# Birth Order and the Decline in College Completion among the Baby Boom Generation

Handy, Christopher and Shester, Katharine

June 2022

# BIRTH ORDER AND THE DECLINE IN COLLEGE COMPLETION AMONG THE BABY BOOM GENERATION

Christopher Handy
University of North Carolina

Katharine L. Shester
Washington and Lee University

June 2022
Forthcoming in *Economic Inquiry*

Abstract: We show that changes in birth order during the U.S. baby boom can explain a substantial share of the decline and recovery in college completion among cohorts born between 1946 and 1974. Combining birth order effects estimated using the Health and Retirement Survey and birth order data from Vital Statistics, we estimate that changes in birth order can explain about 20 percent of the decline in white male college completion rates among the 1946–60 cohorts, and about one third of the rebound among the 1960–74 cohorts.

Keywords: baby boom, birth order, college completion, educational attainment

JEL codes: J13, I20, N32

---

*Handy (corresponding author):* Department of Economics, University of North Carolina, Gardner Hall CB 3305, Chapel Hill, NC 27599. Email: chandy@unc.edu.

*Shester:* Department of Economics, Washington and Lee University, 204 W Washington St, Lexington, VA 24450. Email: shesterk@wlu.edu.

We are grateful to Bill Collins, Andrew Goodman-Bacon, Greg Niemesh, Marianne Wanamaker, and participants at the 2019 Economic History Association Annual Meeting and the 2020 EALE/SOLE/AASLE World Conference for very helpful suggestions. We thank Balen Essak, John Juneau, Margaret Kallus, and Will Schirmer for excellent research assistance.

## I. Introduction

College completion rates in the U.S. more than doubled between the 1925 and 1945 birth cohorts, but abruptly declined toward the beginning of the postwar baby boom (Goldin and Katz, 2008). Among men, college attainment fell more than 6 percentage points between the late 1940s and late 1950s birth cohorts, and among women, college attainment also declined about 2 percentage points in the 1950s. College attainment then resumed a more gradual rise among the 1960s birth cohorts.

The decline in educational attainment across some baby boom cohorts had important consequences for individuals (e.g., lower income) and for society as a whole (e.g., increased inequality and slower economic growth), and this decrease was surprising given that it began at a time when the college premium was increasing.[1] Previous research has attributed some of this reversal in trends to an abnormally high college completion rate for men born in the 1940s due to the Vietnam War (Angrist and Chen, 2011; Card and Lemieux, 2001). However, the estimated effects of the war are much smaller than the observed decline in men's college completion rates, and the abrupt halt in the growth of educational attainment remains largely unexplained.

We add a new explanation for the changes in educational attainment among these cohorts. During the baby boom that followed World War II, the distribution of birth order changed dramatically, with first and second births becoming a smaller share of each birth cohort, and a far higher percentage of children being third-born or later. Because later-born children tend to have lower average educational attainment (Black et al., 2005; Kantarevic and Mechoulan, 2006), this change in birth order would tend to decrease college completion rates.

Figure 1 illustrates our hypothesis.[2] The solid line shows college attainment among white men by birth cohort, which we measure using data from decennial censuses and the American Community Survey. The abrupt decline in college completion in this group begins with the late 1940s cohorts and does not reverse course until about 1960. The dashed line shows changes in birth order across the baby boom

---

[1] Autor, Goldin, and Katz (2020) estimate that the college premium increased by more than 50 percent between 1950 and 1970, declined slightly in the 1970s, then increased steadily after 1980. College completion peaked among the late 1940s birth cohorts, who would have turned 18 while the college premium was still increasing in the mid-1960s.

[2] We present versions of Figure 1 for white and nonwhite men and women in Appendix Figure A1.

generation, using data from Vital Statistics. The share of births third-born or higher increased from just over 30 percent in the late 1940s to almost 50 percent around 1960.[3]

We use the Health and Retirement Survey (HRS) to estimate the effects of birth order on educational attainment for the baby boom generation. To our knowledge, the HRS is by far the largest available nationally representative survey in the U.S. that includes data on completed education for groups of siblings. We use this sibling data to estimate family fixed effects regressions of educational attainment on birth order indicators and cohort fixed effects. We find that birth order effects can be large. For example, among whites, third-born children are 8 percentage points less likely than first-borns to graduate from college. These results echo a large literature on the link between birth order and education, estimated for a variety of countries and educational outcomes (e.g., Black et al., 2005), but we believe we are the first to estimate birth order effects on college completion in the U.S. To learn about the margins on which birth order effects operate, we study the effect of birth order on alternative educational outcomes. We find that birth order effects on high school completion are small, but the birth order effects on college attendance are quite large.

We use our estimates of birth order effects to study the effect of the baby boom on changes in college completion over time. We multiply our estimated birth order effects by national changes in the distribution of birth order, measured from Vital Statistics published tables and birth-level records, to construct counterfactual series of college completion holding constant the distribution of birth order. Our results are particularly effective in explaining the changes in college completion for white men: the baby boom explains about 20 percent of the 6.3 percentage point decrease in college completion across the 1946–60 cohorts, and the end of the boom explains about one third of the 5.9 percentage point rebound in college completion across the 1960–74 cohorts.

[3] If we instead use the share of births fourth-born or higher, or fifth-born or higher, the figure looks very similar, and the inverse relationship between birth order and college completion is still clearly evident.

## II. Background

### *a. The baby boom*

The baby boom was a dramatic change in fertility patterns in the U.S. The fertility rate had been gradually declining for a century until about 1940 (Jones and Tertilt, 2006), when it abruptly increased for nearly twenty years. Figure 2 shows white and nonwhite fertility rates in the United States from 1909 to 2000; among the 1946–74 birth cohorts that are the primary focus in this paper, the vast majority of nonwhite births are black births. The fertility rate suddenly increased for both whites and nonwhites in the 1940s, and peaked around 1960. This substantial reversal of long-term fertility trends was concurrent with major social changes, including increasing educational and labor market opportunities for a generation of young veterans, changing urbanization patterns, and sharply altered labor market prospects for women (Klein, 2006). Women married earlier and at higher rates, had more children, and spaced children closer together (Bailey and Collins, 2011). Fertility rates increased during this period across income levels, racial and ethnic groups, and geographic regions, as well as in both urban and rural areas (Jones and Tertilt, 2006).

There is a rich literature exploring the causes of the baby boom. One early theory is Easterlin's (1976) "relative income" hypothesis, in which fertility is positively related to the gap between realized material resources and material aspirations that were formed during childhood. Easterlin postulated that those born just before and during the Great Depression formed low aspirations of material well-being, which where then greatly exceeded because of strong economic growth following World War II, leading to an increase in fertility.

Another strand of the literature focuses on the role of wealth and price effects in shaping the demand for children. Butz and Ward (1979) attribute the baby boom in the 1950s to rising income for men, and the decline in fertility in the 1960s to better labor market opportunities for women. Schaller (2016), studying a later time period, finds a similar pattern of sex-specific effects of labor demand on fertility. Jones and Schoonbroodt (2016) argue that the large negative income shock of the Great Depression caused a contemporaneous decline in fertility, followed a generation later by a baby boom. Greenwood et al. (2005) hypothesize that productivity growth in the household sector played a large role in the baby boom by lowering the cost of having children. However, Bailey and Collins (2011) use county-level data on electricity and appliance ownership to show that the baby boom did not occur earlier in places that got electricity and adopted household appliances more quickly. Doepke, Hazan, and Maoz (2015) focus on the role of the positive shocks to demand for female labor during World War II. The war effort created a persistent increase in female labor force participation, which they argue crowded out younger cohorts of women, who then tended to marry earlier and have more children. Zhao (2014) argues that increases in marginal tax rates after WWII decreased the cost of having children, particularly for wealthy households, contributing to the rise in fertility. Hill (2014) posits that the expansion of the housing supply after WWII lowered the cost of marriage and children, accounting for up to 10 percent of the baby boom.

In our main results, we treat the baby boom generation as beginning with the 1946 birth cohort and ending with the 1974 birth cohort. Some authors have treated the baby boom as starting as early as 1939 (e.g., Bailey and Collins, 2011), a local minimum in the fertility rate. However, our goal is to explain changes in educational attainment, and the reversal in longstanding trends in college completion only begins with the cohorts born after the war. In the appendix, we show that our estimated birth order estimates are not sensitive to defining the baby boom generation as the 1939–74 cohorts instead of 1946–74. We use the 1974 birth cohort as the end of the baby boom generation because, as Figure 2 shows, after the mid-1970s, fertility rates overall and among whites were very stable for a quarter century.

*b. Birth order and educational attainment*

Within families, later-born children typically have lower educational attainment than earlier-born children. Using data from Norway, Black et al. (2005) find substantial birth order effects: relative to the first-born child, second-, third-, and fourth-born children complete, respectively, 0.34, 0.52, and 0.61 fewer years of education on average. Subsequent work has found similar patterns in other countries, including Britain (Booth and Kee, 2009) and Denmark (Bagger et al., 2021).

In the U.S., Kantarevic and Mechoulan (2006) use the Panel Study of Income Dynamics (PSID) to show that first-born children have higher levels of educational attainment than later-born children, but their estimates for later-born children are imprecise due to the small sample size. De Haan (2010), using the Wisconsin Longitudinal Study, finds the effects of birth order to be approximately linear, such that second-born children complete approximately 0.3 fewer years of schooling than first-born children, and so on. Neither of these papers studies college completion as an outcome of interest, which is our focus in this paper, but our estimates of birth order effects on high school completion and years of schooling are similar to these previous papers. Compared to these earlier papers, the HRS data we use includes far more observations than Kantarevic and Mechoulan (2006) report in the PSID, and the HRS has the advantage of being nationally representative, unlike the state-specific survey used by De Haan (2010).

While we are narrowly interested in educational attainment in this paper, birth order has been shown to have important effects on a number of other outcomes as well, including cognitive test scores in childhood (Hotz and Pantano, 2015; Lehmann et al., 2016), IQ (Black et al., 2011), health (Black et al., 2016), delinquency (Breining et al., 2020), and earnings (Kantarevic and Mechoulan, 2006).

There is a small literature on the sources of birth order effects, much of which uses U.S. data and focuses on the role of parental attitudes and time investments. Price (2008) finds that parents devote equal quality time to each child at any given point in time and that parents decrease total quality time as their children get older, so that earlier-born children accumulate more quality time. On the other hand, Monfardini and See (2016) find little evidence of differences across siblings in parent time investment. Lehmann et al. (2016) find that differences across siblings in home environment scores — based on factors like whether the mother reads to the child, parental attention, discipline patterns, and so on — can explain most of the effect of birth order on cognitive assessment scores in early childhood. Hotz and Pantano (2015) find that parents are stricter with earlier-born children, potentially leading to better outcomes. In addition to differences in parent time investments, older siblings may get more financial investment. Conley and Glauber (2006) find that an increase in family size decreases the likelihood that second-born boys attend private school, but has no effect on first-born boys. Black et al. (2011), using Norwegian data, find limited evidence that selection plays a role; parents may stop childbearing after having a particularly “poor quality” child. Of course, children may also have an impact on their siblings. For example, older siblings could benefit from teaching younger siblings.

## III. Data

We use data from several sources: the HRS, Vital Statistics, and decennial censuses and the ACS. Before describing the data, it is useful to understand how these sources fit together in our analysis. We use the HRS data to estimate the effects of birth order on college completion within families, because the HRS is the largest available dataset with information on completed education for groups of siblings. The Vital Statistics natality data gives us the population distribution of birth order by year of birth, and we combine changes in this distribution with our estimated birth order effects to assess the role of birth order in explaining the aggregate dynamics of college completion across birth cohorts. We use census and ACS data to show actual trends in college completion, as a baseline against which we measure the effects of changes in birth order.

In this section and throughout the paper, we present results separately for whites and nonwhites. In the HRS, we see three race categories (white, black, and other) and an indicator for Hispanic ethnicity, but our sibling fixed effects estimates quickly become noisy in smaller groups when we try to use more than two race categories. In the Vital Statistics data, we can split nonwhite births into black births and other births for some, but not all, years. Therefore, the white and nonwhite categories are the finest categories that we can consistently distinguish in each of our data sources in each birth cohort. For the birth cohorts that form our primary focus in this paper, the Vital Statistics data indicate that the vast majority of nonwhite births are black births: black births were 95 percent of all nonwhite births in 1946 and 88 percent of all nonwhite births in 1974. Note that the Vital Statistics tabulations, on which we rely for most birth cohorts, do not distinguish between Hispanic and non-Hispanic. For consistency, we also do not make use of ethnicity data in the HRS, and most Hispanic respondents are in the white category.

### *a. HRS data on siblings*

Our estimates of the effect of birth order on educational attainment come from the HRS. The HRS is a nationally representative longitudinal survey of individuals over age 50, and we believe it offers the largest available sample of groups of siblings in the U.S. with data on completed education. The survey began in 1992 with an initial cohort of respondents born 1931–41, and younger cohorts have been added every six years. The survey now includes more than 42,000 respondents. Respondents and their spouses are interviewed every two years, from the time of entry into the survey until death. The HRS collects data on a wide variety of topics, including respondents' health, cognition, family structure, income, assets, and employment. The HRS also asks respondents about their siblings, and we use data on respondents' and siblings' age, sex, and education to estimate the effects of birth order using a family fixed effects design.[4]

We report summary statistics for our HRS data in Table 1. We have data for 46,568 siblings (32,566 white and 14,002 nonwhite) from 20,934 HRS respondents (14,932 white and 6,002 nonwhite), where the HRS respondents are also included in our count of siblings. The range of birth cohorts is large, spanning 1912–91, for a variety of reasons: respondents of course have older and younger siblings, successively later birth cohorts are added to the HRS over time, and HRS survey members may marry younger spouses and in some waves spouses were also asked about siblings. To facilitate a comparison between our sibling sample and census data, we also report summary statistics for the subset of siblings in our HRS data born during the baby boom generation, 1946–74, for HRS respondents from any birth cohort, and for HRS respondents born 1946–74. We do not expect the sample of all siblings to match other nationally representative samples, because the inclusion of HRS respondents' siblings leads to overrepresentation of individuals from large households, as can be seen in high average sibship sizes in the all siblings samples. For the census summary statistics, we use the 1946–74 birth cohorts from the same data files described below in subsection (c).

