Birth Order and Public Investments: Evidence from the United States, 1900-1940

Angela Cools and Siobhan O'Keefe*

This Version: April 2022

Abstract

A growing literature demonstrates that birth order affects educational attainment, but the impact of public policy on sibling inequality remains largely unknown. Using linked historical Census data and a family fixed effects model, we examine the impact of birth order for U.S. boys born during the late 1800s and early 1900s, a period of increased public investment in education. Consistent with evidence from recent cohorts in the U.S. and Western Europe, we find that men's educational attainment declines with birth order. Later-born boys obtain 0.2-0.6 years (3-7 percent) fewer years of education than their firstborn brothers. Among whites, later-born boys also have lower earnings and occupation scores. Next, exploiting variation in compulsory schooling laws across states and time, we show that laws requiring eight or more years of schooling substantially compress birth order gaps in educational attainment between white brothers born outside the South.

*Cools and O'Keefe: Davidson College Department of Economics. The authors are grateful to Scott Cunningham and to seminar audiences at Kennesaw State University and the 2021 SEA Conference for comments and suggestions.

1 Introduction

Despite similarities in genetics and childhood circumstances, siblings often experience considerable differences in long run outcomes. One important dimension of inequality is birth order, with firstborn children in the U.S. and Western Europe shown to obtain more schooling, receive higher earnings, and engage in less risky behavior than their later-born counterparts. The magnitude of inequality is substantial: Black, Devereux, and Salvanes (2005) show, for example, that the difference in educational attainment between the first and fifth-born child in a five-child family in Norway is equal to the black-white schooling gap in the year 2000. Despite this considerable inequality, there has been little examination of how public policies, in particular increased educational investments, affect birth order differences. The direction of this effect is theoretically ambiguous: if parental and public inputs are complements, increased public investment could increase sibling inequality. If substitutes, the reverse may occur.

In this paper, we examine birth order effects for U.S. men born in the late 1800s and early 1900s, a period of rapid schooling expansion across the United States. Using information provided by the Census Linking Project (Abramitzky, Boustan, Eriksson, Pérez, and Rashid, 2020) to link men in the 1940 Census to their childhood households, we use family fixed effects models to estimate differences in education and labor market outcomes between earlier- and later-born brothers born between 1880 and 1910. Consistent with recent evidence from the U.S. and Western Europe, we first show that completed educational attainment declines with birth order. The magnitude of these effects is economically meaningful: being a later-born relative to a first-born son decreases years of education by 0.2-0.6 years (3-7 percent). Later-born white sons also have lower earnings and lower occupation scores. Next, we exploit the substantial variation in compulsory schooling laws across location and time to explore how changing public investments affect the birth order gradient. We provide evidence that laws requiring 8 or 9 or more years of schooling compress birth order gaps in educational attainment for white men outside of the South.

This paper adds to a growing literature on birth order effects. For recent cohorts in developed countries, there is evidence of a later-born disadvantage. Black et al. (2005) use administrative data on the Norwegian population ages 16-74 between 1986 and 2000. Using two separate specifications, one controlling for family size and one including family fixed effects, they show that educational attainment declines with birth order. Relative to firstborns, second-borns and third-borns complete 0.34 and 0.53 fewer years of education, respectively, while tenth-borns and later receive 0.94 fewer years. Later-born men also have lower earnings, and later-born women have lower earnings, a lower likelihood of working full time, and a higher likelihood of a teen birth. Other work shows that later-born men in developed countries have lower IQ (Black, Devereux, and Salvanes, 2011), reduced leadership skills (Black, Grönqvist, and Öckert, 2018), and greater disciplinary problems (Breining, Doyle, Figlio, Karbownik, and Roth, 2020) than their earlier-born brothers.

A number of mechanisms have been proposed to explain the later-born disadvantage in Western Europe and the United States. The relationship between family investments in children and later outcomes is well documented (Cunha, Heckman, Lochner, and Masterov, 2006), and time spent with children is one of the most important investments parents can make (Del Boca, Flinn, and Wiswall, 2014). Price (2008) uses the American Time Use Survey to show that firstborns receive 20 minutes more quality time with fathers and 25 minutes more quality time with mothers per day than second-borns of the same age in similar families. Additionally, Lin, Pantano, and Sun (2020) find that later-born children are more likely to be unplanned, which may be harmful if parents made prior decisions to invest in older children based on a smaller expected family size. Finally, later-borns are more likely to be infected with severe respiratory illnesses as infants, potentially through contact with their older siblings. Early life health is critical for child development, so this has negative consequences for later-born's schooling and future earnings (Daysal, Ding, Rossin-Slater, and Schwandt, 2021).

In developing countries, research on birth order has focused primarily on contemporaneous school attendance rather than long-run attainment, and the relationship between birth order and schooling is less clear. Using a family fixed effects model and data from Ecuador in the early 2000s, De Haan, Plug, and Rosero (2014) find that earlier-born children actually have lower levels of preschool cognition and secondary school enrollment than their younger siblings. These effects are most pronounced in lower-income households; as income increases, earlier-borns gain relative to their siblings. In Mexico, Esposito, Kumar, and Villaseñor (2020) instead find an earlier-born advantage in on-track enrollment; like De Haan et al. (2014), however, they also find that earlier-borns do relatively better as family income increases. One mechanism important in the developing countries but less relevant in modern developed countries is child labor. Older children, who have a comparative advantage in home production and market work, may drop out to support the family. For example, in Brazil, Emerson and Souza (2008) find that school attendance is increasing and child labor is decreasing in boys' birth order. In Nepal, Edmonds (2006) shows that older girls' work is increasing in the number of younger siblings and older boys' work is increasing in the number of younger brothers. In Nicaragua and Guatemala, Dammert (2010) finds that older boys do more market work and older sisters do more domestic work than younger siblings. While these studies highlight a first-born disadvantage in school attendance, especially for relatively poor families in developing countries, little is known about whether this translates into long-run differences in completed schooling or labor market earnings.

