandra, Paul Devereux, Petter Lundborg, and Kaveh Majlesi (2020) "Poor little rich kids? The role of nature versus nurture in wealth and other economic outcomes and behaviors," *Review of Economic Studies*, Vol. 87, No. 4, pp. 1683–1725.

Black, Sandra, Paul Devereux, and Kjell Salvanes (2005a) "The more the merrier? The effects of family size and birth order on children's education," *Quarterly Journal of Economics*, Vol. 120, No. 2, pp. 669–700.

——— (2005b) "Why the apple doesn't fall far: Understanding intergenerational transmission of human capita," *American Economic Review*, Vol. 95, No. 1, pp. 437–449.

Black, Sandra, Erik Grönqvist, and Öckert Björn (2018) "Born to lead? The effect of birth order on non-cognitive skills," *Review of Economics and Statistics*, Vol. 100, No. 2, pp. 274–286.

Breining, Sanni, Joseph Doyle, David Figlio, Krzysztof Karbownik, and Jeffrey Roth (2020) "Birth order and delinquency: Evidence from Denmark and Florida," *Journal of Labor Economics*, Vol. 38, No. 1, pp. 95–142.

Brenøe, Anne and Ramona Molitor (2018) "Birth order and health of newborns: What can we learn from Danish registry data?," *Journal of Population Economics*, Vol. 31, pp. 363–395.

Buckles, Kasey, Adrian Haws, Joseph Price, and Haley E.B. Wilbert (2023a) "Breakthroughs in Historical Record Linking Using Genealogy Data: The Census Tree Project," NBER Working Paper 31671.

Buckles, Kasey, Joseph Price, Zachary Ward, and Haley Wilbert (2023b) "Family Trees and Falling Apples: Historical Intergenerational Mobility Estimates for Women and Men," Working Paper.

Bütikofer, Aline, Rita Ginja, Krzysztof Karbownik, and Fanny Landaud (2024) "(Breaking) intergenerational transmission of mental health," *Journal of Human Resources*, Vol. forthcoming.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014) "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *Quarterly Journal of Economics*, Vol. 129, No. 4, pp. 1553–1623.

Clark, Gregory and Neil Cummins (2024) "Birth order and social outcomes, England, 1680-2024."

Conley, Dalton and Rebecca 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," *Journal of Human Resources*, Vol. 41, No. 4, pp. 722–737.

Costa, Dora (2015) "Health and the economy in the United States from 1750 to the present," *Journal of Economic Literature*, Vol. 53, No. 3, pp. 503–570.

Cunha, Flavio, James Heckman, and Susanne Schennach (2010) "Estimating the technology of cognitive and noncognitive skill formation," *Econometrica*, Vol. 78, No. 3, pp. 883–931.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde (2012) "The intergenerational transmission of risk and trust attitudes," *Review of Economic Studies*, Vol. 79, No. 2, pp. 645–677.

Ejrnaes, Mette and Claus Pörtner (2004) "Birth order and the intrahousehold allocation of time and education," *Review of Economics and Statistics*, Vol. 86, No. 4, pp. 1008–1019.

Galor, Oded (2005) "Chapter 4 From Stagnation to Growth: Unified Growth Theory," in Philippe Aghion and Steven N. Durlauf eds. *Handbook of Economic Growth*, Vol. 1: Elsevier, pp. 171–293.

Goldin, Claudia (1990) *Understanding the gender gap: An economic history of American women (NBER Series on Long-Term Factors in Economic Development)*: Oxford University Press.

Greenwood, Jeremy and Ananth Seshadri (2002) "The US demographic transition," *American Economic Review*, Vol. 92, No. 2, pp. 153–159.

Grönqvistt, Erik, Björn Öckert, and Jonas Vlachos (2017) "The intergenerational transmission of cognitive and noncognitive abilities," *Journal of Human Resources*, Vol. 52, No. 4, pp. 887–918.

de Haan, Monique (2010) "Birth order, family size and educational attainment," *Economics of Education Review*, Vol. 29, No. 4, pp. 576–588.

de Haan, Monique, Erik Plug, and Jose Rosero (2014) "Birth order and human capital development: Evidence from Ecuador," *Journal of Human Resources*, Vol. 49, No. 2, pp. 359–392.

Haws, Adrian, David Just, and Joseph Price (2023) "Who (actually) gets the farm? Intergenerational farm succession in the United States," NBER Working Paper 31604.

Helgertz, Jonas, Nesile Ozder, Steven Ruggles, John Robert Warren, Catherine A. Fitch, J. David Hacker, Matt A. Nelson, Joseph P. Price, Evan Roberts, and Matthew Sobek (2024) *IPUMS Multigenerational Longitudinal Panel: Version 1.2 [dataset]*, Minneapolis, MN: IPUMS.

Jayachandran, Seema and Rohini Pande (2017) "Why are Indian children so short? The role of birth order and son preference," *American Economic Review*, Vol. 107, No. 9, pp. 2600–2629.

Jones, Todd, Matthew Millington, and Joseph Price (2023) "Changes in parental gender preference in the USA: Evidence from 1850 to 2019," *Journal of Population Economics*, Vol. 36, No. 4, pp. 3057–3070.

Karbownik, Krzysztof and Anthony Wray (2024) "Lifetime and intergenerational consequences of poor childhood health," *Journal of Human Resources*, Vol. forthcoming.

Lehmann, Jee-Yeon, Ana Nuevo-Chiquero, and Marian Vidal-Fernandez (2018) "The early origins of birth order differences in children's outcomes and parental behavior," *Journal of Human Resources*, Vol. 53, No. 1, pp. 123–156.

Nuevo-Chiquero, Ana, Marian Vidal-Fernandez, and Jee-Yeon Lehmann (2023) "The birth order effect: A modern phenomenon."

Olivetti, Claudia, M Daniele Paserman, Laura Salisbury, and E Anna Weber (2024) "The emergence of the socioeconomic gradient in women's marriage outcomes," *Review of Economics and Statistics*, pp. 1–29.

Price, Joseph (2008) "Parent-child quality time: Does birth order matter?," *Journal of Human Resources*, Vol. 43, No. 1, pp. 240–265.

Price, Joseph, Kasey Buckles, Adrian Haws, , and Haley Wilbert (2023) *The Census Tree [dataset]*, Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Price, Joseph, Kasey Buckles, Jacob Van Leeuwen, and Isaac Riley (2021) "Combining family history and machine learning to link historical records: The Census Tree data set," *Explorations in Economic History*, Vol. 80, p. 101391.

Roher, Julia, Boris Egloff, and Stefan Schmukle (2015) "Examining the effects of birth order on personality," *Proceedings of the National Academy of Sciences*, Vol. 112, No. 46, pp. 14224–14229.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rodgers, and Megan Schouweiler (2024) *IPUMS USA: Version 15.0 [dataset]*, Minneapolis, MN: IPUMS.

Schmidt, Lucie (2008) "Risk preferences and the timing of marriage and childbearing," *Demography*, Vol. 45, No. 2, pp. 439–460.

Shammas, Carole, Marylynn Salomon, and Michel Dahlin (1987) *Inheritance in America: From colonial times to the present*, New Brunswick and London: Rutgers University Press.

Silles, Mary (2010) "The implications of family size and birth order for test scores and behavioral development," *Economic of Education Review*, Vol. 29, No. 5, pp. 795–803.

Snyder, Thomas D (1993) *120 years of American education: A statistical portrait*: US Department of Education.

Song, Xi, Catherine Massey, Karen Rolf, Joseph Ferrie, Jonathan Rothbaum, and Yu Xie (2019) "Long-term decline in intergenerational mobility in the United States since the 1850s," *Proceedings of the National Academy of Sciences*, Vol. 117, No. 1, pp. 251–258.

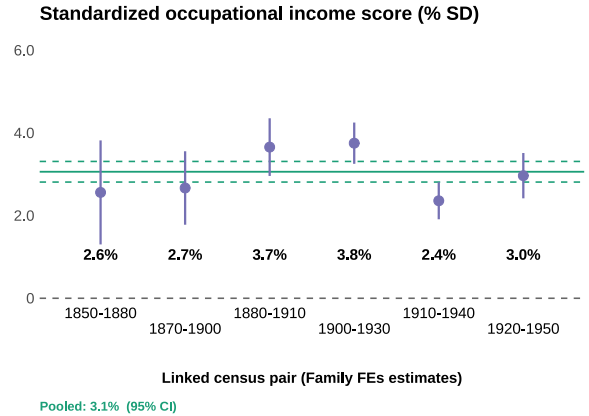
Tan, Hui Ren (2019) "More is less? The impact of family size on education outcomes in the United States, 1850-1940," *Journal of Human Resources*, Vol. 54, No. 4, pp. 1154–1181.

United States. Bureau of the Census (1949) *Historical statistics of the United States, 1789-1945: a supplement to the Statistical abstract of the United States*, Vol. 789: US Government Printing Office.

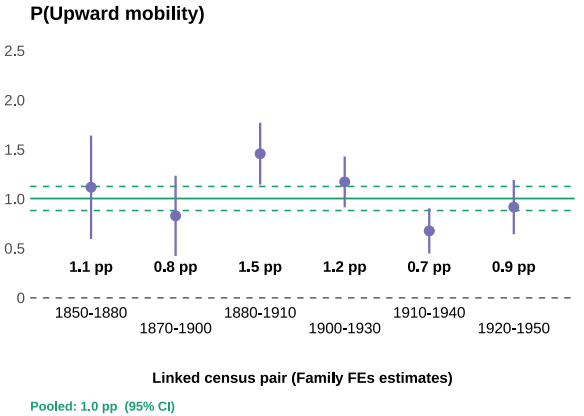
Ward, Zachary (2023) "Intergenerational Mobility in American History: Accounting for Race and Measurement Error," *American Economic Review*, Vol. 113, No. 12, pp. 3213–48.

7 Figures

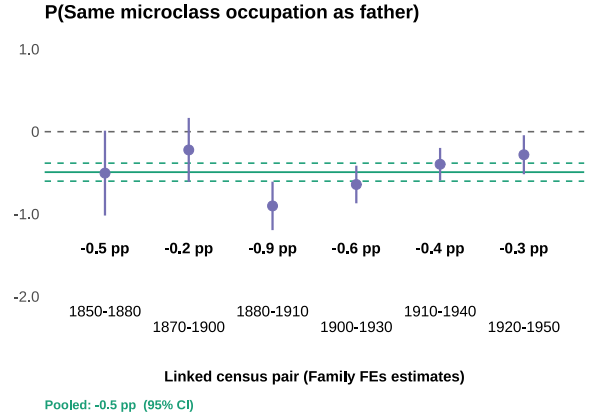
Figure 1. Birth order effects on occupational outcomes for men



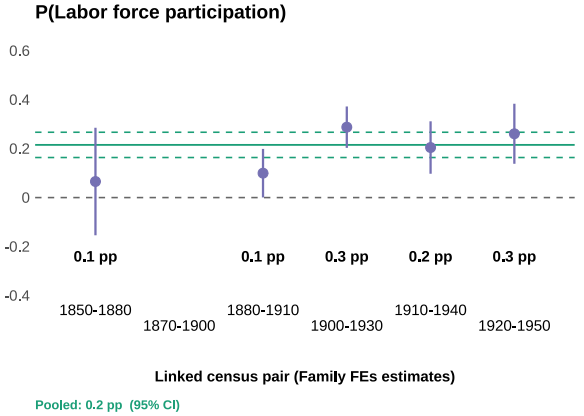
(a) Firstborn premium for occupational income score



(c) Firstborn effect on upward mobility



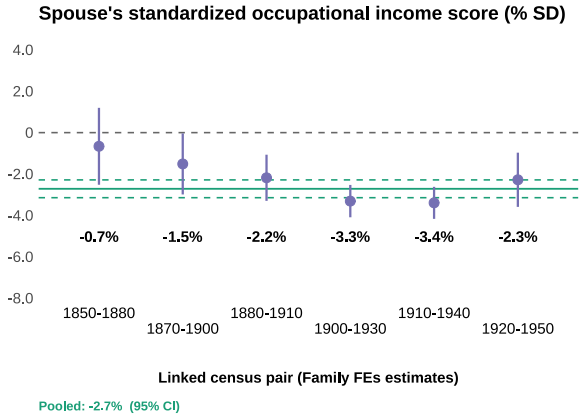
(b) Firstborn effect on occupational immobility



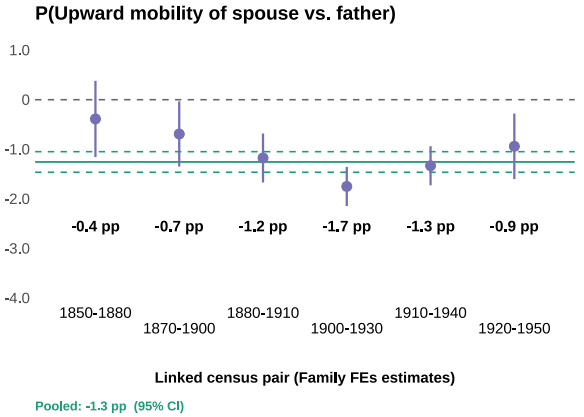
(d) Firstborn effect on labor force participation

Notes: Figure 1a shows the firstborn premium in standardized occupational income score. Figure 1b displays the effect of being firstborn on the probability of having the same microclass occupation as one's father. Figure 1c presents the effect of being firstborn on upward occupational mobility. Figure 1d shows the effect of being firstborn on labor force participation. All figures include 95% confidence intervals and show results from both separate linked census samples and a pooled sample.