Within each race-sex group, our sample of HRS respondents born 1946–74 is slightly below the census sample in high school attainment and somewhat above the census sample in attainment of some college. For whites, our HRS respondents born 1946–74 have college completion rates very similar to the census sample, while for nonwhites, the college completion rates are somewhat higher than in the census sample.

Comparing sibship size between the HRS and census is more difficult, because in the census, young children may have more siblings in the future, and older children may have siblings who have already left the household. However, from 1940 to 1990 the census asked women how many children they ever had, and we use data for women ages 40–59 to measure average fertility by the mother's birth cohort. In the census column in Table 1, we report sibship sizes based on fertility among women born in 1930, because this is approximately when the mother of the typical person in our HRS baby

---

[4] For more information about our construction of the sibling data, see the appendix.

boom sample would have been born. The average sibship sizes are similar in the HRS respondent and census samples: for whites born 1946–74, the average sibship size was 4.02 among HRS respondents and 3.80 in the census data. For nonwhites born 1946–74, the average sibship size was 5.21 among HRS respondents and 5.51 in the census data.

*b. Vital Statistics data on birth order*

For information on the national distribution of birth order, we use 1930–80 natality data from Vital Statistics, measured at the state-race-year level. We collect data for birth cohorts 1930–67 from published summary tables, and we tabulate data for birth cohorts 1968–80 from the individual-level birth data. From these sources, we collect information, by mother's state of residence, on the total number of births and the percentage of total births that are first births, second births, and so on.[5]

Figure 3 shows changes in birth order during this period. The top panel plots the percentage of births that were first-borns, second-borns, etc. Here, one can see changes in fertility on the intensive and extensive margins. At the beginning of the boom in the 1940s, there was initially a spike in the percentage of first births as more women were beginning to have children. Shortly after, the percentage of first births declined and the percentage of higher-order births began to rise as these women had additional children. Around 1960, as the fertility rate started to decrease, the percentage of first births slowly started to increase and higher-order births started to decline. By the late 1960s, first and second births became more prevalent as the two-child household became more common.

The bottom panel of Figure 3 shows a different view of the baby boom. Here, we plot the percentage of births that were third births or higher. This shows a decline in higher-order births in the 1930s, a quick spike around World War II, and then a longer, steady increase from the late 1940s until the early 1960s. The same pattern can be seen for whites and nonwhites, and the racial gap declines beginning in the 1960s.

*c. Census data on college completion*

We use individual-level data for people ages 24–65 from the 1960–2000 censuses and the 2006–17 ACS, accessed through IPUMS, to estimate the college completion rate

[5] The share of births with birth order not reported averages 2.6 percent, with a maximum of 5.7 percent in 1945. We drop these births from our sample when measuring the birth order distribution for each birth cohort. All results are robust to allocating birth order not reported cases among birth orders.

for each birth cohort at the state-race-sex level. We use these measures as the baseline from which we construct counterfactuals that illustrate how much of the change in college completion across cohorts can be explained by changes in the distribution of birth order.

Because of the age range we use, we observe a given birth cohort across multiple censuses, and reported educational attainment in a birth cohort can change over the life cycle. Following Card and Lemieux (2000), we adjust the observed educational attainment by regressing attainment at the state-cohort-census level on state-by-cohort fixed effects, a cubic in age (to account for changes in reported education over the life cycle), and an indicator for observations before 1990, when the census education classification was changed from years of schooling to degree attainment. For college attainment, for cohort $c$ born in state $s$ and observed in year $t$, this model is

$$\text{college}_{sct} = \gamma_{sc} + f(\text{age}_t) + \delta\, 1(t < 1990) + \varepsilon_{sct}\,. \tag{1}$$

Our estimate of college completion for cohort $c$ born in state $s$ is the predicted attainment at age 40 according to the new census education question. We estimate college completion separately for white men, white women, nonwhite men, and nonwhite women, and we repeat the process for high school completion and college attendance for each of the four demographic groups. Throughout the paper, we use the terms "completion" and "attainment" interchangeably.

Appendix Figure A2 presents our estimates of the percentage of men and women in each birth cohort between 1930 and 1980 who graduated from college. Panel a shows trends by sex for whites and nonwhites. For all four race-sex groups, college completion fell noticeably below trend beginning with the late 1940s birth cohorts. Among white men, college attainment peaked in 1948 at 34.7 percent before declining to 27.8 percent in 1960. The college completion rate did not return to its 1948 level until the 1979 birth cohort. College attainment for white women experienced a shorter and smaller decline between 1951 and 1956, then increased much more quickly than for white men over the next two decades. For nonwhite men, the decline was smaller than for white men in percentage points, but similar as a percentage drop. College completion for nonwhite women never decreased notably, but it did stagnate at the same time that college completion for white women was declining. In panels b and c, we show these trends for finer race-ethnicity categories: white non-Hispanic, black non-Hispanic, other non-Hispanic, and Hispanic of any race. For both men (panel b) and women (panel c), the levels and trends in college completion are similar for Hispanic and black non-Hispanic. The college graduation rates for other non-Hispanic men and women are higher, but experience a similar slowdown or decline among the 1950s cohorts.

## IV. The effect of birth order on college completion

### *a. Birth order effects on college completion*

A major threat to the identification of birth order effects if using pooled data without family identifiers is that later born children are more likely to come from larger families, which differ in both observable and unobservable ways from smaller families. Following the existing literature on birth order, we address this by using our HRS data on siblings to run regressions with family fixed effects, which will capture any family-level confounders such as family size. However, birth order effects may still be confounded by other factors. For example, as Figure 3 shows, any particular birth order is more common in some birth years than others, making it important to control for cohort effects. For person *i* in family *j* with birth order *b* (*b* = 1, 2, 3, 4, 5, 6 or above) and born in cohort *c*, we estimate the following family fixed effects model of college completion:

$$\text{college}_{ij} = \alpha + \sum_{b=2}^{6} \beta_b \, 1(\text{BO} = b) + \sum_{c} \gamma_c \, 1(\text{cohort} = c) + \mathbf{x}_{ij}'\delta + \lambda_j + \varepsilon_{ij} \tag{2}$$

In our preferred specification, birth orders 6 or higher are grouped into a single 6+ category. The vector of controls, $\mathbf{x}_{ij}$, includes a female indicator; an indicator for whether the individual is an HRS respondent; and the log cohort size, measured from our Vital Statistics data, for the individual's year and census division of birth (Card and Lemieux, 2000; Bound and Turner, 2007). We interact the birth order indicators with a non-baby boom indicator (individuals born before 1946 or after 1974) so that our sample includes all siblings, but the birth order coefficients we report apply specifically to the baby boom generation; these interactions are omitted in the equation for readability. In all birth order regressions, we cluster standard errors at the family level.

We present our estimated birth order effects in Table 2. In the first column, we estimate common effects for men and women. For whites, second-born children are about 5 percentage points less likely than first-borns to graduate from college, and the disadvantage relative to first-borns grows to 11 percentage points for children who are sixth-born or later. These are large effects, given that the college completion rate among whites in these birth cohorts hovers around 30 percent. We cannot directly compare these results to previous estimates in the birth order literature, because we are not aware of any other estimates of the effect of birth order on college completion in the U.S.[6] However, our results are roughly consistent with Booth and Kee (2009), who find that in Britain, university degree attainment is 14 percent for first-borns but less than 9 percent for middle children.

In the final two columns of Table 2, we present separate estimates of birth order effects for men and women, estimated from a single regression in which we interact sex with the birth order indicators. Birth order generally matters more for white men than white women: compared to first-borns, fifth-born men are 12 percentage points less likely to graduate from college, whereas the corresponding difference for women is 8 percentage points. We can reject the null hypothesis that the birth order effects are the same for men and women at the 10 percent level ($p = 0.08$). The larger effects of birth order that we estimate for men are consistent with a broader literature finding that boys are more sensitive to disadvantage. Aucejo and James (2017), Autor et al. (2019), and Lundberg (2017) all find that boys' educational attainment is disproportionately affected by family background (e.g., family structure, resources, and parental attention) compared to girls. Similarly, Chetty et al. (2016) finds that boys are more adversely affected by growing up in poor neighborhoods.

Results for nonwhites follow a similar pattern: later-born children tend to have progressively lower rates of college completion, and the effects are larger for men than for women. The estimated birth order effects are smaller for nonwhites than for whites. Part of this difference may be due to racial differences in educational attainment during this period. For the 1950 birth cohort, the college completion rates for white men and white women were 32.8 and 26.2 percent, respectively, compared to 14.8 and 14.5 percent for nonwhite men and women. The fourth-born coefficient for white women is consistent with a 27 percent decline in college completion from the average, while the coefficient for nonwhite women suggests a 32 percent decline.

[6] Kantarevic and Mechoulan (2006) study birth order effects in the U.S. on years of schooling and high school completion, but not college completion. We compare our results to this earlier paper in section V.

Our birth order estimates come from our sibling dataset created from the HRS, and we made a variety of decisions in the data construction process that could have influenced our results. We test the sensitivity of our estimates to a number of alternative choices, such as limiting the sample to families with 10 or fewer children, limiting the sample to families with complete education and age data for all siblings, excluding individuals reported by siblings-in-law instead of siblings, and including reported siblings born more than 20 years before or after the respondent. Our estimates are robust to all of these choices. Our estimates are also robust to defining the beginning of the baby boom generation using the 1939 birth cohort instead of 1946. We describe these and other robustness checks in detail in the appendix, and we report the results in Appendix Table A1.

Measurement error in birth order, our key explanatory variable, could bias our results, but the direction of the bias is ambiguous and we expect that any bias toward overstated results is likely to be small. One potential source of measurement error is that HRS respondents may misreport the birth years of their siblings, causing the observed birth order to be mixed up relative to the true birth order. We expect the effects of this would be similar to classical measurement error in a continuous explanatory variable, causing our birth order results to be understated. Another measurement issue is that the HRS asks respondents about living siblings, so sibling mortality could artificially compress the distribution of observed birth order relative to the true distribution, which would cause our estimates to be overstated. However, because of the design of the HRS, the vast majority of sibling reports come from respondents who entered the survey when they were ages 50–56, before sibling mortality is likely to cause major measurement issues. In one sensitivity check that we describe in the appendix and report in Appendix Table A1, we limit the sample to siblings reported by HRS respondents who were no older than 56 when first interviewed, and the estimates are very similar to our preferred results reported above. In section III of the appendix, we construct a simulation to measure how the estimated birth order effects are affected by different rates of sibling mortality. We find that for reasonable sibling mortality rates, the resulting bias is only about 4 or 5 percent. For example, if the true second-born effect on college completion was 4 percentage points, we might expect an estimate of about 4.2 percentage points due to the bias caused by deceased siblings.

In the next section, we estimate the change in aggregate college completion due to changes across birth cohorts in the distribution of birth order. Implicitly, we treat the effects of birth order as stable in a couple of important ways that we see as justified in our data. First, we treat the birth order effects as stable across the cohorts that comprise the baby boom generation. In our birth order regressions, if we interact our birth order indicators with an indicator for being born later in the baby boom (1960–74 instead of 1946–59), we cannot reject the null hypothesis that the birth order effects are the same in the two halves of the generation (the $p$-value is 0.78 among whites and 0.50 among nonwhites). Second, in theory the birth order effects could differ by sibship size, which would be important in the next section because the average sibship size by birth cohort evolved similarly to the fertility rates plotted in Figure 2, increasing early in the baby boom and then decreasing. However, if we interact our birth order indicators with sibship size indicators, we cannot reject the null hypothesis that the birth order effects are the same across sibship sizes; the $p$-value is 0.92 among whites and 0.99 among nonwhites. As an illustration, we plot in Appendix Figure A3 our estimated birth order effects by sibship size if we group all race-sex categories into one sample. The point estimates are monotonic in birth order within each sibship size, and often very similar for a given birth order across sibship sizes.

*b. Contribution of changes in birth order to changes in college completion*

To assess the impact of changes in birth order on changes in college completion rates across birth cohorts, we combine the estimated birth order effects from our HRS data with changes in the distribution of birth order across cohorts, measured from our Vital Statistics data. We have in mind a model in which college completion for a given birth cohort is determined by the distribution of birth order within that cohort as well as by other factors. We begin at the individual level before aggregating to the cohort level. Let $i$ index individuals and let $c$ index birth cohorts. Let college completion for someone with birth order $\mathrm{BO}_{ic} = b$ and sibship (family) size $\mathrm{F}_{ic} = s$ be the sum of a birth order effect, $\beta_b$, a sibship size effect, $\zeta_s$, and an idiosyncratic component, $\eta_{ic}$:

$$\text{college}_{icbs} = \beta_b + \zeta_s + \eta_{ic} = \sum_b \beta_b \, 1(\mathrm{BO}_{ic} = b) + \sum_s \zeta_s \, 1(\mathrm{F}_{ic} = s) + \eta_{ic}. \tag{3}$$

As we discussed above, treating birth order effects as common across sibship sizes appears to be a very reasonable simplification in our HRS sibling data.