This paper contributes to the birth order literature by examining a context with many similarities to developing countries (e.g., lower income per capita, higher fraction in agriculture, less schooling available). Unlike previous work, however, we are able to explore long-run outcomes including completed educational attainment, earnings, and migration. Since we use full-count U.S. Census files, the sample size is large and enables the use of family fixed effects for identification. This paper also contributes by examining the interactions with public investments, a topic not generally explored in either the developing or developed country context.

This paper also adds to the literature on educational investment and the effects of compulsory schooling laws. We build on previous work by incorporating newer panel data techniques that reduce the possibility of bias from variation in treatment timing across states (Goodman-Bacon, 2021). Evidence indicates that compulsory schooling laws increased overall educational attainment for cohorts born in the late 1800s and early 1900s (Clay, Lingwall, and Stephens, 2021). These laws had the largest effect at the bottom of the distribution and therefore decreased overall educational inequality (Lleras-Muney, 2002); however, little is known about how these laws affected within-family inequality. We contribute to this literature by examining how compulsory schooling laws and public educational investments affect intrafamily inequality. Children in larger families are more likely to be in farming households and live in rural areas. If public investments affect the birth-order gradient, this could also have implications for geographic inequalities.

The paper is organized as follows. In Section 2, we provide information on the historical setting of our analysis. In Section 3, we discuss our data and methods used to link individuals between years in complete-count Censuses. Section 4 presents an overview of schooling patterns by birth order, and Section 5 presents our empirical strategy. Sections 6 and 7 present the main results, and Section 8 concludes.

2 Historical Context

The cohort included in the analysis, described further in Section 3, came of age at a time of rapidly changing educational attainment in the United States, sometimes known as the "High School Movement". Between 1910, when we observe many of our sample for the first time, and 1940, when the outcomes are measured, the high school graduation rate increased fourfold, from 9 to 51 percent (Goldin and Katz, 2011). This section briefly discusses the factors behind this increase.

Massachusetts passed the first compulsory schooling law in 1852. By 1890, half of all states had passed compulsory schooling laws and by 1918 every state had passed some form of compulsory schooling law. At its most basic, a compulsory schooling law sets a minimum age at which children must begin school and a maximum age at which they are no longer required to attend. In addition, most states' compulsory schooling laws set a required level of school (in years) after which students were no longer required to attend, even if they were still below the maximum age (Stephens Jr and Yang, 2014). In most states, a child with consistent attendance would reach this number before reaching the maximum age, making this the binding constraint (Goldin and Katz, 2011). All states included a variety of additional exemptions to their compulsory school laws. Exemptions were generally given for families in poverty or who lived very far from the school, as well as children with severe disabilities (Lleras-Muney, 2002).

Although the passage of compulsory school laws was not always formally coordinated with the passage of child labor laws, the actual number of years of school a child was required to attend school was driven by the interaction between the types of laws. Child labor laws were much broader and covered a variety of issues, like what industries children could work in and how to obtain a work permit (Goldin and Katz, 2011).

Previous work has generally found that these laws lead to increases in educational attainment. However, Goldin and Katz (2011) estimate that changes in compulsory schooling laws only explain a small fraction of the overall increase in educational attainment during this time. This is potentially because the laws were only binding for a small subset of children. Most children in the affected age ranges were already attending school, at least part time, beforehand. While small, the increase in attainment was largest for the bottom of the education distribution shrinking educational inequality at this time (Lleras-Muney, 2002). These studies rely on differences in compulsory school laws across states and time for identification of this effect.

However, Stephens Jr and Yang (2014) find that these estimates are driven not by this policy-induced state and time variation but by differences in regional trends in educational attainment by birth cohort, particularly in the South versus the rest of the country. Once region by year of birth fixed effects are included, the statistical relationship between compulsory schooling and educational attainment disappears, except for the earliest cohorts affected by compulsory schooling laws. For the 1895-1912 cohorts (which includes the cohorts in our study), compulsory schooling laws did increase educational attainment. The effect is non-linear: the more years of school required, the larger the effect on the number of years ultimately completed (Clay et al., 2021).

3 Data

To estimate the impact of birth order on adult outcomes, information on childhood family structure needs to be linked to men's completed educational attainment as reported in the 1940 Census. We begin by extracting information from IPUMS on all native-born, non-Hispanic white and black men who were ages 28-59 in the complete-count 1940 Census (Ruggles, Flood, Foster, Goeken, Pacas, Schouweiler, and Sobek, 2021). The 1940 Census provides information on completed educational attainment as well as age, occupation, state of residence, and other demographic outcomes.

To link 1940 men to their childhood households and birth order, we next extract information from IPUMS on all U.S. households in the 1910 or 1900 Census (Ruggles et al., 2021). We classify children as siblings if they live in the same 1900 or 1910 household with the same surviving mother and father.[1] Siblings are then ranked by age and assigned birth order. If two children are the same age, they are both assigned the highest possible birth order. For example, in a household with two children who are 10 years old and one who is 8 years old, they would be assigned birth orders of 1,1, and 3, respectively.[2] Birth order is not assigned if the mother reports more/fewer surviving children than listed in the Census roster which can occur, for example, if older children have already left the parental household. Given the potential for selection into leaving the parental household (e.g., those going to college or forming their own households), we also do not assign birth order if there are any siblings over age 17 still living in the household. Other information obtained from the 1910 or 1900 Census includes race, parental occupation, nativity, and state and county of residence. The reported race of some individuals changes between the 1900 or 1910 and the 1940 Census. In these cases, we classify race as reported in the earlier Census.

We use matches provided by the Census Linking Project to link 1940 records to their childhood (1910 or 1900) records (Abramitzky et al., 2020).[3] We first attempt to match

[1] Non-living parents cannot be linked to a given child.

[2] Note that both male and female children are ranked by birth order, although only male children can be matched to long-run outcomes.