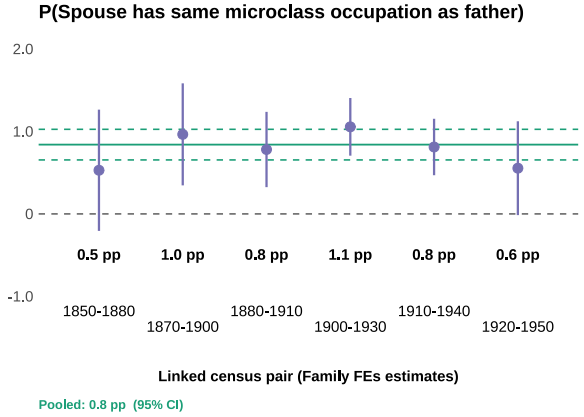
Figure 2. Birth order effects on occupational outcomes for women



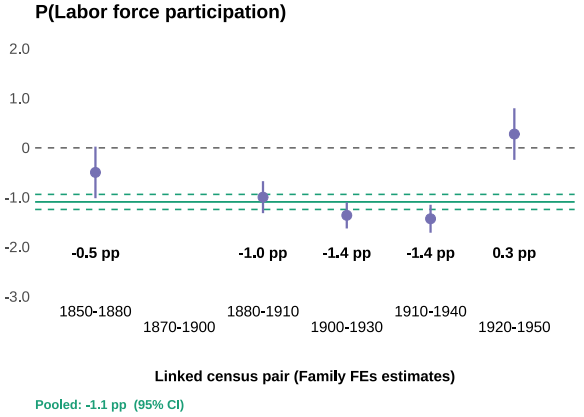
(a) Firstborn premium for spouse's occupational income score



(c) Firstborn effect on spouse's upward mobility



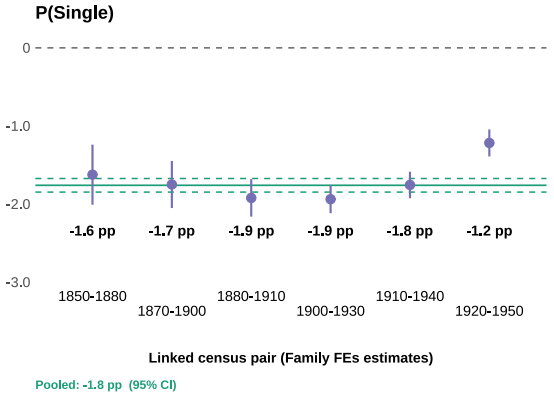
(b) Firstborn effect on spouse's occupational immobility



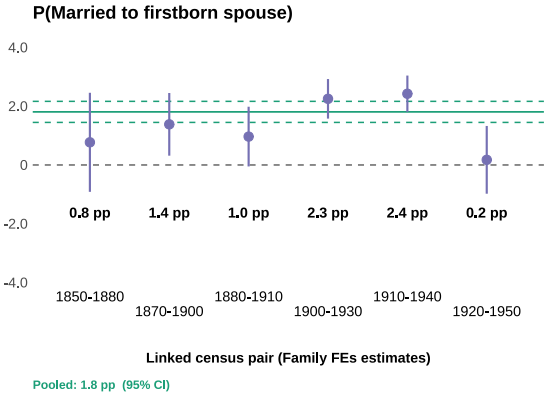
(d) Firstborn effect on labor force participation

Notes: Figure 2a shows the firstborn premium in spouse's standardized occupational income score. Figure 2b displays the effect of being firstborn on the probability of spouse having the same microclass occupation as one's father. Figure 2c presents the effect of being firstborn on spouse's upward occupational mobility. Figure 2d shows the effect of being firstborn on labor force participation. All figures include 95% confidence intervals and show results from both separate linked census samples and a pooled sample.

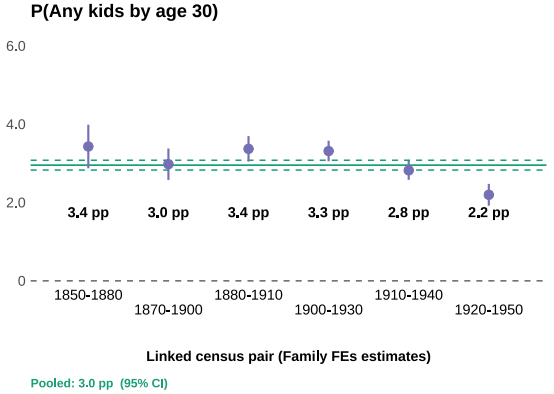
Figure 3. Birth order effects on social outcomes for men



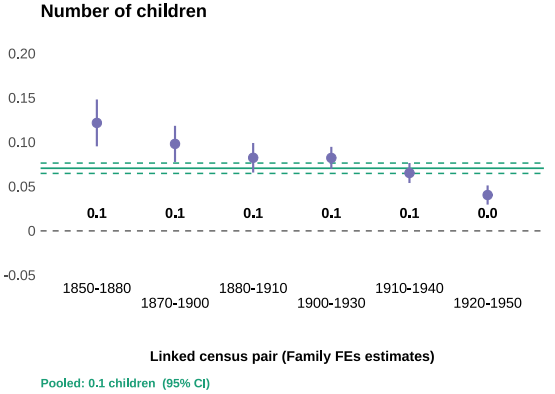
(a) Firstborn effect on P(Single)



(c) Firstborn effect on P(Married to firstborn spouse)



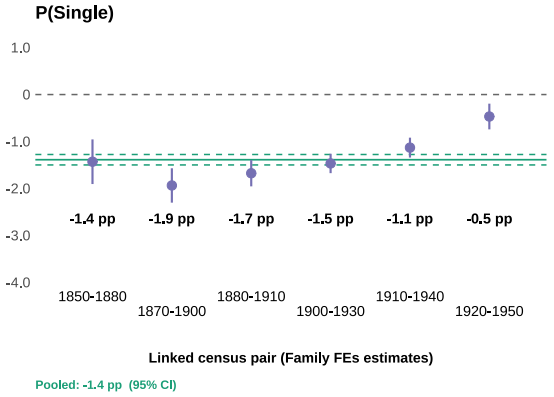
(b) Firstborn effect on P(Any kids by age 30)



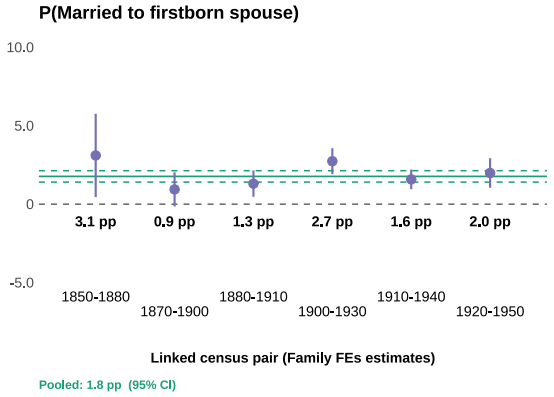
(d) Firstborn premium for number of children

Notes: Figure 3a shows the firstborn effect on P(Single). Figure 3b displays the effect of being firstborn on P(Any kids by age 30). Figure 3c presents the effect of being firstborn on P(Married to firstborn). Figure 3d shows the effect of being firstborn on number of children. All figures include 95% confidence intervals and show results from both separate linked census samples and a pooled sample.

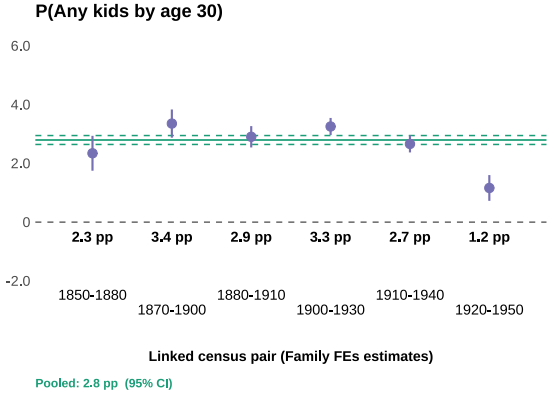
Figure 4. Birth order effects on social outcomes for women



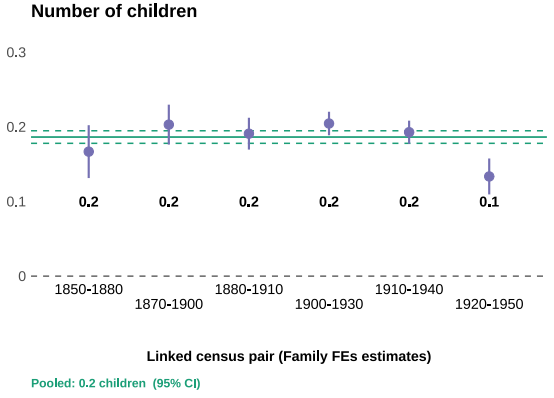
(a) Firstborn effect on P(Single)



(c) Firstborn effect on P(Married to firstborn spouse)



(b) Firstborn effect on P(Any kids by age 30)



(d) Firstborn premium for number of children

Notes: Figure 4a shows the firstborn effect on P(Single). Figure 4b displays the effect of being firstborn on fertility. Figure 4c presents the effect of being firstborn on P(Married to firstborn). Figure 4d shows the effect of being firstborn on number of children. All figures include 95% confidence intervals and show results from both separate linked census samples and a pooled sample.

8 Tables

Table 1. Firstborn Premium for Occupation Status and Absolute Mobility of Men

	(1) 1850 to 1880	(2) 1870 to 1900	(3) 1880 to 1910	(4) 1900 to 1930	(5) 1910 to 1940	(6) 1920 to 1950	(7) Pooled
A. Standardized occupational income score							
Firstborn (=1)	0.026	0.027	0.037	0.038	0.024	0.030	0.031
	(0.006)	(0.005)	(0.004)	(0.003)	(0.002)	(0.003)	(0.001)
Mean of Y	0.189	0.173	0.196	0.202	0.183	0.157	0.184
# of Households	164,142	295,199	474,942	722,628	874,222	563,438	3,094,571
Observations	328,284	590,398	949,884	1,445,256	1,748,444	1,126,876	6,189,142
B. P(Same microclass occupation as father)							
Firstborn (=1)	–0.005	–0.002	–0.009	–0.006	–0.004	–0.003	–0.005
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean of Y	0.425	0.346	0.327	0.258	0.211	0.208	0.265
# of Households	153,347	259,461	445,648	627,770	760,229	512,695	2,759,150
Observations	306,694	518,922	891,296	1,255,540	1,520,458	1,025,390	5,518,300
C. P(Upward mobility)							
Firstborn (=1)	0.011	0.008	0.015	0.012	0.007	0.009	0.010
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Mean of Y	0.319	0.431	0.452	0.523	0.528	0.550	0.498
# of Households	153,347	259,486	445,671	627,783	760,261	512,714	2,759,262
Observations	306,694	518,972	891,342	1,255,566	1,520,522	1,025,428	5,518,524
D. P(Downward mobility)							
Firstborn (=1)	–0.005	–0.006	–0.006	–0.007	–0.004	–0.007	–0.006
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean of Y	0.262	0.226	0.228	0.227	0.273	0.255	0.247
# of Households	153,347	259,486	445,671	627,783	760,261	512,714	2,759,262
Observations	306,694	518,972	891,342	1,255,566	1,520,522	1,025,428	5,518,524
E. P(Labor force participation)							
Firstborn (=1)	0.001		0.001	0.003	0.002	0.003	0.002
	(0.001)		(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Mean of Y	0.964		0.979	0.976	0.954	0.956	0.965
# of Households	172,814		518,480	793,051	930,158	682,804	3,097,307
Observations	345,628		1,036,960	1,586,102	1,860,316	1,365,608	6,194,614

Notes: Each coefficient comes from a separate regression on male-only samples. Columns 1 to 6 present estimates for each linked census pair, and column 7 presents the pooled estimate across all census waves. The dependent variable for each panel are described in the panel titles. All regressions include family (sibling group) and birth year fixed effects. Standard errors (in parentheses) are clustered at the family level.

Table 2. Firstborn Premium for Occupation Status and Absolute Mobility of Women

	(1) 1850 to 1880	(2) 1870 to 1900	(3) 1880 to 1910	(4) 1900 to 1930	(5) 1910 to 1940	(6) 1920 to 1950	(7) Pooled
A. Spouse's standardized occupational income score							
Firstborn (=1)	–0.007	–0.015	–0.022	–0.033	–0.034	–0.023	–0.027
	(0.009)	(0.008)	(0.006)	(0.004)	(0.004)	(0.007)	(0.002)
Mean of Y	0.100	0.086	0.126	0.137	0.147	0.057	0.123
# of Households	69,280	106,167	193,463	317,416	311,815	110,625	1,108,766
Observations	138,560	212,334	386,926	634,832	623,630	221,250	2,217,532
B. P(Spouse has same microclass occupation as father)							
Firstborn (=1)	0.005	0.010	0.008	0.011	0.008	0.006	0.008
	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)
Mean of Y	0.495	0.427	0.391	0.299	0.239	0.224	0.317
# of Households	66,436	96,859	184,615	280,607	274,875	101,517	1,004,909
Observations	132,872	193,718	369,230	561,214	549,750	203,034	2,009,818
C. P(Upward mobility of spouse vs. father)							
Firstborn (=1)	–0.004	–0.007	–0.012	–0.017	–0.013	–0.009	–0.013
	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.001)
Mean of Y	0.280	0.387	0.422	0.506	0.539	0.570	0.480
# of Households	66,436	96,864	184,640	280,613	274,889	101,526	1,004,968
Observations	132,872	193,728	369,280	561,226	549,778	203,052	2,009,936
D. P(Downward mobility of spouse vs. father)							
Firstborn (=1)	–0.002	–0.003	0.004	0.007	0.006	0.003	0.004
	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.001)
Mean of Y	0.228	0.185	0.189	0.196	0.227	0.212	0.206
# of Households	66,436	96,864	184,640	280,613	274,889	101,526	1,004,968
Observations	132,872	193,728	369,280	561,226	549,778	203,052	2,009,936
E. P(Labor force participation)							
Firstborn (=1)	–0.005		–0.010	–0.014	–0.014	0.003	–0.011
	(0.003)		(0.002)	(0.001)	(0.001)	(0.003)	(0.001)
Mean of Y	0.167		0.193	0.212	0.263	0.306	0.231
# of Households	110,220		321,550	500,096	501,951	169,825	1,603,642
Observations	220,440		643,100	1,000,192	1,003,902	339,650	3,207,284

Notes: Each coefficient comes from a separate regression on female-only samples. Columns 1 to 6 present estimates for each linked census pair, and column 7 presents the pooled estimate across all census waves. The dependent variable for each panel are described in the panel titles. All regressions include family (sibling group) and birth year fixed effects. Standard errors (in parentheses) are clustered at the family level.