We are interested in aggregating equation 3 to the cohort level. Let $\mathrm{pBO}b_c$ be the fraction of birth cohort $c$ that has birth order $b$, and let $\mathrm{pF}s_c$ be the fraction of birth cohort $c$ that has sibship size $s$. We normalize to zero the effects for the first birth order ($\beta_1 = 0$)

and sibship size ($\zeta_1 = 0$), and as with the birth order effects presented in the previous section, we group both birth orders and sibship sizes 6 or higher into a single 6+ category. Then the college completion rate for birth cohort $c$ can be written as

$$\text{college}_c = \sum_{b=2}^{6} \beta_b \, \text{pBO}b_c + \sum_{s=2}^{6} \zeta_s \, \text{pF}s_c + \overline{\eta}_c. \tag{4}$$

We are able to obtain causal effects of birth order, $\beta_b$, using family fixed effects estimation with our sibling data, but we cannot estimate sibship size effects. Therefore, we group the sibship size effects and $\overline{\eta}_c$ in equation 4 into a single term, $\theta_c$, that captures all determinants of college completion other than birth order:

$$\text{college}_c = \theta_c + \sum_{b=2}^{6} \beta_b \, \text{pBO}b_c. \tag{5}$$

It is important to note that none of the equations in this subsection are regression specifications. Correlation at the cohort level between the distribution of birth order and other determinants of college completion does not threaten the validity of our counterfactual exercise here, because we are not estimating the birth order effects, $\beta_b$, using cohort-level data on college completion and the distribution of birth order. Instead, for the counterfactual exercise we use our estimated birth order effects presented above, which we take to be causal based on the family fixed effects design.

We wish to construct a counterfactual series of educational attainment in various birth cohorts in which the distribution of birth order is held constant as it was in some base year. We fix 1946, the beginning of the post-war baby boom, as our base year. Counterfactual college attainment in cohort $c$, using the distribution of birth order from the 1946 cohort, is

$$\widetilde{\text{college}}_c = \theta_c + \sum_{b=2}^{6} \beta_b \, \text{pBO}b_{1946}. \tag{6}$$

We can rewrite equation 5 in terms of $\theta_c$ and substitute this into equation 6 to express counterfactual college attainment as actual college attainment minus a birth order effect that is due to the changes in the distribution of birth order since 1946:

$$\widetilde{\text{college}}_c = \text{college}_c - \sum_{b=2}^{6} \beta_b \, (\text{pBO}b_c - \text{pBO}b_{1946}). \tag{7}$$

To estimate these counterfactuals, we compute the changes in birth order, $\text{pBO}b_c - \text{pBO}b_{1946}$, for each birth cohort after 1946 using our Vital Statistics data. For the birth order effects, $\beta_b$, we use the estimates presented above from our HRS sibling data.[7]

In Figure 4, we plot the actual college completion rates ($\text{college}_c$) and the counterfactual college completion rates ($\widetilde{\text{college}}_c$) for all four race-sex groups. By construction, the counterfactual series coincide with observed college completion in 1946, our chosen base year. Choosing a different base year would shift the counterfactual series up or down, but would not affect the year-to-year changes in this series. For white men, college completion decreased more than six percentage points between the late 1940s and late 1950s birth cohorts, before rebounding during the 1960s. The counterfactual series in which birth order is held constant decreases noticeably less, indicating that birth order can explain an important share of the changes in college completion among this generation.

The other race-sex groups did not experience the large decline in college completion we see among white men, but for each of these other groups, college completion was well below trend for the 1950s and 1960s birth cohorts, stagnating or declining early in the baby boom generation before catching up later in the generation. In each case, the counterfactual series does not fall as far below trend as the actual series, indicating that birth order explains some of the changes in college completion.

In Appendix Table A3, we quantify the effects of birth order by race-sex group. We split the baby boom generation into two periods: 1946–60, during which fertility and average birth order were increasing (see Figures 2 and 3) and college completion was falling below trend, and 1960–74, during which fertility and average birth order were decreasing and college completion was accelerating. For white men, college completion fell by 6.3 percentage points across the 1946–60 cohorts. Comparing changes in the actual and counterfactual college completion rates for white men in 1946 and 1960, we find that changes in birth order can explain 1.3 percentage points of this decline, or about 20 percent. This effect of 1.3 percentage points can be seen on the plot in Figure 4: it is the difference between the 1946–60 change in actual college completion and the change in in the counterfactual series that holds constant the distribution of birth order. For the 1960–74 birth cohorts, college completion among white men rose 5.9 percentage points, and

[7] We report the birth order distribution separately for whites and nonwhites in Appendix Table A2. For the birth order effects, we use the sex-specific birth order effects reported in Table 2.

changes in birth order can explain 1.9 percentage points, or about one third, of this increase.

For race-sex groups other than white men, the interpretation of the effects in Appendix Table A3 is more subtle. Although white women and nonwhite men experienced decreases in college completion across the 1950s birth cohorts, and college completion stagnated across these cohorts for nonwhite women, each group had higher college completion at the end of the baby boom generation than at the beginning. Other forces, such as increasing labor market opportunities for women, likely played a large role in the longer-term dynamics of educational attainment, so we prefer to think of changes in birth order as providing a drag on or a boost to these longer-term trends. For example, across the 1946–60 birth cohorts, college completion rose by 4.4 percentage points among white women, but our estimates suggest that changes in birth order prevented this increase from being 0.9 percentage points larger.[8]

**V. When does birth order matter?**

The effect of birth order on college completion could be realized at many different points on the path to a college degree. Here, we focus two of these possible earlier margins: high school completion and college attendance. In each case, we apply the analysis described above, simply changing the dependent variable.

*a. Birth order and high school completion*

In Table 3 we present estimates of the effects of birth order on high school completion. The results indicate that later-born children are less likely to graduate from high school, but the magnitudes of the birth order effects are smaller for high school completion than for college completion. In the first column, we estimate common effects for men and women. The point estimates indicate that white second- and third-born children are about 3 percentage points less likely to complete high school than first-

---

[8] We have also explored ways to estimate the proportion of the "lost" college completion during the baby boom generation that can be accounted for by changes in birth order. But this necessarily involves specifying a hypothetical trend in college completion against which the slowdown can be judged, which is not something we feel confident doing. As a very rough approximation, if we assume a linear trend for each group's college completion connecting the 1946 and 1974 birth cohorts, and then compute the average per-cohort shortfall in college completion relative to this hypothetical trend, we estimate that changes in birth order can explain 20–30 percent of the lost college completion for white men and white women.

borns, but the deficit relative to first-borns stabilizes at 3–5 percentage points for children born fourth or later. In the final two columns, we present separate estimates of birth order effects for men and women, estimated from a single regression in which we interact sex with the birth order indicators. In contrast to the results for college completion, the point estimates indicate that birth order matters more for white women than for white men, but we cannot reject the null hypothesis that the effects are the same for men and women. For nonwhites, the estimated birth order effects are large for men (e.g., 9 p.p. for fourth-borns) but smaller and statistically insignificant for women.

These results are similar to the estimates in Kantarevic and Mechoulan (2006), which, to our knowledge, is the only other paper that estimates birth order effects on high school completion in the U.S. Kantarevic and Mechoulan (2006) use the PSID and report a difference in high school completion between first- and second-borns of about 3 percentage points in their family fixed effects regressions, although they do not estimate this effect by race and sex as we do. Like us, they find that the gap between first- and second-borns is larger than the gap between other pairs of adjacent birth orders. They do not report estimates of the effects of birth order on college attainment — our results above are the first such estimates for college completion in the U.S., as far as we know — but they do estimate the effects of birth order on years of schooling. In results not reported, we find that the difference between first- and second-borns is about 0.3 years of education, slightly larger than the estimate of roughly 0.2 years in Kantarevic and Mechoulan (2006), but very similar to the estimates reported by De Haan (2010) from the Wisconsin Longitudinal Study. As with high school completion, both our estimates and those of Kantarevic and Mechoulan (2006) suggest that the effect of a marginal increment in birth order on years of education declines as birth order increases.

Appendix Figure A4 shows graphically the estimated contribution of birth order to changes in high school completion, using the sex-specific birth order estimates from Table 3 and the counterfactual exercise described above for college completion. High school completion slowed down for all four race-sex groups across the 1950s birth cohorts, but only white men experienced an actual decline in high school completion. The counterfactual series indicate that changes in birth order do explain some of the slowdown in high school completion for all groups, but Appendix Table A3 shows that the effects are small. For the 1960–74 birth cohorts, changes in birth order boosted high school completion for nonwhite men by about 1.6 percentage points, but in all other cases the effects of changes in birth order are less than 1 percentage point.

*b. Birth order and college attendance*

We also estimate the effects of birth order on college attendance, which we define in our HRS data as completing at least 13 years of schooling. In Table 4, we find that the effects of birth order on college attendance, compared to our estimated effects on college completion, are larger for white men and nonwhite women, about the same for white women, and slightly smaller for nonwhite men.

In Appendix Figure A5, we plot actual and counterfactual trends in college attendance. As with college completion, most groups experienced at least some decline in college attendance across the 1950s birth cohorts, and the counterfactual trends indicate that changes in birth order can help explain this phenomenon. Appendix Table A3 quantifies these counterfactual estimates. We find that among white men, changes in birth order can explain almost 20 percent of the decrease in college attendance across the 1946–60 birth cohorts, and almost 40 percent of the increase across the 1960–74 birth cohorts. For the other race-sex groups, changes in birth order were a small drag on college attendance across the 1946–60 birth cohorts, but boosted college attendance by 1.4–2.4 percentage points across the 1960–74 birth cohorts.

## VI. Discussion

We find that changes in the distribution of birth order caused by the baby boom is an important new explanation for the surprising decline in college completion experienced during that generation. Birth order does an especially good job explaining changes in college attainment for white men, accounting for about 20 percent of the decline in college completion across the 1946–60 cohorts, and about one third of the increase in college completion across the 1960–74 cohorts. Birth order effects are large for some other groups, as well. For example, changes in birth order can explain more than 40 percent of the increase in college completion for nonwhite men across the 1960–74 cohorts.

To put the magnitude of our results in perspective, another contributing explanation that has been offered for the decline in college attainment among men is that college attainment for the late 1940s cohorts was unusually high due to the Vietnam War, either because of draft avoidance or the GI Bill. Card and Lemieux (2001) find that the impact of Vietnam draft avoidance behavior on male college completion for the 1947 birth cohort was 2.2 percentage points. Angrist and Chen (2011) argue that the link between increased educational attainment and the Vietnam War was due to the GI Bill, rather than draft avoidance. They find that serving in the military in the Vietnam War increased college completion among white male veterans born 1948–52 by 5 percentage points. Multiplying this effect by the share of these cohorts that were veterans (about 30 percent) yields an estimate of the effect of wartime service on college completion of 1.5 percentage points. We find that changes in birth order during the baby boom are roughly as important: changes in birth order can explain 1.3 percentage points of the decrease in college completion among white men across the 1946–60 birth cohorts. Changes in birth order can also explain some of the slowdown in educational attainment among women during the baby boom generation, giving us an important new explanation for the evolution of human capital during this period of U.S. history.

**References**

Angrist, Joshua D. and Stacey H. Chen. 2011. "Schooling and the Vietnam-era GI Bill: Evidence from the draft lottery." American Economic Journal: Applied Economics, 3(2), 96–118.

Aucejo, Esteban M., and Jonathan James. 2019. "Catching Up to Girls: Understanding the Gender Imbalance in Educational Attainment within Race." Journal of Applied Econometrics, 34(4), 502–525.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2019. "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes." American Economic Journal: Applied Economics, 11(3): 338–381.

Autor, David, Claudia Goldin, and Lawrence F. Katz. 2020. "Extending the Race between Education and Technology." AEA Papers and Proceedings, 110, 347–351.

Bagger, Jesper, Javier A. Birchenall, Hani Mansour, and Sergio Urzúa. 2021. "Education, Birth Order and Family Size." The Economic Journal, 131(633), 33–69.

Bailey, Martha J., and William J. Collins. 2011. "Did Improvements in Household Technology Cause the Baby Boom? Evidence from Electrification, Appliance Diffusion, and the Amish." American Economic Journal: Macroeconomics, 3, 189–217.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "The more the merrier? The effect of family size and birth order on children's education." The Quarterly Journal of Economics, 120(2), 669–700.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2011. "Older and Wiser? Birth Order and IQ of Young Men." CESifo Economic Studies, 57, 103–120.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2016. "Healthy(?), wealthy, and wise: Birth order and adult health." Economics and Human Biology, 23, 27–45.

Booth, Alison L., and Hiau Joo Kee. 2009. "Birth order matters: the effect of family size and birth order on educational attainment." Journal of Population Economics, 22, 367–397.

Bound, John, and Sarah Turner. 2007. "Cohort crowding: How resources affect collegiate attainment." Journal of Public Economics, 91, 877–899.

Breining, Sanni, Joseph Doyle, David N. Figlio, Krzysztof Karbownik, and Jeffrey Roth. 2020. "Birth Order and Delinquency: Evidence from Denmark and Florida." Journal of Labor Economics, 38(1), 95–142.

Butz, William P., and Michael P. Ward. 1979. "The Emergence of Countercyclical U.S. Fertility." American Economic Review, 69(3), 318–328.

Card, David and Thomas Lemieux. 2000. "Dropout and Enrollment Trends in the Post-War Period: What Went Wrong in the 1970s?" NBER Working Paper 7658.

Card, David and Thomas Lemieux. 2001. "Going to College to Avoid the Draft: The Unintended Legacy of the Vietnam War." American Economic Review: Papers and Proceedings, 91(2), 97–102.

Chetty, Raj, Nathaniel Hendren, Frina Lin, Jeremy Majerovitz, and Benjamin Scuderi. 2016. "Childhood Environment and Gender Gaps in Adulthood." American Economic Review Papers and Proceedings, 106(5): 282–88.

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, 41(4), 722–737.

De Haan, Monique. 2010. "Birth order, family size and educational attainment." Economics of Education Review, 29, 576–588.

Doepke, Matthias, Moshe Hazan, and Yishay D. Maoz. 2015. "The Baby Boom and World War II: A Macroeconomic Analysis." Review of Economic Studies, 82, 1031–1073.

Easterlin, Richard A. 1976. "The Conflict between Aspirations and Resources." Population and Development Review, 2(3/4), 417–425.

Goldin, Claudia and Lawrence F. Katz. 2008. The Race Between Education and Technology. Cambridge, MA: Harvard University Press.