[3] We use the matches generated by the ABE algorithm with NYSIIS standardized names. See Abramitzky, Boustan, Eriksson, Feigenbaum, and Pérez (2021) for details.

the 1940 sample to birth order and other household information in 1910. If birth order information is unavailable in 1910 (for example, if older siblings have left the household or the individual is 19 in 1910), we match birth order and other information from the 1900 Census. If birth order is not available in either year, we first match 1910 household information and then, if unavailable, 1900 household information. We are able to match 5,648,498 (27 percent of the 1940 Census) native-born white men and 400,105 (17 percent of the 1940 Census) native-born black men to their childhood households. Of these matched observations, we have non-missing 1940 educational attainment and birth order for 3,475,796 white men and 175,021 black men. We refer to the sample of men who are matched, have birth order information, and have non-missing information on 1940 educational attainment as the analysis sample.

Table 1 displays summary statistics for the analysis sample in columns (1) and (2).[4] Columns (3) and (4) compare the analysis sample to other men in the 1940, 1910, or 1900 Censuses by showing the coefficients (column (3)) and standard errors (column (4)) from regressions of the given characteristic on a dummy variable for being in the analysis sample. Differences between the analysis sample and the remainder of the population come from three sources. First, since matching relies on unique names within state of birth and age cells, men with unusual names and those who are numerate (and can record a correct age) may be more likely to be matched. If there is a relationship between socioeconomic status and naming conventions, the sample will reflect that. Second, birth order information is missing if a parent is deceased or not living in the household, which may be less likely in families with higher socioeconomic status. Finally, birth order is also missing if some siblings are not living in the household, which is more common for older children and those in large families.

As expected, the analysis sample differs from the population. In 1940, men in the analysis sample display characteristics associated with higher socioeconomic status. They are more likely to be married, have more years of education, and more likely to have completed eighth grade, high school, or any college. They are also somewhat younger.

[4]Table A1 in the Appendix displays summary statistics for the full matched sample (including those without birth order information).

For native-born whites, however, the magnitude of these differences is relatively small. For example, there is only a 0.5 difference in years of education. For native-born blacks, the analysis group differs more substantially from the population. This may be in part driven by lower overall match rates among the black population.

Although the analysis sample is not perfectly representative of the U.S. population, differences are relatively small, especially for native-born whites, suggesting these results can provide insight into the larger native-born white population. However, results may be considerably different for certain groups excluded from our analysis. In particular, our results are less representative for black men and we exclude women as well as those who are Hispanic, Asian, Native American, and/or immigrants. Finally, given our construction of the birth order variable, our results are unable to capture effects among boys with a deceased parent or a parent not living in the household; birth order effects may be substantially different in these groups.

4 Birth Order and Schooling

We begin our analysis with descriptive statistics on the relationship between birth order and completed educational attainment. Figure 1 shows the age-adjusted distribution of completed attainment by birth order for native-born whites (panel (a)) and blacks (panel (b)), ranging from 0 years to 13 years or more. To adjust for age, we run a regression where the dependent variable is a dummy for each level of completed schooling and the independent variables are birth order and year of birth dummies. We then obtain predicted values for individuals of each birth order at year of birth equal to 1897 (roughly the median of the sample).[5] The figures are separated by region given wide geographic disparities in schooling access the early 1900s (see, e.g., Goldin and Katz (2010)).[6] As shown in

[5]In each Census year, we calculate year of birth by taking the Census year minus the reported age.

[6]We follow Clay et al. (2021) and place non-Confederate Southern states in the Midwest or Northeast. Regions are defined as follows. The Northeast includes Connecticut, Delaware, the District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont. The Midwest includes Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, West Virginia, and Wisconsin. The South includes Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia. The West includes Arizona, California, Colorado, Idaho, Montana,

panel (a), the modal level of schooling for whites in all regions is 8 years, reflecting the widespread access to grammar school by the early 1900s. Completion of grammar school was less common in the South, driven in part by the region's rural population and lower school funding (Goldin and Katz, 2010). Completion of high school and post-secondary education is also common, especially in the Northeast and West.

Panel (a) shows evidence of substantial disparities in educational attainment by birth order. While later-borns were more likely to finish their schooling at eight years of education or less, earlier-borns were more likely to complete high school and to go on to college. Interestingly, there are not substantial differences in the likelihood of completing some high school (9th, 10th, or 11th grade): differences by birth order are concentrated in the likelihood of going on to high school at all rather than differences in dropout during high school.

Panel (b) shows the same figure for native-born blacks. Given that over 80 percent of the black population lived in the South as of the early 1900s (see Table 1), the sample size is small and estimates are imprecise outside of the South. However, in the Northeast and Midwest, the distribution of completed schooling looks similar to the figures for whites, although birth order effects are less clear. In the South, schooling levels are much lower; less than a quarter of the population reports schooling beyond eighth grade. The figure also suggests that later-born black children are somewhat more likely to drop out before eighth grade, although differences are small. These smaller differences likely reflect institutional factors limiting access to education for black children (Aaronson and Mazumder, 2011).

5 Empirical Strategy

While Figure 1 presents evidence of differential attainment by birth order, these differences are not necessarily causal if other characteristics differ between earlier-born and later-born men. One likely difference is family size: in larger families, children are mechanically more likely to have a later birth order. For example, if a child has a birth

Nevada, New Mexico, Oklahoma, Oregon, Utah, Washington, and Wyoming.

order of 4, he must be in a family with at least 4 children. If larger families have different socioeconomic status than smaller families, differences in outcomes across birth order may be driven by family characteristics. To circumvent this problem and obtain a causal estimate of birth order, we use a family fixed effects model and compare outcomes between earlier-born and later-born brothers within the same family. We must also account for the fact that, conditional on family fixed effects, later-borns will mechanically be in later cohorts. If educational attainment is increasing over time, failing to control for age would bias the birth order effect up as later-borns receive more education due to their cohort. As a result, we include a full set of cohort dummies in all specifications.