Table 3. Firstborn Premium for Social Outcomes of Men

	(1) 1850 to 1880	(2) 1870 to 1900	(3) 1880 to 1910	(4) 1900 to 1930	(5) 1910 to 1940	(6) 1920 to 1950	(7) Pooled
A. P(Single)							
Firstborn (=1)	−0.016	−0.017	−0.019	−0.019	−0.018	−0.012	−0.018
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Mean of Y	0.132	0.182	0.166	0.139	0.148	0.109	0.144
# of Households	173,078	353,909	518,923	793,629	931,865	683,164	3,454,568
Observations	346,156	707,818	1,037,846	1,587,258	1,863,730	1,366,328	6,909,136
B. P(Any kids by age 30)							
Firstborn (=1)	0.034	0.030	0.034	0.033	0.028	0.022	0.030
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Mean of Y	0.630	0.489	0.528	0.545	0.530	0.576	0.543
# of Households	173,078	353,909	518,923	793,629	931,865	683,164	3,454,568
Observations	346,156	707,818	1,037,846	1,587,258	1,863,730	1,366,328	6,909,136
C. Age at first birth conditional on having any children							
Firstborn (=1)	−0.450	−0.413	−0.433	−0.421	−0.360	−0.327	−0.388
	(0.039)	(0.028)	(0.023)	(0.018)	(0.018)	(0.019)	(0.009)
Mean of Y	25.463	26.975	26.357	26.094	25.826	26.169	26.137
# of Households	102,596	170,673	249,150	375,807	400,616	373,885	1,672,727
Observations	205,192	341,346	498,300	751,614	801,232	747,770	3,345,454
D. P(Any children)							
Firstborn (=1)	0.017	0.015	0.018	0.019	0.018	0.010	0.017
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Mean of Y	0.758	0.675	0.672	0.672	0.638	0.727	0.678
# of Households	173,078	353,909	518,923	793,629	931,865	683,164	3,454,568
Observations	346,156	707,818	1,037,846	1,587,258	1,863,730	1,366,328	6,909,136
E. Number of children conditional on having any children							
Firstborn (=1)	0.122	0.098	0.082	0.082	0.065	0.040	0.071
	(0.013)	(0.010)	(0.008)	(0.006)	(0.006)	(0.006)	(0.003)
Mean of Y	3.462	3.262	3.147	2.820	2.574	2.482	2.819
# of Households	108,544	183,735	268,924	406,984	436,581	393,731	1,798,499
Observations	217,088	367,470	537,848	813,968	873,162	787,462	3,596,998
F. P(Married to firstborn spouse conditional on being married)							
Firstborn (=1)	0.008	0.014	0.010	0.023	0.024	0.002	0.018
	(0.009)	(0.005)	(0.005)	(0.003)	(0.003)	(0.006)	(0.002)
Mean of Y	0.346	0.340	0.340	0.395	0.374	0.367	0.369
# of Households	96,350	198,866	256,618	473,884	525,742	251,763	1,803,223
Observations	117,869	250,567	315,173	606,784	678,597	298,744	2,267,734

Notes: Each coefficient comes from a separate regression on male-only samples. Columns 1 to 6 present estimates for each linked census pair, and column 7 presents the pooled estimate across all census waves. The dependent variable for each panel are described in the panel titles. All regressions include family (sibling group) and birth year fixed effects. Standard errors (in parentheses) are clustered at the family level.

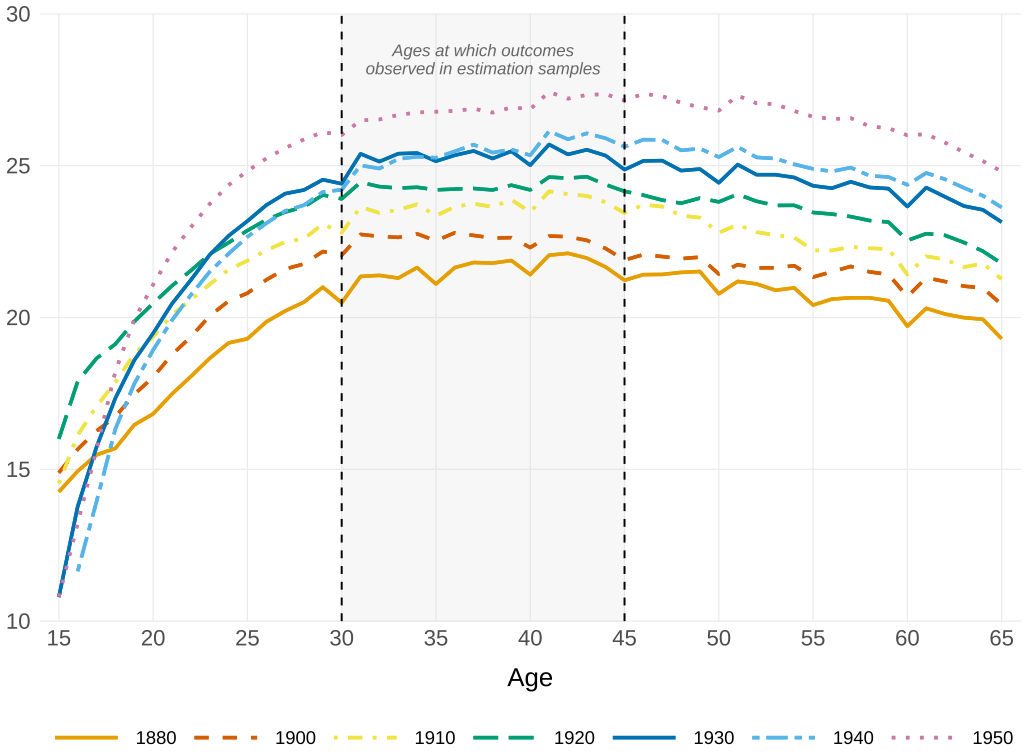
Table 4. Firstborn Premium for Social Outcomes of Women

	(1) 1850 to 1880	(2) 1870 to 1900	(3) 1880 to 1910	(4) 1900 to 1930	(5) 1910 to 1940	(6) 1920 to 1950	(7) Pooled
A. P(Single)							
Firstborn (=1)	−0.014	−0.019	−0.017	−0.015	−0.011	−0.005	−0.014
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean of Y	0.150	0.180	0.167	0.140	0.155	0.110	0.151
# of Households	110,348	203,055	321,966	500,437	503,119	170,021	1,808,946
Observations	220,696	406,110	643,932	1,000,874	1,006,238	340,042	3,617,892
B. P(Any kids by age 30)							
Firstborn (=1)	0.023	0.034	0.029	0.033	0.027	0.012	0.028
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Mean of Y	0.752	0.644	0.673	0.684	0.670	0.769	0.686
# of Households	110,348	203,055	321,966	500,437	503,119	170,021	1,808,946
Observations	220,696	406,110	643,932	1,000,874	1,006,238	340,042	3,617,892
C. Age at first birth conditional on having any children							
Firstborn (=1)	−0.550	−0.585	−0.561	−0.568	−0.499	−0.295	−0.520
	(0.048)	(0.033)	(0.027)	(0.020)	(0.020)	(0.032)	(0.011)
Mean of Y	21.932	23.805	22.755	23.112	22.703	22.225	22.845
# of Households	72,412	117,302	186,180	293,240	275,239	110,810	1,055,183
Observations	144,824	234,604	372,360	586,480	550,478	221,620	2,110,366
D. P(Any children)							
Firstborn (=1)	0.016	0.021	0.019	0.021	0.016	0.003	0.018
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Mean of Y	0.791	0.731	0.731	0.742	0.712	0.779	0.737
# of Households	110,348	203,055	321,966	500,437	503,119	170,021	1,808,946
Observations	220,696	406,110	643,932	1,000,874	1,006,238	340,042	3,617,892
E. Number of children conditional on having any children							
Firstborn (=1)	0.167	0.203	0.191	0.205	0.193	0.134	0.186
	(0.018)	(0.014)	(0.011)	(0.008)	(0.008)	(0.012)	(0.004)
Mean of Y	4.005	3.670	3.581	3.189	2.930	2.892	3.269
# of Households	75,500	122,800	195,240	307,947	289,409	113,110	1,104,006
Observations	151,000	245,600	390,480	615,894	578,818	226,220	2,208,012
F. P(Married to firstborn spouse conditional on being married)							
Firstborn (=1)	0.031	0.009	0.013	0.027	0.016	0.020	0.018
	(0.014)	(0.006)	(0.004)	(0.004)	(0.003)	(0.005)	(0.002)
Mean of Y	0.419	0.346	0.342	0.439	0.385	0.370	0.385
# of Households	55,442	141,698	230,837	327,509	380,365	138,676	1,274,527
Observations	65,084	191,438	314,808	419,264	532,123	207,671	1,730,388

Notes: Each coefficient comes from a separate regression on female-only samples. Columns 1 to 6 present estimates for each linked census pair, and column 7 presents the pooled estimate across all census waves. The dependent variable for each panel are described in the panel titles. All regressions include family (sibling group) and birth year fixed effects. Standard errors (in parentheses) are clustered at the family level.

A Appendix Figures: Descriptive Statistics

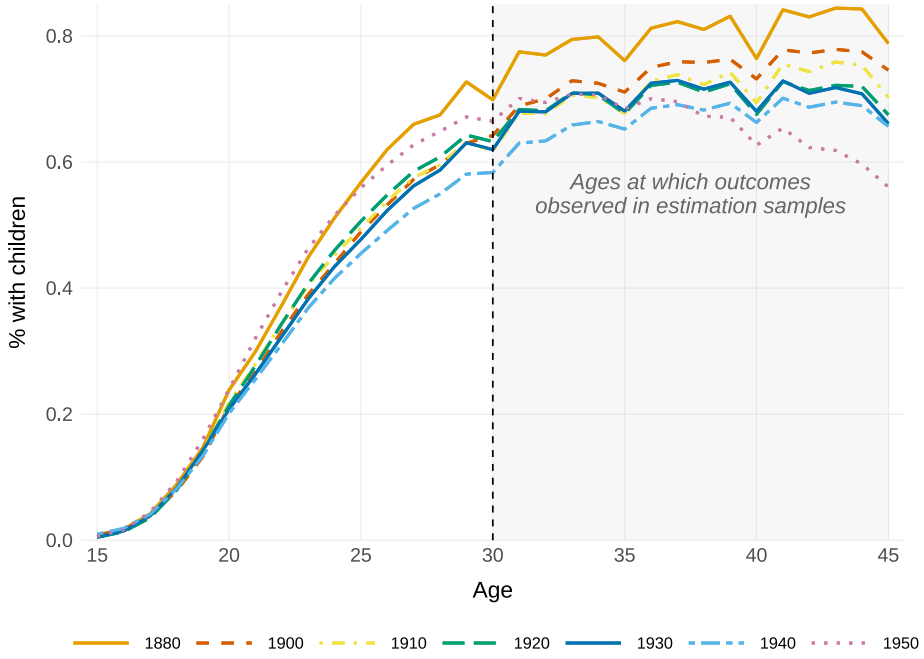
Figure A.1. Occupational income score by age and census year



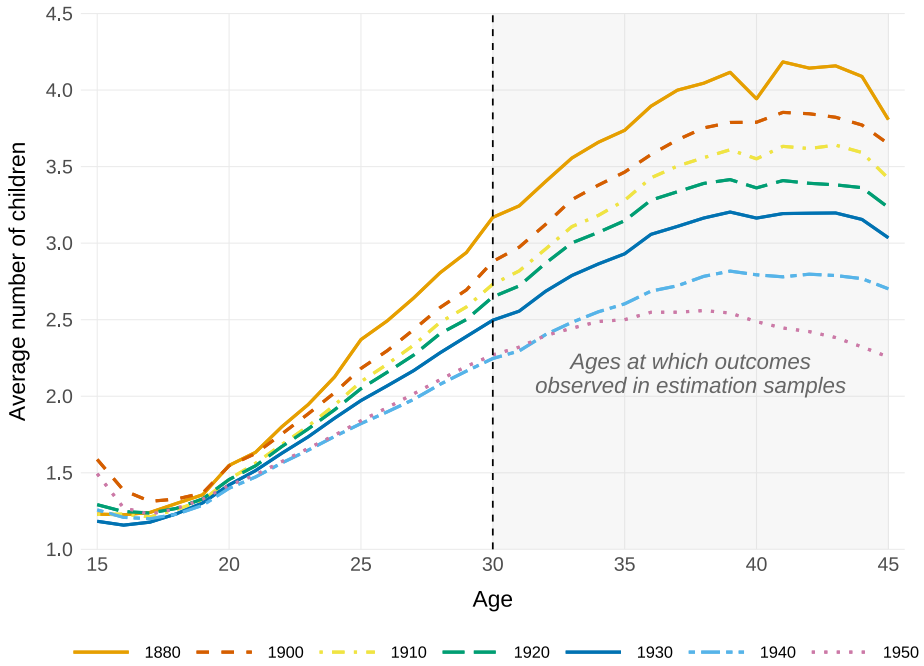
Notes: Mean occupational income score by age for males with non-zero scores, by census year. Dashed lines at ages 30 and 45 mark the estimation sample age window.

Figure A.2. Co-resident children by age and census year (women)

(a) Share with any children



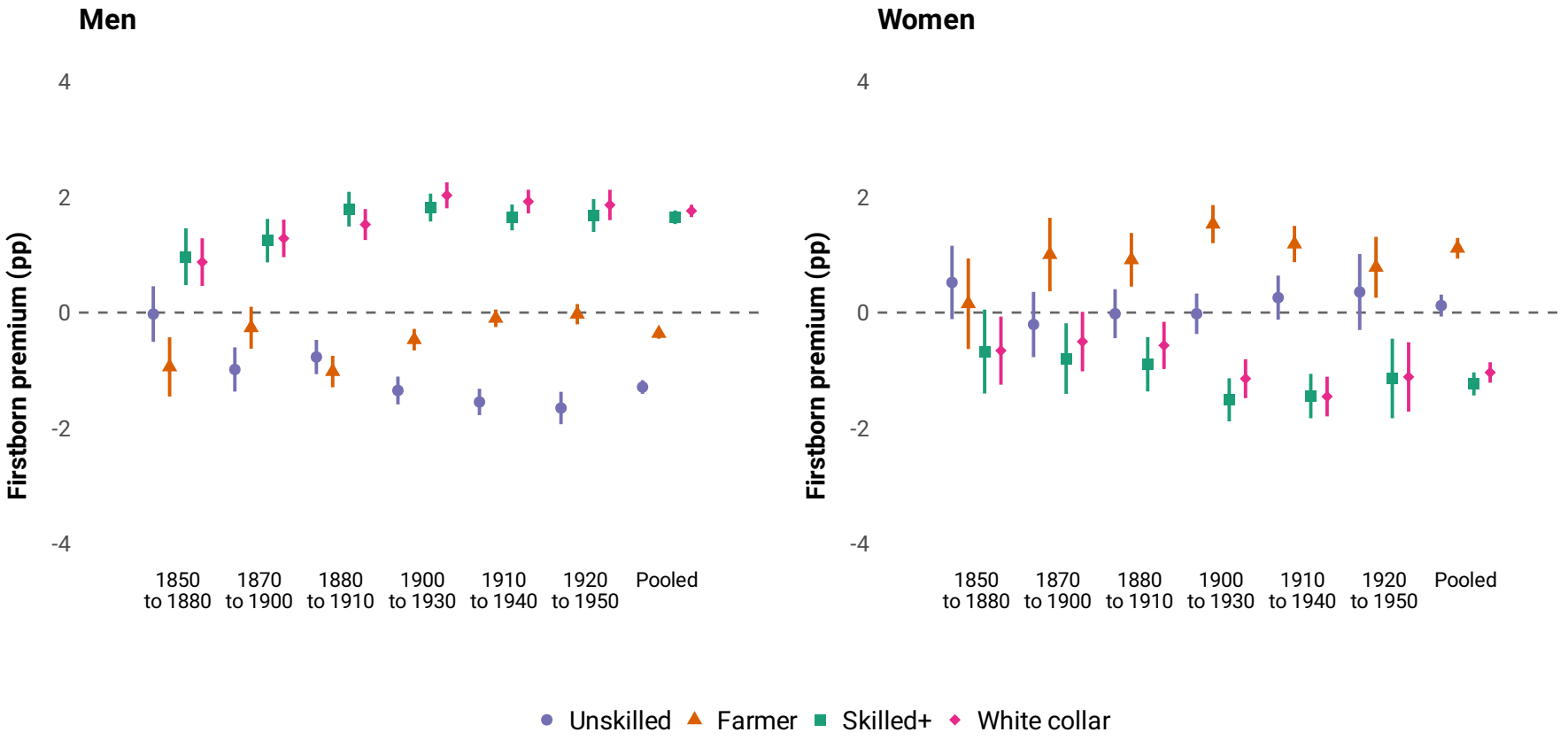
(b) Average number of children



Notes: Co-resident children identified via IPUMS `momloc`. Dashed line at age 30 marks the estimation sample threshold.