Greenwood, Jeremy, Ananth Seshadri, and Guillaume Vandenbroucke. 2005. "The Baby Boom and Baby Bust." American Economic Review, 95(1), 183–207.

Health and Retirement Study public use dataset. 2019. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI.

Hill, Matthew J. 2014. “Homes and Husbands for All: Marriage, Housing, and the Baby Boom.” Working paper.

Hotz, V. Joseph, and Juan Pantano. 2015. “Strategic parenting, birth order, and school performance.” Journal of Population Economics, 28, 911–936.

Jones, Larry E., and Alice Schoonbroodt. 2016. “Baby busts and baby booms: The fertility response to shocks in dynastic models.” Review of Economic Dynamics, 22, 157–178.

Jones, Larry E., and Michele Tertilt. 2006. “An Economic History of Fertility in the U.S.: 1826–1960.” NBER Working Paper No. 12796.

Kantarevic, Jasmin, and Stéphane Mechoulan. 2006. “Birth Order, Educational Attainment, and Earnings: An Investigation Using the PSID.” Journal of Human Resources, 41(4), 755–777.

Klein, Herbert S. 2006. “The U.S. Baby Bust in Historical Perspective.” In Fred R. Harris (ed.), *The baby bust: Who will do the work? Who will pay the taxes?*, Rowman & Littlefield.

Lehmann, Jee-Yeon K., Ana Nuevo-Chiquero, and Marian Vidal-Fernandez. 2016. “The Early Origins of Birth Order Differences in Children’s Outcomes and Parental Behavior.” Journal of Human Resources, 53(1), 123–156.

Lundberg, Shelly. 2017. “Father Absence and the Educational Gender Gap.” IZA Discussion Paper 10814.

Monfardini, Chiara, and Sarah Grace See. 2016. “Birth order and child cognitive outcomes: an exploration of the parental time mechanism.” Education Economics, 24(5), 481–495.

National Center for Health Statistics (NCHS), Vital Statistics of the United States, Volume 1 – Natality, 1930–67.

National Center for Health Statistics (NCHS) Data File Documentations, Natality, 1968–80 (machine readable data file and documentation), National Center for Health Statistics, Hyattsville, Maryland.

Price, Joseph. 2008. “Parent-Child Quality Time: Does Birth Order Matter?” Journal of Human Resources, 43(1), 240–265.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. IPUMS USA: Version 8.0 [1960, 1980, 1990, and 2000 5% samples; 1970 1% samples; 2006–2017 ACS]. Minneapolis, MN: IPUMS, 2018.

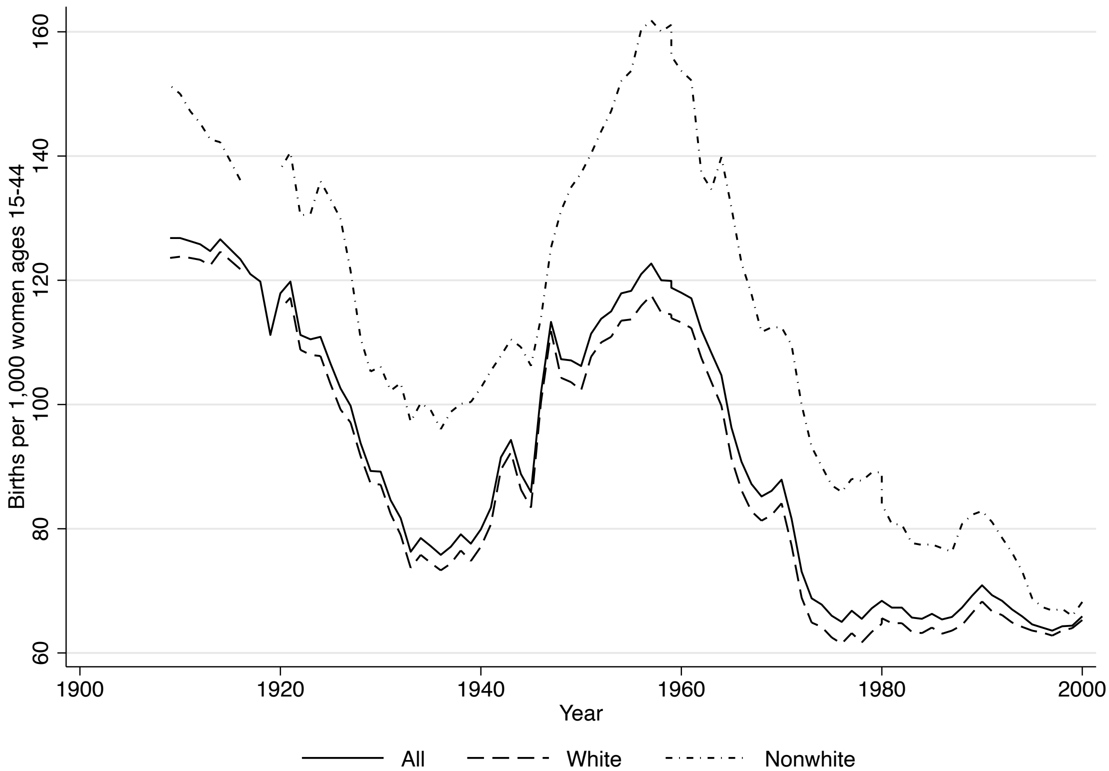
Schaller, Jessamyn. 2016. "Booms, Busts, and Fertility: Testing the Becker Model Using Gender-Specific Labor Demand." Journal of Human Resources, 51(1), 1–29.

Zhao, Kai. 2014. "War finance and the baby boom." Review of Economic Dynamics, 17, 459–473.

Figure 1: White male college completion and the share of individuals third-born or higher


Notes: Educational attainment for each cohort is measured using data for 36–45 year olds from the 1970–2000 censuses and the 2006–17 ACS. Birth order data is from Vital Statistics and is aggregated across all states (including Washington, DC) with reported birth data except for Alaska and Hawaii. South Dakota is missing from the Vital Statistics data in 1930–31 and Texas is omitted in 1930–32.

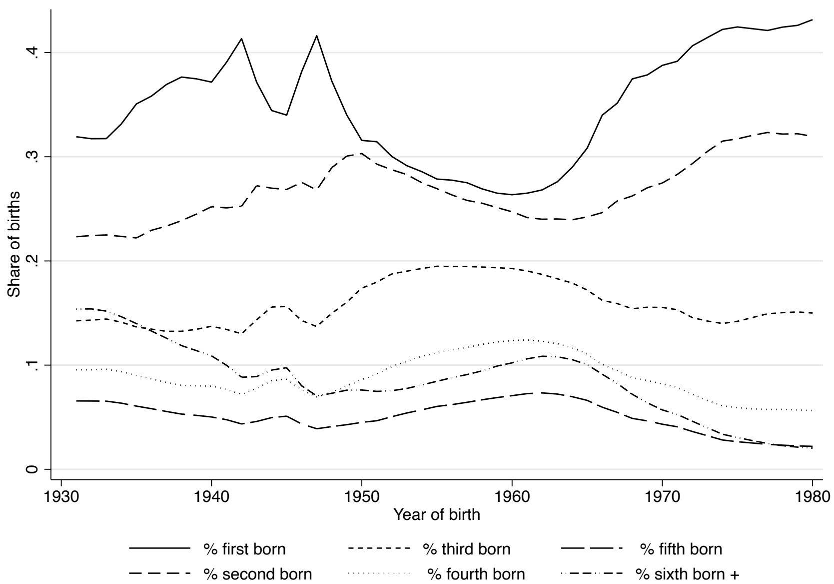
Figure 2: Fertility rates by year, 1909–2000



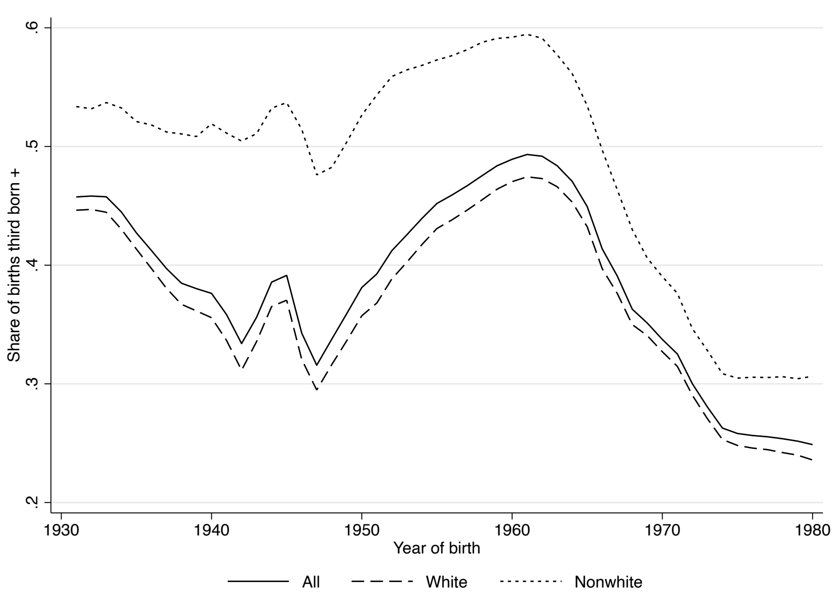

Notes: Data is from https://www.cdc.gov/nchs/data/statab/t001x01.pdf. Fertility rates are defined as births per 1,000 women aged 15–44. Births to nonresidents are excluded beginning in 1970. Race is defined by mother from 1980–2000 and by child before 1980. Birth counts before 1959 are adjusted for underregistration.

Figure 3: Distribution of birth order by birth year, 1930–80

a. Share of births in each cohort with a given birth order

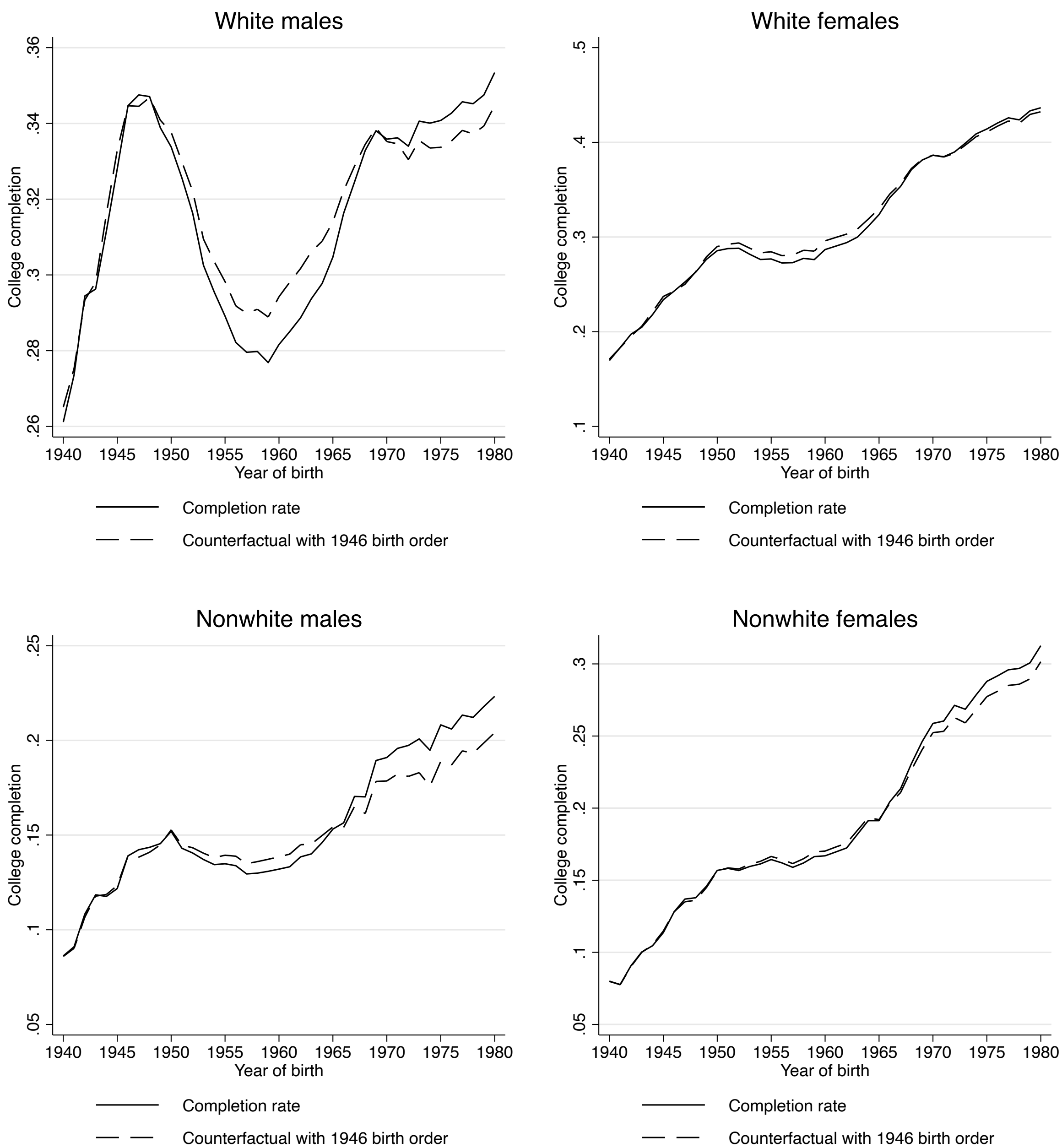



b. Share of births in each cohort that are third births or higher




Notes: Data is from Vital Statistics and is aggregated across all states (including Washington, DC) with reported birth data except for Alaska and Hawaii. South Dakota is missing from the Vital Statistics data in 1930–31 and Texas is omitted in 1930–32.

Figure 4: Actual and counterfactual trends in college completion




Notes: We plot estimated educational attainment at age 40 from the census and ACS for birth cohorts 1940–80. We use the sex-specific birth order coefficients in Table 2 to compute counterfactual educational attainment for each cohort using the 1946 birth order distribution; see text for details about these counterfactual series.