Specifically, the following equation is estimated:

$$y_{i,f} = \alpha + \sum_{k=2}^{5} \beta^k ORDER_{i,f}^k + \sum_{j=1}^{27} \gamma^j YOB_{i,f}^j + \eta_f + \epsilon_{i,f} \tag{1}$$

where $y_{i,f}$ is a variable measuring completed education in 1940 for individual i in family f, $ORDER^k$ are dummies measuring birth order (second, third, fourth, fifth or higher), and η_f is a family fixed effect. Year of birth fixed effects, $YOB_{i,f}^j$ account for the fact that the average level of schooling is increasing over time. Completed education is measured in years, which ranges from 0 (no schooling) to 20 (8+ years of college), and at three levels: at least eighth grade, at least high school graduation, and any college. These levels correspond to the most frequent stopping points of educational attainment (see Figure 1). Estimates are presented separately for non-Southern and Southern males, given the differences attainment patterns documented in Figure 1 and the documented differences in educational trends in the South as opposed to the rest of the county (Clay et al., 2021).

Given the use of family fixed effects, the outcomes for men without a brother in the analysis sample will be absorbed by the fixed effect. We therefore restrict the regression sample to men with birth order information and at least one matched brother. Summary statistics on the regression sample are shown in Appendix Table A2.

Since brothers have the same family characteristics (size, parental education, and other unobserved factors), the remaining variation in brothers' outcomes can be attributed to position within the family (i.e. birth order). The estimates of $ORDER^k$ can be interpreted as causal if, after family fixed effects and age controls, remaining variation in outcomes between brothers can be attributed to birth order. This interpretation is consistent with prior work on birth order effects (Lehmann, Nuevo-Chiquero, and Vidal-Fernandez, 2018).

We use this same specification to document the relationship between birth order and labor market outcomes. We begin with annual earnings in the prior year, reported by wage and salary workers. Following Acemoglu and Angrist (2000), we construct weekly earnings by using the annual earnings measure and replacing earnings above the 98th percentile with 1.5 times the 98th percentile value. We then divide earnings by the number of weeks worked and, following Schmick and Shertzer (2019), drop weekly earnings in the top and bottom 1 percent. We also consider other measures of work: whether a man has the same occupation as his father (1=yes, 0=no, using 1950 occupation codes for both), whether a man is working in a farm occupation (1=yes, 0=no), and the occupational income score calculated by IPUMS. By using the occupational income score we are able to capture a measure of permanent income for all working men, even the self employed, whose earnings are not captured by the wage variable. Finally, given the importance of migration to men's earnings growth in the twentieth century (Collins and Wanamaker, 2014), we develop a measure of whether an individual is a migrant. This takes a value of 1 if the man is in the same state in 1940 as he was in childhood and a value of 0 otherwise.

6 Birth Order Results

This section begins by documenting within-family differences in educational attainment, then follows these patterns beyond school into the labor market.

6.1 Educational Attainment

Figure 2 presents the coefficients on the birth order terms from the estimation of Equation 1 for native-born whites. Table A3 in the Appendix also presents the results. We also perform a regression with a "later-born" dummy, which takes a value of 1 if the individual is second born or higher and takes a zero otherwise. Results are shown in Table A4 in the Appendix.

Consistent with results from more recent cohorts, we find that years of educational attainment declines with birth order. Outside of the South, white boys who are second, third, fourth, or fifth born or higher receive 0.27, 0.34, 0.35, and 0.32 fewer years of education, respectively, than their first-born brothers. Compared to a mean of a mean of 9.4 years of education, this equates to 2.9-3.7 percent fewer years of school.

The effects are even larger in the South, where boys who are second, third, fourth, or fifth born or higher receive 0.37, 0.54, 0.54, and 0.60 fewer years of education, respectively, than their first-born brothers. Not only is the magnitude of the coefficient larger, the relative penalty for later-borns is higher in the South. This equates to 4.5-7.2 percent on the smaller mean attainment in this region (8.3 years). In the South, the effect of birth order on years of education is larger than other shocks to educational attainment at this time. For example, Baker, Blanchette, and Eriksson (2020) find early exposure to the boll weevil increased educational attainment 0.24 to 0.36 years, a slightly smaller effect than the difference between firstborns and later-born siblings. While large, the raw effect on years of education is remarkably similar to the estimates for men in recent Norwegian cohorts found by Black et al. (2005). They find a 0.30 and 0.46 decrease in years of educational attainment for second-born and third-borns versus first-borns, although the relative effects are larger in our setting due to a higher mean educational attainment among recent Norwegian cohorts.

Turning to common levels of completion, birth order effects display considerable heterogeneity across regions. Outside of the South, birth order effects on eighth grade completion are relatively small. This is consistent with the high overall rates of primary school completion in these regions (82 percent in this sample). Since so many children are already going to school through eighth grade, this does not seem to be a margin where potential differences in parental investment have a strong effect. Although second-, third-, and fourth-borns are slightly less like to complete primary school, most children, regardless of birth order, comply with the standard of eighth grade attainment.

In the South, where grammar school completion is less common (only 57 percent of the Southern sample completes eighth grade), eighth grade completion declines with birth order. The effects are large: second, third, fourth, and fifth-borns and higher are 4.7, 7.0, 6.7, and 7.6 percentage points (8.2-13.3 percent relative to the mean) less likely to complete eighth grade than their first-born brothers, respectively. In a lower-completion environment, parental investments are more likely to have an effect at this margin. Again, this is a large effect relative other interventions in this time period. For example, Schmick and Shertzer (2019) find that a ten percent increase in education funding leads to a two percent increase in eighth grade completion; this is equivalent to moving from being the third to the second born in our context and less than half the gap between the first and later-borns.

The cross-region differences documented for eighth grade are less pronounced at higher levels of attainment. Outside the South, fifth-borns and higher are 6.2 percentage points less likely to complete high school than their first-born brothers, about a 23 percent decrease on a mean of 27 percent. In the South, fifth-borns or higher are 5.7 percentage points less likely to complete high school, a similar 26 percent decrease on a mean of 22 percent. The percentage magnitude of these differences in high school completion are similar to overall differences between the South and the rest of the country. Completion of any college follows the same pattern, with fifth-borns and higher being 4.9 percentage points (38 percent) to complete any college outside the South and 3.6 percentage points (32 percent) less likely to complete any college in the South. These extremely large effects suggest that, at levels with substantial direct or opportunity costs, earlier born men are substantially more likely to stay in school.