B Appendix Figures: Additional Outcomes

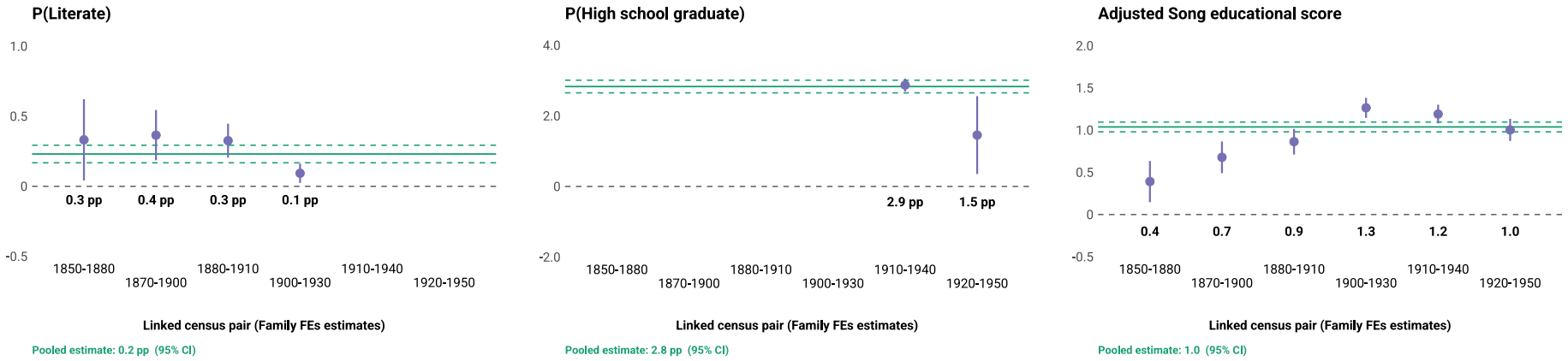
Figure B.1. Occupational categories (men and women)



Notes: Each panel shows the firstborn premium for four binary occupational category indicators: P(Farmer), P(Unskilled), P(Skilled or white collar), and P(White collar). Estimates are from separate linked census samples 30 years apart with family and birth year fixed effects. All panels include 95% confidence intervals.

C Appendix Figures: Mechanisms

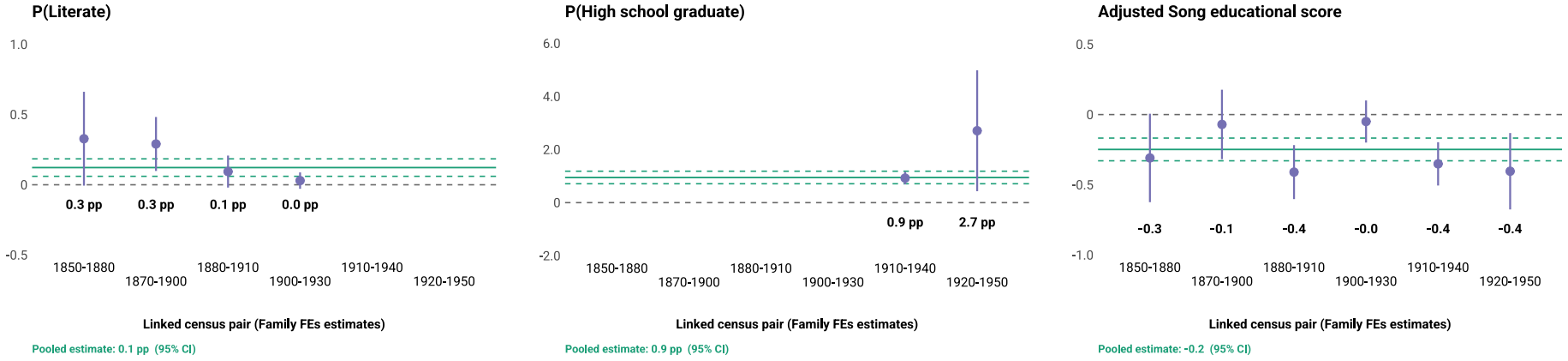
Figure C.1. Birth order effects on human capital outcomes



(a) Men: P(Literate)

(b) Men: P(High school graduate)

(c) Men: Adjusted Song score



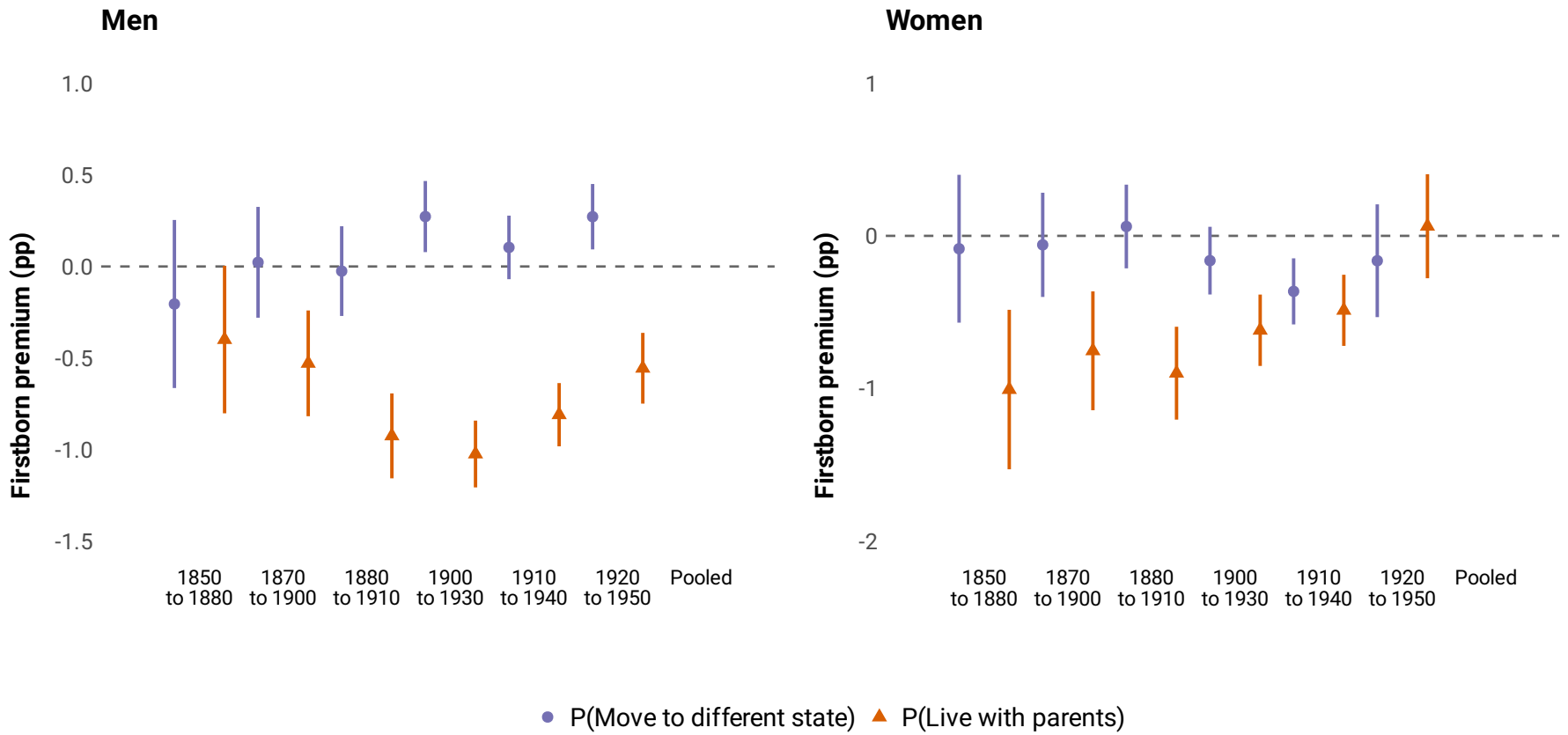
(d) Women: P(Literate)

(e) Women: P(High school graduate)

(f) Women: Adjusted Song score

Notes: Top row presents results for men, bottom row for women. Figures C.1a and C.1d show the firstborn effect on the probability of being literate (able to read and write), available for outcome censuses 1880–1930. Figures C.1b and C.1e show the firstborn effect on the probability of high school graduation, available for outcome censuses 1940 and 1950. Figures C.1c and C.1f show the firstborn effect on the adjusted Song occupational education score. All figures include 95% confidence intervals and show results from both separate linked census samples and a pooled sample.

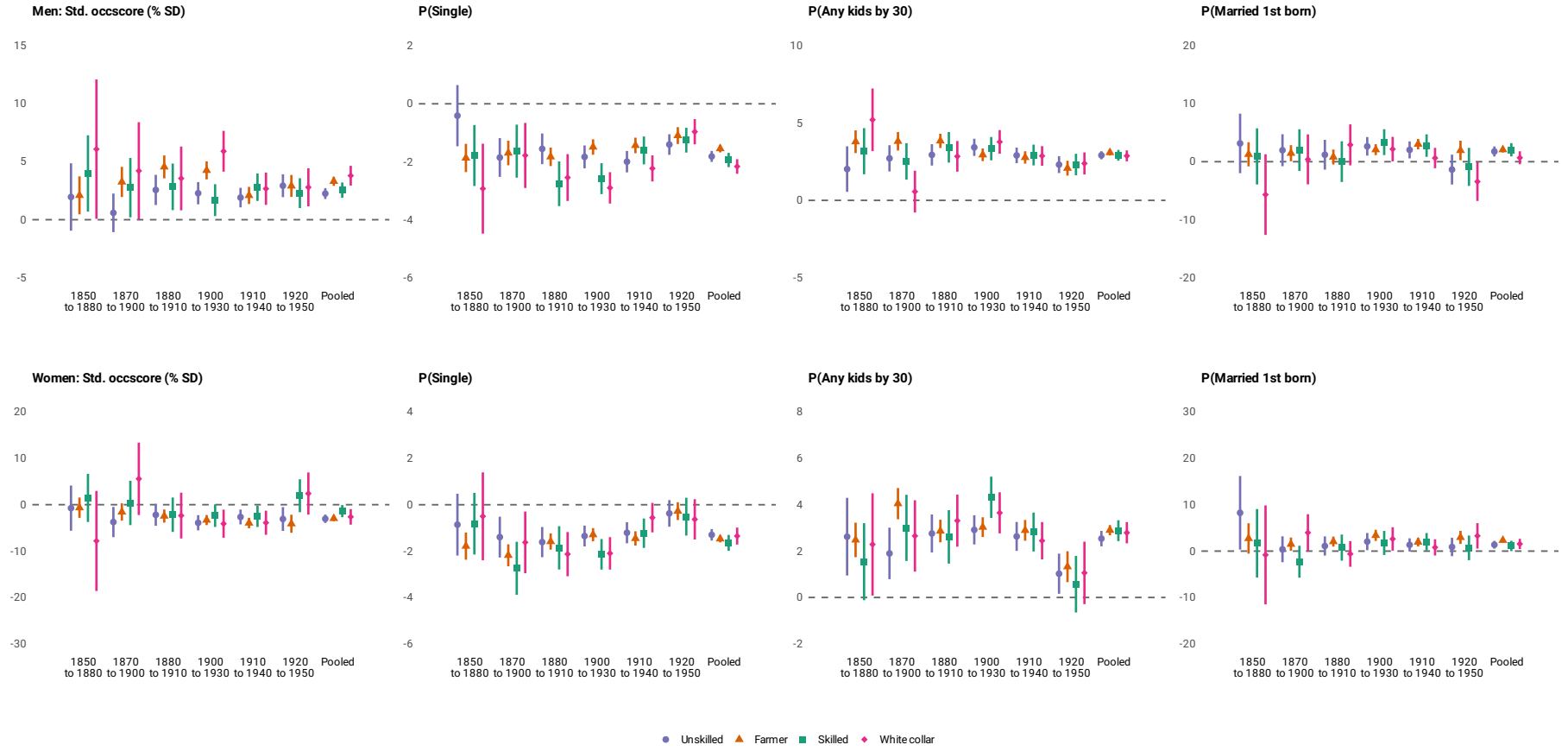
Figure C.2. Geographic mobility (men and women)



Notes: State migration: living in a different state at T2 than at T1. Live with parents: having a non-zero mother or father location pointer (momloc or poploc) in the T2 census. Estimates are from separate linked census samples 30 years apart with family and birth year fixed effects. All panels include 95% confidence intervals.