Table 1: Summary statistics

| | HRS respondents and siblings | | HRS respondents only | | Census |
|---|---|---|---|---|---|
| | All cohorts | Born 1946–74 | All cohorts | Born 1946–74 | Born 1946–74 |
| *White* | | | | | |
| Birth year | 1947.7 | 1956.2 | 1944.0 | 1956.3 | 1959.7 |
| Female | 0.534 | 0.534 | 0.567 | 0.583 | 0.504 |
| Sibship size | 4.22 | 4.61 | 3.46 | 4.02 | 3.80 |
| Men | | | | | |
| High school grad | 0.846 | 0.871 | 0.835 | 0.881 | 0.902 |
| Some college | 0.477 | 0.512 | 0.545 | 0.623 | 0.566 |
| College grad | 0.288 | 0.309 | 0.304 | 0.340 | 0.312 |
| Women | | | | | |
| High school grad | 0.847 | 0.878 | 0.834 | 0.884 | 0.922 |
| Some college | 0.463 | 0.527 | 0.499 | 0.604 | 0.589 |
| College grad | 0.253 | 0.308 | 0.246 | 0.321 | 0.309 |
| # siblings (*N*) | 32,566 | 18,732 | 14,900 | 6,893 | |
| # families | 14,932 | 8,408 | 14,900 | 6,893 | |
| *Nonwhite* | | | | | |
| Birth year | 1954.5 | 1958.2 | 1952.3 | 1958.2 | 1961.0 |
| Female | 0.552 | 0.545 | 0.589 | 0.576 | 0.530 |
| Sibship size | 5.75 | 5.89 | 4.93 | 5.21 | 5.51 |
| Men | | | | | |
| High school grad | 0.742 | 0.766 | 0.713 | 0.754 | 0.804 |
| Some college | 0.357 | 0.371 | 0.432 | 0.466 | 0.414 |
| College grad | 0.180 | 0.186 | 0.192 | 0.210 | 0.158 |
| Women | | | | | |
| High school grad | 0.773 | 0.807 | 0.738 | 0.797 | 0.839 |
| Some college | 0.417 | 0.447 | 0.469 | 0.531 | 0.490 |
| College grad | 0.209 | 0.222 | 0.205 | 0.230 | 0.193 |
| # siblings (*N*) | 14,002 | 10,482 | 5,959 | 4,078 | |
| # families | 6,002 | 4,522 | 5,959 | 4,078 | |

Notes: For the HRS sample, we use sibling data for all HRS waves, 1992–2016. For more details about our construction of this sibling sample, see the appendix. The sample size in the first column exceeds what we report for the birth order regressions because only children are included in this table but not in regressions with family fixed effects. For the census sample, we use the 1960–2000 censuses and 2006–17 ACS. Educational attainment in the census sample is adjusted for age, as described in section III(c). Sibship size in the census sample is reported for children of mothers who were born in 1930, approximately the typical mothers' birth cohort for individuals in our HRS sibling sample who were born 1946–74. All statistics in both data sets are computed using sampling weights.

Table 2: The effect of birth order on college completion

| White | | | |
|---|---|---|---|
| | All | Male | Female |
| birth order 2 | –0.049*** | –0.037* | –0.059*** |
| | (0.013) | (0.020) | (0.018) |
| birth order 3 | –0.076*** | –0.084*** | –0.069*** |
| | (0.016) | (0.022) | (0.021) |
| birth order 4 | –0.070*** | –0.069*** | –0.070** |
| | (0.019) | (0.025) | (0.024) |
| birth order 5 | –0.095*** | –0.117*** | –0.076** |
| | (0.023) | (0.030) | (0.028) |
| birth order 6+ | –0.109*** | –0.141*** | –0.082** |
| | (0.027) | (0.032) | (0.031) |
| Birth year fixed effects | x | x | |
| Family fixed effects | x | x | |
| # siblings ($N$) | 29,041 | 29,041 | |
| # families | 11,407 | 11,407 | |

| Nonwhite | | | |
|---|---|---|---|
| | All | Male | Female |
| birth order 2 | –0.041** | –0.075*** | –0.014 |
| | (0.018) | (0.026) | (0.025) |
| birth order 3 | –0.037* | –0.063** | –0.015 |
| | (0.020) | (0.029) | (0.027) |
| birth order 4 | –0.065*** | –0.090*** | –0.046 |
| | (0.024) | (0.032) | (0.030) |
| birth order 5 | –0.074** | –0.107*** | –0.049 |
| | (0.030) | (0.037) | (0.036) |
| birth order 6+ | –0.088*** | –0.125*** | –0.059 |
| | (0.033) | (0.040) | (0.037) |
| Birth year fixed effects | x | x | |
| Family fixed effects | x | x | |
| # siblings ($N$) | 13,065 | 13,065 | |
| # families | 5,062 | 5,062 | |

Notes: Data are from the HRS. All regressions include an indicator for sex and an indicator for whether the individual is an HRS respondent, as well as the log cohort size, measured from Vital Statistics data, for the individual's year and census division of birth. We interact each birth order indicator with an indicator for being born outside the baby boom generation (defined as the 1946–74 birth cohorts), and report the coefficients for the baby boom generation. Male and female coefficients are estimated in a single regression in which sex is interacted with birth order. Regressions are weighted using HRS sampling weights. Standard errors are clustered at the family level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: The effect of birth order on high school completion

| White | | | |
|---|---|---|---|
| | All | Male | Female |
| birth order 2 | –0.030*** | –0.029** | –0.031*** |
| | (0.008) | (0.012) | (0.012) |
| birth order 3 | –0.029*** | –0.022 | –0.036** |
| | (0.011) | (0.014) | (0.014) |
| birth order 4 | –0.048*** | –0.037** | –0.058*** |
| | (0.015) | (0.018) | (0.017) |
| birth order 5 | –0.032* | –0.030 | –0.034* |
| | (0.018) | (0.022) | (0.021) |
| birth order 6+ | –0.042* | –0.034 | –0.048* |
| | (0.021) | (0.025) | (0.025) |
| Birth year fixed effects | x | x | |
| Family fixed effects | x | x | |
| # Siblings (N) | 29,041 | 29,041 | |
| # Families | 11,407 | 11,407 | |

| Nonwhite | | | |
|---|---|---|---|
| | All | Male | Female |
| birth order 2 | –0.033** | –0.064*** | –0.007 |
| | (0.015) | (0.024) | (0.021) |
| birth order 3 | –0.043** | –0.077*** | –0.015 |
| | (0.018) | (0.026) | (0.023) |
| birth order 4 | –0.055** | –0.089*** | –0.028 |
| | (0.023) | (0.030) | (0.028) |
| birth order 5 | –0.062** | –0.071* | –0.058* |
| | (0.027) | (0.034) | (0.032) |
| birth order 6+ | –0.018 | –0.071* | 0.025 |
| | (0.032) | (0.038) | (0.034) |
| Birth year fixed effects | x | x | |
| Family fixed effects | x | x | |
| # siblings ($N$) | 13,065 | 13,065 | |
| # families | 5,062 | 5,062 | |

Notes: Data are from the HRS. All regressions include an indicator for sex and an indicator for whether the individual is an HRS respondent, as well as the log cohort size, measured from Vital Statistics data, for the individual's year and census division of birth. We interact each birth order indicator with an indicator for being born outside the baby boom generation (defined as the 1946–74 birth cohorts), and report the coefficients for the baby boom generation. Male and female coefficients are estimated in a single regression in which sex is interacted with birth order. Regressions are weighted using HRS sampling weights. Standard errors are clustered at the family level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: The effect of birth order on college attendance

| White | All | Male | Female |
|---|---|---|---|
| birth order 2 | –0.049*** | –0.055*** | –0.044*** |
| | (0.014) | (0.021) | (0.019) |
| birth order 3 | –0.088*** | –0.111*** | –0.068*** |
| | (0.017) | (0.023) | (0.022) |
| birth order 4 | –0.096*** | –0.107*** | –0.086*** |
| | (0.021) | (0.027) | (0.025) |
| birth order 5 | –0.112*** | –0.159*** | –0.071** |
| | (0.026) | (0.032) | (0.030) |
| birth order 6+ | –0.134*** | –0.187*** | –0.088** |
| | (0.030) | (0.035) | (0.035) |
| Birth year fixed effects | x | x | |
| Family fixed effects | x | x | |
| # siblings (*N*) | 29,041 | 29,041 | |
| # families | 11,407 | 11,407 | |

| Nonwhite | All | Male | Female |
|---|---|---|---|
| birth order 2 | –0.027 | –0.028 | –0.026 |
| | (0.020) | (0.030) | (0.029) |
| birth order 3 | –0.037 | –0.040 | –0.035 |
| | (0.023) | (0.031) | (0.031) |
| birth order 4 | –0.063** | –0.072* | –0.056 |
| | (0.029) | (0.037) | (0.035) |
| birth order 5 | –0.062* | –0.053 | –0.071* |
| | (0.034) | (0.042) | (0.041) |
| birth order 6+ | –0.099** | –0.120*** | –0.080* |
| | (0.039) | (0.045) | (0.044) |
| Birth year fixed effects | x | x | |
| Family fixed effects | x | x | |
| # siblings (*N*) | 13,065 | 13,065 | |
| # families | 5,062 | 5,062 | |

Notes: Data are from the HRS. All regressions include an indicator for sex and an indicator for whether the individual is an HRS respondent, as well as the log cohort size, measured from Vital Statistics data, for the individual's year and census division of birth. We interact each birth order indicator with an indicator for being born outside the baby boom generation (defined as the 1946–74 birth cohorts), and report the coefficients for the baby boom generation. Male and female coefficients are estimated in a single regression in which sex is interacted with birth order. Regressions are weighted using HRS sampling weights. Standard errors are clustered at the family level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

## Online appendix for Handy and Shester, "Birth order and the decline in college completion among the baby boom generation," *Economic Inquiry*

### I. Details on data construction

*a. Health and Retirement Survey (HRS)*

We use data from the HRS on respondents and their siblings. Data on respondents for all waves, 1992–2016, is from the RAND HRS Longitudinal File, and data on siblings comes from the HRS files for each wave.

Sibling age was collected in the 1992 and 1994 HRS waves, but not in the 1993 and 1995 waves for the oldest cohort. Beginning in 1996, sibling age and educational attainment is collected in each biannual wave. In 1992–2000, the household's family respondent reports on both siblings and siblings-in-law. We use relationship codes to assign the sibling to the appropriate respondent. If the family respondent reports a sibling-in-law but is not married in that wave, we assign the sibling to the respondent's most recent spouse. If a sibling cannot be assigned in this way, they are dropped from the sample. Beginning with the 2002 wave, the family respondent reports only on his or her own siblings.

We set reported education to missing for any siblings younger than 25 when education is reported, which happens in a small number of cases in which an HRS respondent is married to someone younger who reports younger siblings. For some siblings, there is data on age and/or education from multiple waves. We infer the sibling's birth year from the reported age. If two siblings have the same birth year, we assume that the sibling reported first was born earlier. We then drop siblings with any inconsistent reports about sex, a range of reported birth years greater than 2, or a range of reported years of education greater than 1. Finally, we resolve the remaining small inconsistencies by taking the median of the sibling's reported birth years and completed years of education, rounded to the nearest integer.

*b. Vital Statistics*

Data on total births and the population distribution of birth order by year of birth come from Vital Statistics natality files. Data for the 1931–67 birth cohorts come from summary tables in published pdfs. Data for the 1968–2000 birth cohorts come from individual-level natality data from the NBER Vital Statistics page.

The share of births with birth order not reported averages 2.6 percent, with a maximum of 5.7 percent in 1945. We drop these births from our sample when measuring the birth order distribution for each birth cohort. If we instead allocate birth order not reported cases among birth order — by estimating separate regressions of the percentage of births of each birth order on the percentage of births without reported birth order, state fixed effects, and year fixed effects, and allocating birth order not reported cases across birth order bins based on these coefficients — our results are very similar.

We report annual data on total births and the distribution of birth order by race in Table A2.

**II. Sensitivity checks for estimates of birth order effects**

We test the sensitivity of our reported birth order effects to changes in our definition of the baby boom generation, and to changes in our construction of the sibling sample in the HRS. We describe the sensitivity checks here, and we show the results in Table A1. We report results of these checks only for white male college completion, because that is the key result in our paper, but results of the same sensitivity checks for other groups or other outcomes are available on request.

In Table A1, specification 1 is identical to the birth order effects reported in Table 2, and the effect of changes in birth order on college completion is identical to what is reported in Table A3.

In specification 2, we expand our definition of the baby boom generation to birth years 1939–74, instead of 1946–74 as in our base specification, because some authors (e.g., Bailey and Collins, 2011) use 1939 as the beginning of the baby boom. This affects our birth order estimates because we interact our birth order indicators with a non-baby boom indicator, and report estimates specific to the baby boom generation. Here, our sample stays the same, but the reported birth order effects apply to the 1939–74 cohorts.

Specifications 3–9 address choices we made in the construction of our sibling data. In our base sample, we include some HRS respondents who report many siblings, but some of these reports could be erroneous or duplicates. Specification 3 limits our sample to respondents and their siblings from families with 10 or fewer children (9 or fewer reported siblings). In our base sample, we drop reported siblings who are more than 20 years older or younger than the respondent. In specification 4, we include these possible siblings.

In our base sample, when siblings are missing age data, we assign birth order among the siblings with valid age data, implicitly assuming the siblings with valid data were born first. In specification 5, we instead assume the siblings with valid age data were born last. In specification 6, if any sibling is missing data on age or education, we drop the entire family from the sample.

In our base sample, we drop siblings with any inconsistent reports about sex, a range of reported birth years greater than 2, or a range of reported years of education greater than 1. In specification 7, we include these siblings in the sample. When we do so, we resolve inconsistent reports about sex by assuming the most commonly reported sex, and we resolve inconsistent reports about birth year and education in the same way as we do for our main sample, by taking the median of the sibling's reported birth years and completed years of education, rounded to the nearest integer.