Results for native-born blacks are given in Figure 2 and panel (b) of Tables A3 and A4 in the Appendix. Sample sizes are much smaller, especially outside the South, and the results are less precise. However, there is evidence that earlier borns attain more education. For example, Table A4 shows that, relative to their firstborn brothers, later-borns obtain 0.2 fewer years of education and are 3.3 percentage points (11.5 percent) less likely to complete eighth grade in the South. However, there are no birth order effects on high school or college completion in the South. This is likely because many black families in the South lacked access or resources for education beyond eighth grade, regardless of birth order (Aaronson and Mazumder, 2011). Differences in parental investment can have a strong effect on children's educational attainment but are unable to overcome the structural barriers to school at this time. For black men outside the South, where sample sizes are very small, there is no statistically significant relationship between birth order and educational attainment.

6.2 Other Long Run Outcomes

To what extent does birth order affect men's long-run outcomes outside of educational attainment? Men's later life outcomes such as earnings and occupation may be affected by educational attainment but also other characteristics that may differ with birth order such as IQ and parental expectations. To examine other long-run outcomes, we run the specification from Equation 1 with dependent variables reported in the 1940 Census.

Tables 2 and 3 show the results on these other long-run outcomes. Focusing first on native-born whites (Table 2), later-born men are less likely to have any wage or salary income and those who report income earn 1-2 percent less. Because the earnings of the self-employed are not fully reported in the 1940 census, we also look at occupational income score. As discussed above, this measure is calculated using the median total income observed for individuals with that occupation in the 1950 census. Beyond allowing us to examine outcomes for the self employed, this measure is not subject to idiosyncratic labor market shocks and can be thought of as closer to permanent income. We exclude farmers from this measure as occupation scores cannot capture the diversity of economic circumstances among this group (Feigenbaum, 2018). Using this measure, we observe a similar pattern: later-born brothers have lower occupational scores than their earlier-born brothers, although the results are only statistically significant outside the South. This shows that the differences in educational attainment translate into occupational updates in the labor market.

Relative to firstborns, later-borns are more likely to be working in the same occupation as their father. Firstborn's relatively larger wages and occupational upgrading are not solely driven by direct family transfers, like inheriting the family business. Firstborn men appear to be moving into better occupations than their fathers, while second borns stay behind. This may be partially driven by firstborns moving away from farming, while second, third, and fourth-borns are more likely to remain in farming. Some of the wage disadvantage of later borns outside the South may also come from the fact that they are less likely to migrate. Relative to the mean, second-born brothers in the non-South are 1.4 percent less likely to move away from their state of birth.

For native-born blacks (Table 3), the educational advantages obtained by earlier-born men do not translate into improved labor market outcomes. There is no significant impact of birth order on wage or salary employment, earnings, or various measures of occupational status. Furthermore, despite the high levels and substantial economic benefits of black migration between the early 1900s and 1940 (Collins and Wanamaker, 2014), we also find no impact of birth order on the likelihood of interstate migration overall or migration from the South to the North.

6.3 Heterogeneity by Family SES and Sibling Gender

We next examine heterogeneity in the birth order gradient. In this section, we focus on family characteristics: socioeconomic status and sibling gender. From prior work, the relationship between socioeconomic status and the birth order gradient in education is not clear. On the one hand, the most resource-constrained families may be more likely to send their oldest children to work (reversing the birth order gradient). On the other, they might also focus limited resources such as money, time, and expectations on the highest-return children (reinforcing the gradient). Additionally, if parental investment is especially important to push children to higher levels of educational attainment than the "norm" (eighth grade in this setting), the effects could be larger in lower-income families with lower average attainment.

Because family income or father's earnings are not reported in the 1900 or 1910 censuses, we examine father's occupational income score. This measure is calculated using the median total income observed for individuals with that occupation in the 1950 census; a higher occupational score is associated with a higher paying profession. We separate the sample into three groups based on father's occupational income score: below the 25th percentile (low income), 25th-75th percentile (middle income), and above the 75th percentile (high income). We also examine farmers separately as occupational income scores are artificially low for farmers, especially in the early 1900s (Feigenbaum, 2018). We focus this analysis on white men since the vast majority of black men have a father working in farming.

Figure 4 and Appendix Table A5 show the results. We combine regions for brevity, but the patterns by father's occupation are similar inside and outside the South. Panel (a) of Figure 4 shows that the coefficients on years of education are remarkably similar across the four groups: being a fifth-born relative to a first-born decreases attainment by about 0.4 years regardless of father's occupation. The magnitude of effects is slightly larger for those in farming and those in low-income households given the lower mean educational attainment for these groups. For example, the later-born brothers in a lower income family complete 3–10% fewer years of schooling as compared to 2–5% for those in the most well-off families.

Focusing on completion tells a more varied story across the different status families. Looking at eighth grade completion in panel (b), the birth-order gradient is strongest for farming families and low to middle income families. However, the effect is attenuated for the highest-status families. While the point estimates are not statistically different from each other, this suggests that for families with the most resources, eighth grade completion is not a margin where the first-born advantage plays a strong role. By panel (d) the pattern has reversed, and the birth-order effect for attending any college is strongest for the highest-income families. Given the very low rates of college attendance for the other groups (8-14%) there is less room for family investments to increase college-going. However, almost 30% of children born to the highest-income families attend some college. Unlike eighth grade, this is a margin where the first-born advantage can play a role.