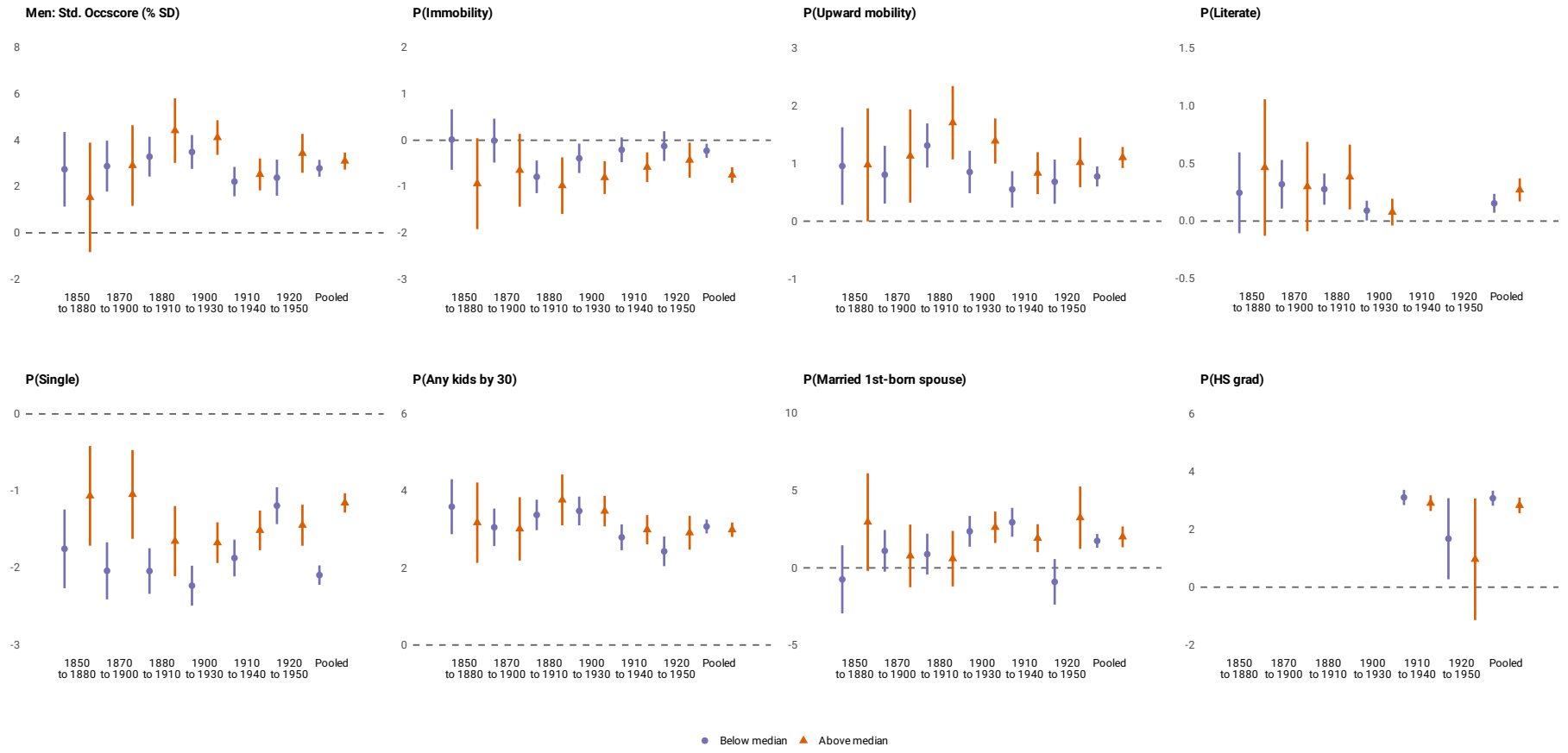
D Appendix Figures: Heterogeneity

Figure D.1. Heterogeneity by father's occupation (men and women)



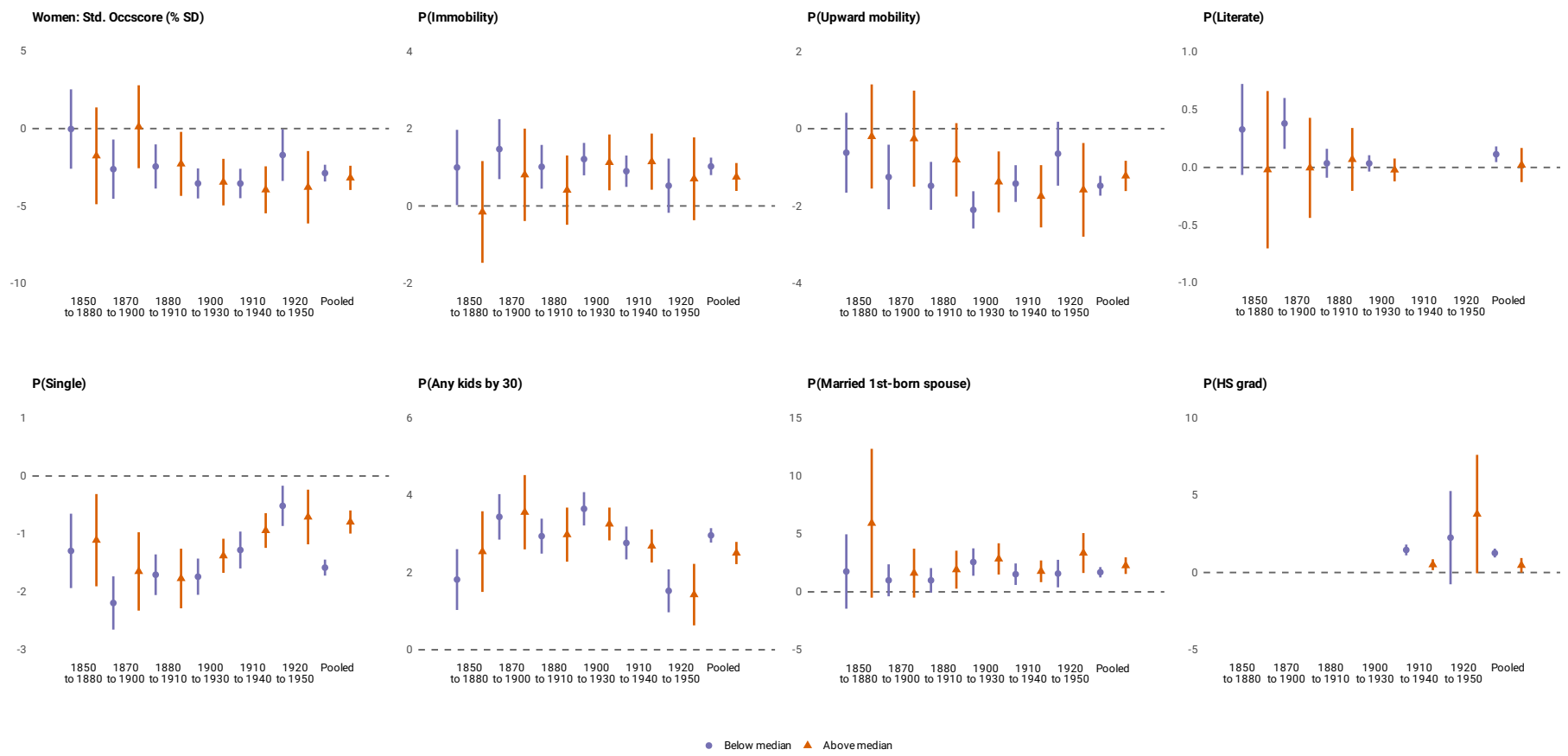
Notes: Each panel shows the firstborn premium estimated separately within four father's occupation categories: unskilled, farmer, skilled (craftsmen), and white collar. Outcomes: standardized occupational income score, P(Single), P(Any kids by age 30), and P(Married to firstborn spouse). Estimates are from separate linked census samples 30 years apart with family and birth year fixed effects. All panels include 95% confidence intervals.

Figure D.2. Heterogeneity by family size (Men)



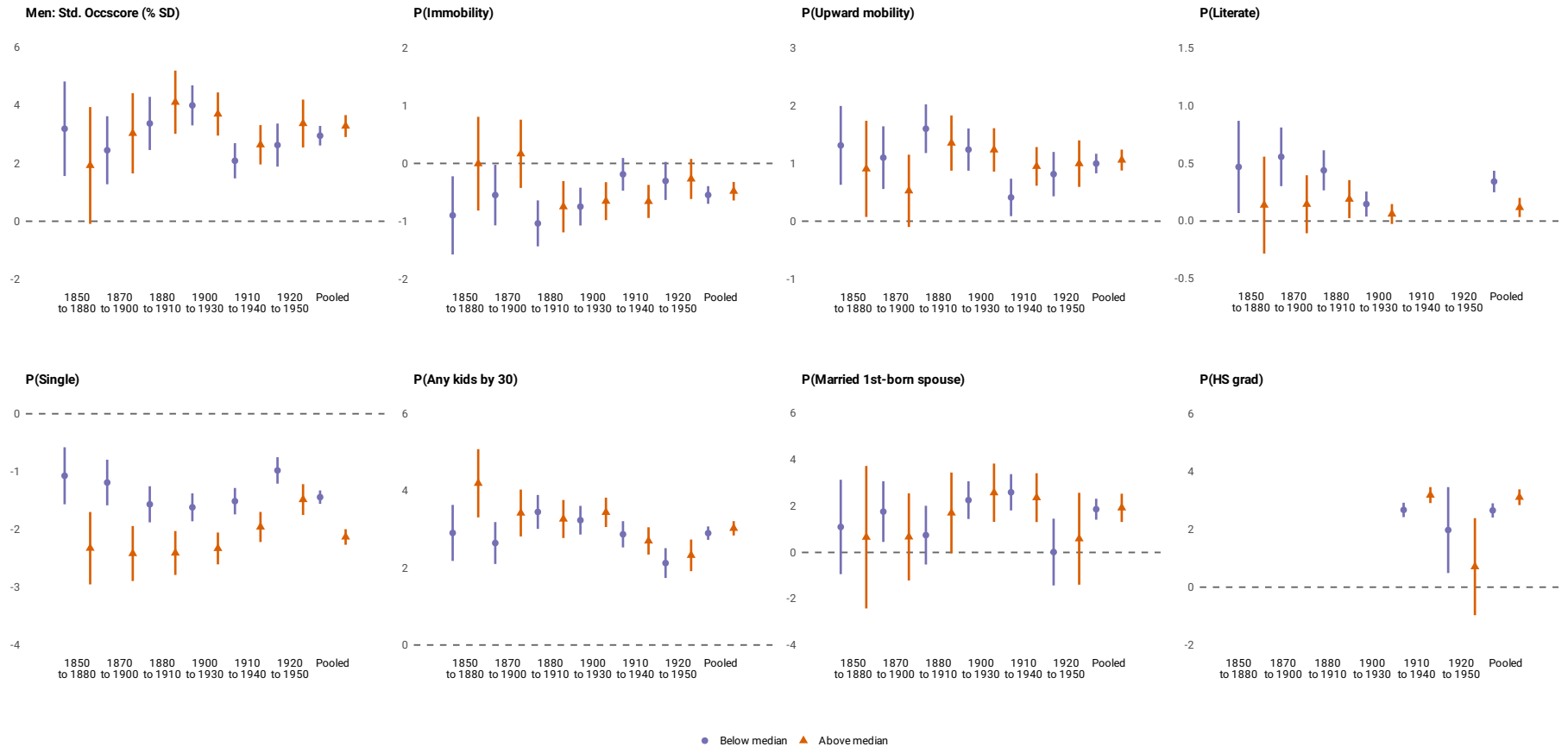
Notes: Each panel shows the firstborn premium estimated separately for families below and above the median family size. Row 1: standardized occupational income score, P(Immobility), P(Upward mobility), P(Literate). Row 2: P(Single), P(Any kids by age 30), P(Married to firstborn spouse), P(HS grad). Estimates are from separate linked census samples 30 years apart with family and birth year fixed effects. All panels include 95% confidence intervals.

Figure D.3. Heterogeneity by family size (Women)



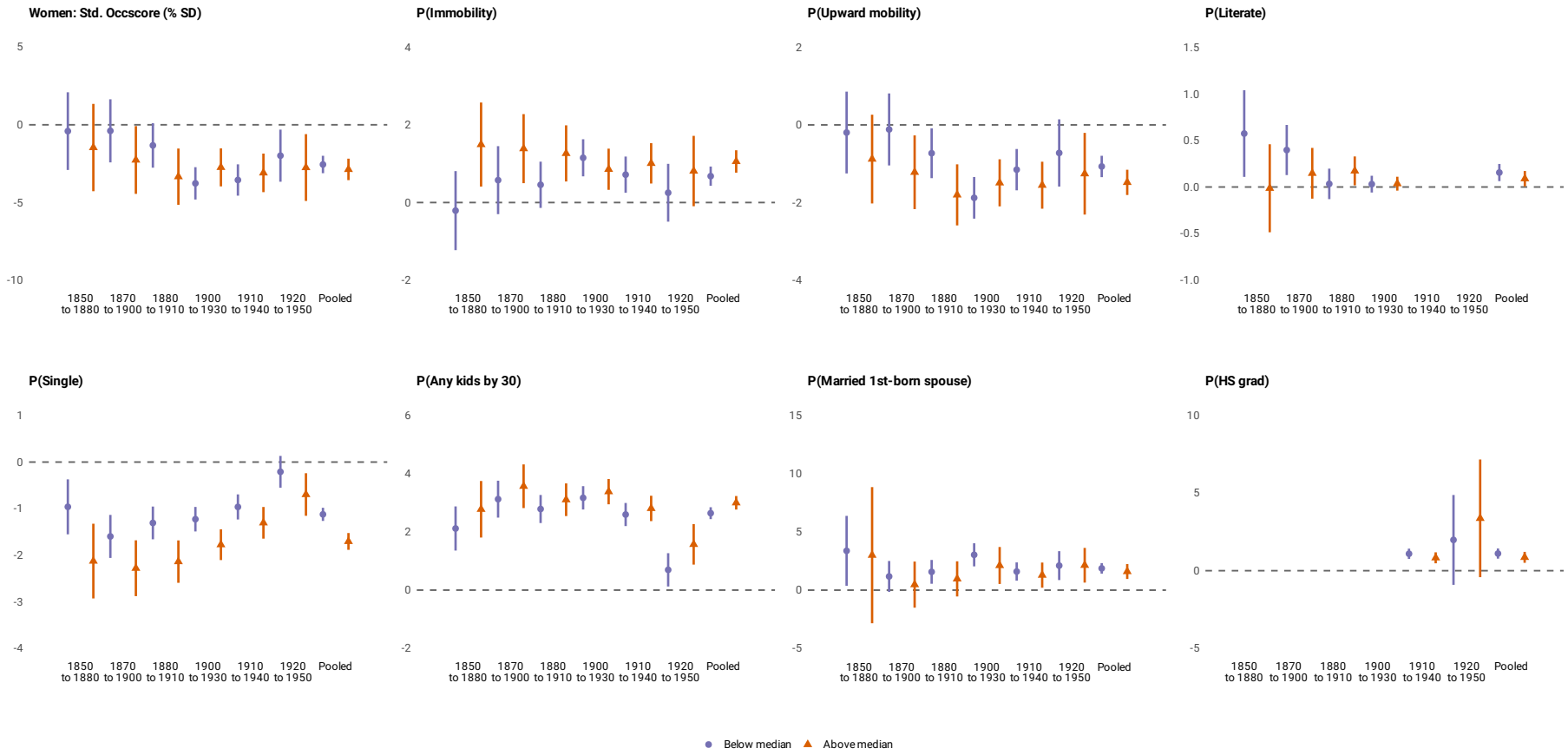
Notes: See notes to previous figure. Results shown for women.