The HRS asks about living siblings, and there is a concern that deceased siblings could artificially compress the distribution of birth order within a family and inflate the magnitude of our estimated birth order effects. Most HRS respondents entered the survey at ages 50–56 and report data on siblings in the first wave they are interviewed, and we expect any bias from sibling mortality to be greater among families in which the HRS respondent entered the survey at a later age. In our base sample, we keep all sibling data regardless of the age at which the HRS respondent entered the survey. In specification 8, we limit the sample to siblings reported by HRS respondents who were age 56 or younger when they entered the survey. We further explore the potential effects of sibling mortality in section III below.

In earlier HRS waves, the family respondent reports both on his or her own siblings and his or her spouse's siblings, but reports about siblings-in-law might be of lower quality. In specification 9, we exclude individuals reported by siblings-in-law.

**III. Potential bias in birth order effects due to deceased siblings**

As we noted above, the HRS asks about living siblings, and this could create a problematic bias: because we derive birth order from the reported ages of the siblings, missing (deceased) siblings will cause the measured distribution of birth order to be too compressed, inflating the magnitude of the birth order estimates. In a sensitivity check above, we limit the sample to siblings reported by HRS respondents who were age 56 or younger when they entered the survey, as a way of minimizing the potential effects of sibling mortality. In this section, we further explore this issue by simulating the effects of sibling mortality on our birth order estimates.

We can easily quantify the degree of the bias in a very simplified case that may provide more intuition than the simulations we discuss below. Let $y$ be college completion and let $x$ be birth order, and suppose the true regression model is $y = \alpha + \beta x + \varepsilon$. Suppose that instead of including $x$ in the regression, we include the mismeasured variable $\widetilde{x} = px$, where $0 < p < 1$ represents the survival probability. This is an unrealistic setup, but captures in a simple way the idea that sibling mortality will artificially compress the observed distribution of birth order. In a regression of $y$ on $\widetilde{x}$, which is a scaled-down version of the true birth order $x$, the slope coefficient, $\widetilde{\beta}$, will be inflated relative to the true slope $\beta$ by a factor of $1/p$: $\widetilde{\beta} = \dfrac{\text{Cov}(\widetilde{x}, y)}{\text{Var}(\widetilde{x})} = \dfrac{p\,\text{Cov}(x, y)}{p^2\text{Var}(x)} = \dfrac{1}{p}\beta$.

Every six years, the HRS expands by adding a sample of people born 51–56 years prior. (See "Longitudinal Cohort Sample Design" at the HRS methodology page.) In 2014, CDC life tables indicate that, among whites, survivorship to age 53 was about 93 percent (National Vital Statistics Reports, volume 66, number 4, table 4). For a survival rate of $p = 0.93$, the $1/p$ scaling factor implies that the birth order estimates would be overstated by (1/0.93) – 1 = 7.5 percent. However, in simulations that better match our empirical strategy, we find that the actual bias is even smaller than this.

In our simulations, we first generated 100,000 families of a given sibship size. For clarity in interpreting the results, we assume the birth order effects are linear and normalize the effect to 1 for each increment of birth order. For example, we set the outcome to 1 for first-borns and 3 for third-borns, so that the correct estimate for the effect of being third-born, relative to first-born, is $\hat{\beta}_3 = 2$. We then randomly keep a fraction $p$ of all the siblings. Finally, we regress the outcome on *observed* birth order and family fixed effects. We repeat this simulation for different survival probabilities, using values of $p$ that include survival probabilities both above and below the 93 percent we mentioned above.

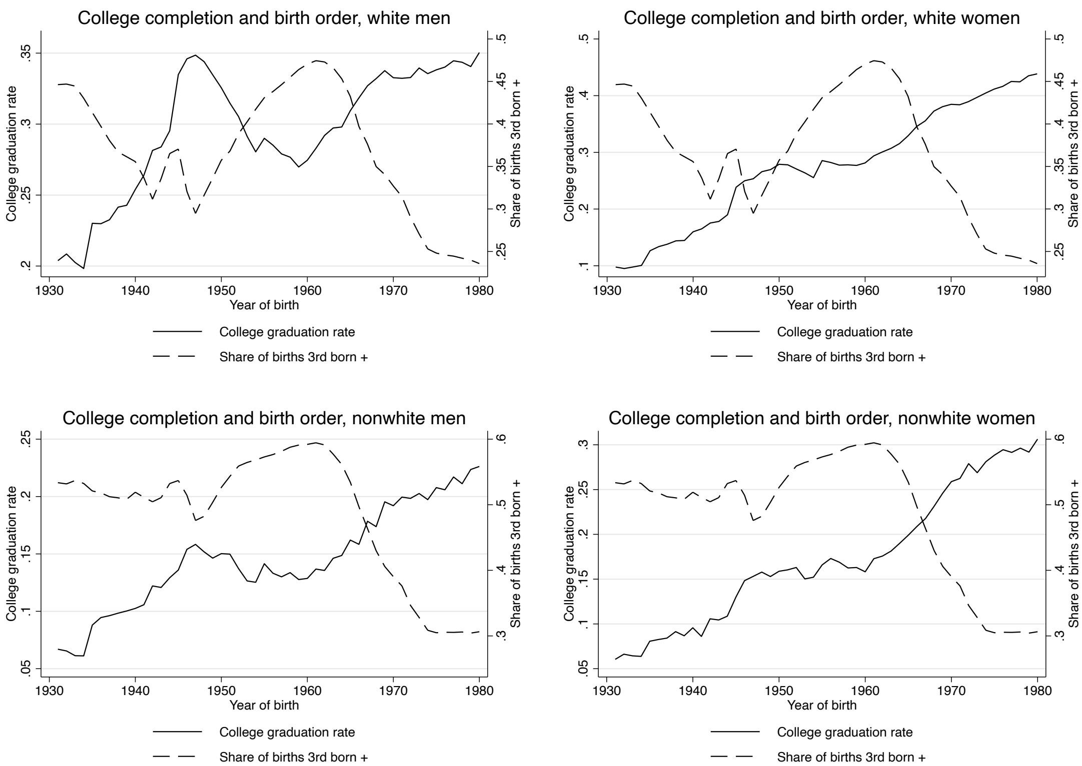
The results for sibship sizes 3 and 4 are presented in Table A4. (We cannot run this simulation for sibship size 2, because any deceased siblings would cause the remaining sibling to be omitted from a family fixed effects regression.) The top row shows estimated effects when all siblings survive ($p = 1$); in this scenario, the estimated effects are the true effects. When the probability of survival is less than 1, the estimated birth order effects are biased upwards. The bias in the estimate of the second-born effect, $\hat{\beta}_2$, is close to the simple $1/p$ scaling factor we mentioned above, but the bias is smaller for higher-order births within a given sibship size. For example, when 90 percent of siblings survive to appear in the survey (so that $1/p$ implies the estimate is inflated by (1/0.9) – 1 = 11 percent), we find that among sibship size 3, the second-born effect is inflated by 8 percent ((1.08 – 1)/1), and the third-born effect is inflated by 2 percent ((2.04 – 2)/2).

To get a sense of how this might affect the mix of sibship sizes we see in the baby boom generation, we modified our simulation to include a mix of sibship sizes. To estimate the appropriate distribution of sibship sizes, we used data from the June 1980 Current Population Survey fertility supplement on completed fertility among a sample of white women who had their first child between 1945 and 1965. We found that 13.0 percent of women in this sample had one child, 28.5 percent had two children, and so on, and we constructed our simulation accordingly. The resulting family fixed effects estimates of birth order effects are reported in Table A5. (Note that the CPS top-coded number of children at 10, but we top-coded the observed birth order at 6 to match the estimates in our paper.)

In our simulation, with a survival probability of $p = 0.93$, the estimated second-born effect is overstated by about 5 percent, the third-born effect is overstated by about 4 percent (2.08/2), the third-born effect is also overstated by about 4 percent (3.12/3), and so on. Therefore, if the true second-born effect on college completion was 4 percentage points ($\beta_2 = -0.04$), we might expect an estimate of about 4.2 percentage points due to the bias caused by deceased siblings, and if the true fifth-born effect was 8 percentage points, we might expect an estimate of about 8.3 percentage points.

As we discuss in the paper, sibling mortality is almost certainly not the only source of measurement error in birth order. In particular, it is likely that some HRS respondents misstate the birth years of their siblings in ways that mix up the birth order we observe relative to the true order. We expect that this would attenuate our birth order effect estimates toward zero, similar to the effects of classical measurement error. Without knowing how common this type of measurement error is, we cannot say whether our estimates are overstated or understated.

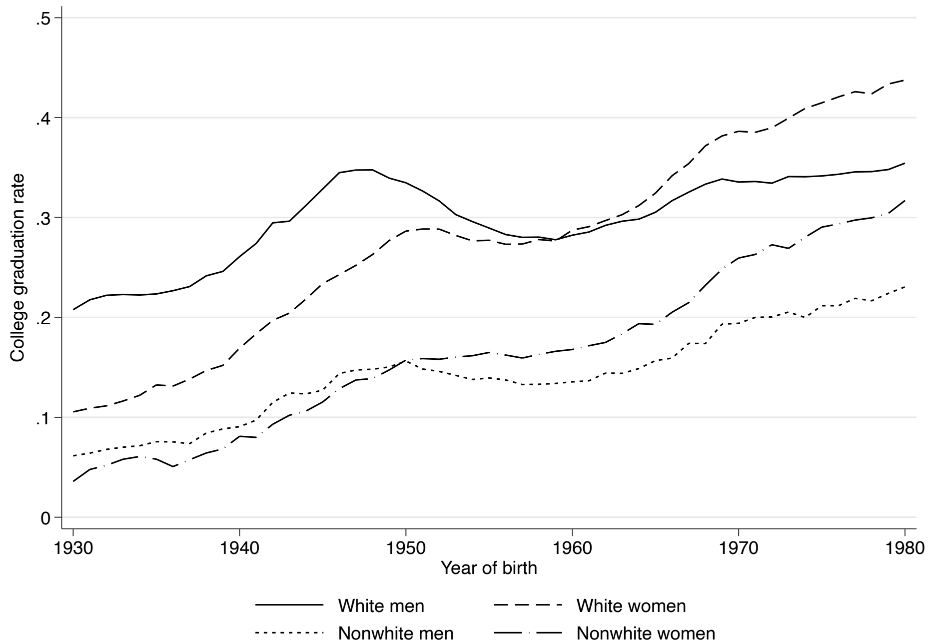
Figure A1: College completion and the percentage of individuals third-born or higher



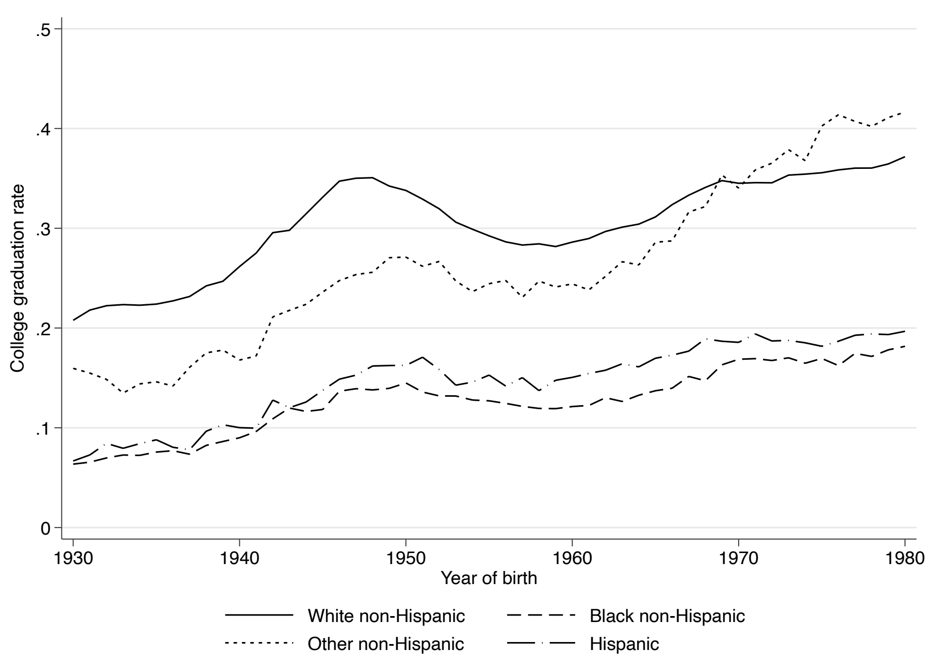

Notes: Educational attainment for each cohort is measured using data for 36–45 year olds from the 1970–2000 censuses and the 2006–17 ACS. Birth order data is from Vital Statistics and is aggregated across all states (including Washington, DC) with reported birth data except for Alaska and Hawaii. South Dakota is missing from the Vital Statistics data in 1930–1931 and Texas is omitted in 1930–1932.