Birth order effects may also differ based on the gender of siblings. In particular, if parents make investments based on children's future earnings potential, a later-born boy may be more disadvantaged by older brothers than older sisters. We repeat the main results separating birth order by sibling gender. For example, if a boy has two older sisters and one older brother, he is considered second born among brothers and third born among sisters. Appendix Tables A6 and A7 show the results for order among brothers and sisters, respectively. In both sets of tables, there is a consistent pattern of later-born disadvantage. This suggests that part of the birth order differential is driven by a general resource story. Assuming families have a binding budget constraint, the amount of parental time and resources available per child will decrease as the family grows. Recent work has found parental time is an important mechanism for explaining long-run criminal outcomes between first and later-borns (Breining et al., 2020). Even in a different historical setting, first borns will likely have more undivided parental attention. As a result, earlier born children will spend their early childhood with more resources available than later born children. Given the importance of early childhood for later life outcomes, this resource disparity may explain why earlier born children have better educational outcomes, regardless of their gender (Black et al., 2005).

However, the birth order effects are stronger when we focus on the number of older brothers an individual has. Particularly for second and third borns, the coefficients are generally statistically different from each other: the later-born disadvantage is larger if the older siblings in the household are males. This implies that strategic investments also play a role in the birth order gradient.

7 Birth Order and Public Education

Given the vast disparities in sibling outcomes in both historical and contemporary settings, it is important to know how public policy interacts with birth order. As discussed in section 2, the early twentieth century was a time of rapidly increasing public investments in education. If public investments can substitute for familial investments, then the rise of public investments should shrink the gap between earlier and later borns discussed in section 6. However, public and private investments could be complementary, particularly if early life investments improve children's ability to develop human capital later in life (Johnson and Jackson, 2019). In our setting, this dynamic complementarity would lead firstborns to benefit more from public investments in education. The relationship between birth order and public human capital investments has not been previously studied.

We explore the impacts of required schooling laws passed at the state level in the late 1800s and early 1900s. Given the costly nature of enforcement mechanisms (e.g., truant officers) as well as the short-run economic costs of children not being at work, these laws can be viewed as a form of public investment. Previous studies of compulsory schooling have relied on regressions with dummy variables for various laws combined with state and cohort fixed effects (referred to as two-way fixed effects regressions), but a recent literature has highlighted that two way fixed effects (TWFE) specifications may produce biased estimates of treatment effects. As discussed in Appendix B, these biases may be especially problematic in cases with few or no "untreated" units. Estimates relying on Gardner's (2021) two-stage difference-in-differences, an alternative estimator without the TWFE bias, reveal strong, positive effects of laws requiring eight or nine or more years of schooling for whites outside of the South.

We employ Gardner (2021)'s two-stage difference-in-differences strategy to examine how compulsory schooling laws affect the birth order gradient. To simplify our regression, we focus on interaction of various compulsory schooling laws with a dummy variable for being a later-born (not firstborn) child.[7] We first focus our analysis on whites living outside of the South. Because the level of schooling required is much lower in Southern states than the rest of the country at this time (Clay et al., 2021), it is not appropriate to specify the compulsory schooling cutoff points at the same level for the Southern and non-Southern cohorts.

In the first stage, we perform the following regression for the *untreated* groups (those with less than eight years of schooling required):

[7] We reproduce our regressions on the relationship between birth order and education from Figure 2 with a later-born dummy (rather than individual birth order dummies) in Appendix Table A4.

$$y_{iscf} = \alpha + \gamma_{sb} + \delta_{bc} + \eta_{rc} + \epsilon_{iscf} \tag{2}$$

where y_{iscf} is completed years of schooling reported in 1940 for individual i born in state s in year c to family f. Birth order (first vs later-born) by state interactions, γ_{sb}, account for initial differences across location and year of birth by birth order interactions. Birth order by cohort interactions, δ_{bc}, account for national changes in the birth order gradient over time. Additionally, Stephens Jr and Yang (2014) highlight the importance of controlling for regional trends when estimating the impact of compulsory schooling laws; therefore, we also include year-of-birth by region fixed effects η_{rc}.

This first stage generates group and time fixed effects that are uncontaminated by treatment effects. Using these estimates, we then obtain residualized years of education for both the treated and untreated units in the sample: $\tilde{y}_{isc} = y_{isc} - \hat{\gamma}_{sb} - \hat{\delta_{bc}} - \hat{\eta}_{rc}$. Note that since fixed effects are based on untreated units, we are not able to obtain residuals for time periods in which all units in the region are treated or for always treated states. These are omitted from the analysis. We then regress these residuals on treatment dummies and their interactions with birth order:

$$\begin{aligned} \tilde{y}_{iscf} = \omega + \phi_1 SCHYRS^8_{iscf} + \phi_2 SCHYRS^8_{iscf} * LATERBORN_{iscf} + \\ \phi_3 SCHYRS^9_{iscf} + \phi_4 SCHYRS^9_{iscf} * LATERBORN_{iscf} + \beta f + \mu_{iscf} \end{aligned} \tag{3}$$

where $SCHYRS^8_{isc}$ and $SCHYRS^9_{isc}$ are dummies representing a state requirement of eight and nine or more years of schooling, respectively. Developed by Clay et al. (2021), this measure captures the effects of compulsory attendance laws, taking into account exemptions for child labor and years of schooling as well as continuation schools. We also include family fixed effects to focus on comparing effects between brothers.

In Table 4, panel (a), we display the coefficients on the interactions of compulsory schooling laws with being a later-born brother for non-Southern whites. As shown in column (1), these laws compress the first vs later-born gap in years of education by about

0.4 years. Laws requiring eight years of schooling also compress differences at the eighth grade, high school, and any college thresholds by 0.03, 0.05, and 0.03 percentage points, respectively. These results are most consistent with a substitution effect between familial and public investments. The positive coefficients in panel (a) of Table 4 show that public investments in human capital are able to reduce birth order differences in educational attainment between earlier and later born white brothers living outside of the South. We perform similar estimates for black men living outside of the South in Table 5, panel (a), but find no significant effects of these laws for this group.