Figure D.4. Heterogeneity by mother's age at first birth (Men)



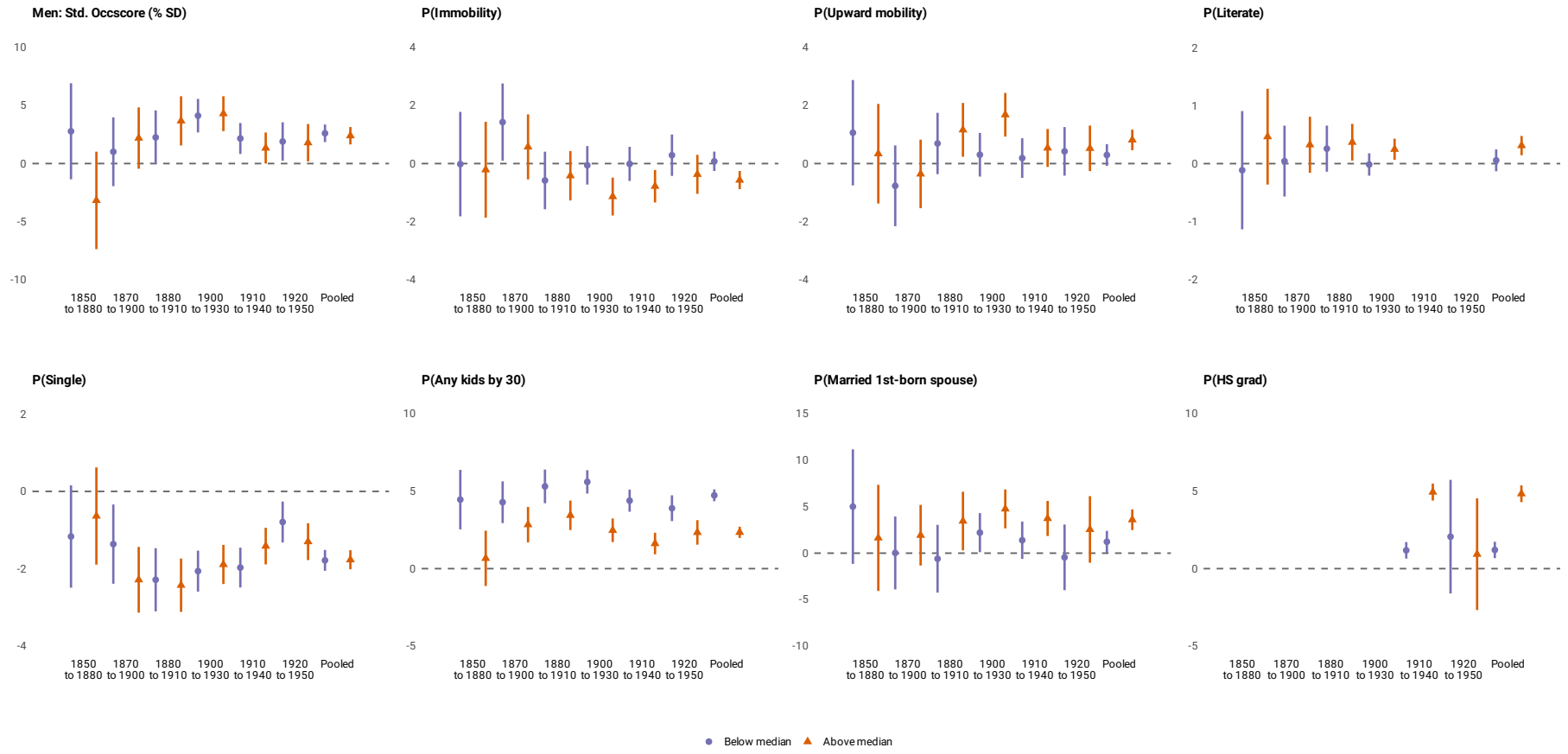
Notes: Each panel shows the firstborn premium estimated separately for families where the mother's age at first birth is below or above the median. Row 1: standardized occupational income score, P(Immobility), P(Upward mobility), P(Literate). Row 2: P(Single), P(Any kids by age 30), P(Married to firstborn spouse), P(HS grad). Estimates are from separate linked census samples 30 years apart with family and birth year fixed effects. All panels include 95% confidence intervals.

Figure D.5. Heterogeneity by mother's age at first birth (Women)



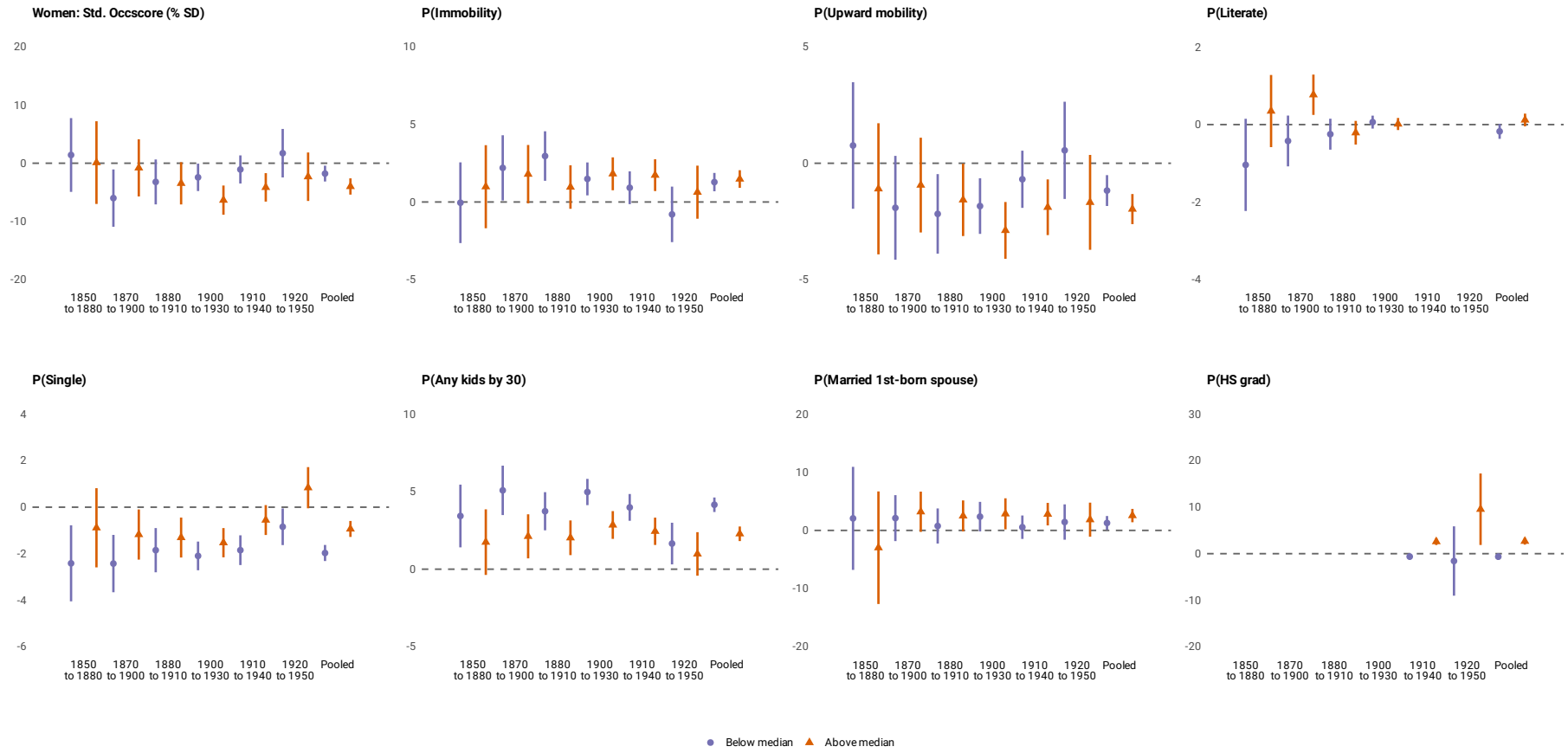
Notes: See notes to previous figure. Results shown for women.

Figure D.6. Heterogeneity by birth spacing (Men)



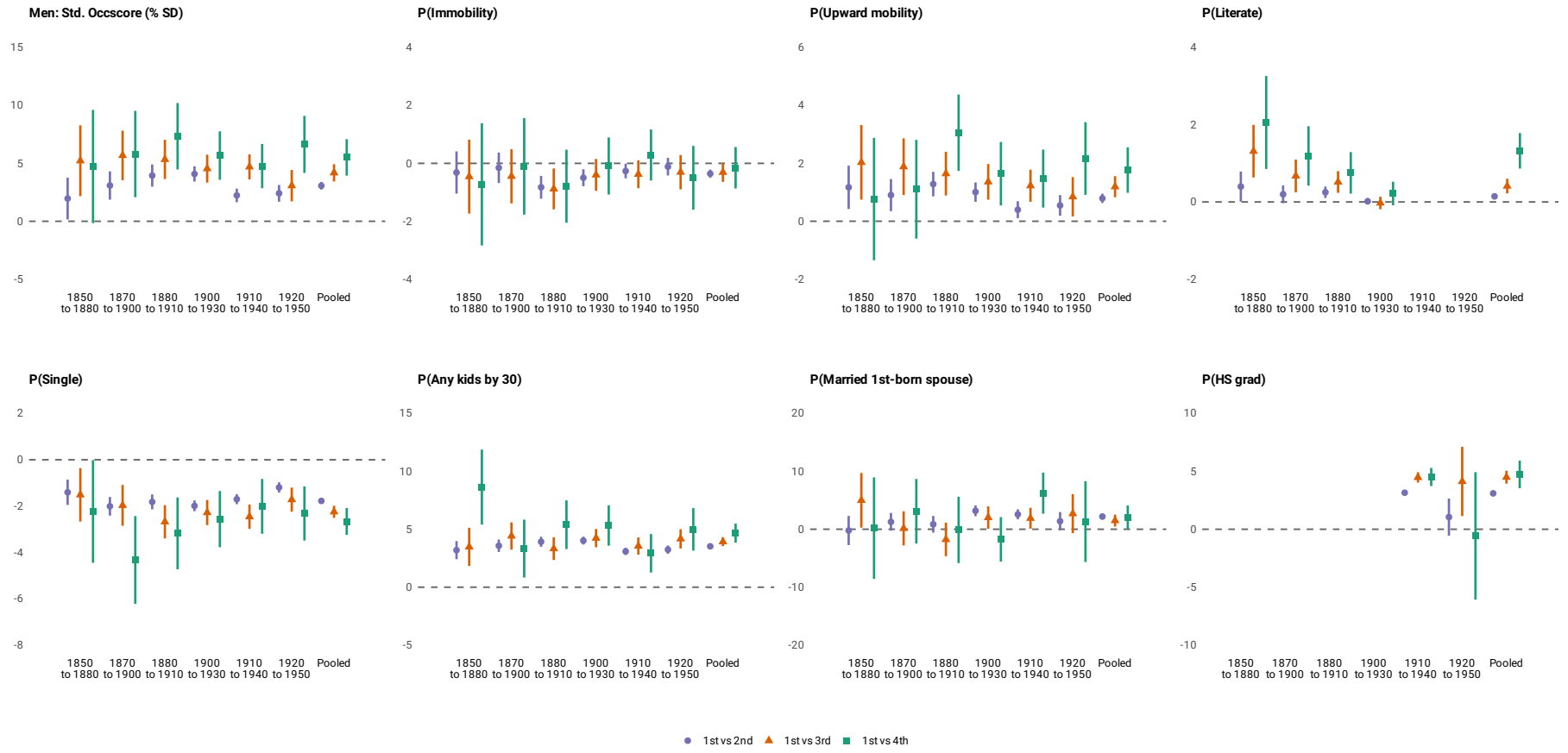
Notes: Each panel shows the firstborn premium estimated separately for families where the age gap between the first- and second-born child is below or above the median. Row 1: standardized occupational income score, P(Immobility), P(Upward mobility), P(Literate). Row 2: P(Single), P(Any kids by age 30), P(Married to firstborn spouse), P(HS grad). Estimates are from separate linked census samples 30 years apart with family and birth year fixed effects. All panels include 95% confidence intervals.