Figure A2: College graduation by birth year, 1930–80

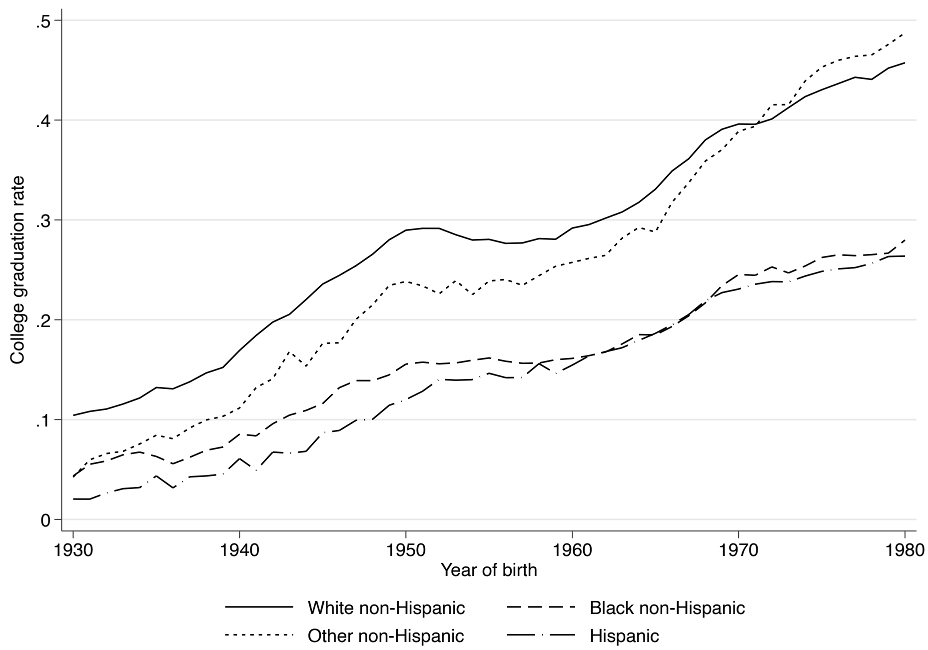
a. White and Nonwhite




b. Men: White non-Hispanic, Hispanic, Black non-Hispanic, and Other non-Hispanic

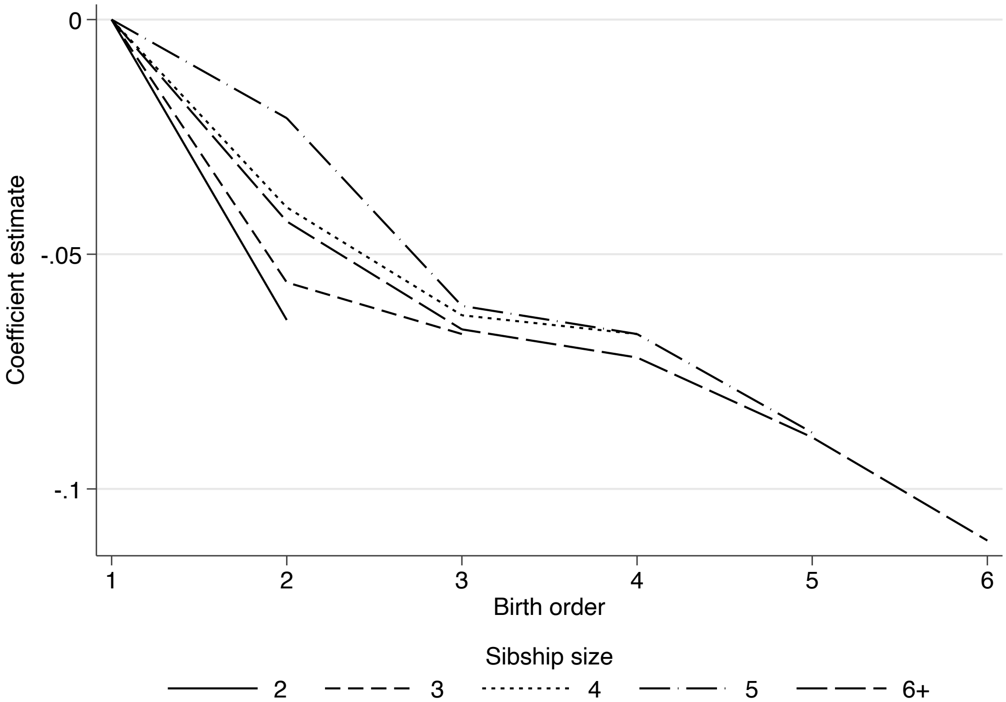


c. Women: White non-Hispanic, Hispanic, Black non-Hispanic, and Other non-Hispanic



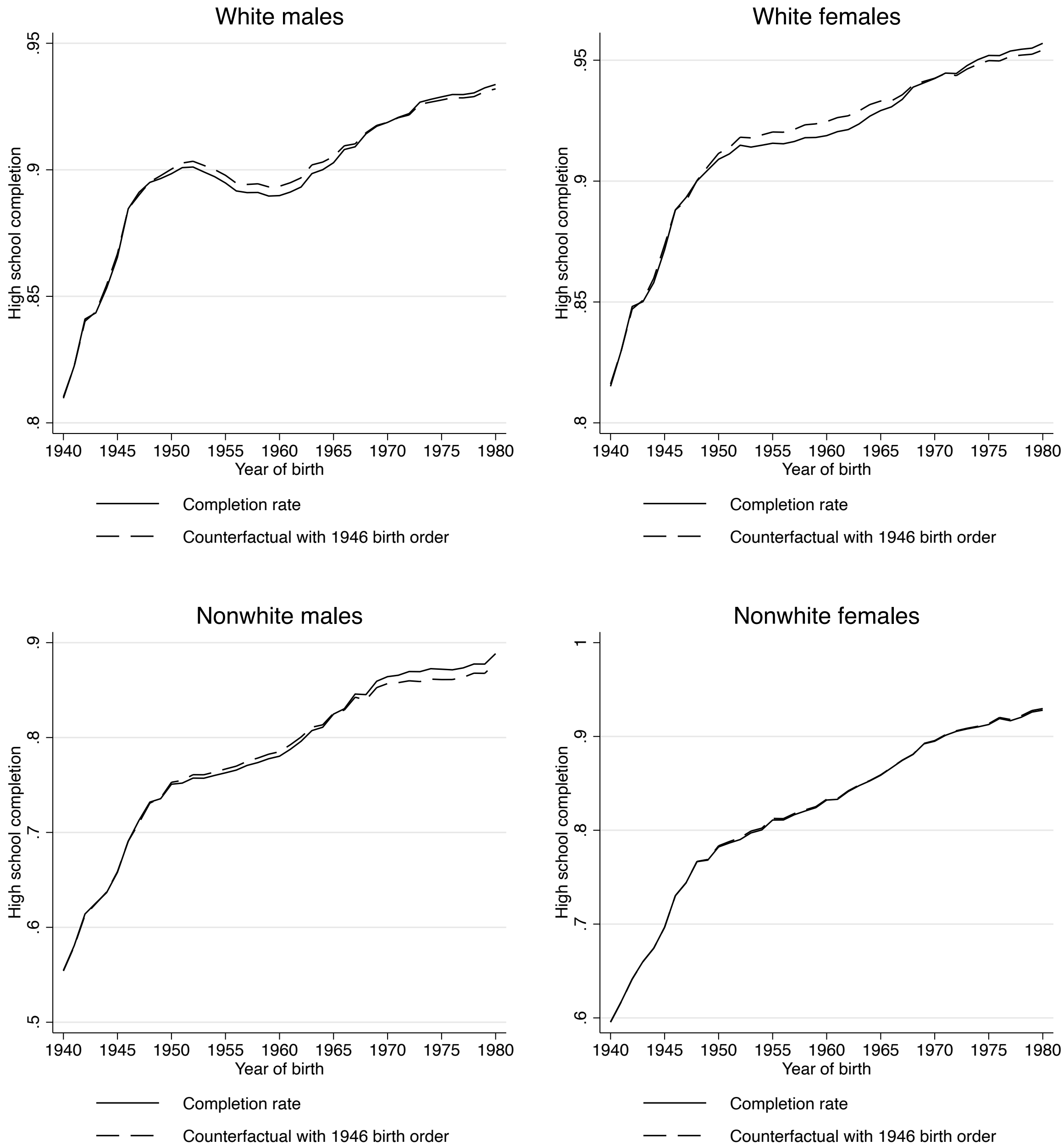

Notes: Educational attainment for each cohort is estimated using data for 25–64 year olds from the 1980–2000 censuses and the 2006–17 ACS. We regress educational attainment on an age cubic, state-by-cohort fixed effects, and an indicator for whether the sample was before 1990, as the education question changed slightly at that time. In the figures above, we plot predicted educational attainment for each birth cohort at age 40 according to the new census education question.

Figure A3: Estimated effects of birth order on college completion, by family size



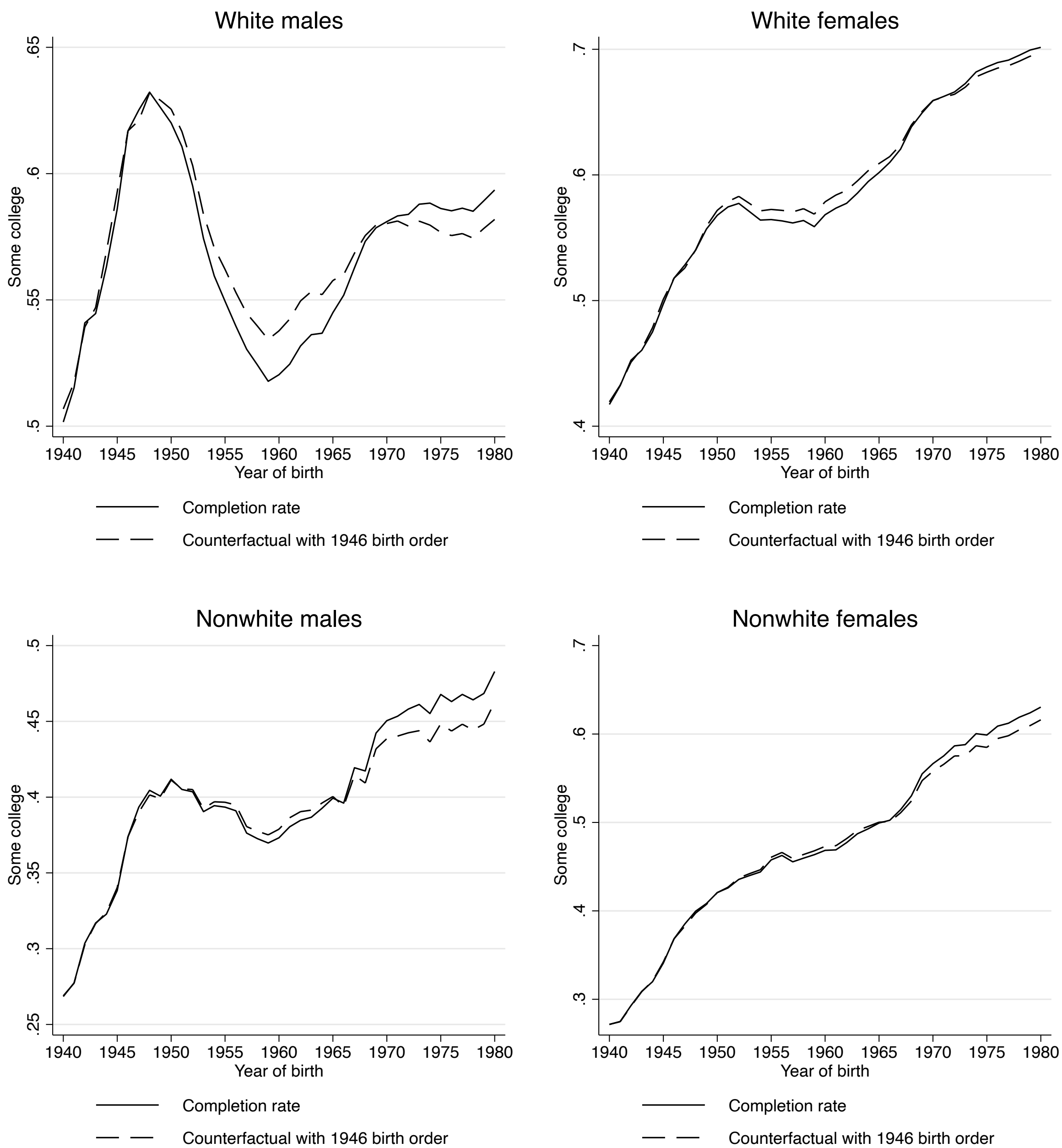

Notes: Data are from the HRS. We plot coefficients from a modified version of our preferred birth order regression in which we interact the birth order indicators with sibship size indicators. The regression includes an indicator for sex and an indicator for whether the individual is an HRS respondent, as well as the log cohort size, measured from Vital Statistics data, for the individual's year and census division of birth. We interact each birth order variable with an indicator for being born outside the baby boom generation (defined as the 1946–74 birth cohorts), and report the coefficients for the baby boom generation. The regression is weighted using HRS sampling weights.

Figure A4: Actual and counterfactual trends in high school completion




Notes: We plot estimated educational attainment at age 40 from the census and ACS for birth cohorts 1940–80. We use the sex-specific birth order coefficients in Table 3 to compute counterfactual educational attainment for each cohort using the 1946 birth order distribution; see text for details about these counterfactual series.

Figure A5: Actual and counterfactual trends in college attendance




Notes: We plot estimated educational attainment at age 40 from the census and ACS for birth cohorts 1940–80. We use the sex-specific birth order coefficients in Table 4 to compute counterfactual educational attainment for each cohort using the 1946 birth order distribution; see text for details about these counterfactual series.