We next perform a similar analysis for Southern men. Since compulsory schooling laws required extremely low levels of education in this region during most of the sample period (see Appendix B), we focus on laws requiring one to five or six years of schooling. Results for white and black brothers are shown in panel (b) of Tables 4 and 5, respectively. The results provide no significant evidence that compulsory schooling laws affect the birth order gap in the South.

8 Conclusion

Birth order differences are a substantial source of educational and labor market inequality. In this paper, we use linked historical Census data to provide the first evidence documenting birth order effects for U.S. men born in the late 1800s and early 1900s. Relative to their firstborn brothers, later-born white men complete 0.3-0.6 fewer years of education, have lower earnings, and have lower occupational scores. Later-born black men complete 0.2 fewer years of education than their firstborn brothers, but this does not translate into differences in labor market outcomes. Our point estimates for white men are remarkably similar to those obtained by Black et al. (2005) using more recent cohorts in Norway, but are larger in magnitude considering the lower overall levels of educational attainment in the early twentieth century.

We then incorporate compulsory schooling laws into the analysis to examine whether public investments can help mitigate within-family inequality in educational outcomes.

Previous work has not considered the relationship between within-family inequality and public investments. We find that, for white men outside of the South, compulsory schooling laws substantially reduce birth order gaps in educational attainment. This indicates that public investments can play a role in reducing birth order gaps by potentially serving as a substitute to parental investment.

References

Aaronson, D. and B. Mazumder (2011). The Impact of Rosenwald Schools on Black Achievement. *Journal of Political Economy 119*(5), 821–888.

Abramitzky, R., L. Boustan, K. Eriksson, J. Feigenbaum, and S. Pérez (2021, September). Automated linking of historical data. *Journal of Economic Literature 59*(3), 865–918.

Abramitzky, R., L. Boustan, K. Eriksson, S. Pérez, and M. Rashid (2020). Census linking project: Version 2.0 [dataset]. `https://censuslinkingproject.org`.

Acemoglu, D. and J. Angrist (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER macroeconomics annual 15*, 9–59.

Baker, R. B., J. Blanchette, and K. Eriksson (2020). Long-run impacts of agricultural shocks on educational attainment: Evidence from the boll weevil. *The Journal of Economic History 80*(1), 136–174.

Black, S. E., P. J. Devereux, and K. 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, S. E., P. J. Devereux, and K. G. Salvanes (2011). Older and wiser? Birth order and IQ of young men. *CESifo Economic Studies 57*(1), 103–120.

Black, S. E., E. Grönqvist, and B. Öckert (2018). Born to lead? The effect of birth order on non-cognitive abilities. *The Review of Economics and Statistics 100*(2), 274–286.

Breining, S., J. Doyle, D. N. Figlio, K. Karbownik, and J. Roth (2020). Birth order and delinquency: Evidence from Denmark and Florida. *Journal of Labor Economics 38*(1), 95–142.

Butts, K. (2022). did2s_stata: Two-stage difference-in-differences following gardner (2021). `https://github.com/kylebutts/did2s_stata/`.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics 225*(2), 200–230.

Clay, K., J. Lingwall, and M. J. Stephens (2021). Laws, educational outcomes, and returns to schooling evidence from the first wave of U.S. state compulsory attendance laws. *Labour Economics 68*, 101935.

Collins, W. J. and M. H. Wanamaker (2014). Selection and economic gains in the great migration of African Americans: new evidence from linked census data. *American Economic Journal: Applied Economics 6*(1), 220–52.

Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education 1*, 697–812.

Dammert, A. C. (2010). Siblings, child labor, and schooling in Nicaragua and Guatemala. *Journal of Population Economics 23*(1), 199–224.

Daysal, N. M., H. Ding, M. Rossin-Slater, and H. Schwandt (2021). Germs in the family: The long-term consequences of intra-household endemic respiratory disease spread. nber working paper no. 29524.

De Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review 110*(9), 2964–96.

De Haan, M., E. Plug, and J. Rosero (2014). Birth order and human capital development evidence from Ecuador. *Journal of Human Resources 49*(2), 359–392.

Del Boca, D., C. Flinn, and M. Wiswall (2014). Household choices and child development. *Review of Economic Studies 81*(1), 137–185.

Edmonds, E. V. (2006). Understanding sibling differences in child labor. *Journal of Population Economics 19*(4), 795–821.

Emerson, P. M. and A. P. Souza (2008). Birth order, child labor, and school attendance in Brazil. *World Development 36*(9), 1647–1664.

Esposito, L., S. M. Kumar, and A. Villaseñor (2020). The importance of being earliest: birth order and educational outcomes along the socioeconomic ladder in Mexico. *Journal of Population Economics*, 1–31.

Feigenbaum, J. J. (2018). Multiple measures of historical intergenerational mobility: Iowa 1915 to 1940. *The Economic Journal 128*(612), F446–F481.

Gardner, J. (2021). Two-stage differences in differences. working paper.

Goldin, C. and L. F. Katz (2010). *The Race Between Education and Technology.* harvard university press.

Goldin, C. and L. F. Katz (2011, August). *Understanding Long-Run Economic Growth: Geography, Institutions, and the Knowledge Economy*, Chapter Mass Secondary Schooling and the State: The Role of State Compulsion in the High School Movement. University of Chicago Press.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics 225*(2), 254–277.

Johnson, R. C. and C. K. Jackson (2019). Reducing inequality through dynamic complementarity: Evidence from Head Start and public school spending. *American Economic Journal: Economic Policy 11*(4), 310–49.

Lehmann, J.-Y. K., A. Nuevo-Chiquero, and M. Vidal-Fernandez (2018). The early origins of birth order differences in children's outcomes and parental behavior. *Journal of Human Resources 53*(1), 123–156.

Lin, W., J. Pantano, and S. Sun (2020). Birth Order and Unwanted Fertility. *Journal of Population Economics 33*(2), 413–440.

Lleras-Muney, A. (2002). Were compulsory attendance and child labor laws effective? an analysis from 1915 to 1939. *The Journal of Law and Economics 45*(2), 401–435.