Figure D.7. Heterogeneity by birth spacing (Women)



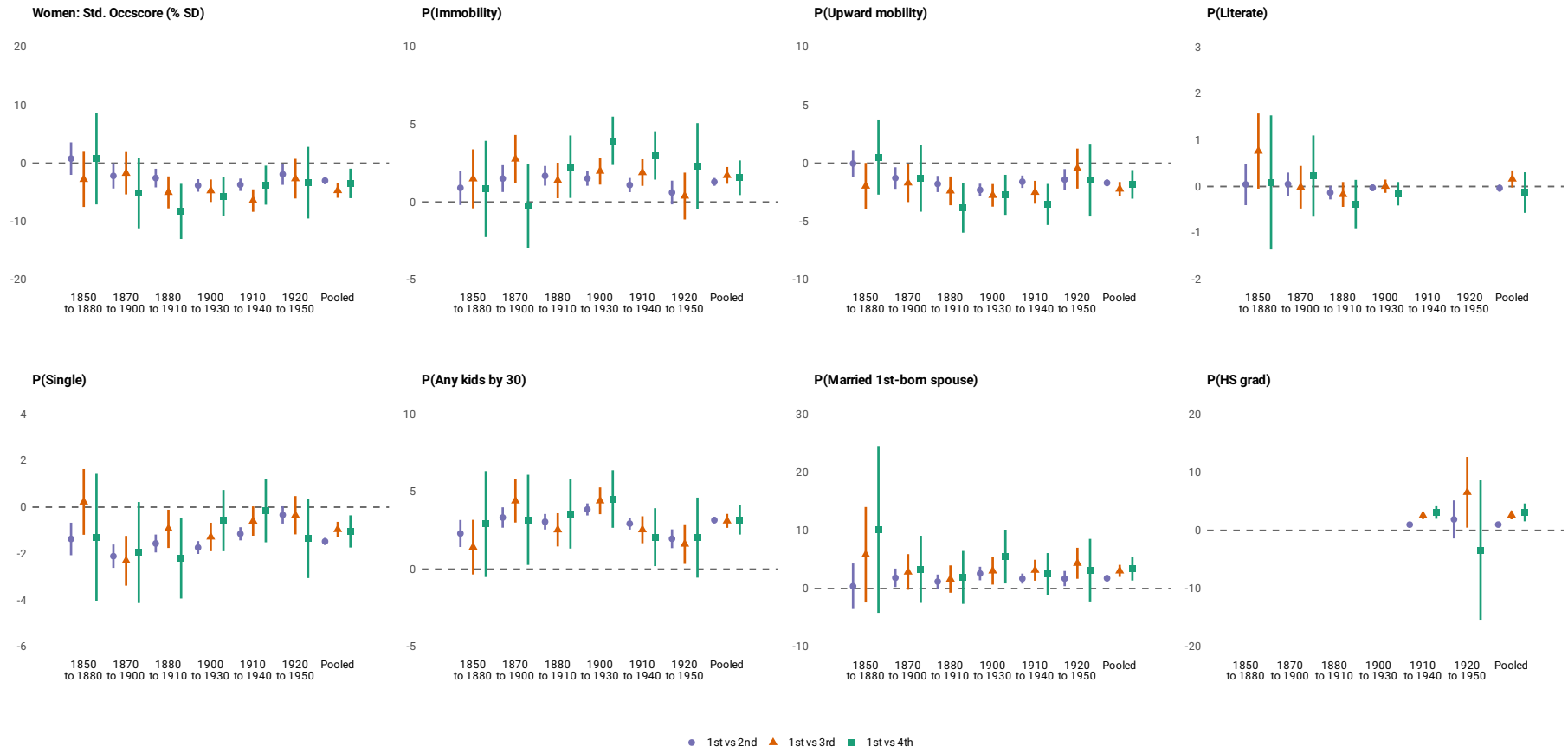
Notes: See notes to previous figure. Results shown for women.

Figure D.8. Firstborn premium by comparison birth order (Men)



Notes: Each panel shows the firstborn premium estimated from separate regressions restricting the control group to 2nd, 3rd, or 4th borns only. Row 1: standardized occupational income score, P(Immobility), P(Upward mobility), P(Literate). Row 2: P(Single), P(Any kids by age 30), P(Married to firstborn spouse), P(HS grad). Estimates are from separate linked census samples 30 years apart with family and birth year fixed effects. All panels include 95% confidence intervals.

Figure D.9. Firstborn premium by comparison birth order (Women)



Notes: See notes to previous figure. Results shown for women.

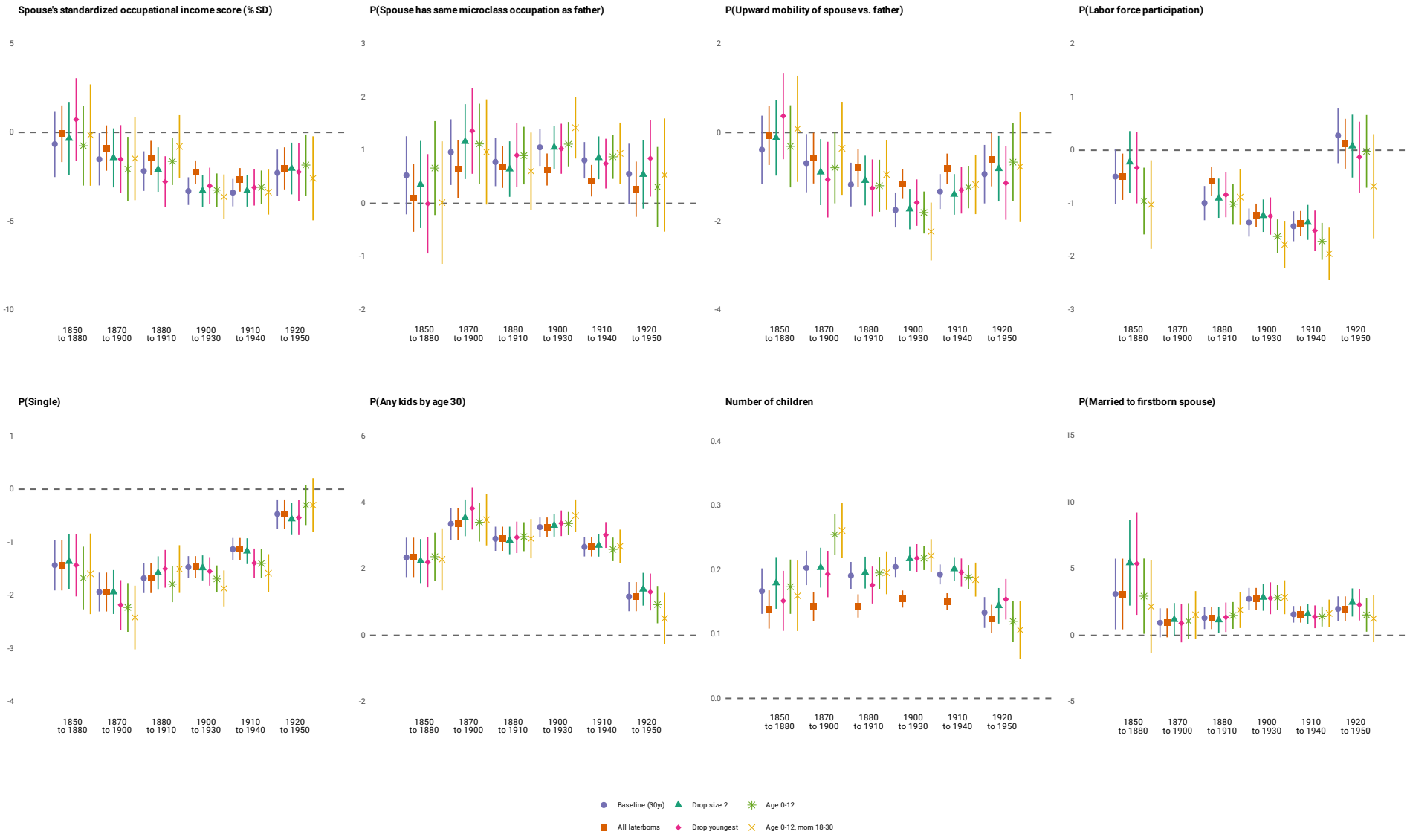
E Appendix Figures: Robustness

Figure E.1. Robustness: Sample restrictions for men



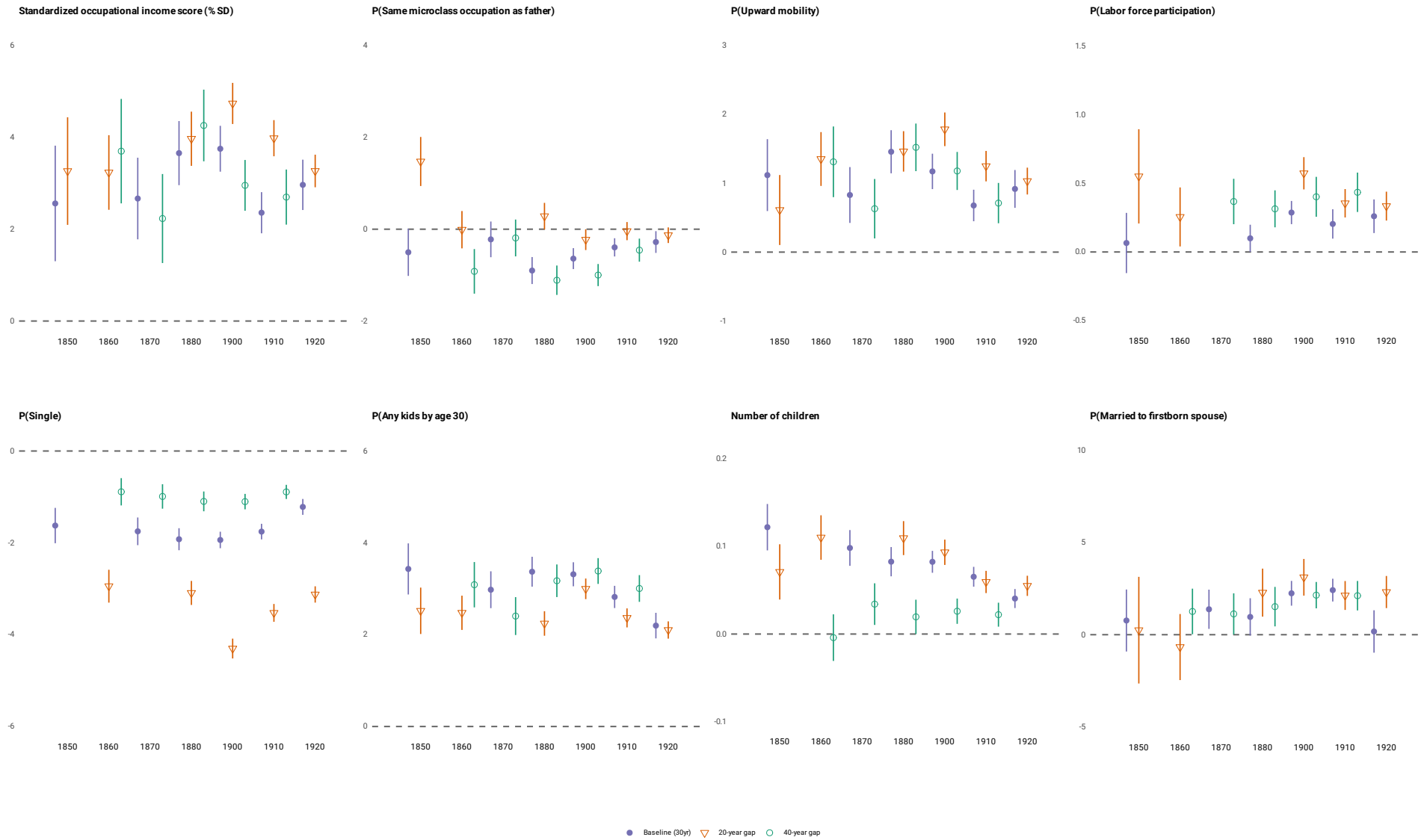
Notes: Each panel shows the firstborn premium from separate linked census samples under alternative sample restrictions: baseline (30-year gap), all laterborns, drop families of size 2, drop youngest children, children aged 0–12 only, and children aged 0–12 with mothers aged 18–30. All panels include 95% confidence intervals.

Figure E.2. Robustness: Sample restrictions for women



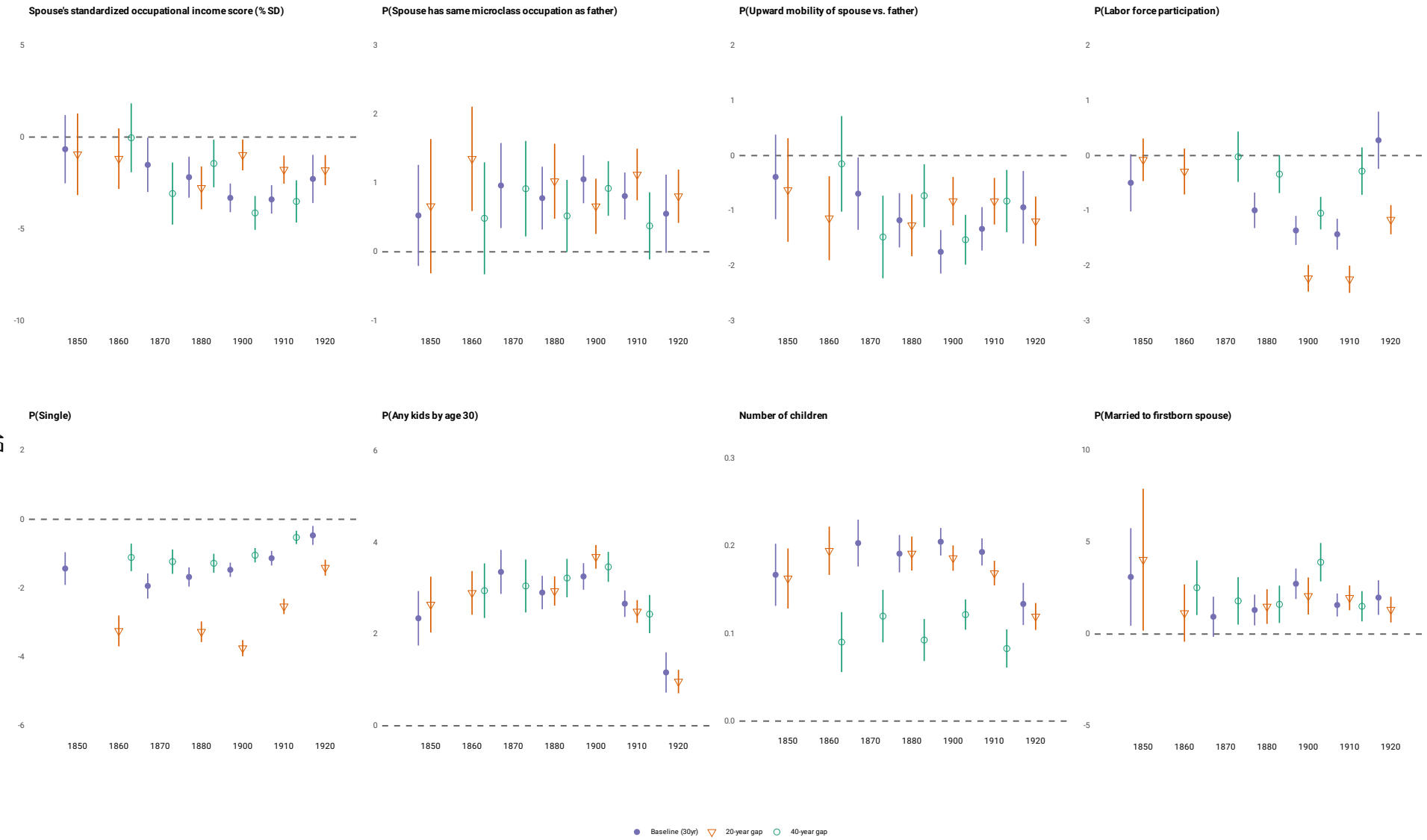
Notes: See notes to previous figure. Results shown for women.