Table A1: Robustness of birth order effects: White male college completion

| | Birth order estimates | | | | | *Δ in college explained by Δ in birth order* | | # siblings |
|---|---|---|---|---|---|---|---|---|
| | 2nd | 3rd | 4th | 5th | 6th + | *1946–60* | *1960–74* | # families |
| (1) Base specification, 1946–74 birth cohorts | | | | | | | | |
| | –0.037* | –0.084*** | –0.069*** | –0.117*** | –0.141*** | *–0.013**** | *0.019**** | 29,041 |
| | (0.020) | (0.022) | (0.025) | (0.030) | (0.032) | *(0.003)* | *(0.004)* | 11,407 |
| (2) Extend baby boom back to 1939 birth cohort (1939–74) | | | | | | | | |
| | –0.040** | –0.081*** | –0.074*** | –0.118*** | –0.137*** | *–0.013**** | *0.019**** | 29,041 |
| | (0.017) | (0.019) | (0.022) | (0.027) | (0.029) | *(0.003)* | *(0.004)* | 11,407 |
| (3) Limit sample to families with 10 or fewer children | | | | | | | | |
| | –0.039* | –0.085*** | –0.072*** | –0.119*** | –0.133*** | *–0.013**** | *0.019**** | 27,957 |
| | (0.021) | (0.023) | (0.026) | (0.031) | (0.034) | *(0.003)* | *(0.004)* | 11,076 |
| (4) Include siblings more than 20 years older or younger than respondent | | | | | | | | |
| | –0.034* | –0.080*** | –0.062** | –0.114*** | –0.133*** | *–0.012**** | *0.018**** | 29,262 |
| | (0.020) | (0.022) | (0.024) | (0.029) | (0.031) | *(0.003)* | *(0.004)* | 11,425 |
| (5) Assume siblings with valid age data are youngest | | | | | | | | |
| | –0.033 | –0.084*** | –0.070*** | –0.113*** | –0.141*** | *–0.013**** | *0.019**** | 29,041 |
| | (0.020) | (0.022) | (0.025) | (0.030) | (0.032) | *(0.003)* | *(0.004)* | 11,407 |
| (6) Limit sample to families with education and age data for all siblings | | | | | | | | |
| | –0.034* | –0.083*** | –0.068*** | –0.114*** | –0.139*** | *–0.012**** | *0.019**** | 28,794 |
| | (0.020) | (0.022) | (0.025) | (0.030) | (0.032) | *(0.003)* | *(0.004)* | 11,277 |
| (7) Include siblings with inconsistently reported sex, age, and/or education | | | | | | | | |
| | –0.035* | –0.082*** | –0.070*** | –0.109*** | –0.138*** | *–0.012**** | *0.019**** | 29,883 |
| | (0.020) | (0.022) | (0.024) | (0.029) | (0.030) | *(0.003)* | *(0.005)* | 11,458 |
| (8) Exclude sibling data if respondent was older than 56 when first interviewed | | | | | | | | |
| | –0.036* | –0.082*** | –0.072*** | –0.115*** | –0.142*** | *–0.013**** | *0.019**** | 25,435 |
| | (0.021) | (0.022) | (0.025) | (0.030) | (0.033) | *(0.003)* | *(0.004)* | 9,854 |
| (9) Limit sample to families in which siblings are self-reported (not by spouse) | | | | | | | | |
| | –0.052* | –0.093*** | –0.067* | –0.116*** | –0.144*** | *–0.013**** | *0.019**** | 22,751 |
| | (0.029) | (0.032) | (0.036) | (0.043) | (0.045) | *(0.004)* | *(0.006)* | 11,389 |

Notes: Data are from the HRS. Each row reports results from a separate regression. Specification 1 is identical to the one reported for white men in Table 2. All regressions include family fixed effects, birth year fixed effects, an indicator for sex and an indicator for whether the individual is an HRS respondent, as well as the log cohort size, measured from Vital Statistics data, for the individual's year and census division of birth. We interact each birth order indicator with an indicator for being born outside the baby boom generation (defined as the 1946–74 birth cohorts), and report the coefficients for the baby boom generation. We estimate male and female coefficients in a single regression in which sex is interacted with birth order, and here report only the coefficients for men. Regressions are weighted using HRS sampling weights. Standard errors are clustered at the family level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2: National births and birth order distribution by race

| Year | Total births | White birth order distribution (%) 1 | 2 | 3 | 4 | 5 | 6+ | Nonwhite birth order distribution (%) 1 | 2 | 3 | 4 | 5 | 6+ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1946 | 3,288,672 | 39.3 | 28.3 | 14.4 | 7.4 | 4.0 | 6.5 | 28.2 | 21.1 | 14.3 | 10.2 | 7.4 | 18.8 |
| 1947 | 3,699,940 | 42.6 | 27.5 | 13.8 | 6.7 | 3.6 | 5.7 | 31.5 | 21.3 | 13.9 | 9.4 | 6.8 | 17.1 |
| 1948 | 3,535,068 | 38.4 | 29.7 | 15.1 | 7.2 | 3.8 | 5.9 | 29.3 | 22.9 | 14.8 | 9.7 | 6.8 | 16.6 |
| 1949 | 3,559,529 | 35.2 | 31.0 | 16.1 | 7.7 | 3.9 | 6.1 | 26.7 | 23.0 | 15.7 | 10.3 | 7.1 | 17.2 |
| 1950 | 3,553,688 | 32.7 | 31.5 | 17.5 | 8.2 | 4.0 | 6.0 | 25.2 | 22.3 | 16.7 | 11.0 | 7.5 | 17.3 |
| 1951 | 3,750,850 | 32.7 | 30.5 | 18.1 | 8.7 | 4.2 | 5.8 | 24.4 | 21.3 | 16.9 | 11.8 | 7.8 | 17.6 |
| 1952 | 3,846,986 | 31.2 | 30.0 | 19.0 | 9.4 | 4.5 | 5.9 | 23.6 | 20.6 | 16.7 | 12.5 | 8.5 | 18.0 |
| 1953 | 3,902,120 | 30.3 | 29.5 | 19.4 | 10.0 | 4.8 | 6.0 | 23.5 | 20.2 | 16.3 | 12.5 | 9.0 | 18.5 |
| 1954 | 4,017,362 | 29.6 | 28.8 | 19.8 | 10.5 | 5.1 | 6.2 | 23.6 | 19.7 | 15.9 | 12.4 | 9.2 | 19.2 |
| 1955 | 4,047,295 | 29.0 | 28.1 | 20.0 | 11.0 | 5.5 | 6.5 | 22.9 | 20.0 | 15.8 | 12.3 | 9.2 | 19.7 |
| 1956 | 4,163,090 | 28.9 | 27.5 | 20.0 | 11.2 | 5.7 | 6.7 | 22.7 | 19.8 | 15.7 | 12.2 | 9.2 | 20.5 |
| 1957 | 4,254,784 | 28.7 | 26.9 | 20.0 | 11.6 | 5.9 | 7.0 | 22.5 | 19.6 | 15.9 | 12.0 | 9.2 | 20.8 |
| 1958 | 4,203,812 | 28.1 | 26.6 | 20.0 | 11.9 | 6.2 | 7.3 | 21.9 | 19.5 | 15.9 | 12.2 | 9.2 | 21.3 |
| 1959 | 4,238,504 | 27.6 | 26.2 | 19.9 | 12.1 | 6.4 | 7.7 | 21.9 | 19.2 | 15.7 | 12.2 | 9.3 | 21.8 |
| 1960 | 4,233,082 | 27.5 | 25.8 | 19.9 | 12.3 | 6.6 | 8.0 | 21.9 | 19.0 | 15.6 | 12.1 | 9.2 | 22.2 |
| 1961 | 4,248,814 | 27.6 | 25.2 | 19.6 | 12.4 | 6.8 | 8.3 | 22.0 | 18.8 | 15.4 | 12.2 | 9.2 | 22.5 |
| 1962 | 4,012,710 | 28.0 | 25.0 | 19.3 | 12.2 | 6.9 | 8.6 | 22.3 | 18.8 | 15.3 | 11.9 | 9.2 | 22.5 |
| 1963 | 3,943,662 | 28.5 | 25.0 | 18.9 | 12.1 | 6.9 | 8.7 | 23.3 | 19.0 | 15.0 | 11.7 | 9.0 | 22.0 |
| 1964 | 4,002,864 | 29.8 | 24.9 | 18.5 | 11.7 | 6.7 | 8.4 | 24.6 | 19.3 | 14.8 | 11.3 | 8.6 | 21.5 |
| 1965 | 3,736,940 | 31.7 | 25.1 | 17.7 | 11.1 | 6.3 | 8.1 | 26.8 | 19.8 | 14.6 | 10.9 | 8.1 | 19.8 |
| 1966 | 3,584,722 | 34.9 | 25.4 | 16.6 | 10.0 | 5.7 | 7.4 | 29.7 | 20.6 | 14.3 | 10.2 | 7.4 | 17.8 |
| 1967 | 3,504,803 | 35.8 | 26.6 | 16.3 | 9.5 | 5.2 | 6.7 | 32.1 | 21.5 | 14.0 | 9.6 | 6.8 | 16.0 |
| 1968 | 3,480,602 | 37.9 | 27.1 | 15.7 | 8.7 | 4.6 | 5.9 | 35.1 | 21.9 | 13.7 | 9.0 | 6.1 | 14.1 |
| 1969 | 3,577,684 | 38.1 | 27.8 | 15.9 | 8.5 | 4.4 | 5.3 | 36.6 | 22.9 | 13.8 | 8.8 | 5.8 | 12.1 |
| 1970 | 3,707,422 | 39.0 | 28.2 | 15.8 | 8.1 | 4.1 | 4.7 | 37.2 | 23.7 | 14.1 | 8.7 | 5.5 | 10.8 |
| 1971 | 3,532,846 | 39.4 | 29.1 | 15.6 | 7.8 | 3.9 | 4.3 | 37.9 | 24.5 | 14.2 | 8.4 | 5.2 | 9.9 |
| 1972 | 3,236,071 | 40.6 | 30.2 | 14.8 | 7.1 | 3.4 | 3.9 | 40.2 | 25.0 | 13.9 | 7.9 | 4.7 | 8.3 |
| 1973 | 3,114,997 | 41.3 | 31.3 | 14.4 | 6.5 | 3.1 | 3.5 | 40.8 | 26.0 | 14.2 | 7.6 | 4.3 | 7.1 |
| 1974 | 3,137,402 | 42.2 | 32.3 | 14.0 | 5.9 | 2.7 | 2.9 | 41.6 | 27.4 | 14.1 | 7.1 | 3.8 | 6.0 |

Notes: Data are from Vital Statistics published tables and birth-level records.

Table A3: Changes in educational attainment and amount explained by birth order

| | White | | Nonwhite | |
|---|---|---|---|---|
| | Male | Female | Male | Female |
| *Panel A: Changes in college completion* | | | | |
| 1946–60 Δ in college completion | –0.063 | 0.044 | –0.007 | 0.039 |
| Effect of Δ in birth order | –0.013*** | –0.009*** | –0.007*** | –0.003 |
| | (0.003) | (0.003) | (0.002) | (0.002) |
| 1960–74 Δ in college completion | 0.059 | 0.122 | 0.063 | 0.111 |
| Effect of Δ in birth order | 0.019*** | 0.012*** | 0.026*** | 0.014* |
| | (0.004) | (0.004) | (0.009) | (0.008) |
| *Panel B: Changes in college attendance* | | | | |
| 1946–60 Δ in college attendance | –0.097 | 0.050 | –0.001 | 0.099 |
| Effect of Δ in birth order | –0.017*** | –0.010*** | –0.006** | –0.005* |
| | (0.003) | (0.003) | (0.002) | (0.002) |
| 1960–74 Δ in college attendance | 0.068 | 0.114 | 0.082 | 0.131 |
| Effect of Δ in birth order | 0.026*** | 0.014*** | 0.024** | 0.018* |
| | (0.005) | (0.005) | (0.010) | (0.009) |
| *Panel C: Changes in high school completion* | | | | |
| 1946–60 Δ in high school completion | 0.005 | 0.031 | 0.089 | 0.102 |
| Effect of Δ in birth order | –0.004* | –0.006*** | –0.005** | –0.001 |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| 1960–74 Δ in high school completion | 0.038 | 0.031 | 0.092 | 0.078 |
| Effect of Δ in birth order | 0.005 | 0.008** | 0.016* | 0.0001 |
| | (0.003) | (0.003) | (0.008) | (0.008) |

Notes: Educational attainment data are from the 1960–2000 censuses and 2006–17 ACS. Birth order data are from Vital Statistics. We multiply changes in birth order by sex-specific coefficients reported in Table 2, then sum over birth orders to obtain the estimated effect of changes in birth order. These estimates correspond to changes in the appropriate counterfactual series plotted in Figure 4 and Appendix Figures A4 and A5.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Simulated estimates of birth order effects for various sibling survival probabilities, separate for sibship sizes 3 and 4

| | Estimated birth order effects (relative to first-born) | | | | |
|---|---|---|---|---|---|
| | Sibship size 3 | | Sibship size 4 | | |
| Survival probability, $p$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\beta}_4$ |
| 1 | 1 | 2 | 1 | 2 | 3 |
| 0.95 | 1.04 | 2.02 | 1.05 | 2.09 | 3.05 |
| 0.90 | 1.08 | 2.04 | 1.10 | 2.17 | 3.09 |
| 0.85 | 1.12 | 2.06 | 1.16 | 2.23 | 3.13 |

Notes: We report the results of the simulation exercise described in section III of the appendix, in which we estimate birth order effects for various rates of sibling mortality. In the first row, all siblings survive, and we set the "true" effects of birth order, relative to first-born, to be 1 for second-born, 2 for third-born, etc. In the following rows, we report the estimated birth order effects for different sibling survival probabilities, where the observed birth order will now be mismeasured for sibling groups with any deceased siblings. The proportional difference between the "true" effect in the first row and the simulated effect is the estimated degree of bias due to sibling mortality.

Table A5: Simulated estimates of birth order effects for various sibling survival probabilities

| | Estimated birth order effects (relative to first-born) | | | | |
|---|---|---|---|---|---|
| Survival probability, $p$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\beta}_4$ | $\hat{\beta}_5$ | $\hat{\beta}_{6+}$ |
| 1 | 1 | 2 | 3 | 4 | 5.81 |
| 0.95 | 1.03 | 2.06 | 3.08 | 4.11 | 5.92 |
| 0.90 | 1.07 | 2.12 | 3.17 | 4.23 | 6.03 |
| 0.85 | 1.11 | 2.19 | 3.27 | 4.34 | 6.14 |

Notes: We report the results of the simulation exercise described in section III of the appendix, in which we estimate birth order effects for various rates of sibling mortality. In the first row, all siblings survive, and we set the "true" effects of birth order, relative to first-born, to be 1 for second-born, 2 for third-born, etc. In the following rows, we report the estimated birth order effects for different sibling survival probabilities, where the observed birth order will now be mismeasured for sibling groups with any deceased siblings. The proportional difference between the "true" effect in the first row and the simulated effect is the estimated degree of bias due to sibling mortality.