Figure E.3. Robustness: Gap variants for men



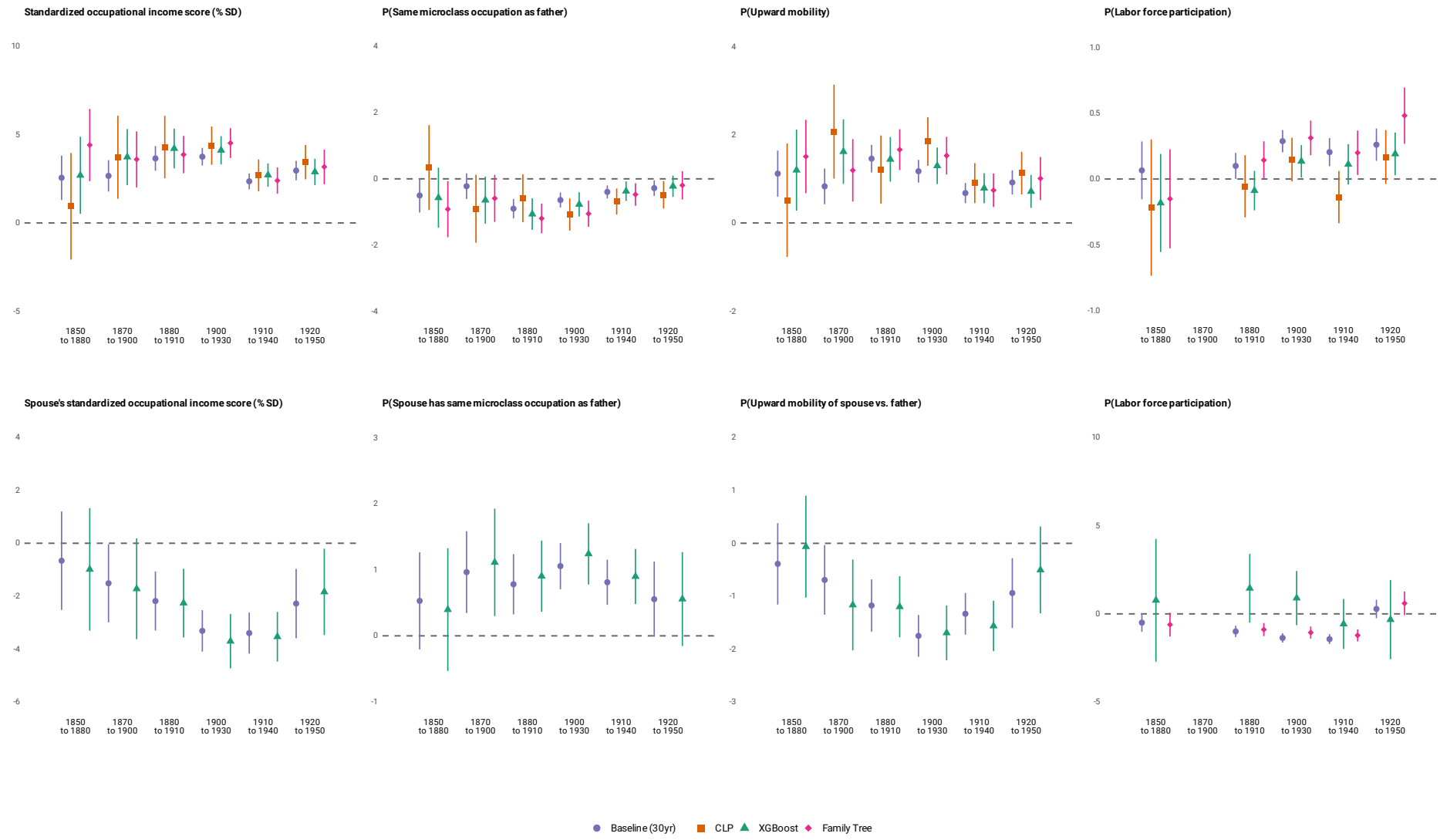
Notes: Each panel shows the firstborn premium using census pairs 20, 30, and 40 years apart. The x-axis shows the baseline (childhood) census year. All panels include 95% confidence intervals.

Figure E.4. Robustness: Gap variants for women



Notes: See notes to previous figure. Results shown for women.

Figure E.5. Robustness: Alternative linking algorithms (occupational outcomes)



Notes: Top row: men. Bottom row: women. Each panel shows the firstborn premium when restricting the sample to links from specific algorithms: baseline (all Census Tree links), Census Linking Project (CLP, men only), XGBoost machine learning, and Family Tree (identified by FamilySearch users). All panels include 95% confidence intervals.

F Appendix Tables

Table F.1. Occupational Categories

	Men				Women			
	(1) P(Unskilled)	(2) P(Farmer)	(3) P(Skilled+)	(4) P(White collar)	(5) P(Unskilled)	(6) P(Farmer)	(7) P(Skilled+)	(8) P(White collar)
A. 1850 to 1880								
Firstborn (=1)	−0.000	−0.009	0.010	0.009	0.005	0.002	−0.007	−0.007
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)
Mean of Y	0.235	0.459	0.306	0.178	0.154	0.582	0.264	0.153
# of Households			164,142				69,280	
Observations			328,284				138,560	
B. 1870 to 1900								
Firstborn (=1)	−0.010	−0.003	0.012	0.013	−0.002	0.010	−0.008	−0.005
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Mean of Y	0.291	0.359	0.349	0.204	0.191	0.497	0.312	0.185
# of Households			295,199				106,167	
Observations			590,398				212,334	
C. 1880 to 1910								
Firstborn (=1)	−0.008	−0.010	0.018	0.015	−0.000	0.009	−0.009	−0.006
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Mean of Y	0.277	0.315	0.407	0.249	0.196	0.432	0.372	0.231
# of Households			474,942				193,463	
Observations			949,884				386,926	
D. 1900 to 1930								
Firstborn (=1)	−0.013	−0.005	0.018	0.020	−0.000	0.015	−0.015	−0.011
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Mean of Y	0.297	0.205	0.498	0.312	0.238	0.304	0.458	0.284
# of Households			722,628				317,416	
Observations			1,445,256				634,832	
E. 1910 to 1940								
Firstborn (=1)	−0.015	−0.001	0.016	0.019	0.003	0.012	−0.014	−0.015
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Mean of Y	0.368	0.145	0.487	0.316	0.310	0.228	0.462	0.288
# of Households			874,222				311,815	
Observations			1,748,444				623,630	
F. 1920 to 1950								
Firstborn (=1)	−0.017	−0.000	0.017	0.019	0.004	0.008	−0.011	−0.011
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.004)	(0.003)
Mean of Y	0.318	0.135	0.547	0.326	0.293	0.212	0.495	0.272
# of Households			563,438				110,625	
Observations			1,126,876				221,250	
G. Pooled								
Firstborn (=1)	−0.013	−0.004	0.017	0.018	0.001	0.011	−0.012	−0.010
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean of Y	0.314	0.221	0.465	0.289	0.247	0.332	0.422	0.257
# of Households			3,094,571				1,108,766	
Observations			6,189,142				2,217,532	

Notes: Binary occupational category indicators derived from 1950 occupation codes. White collar includes professional, managerial, clerical, and sales occupations. Skilled+ includes white collar plus craftsmen. Columns (1)–(4) report estimates for men's own occupation; columns (5)–(8) report estimates for women's spouse's occupation. All specifications include family and birth year fixed effects. Standard errors clustered at the family level in parentheses.

Table F.2. Firstborn Premium for Human Capital Outcomes of Men

	(1) 1850 to 1880	(2) 1870 to 1900	(3) 1880 to 1910	(4) 1900 to 1930	(5) 1910 to 1940	(6) 1920 to 1950	(7) Pooled
A. P(Literate: can read and write)							
Firstborn (=1)	0.003	0.004	0.003	0.001			0.002
	(0.001)	(0.001)	(0.001)	(0.000)			(0.000)
Mean of Y	0.923	0.939	0.959	0.980			0.961
# of Households	172,936	353,053	518,730	793,329			1,838,048
Observations	345,872	706,106	1,037,460	1,586,658			3,676,096
B. P(Any high school)							
Firstborn (=1)					0.032	0.017	0.031
					(0.001)	(0.006)	(0.001)
Mean of Y					0.490	0.555	0.492
# of Households					901,154	30,168	931,322
Observations					1,802,308	60,336	1,862,644
C. P(High school graduate)							
Firstborn (=1)					0.029	0.015	0.028
					(0.001)	(0.006)	(0.001)
Mean of Y					0.291	0.359	0.293
# of Households					901,154	30,168	931,322
Observations					1,802,308	60,336	1,862,644
D. Adjusted Song educational score							
Firstborn (=1)	0.391	0.677	0.863	1.264	1.190	1.002	1.038
	(0.124)	(0.096)	(0.077)	(0.061)	(0.056)	(0.067)	(0.029)
Mean of Y	56.029	54.407	55.244	55.218	52.758	55.391	54.524
# of Households	164,086	295,120	474,619	722,268	873,706	563,153	3,092,952
Observations	328,172	590,240	949,238	1,444,536	1,747,412	1,126,306	6,185,904

Notes: Each coefficient comes from a separate regression on male-only samples. Columns 1 to 6 present estimates for each linked census pair, and column 7 presents the pooled estimate across all census waves. The dependent variable for each panel are described in the panel titles. Panel A reports the firstborn effect on the probability of being literate (able to read and write), available for outcome censuses 1880–1930. Panels B and C report effects on high school attendance and completion using detailed education codes, available for outcome censuses 1940 and 1950. Panel D reports the effect on the adjusted Song occupational education score. All regressions include family (sibling group) and birth year fixed effects. Standard errors (in parentheses) are clustered at the family level.

Table F.3. Firstborn Premium for Human Capital Outcomes of Women

	(1) 1850 to 1880	(2) 1870 to 1900	(3) 1880 to 1910	(4) 1900 to 1930	(5) 1910 to 1940	(6) 1920 to 1950	(7) Pooled
A. P(Literate: can read and write)							
Firstborn (=1)	0.003	0.003	0.001	0.000			0.001
	(0.002)	(0.001)	(0.001)	(0.000)			(0.000)
Mean of Y	0.922	0.955	0.975	0.991			0.973
# of Households	110,281	202,557	321,707	500,265			1,134,810
Observations	220,562	405,114	643,414	1,000,530			2,269,620
B. P(Any high school)							
Firstborn (=1)					0.012	0.017	0.012
					(0.001)	(0.012)	(0.001)
Mean of Y					0.575	0.567	0.575
# of Households					488,443	7,019	495,462
Observations					976,886	14,038	990,924
C. P(High school graduate)							
Firstborn (=1)					0.009	0.027	0.009
					(0.001)	(0.012)	(0.001)
Mean of Y					0.368	0.350	0.368
# of Households					488,443	7,019	495,462
Observations					976,886	14,038	990,924
D. Adjusted Song educational score							
Firstborn (=1)	–0.309	–0.070	–0.410	–0.050	–0.351	–0.404	–0.248
	(0.161)	(0.126)	(0.098)	(0.076)	(0.079)	(0.139)	(0.041)
Mean of Y	57.101	58.443	58.608	58.636	56.150	54.078	57.359
# of Households	80,277	138,291	249,829	417,150	432,610	137,466	1,455,623
Observations	160,554	276,582	499,658	834,300	865,220	274,932	2,911,246

Notes: Each coefficient comes from a separate regression on female-only samples. Columns 1 to 6 present estimates for each linked census pair, and column 7 presents the pooled estimate across all census waves. The dependent variable for each panel are described in the panel titles. Panel A reports the firstborn effect on the probability of being literate (able to read and write), available for outcome censuses 1880–1930. Panels B and C report effects on high school attendance and completion using detailed education codes, available for outcome censuses 1940 and 1950. Panel D reports the effect on the adjusted Song occupational education score. All regressions include family (sibling group) and birth year fixed effects. Standard errors (in parentheses) are clustered at the family level.

Table F.4. Geographic Mobility

	Men		Women	
	(1) P(State migrant)	(2) P(Live w/ parents)	(3) P(State migrant)	(4) P(Live w/ parents)
A. 1850 to 1880				
Firstborn (=1)	–0.002	–0.004	–0.001	–0.010
	(0.002)	(0.002)	(0.002)	(0.003)
Mean of Y	0.350	0.147	0.285	0.183
# of Households	173,078		110,348	
Observations	346,156		220,696	
B. 1870 to 1900				
Firstborn (=1)	0.000	–0.005	–0.001	–0.008
	(0.002)	(0.001)	(0.002)	(0.002)
Mean of Y	0.279	0.165	0.228	0.203
# of Households	353,909		203,055	
Observations	707,818		406,110	
C. 1880 to 1910				
Firstborn (=1)	–0.000	–0.009	0.001	–0.009
	(0.001)	(0.001)	(0.001)	(0.002)
Mean of Y	0.280	0.153	0.238	0.192
# of Households	518,923		321,966	
Observations	1,037,846		643,932	
D. 1900 to 1930				
Firstborn (=1)	0.003	–0.010	–0.002	–0.006
	(0.001)	(0.001)	(0.001)	(0.001)
Mean of Y	0.285	0.144	0.249	0.170
# of Households	793,629		500,437	
Observations	1,587,258		1,000,874	
E. 1910 to 1940				
Firstborn (=1)	0.001	–0.008	–0.004	–0.005
	(0.001)	(0.001)	(0.001)	(0.001)
Mean of Y	0.266	0.153	0.234	0.181
# of Households	931,865		503,119	
Observations	1,863,730		1,006,238	
F. 1920 to 1950				
Firstborn (=1)	0.003	–0.006	–0.002	0.001
	(0.001)	(0.001)	(0.002)	(0.002)
Mean of Y	0.205	0.138	0.213	0.147
# of Households	683,164		170,021	
Observations	1,366,328		340,042	
G. Pooled				
Firstborn (=1)				
Mean of Y				
# of Households				
Observations				

Notes: State migration is defined as living in a different state at T2 than at T1. Living with parents is defined as having a non-zero mother or father location pointer (momloc or poploc) in the T2 census. Columns (1)–(2) report estimates for men; columns (3)–(4) report estimates for women. All specifications include family and birth year fixed effects. Standard errors clustered at the family level in parentheses.