Price, J. (2008). Parent-child quality time does birth order matter? *Journal of Human Resources 43*(1), 240–265.

Ruggles, S., S. Flood, S. Foster, R. Goeken, J. Pacas, M. Schouweiler, and M. Sobek (2021). IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021. https://doi.org/10.18128/D010.V11.0.

Schmick, E. J. and A. Shertzer (2019). The Impact of Early Investments in Urban School Systems in the United States. NBER Working Paper No. w25663.

Stephens Jr, M. and D.-Y. Yang (2014). Compulsory education and the benefits of schooling. *American Economic Review 104*(6), 1777–92.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics 225*(2), 175–199.

Table 1: Summary Statistics for Analysis Sample

(a) Native-Born Whites

	(1) Analysis Sample	(2) Std Dev	(3) Diff vs Population	(4) Std Err
1940 Characteristics				
Age	41.058	(8.961)	-0.266***	(0.004)
Any Earnings	0.755	(0.428)	-0.004***	(0.000)
Years of Education	9.528	(3.404)	0.473***	(0.002)
Completed Grade 8	0.796	(0.431)	0.052***	(0.000)
Completed High School	0.304	(0.444)	0.041***	(0.000)
Any College	0.151	(0.335)	0.027***	(0.000)
Married	0.838	(0.397)	0.041***	(0.000)
Urban	0.549	(0.497)	-0.008***	(0.000)
Number of Observations	3,475,796		20,643,376	
1900 or 1910 Characteristics				
Number of Children under 18 in Household	3.680	(2.143)	0.041***	(0.001)
First Born	0.374	(0.485)	-0.007***	(0.000)
Second Born	0.262	(0.440)	-0.001***	(0.000)
Third Born	0.165	(0.371)	0.001***	(0.000)
Fourth Born	0.099	(0.296)	0.002***	(0.000)
Fifth Born or Higher	0.100	(0.296)	0.004***	(0.000)
Father is Farm Owner or Worker	0.464	(0.499)	-0.012***	(0.000)
Northeast	0.269	(0.453)	-0.022***	(0.000)
Midwest	0.470	(0.493)	0.058***	(0.000)
South	0.173	(0.413)	-0.052***	(0.000)
West	0.089	(0.265)	0.015***	(0.000)
Number of Observations	3,475,796		26,782,211	

(b) Native-Born Blacks

	(1) Analysis Sample	(2) Std Dev	(3) Diff vs Population	(4) Std Err
1940 Characteristics				
Age	41.003	(8.731)	0.117***	(0.017)
Any Earnings	0.747	(0.431)	-0.007***	(0.001)
Years of Education	6.357	(3.498)	0.975***	(0.009)
Completed Grade 8	0.363	(0.442)	0.104***	(0.001)
Completed High School	0.116	(0.256)	0.049***	(0.001)
Any College	0.057	(0.181)	0.025***	(0.001)
Married	0.828	(0.408)	0.042***	(0.001)
Urban	0.526	(0.498)	-0.023***	(0.001)
Number of Observations	175,021		2,315,711	
1900 or 1910 Characteristics				
Number of Children under 18 in Household	4.186	(2.342)	0.220***	(0.005)
First Born	0.338	(0.478)	-0.016***	(0.001)
Second Born	0.241	(0.428)	-0.001***	(0.001)
Third Born	0.168	(0.370)	0.004***	(0.001)
Fourth Born	0.109	(0.309)	0.003***	(0.001)
Fifth Born or Higher	0.144	(0.343)	0.010***	(0.001)
Father is Farm Owner or Worker	0.671	(0.444)	-0.063***	(0.001)
Northeast	0.084	(0.236)	0.025***	(0.001)
Midwest	0.091	(0.251)	0.025***	(0.001)
South	0.811	(0.345)	-0.054***	(0.001)
West	0.014	(0.104)	0.003***	(0.000)
Number of Observations	175,021		4,142,369	

Statistics are from the population of native-born males ages 28-59 in 1940 or ages 0-17 in 1910 or 1900. Columns (1) and (2) are the mean and standard deviation, respectively, for the analysis sample. Column (3) displays the coefficient from a regression where the dependent variable is a given characteristic and the independent variable is a dummy that equals 1 for analysis sample observations. Column (4) is the standard error of the coefficient. Number of observations is the number in the analysis sample in column (1) and the population in column (3). * p<0.1 ** p<0.05 *** p<0.01

Table 2: Birth Order and Other Long-Run Outcomes: Native-Born Whites

(a) Non-South

	Non-South					
	(1) Any Earnings	(2) Log Earnings	(3) Log Occ Score	(4) Farm Occ	(5) Same Occ as Father	(6) Migrant
Second	-0.007***	-0.010**	-0.011***	0.009***	0.011***	-0.005***
	(0.002)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)
Third	-0.007**	-0.021***	-0.015***	0.010***	0.010***	-0.003
	(0.003)	(0.007)	(0.003)	(0.003)	(0.003)	(0.003)
Fourth	-0.006	-0.018*	-0.021***	0.008**	0.008**	-0.004
	(0.004)	(0.010)	(0.004)	(0.003)	(0.004)	(0.004)
Fifth+	-0.004	-0.010	-0.025***	0.006	0.005	-0.002
	(0.005)	(0.013)	(0.005)	(0.005)	(0.005)	(0.005)
Observations	1053152	711008	808063	1010758	878683	1053152
R^2	0.553	0.717	0.609	0.667	0.625	0.677
Adjusted R^2	0.188	0.327	0.170	0.376	0.298	0.413
Dep Var Mean	0.749	3.272	3.338	0.201	0.203	0.280

(b) South

	South						
	(1) Any Earnings	(2) Log Earnings	(3) Log Occ Score	(4) Farm Occ	(5) Same Occ as Father	(6) Migrant	(7) Northern Migrant
Second	-0.010**	-0.018	-0.004	0.015***	0.012**	-0.001	0.003
	(0.005)	(0.015)	(0.006)	(0.005)	(0.005)	(0.004)	(0.003)
Third	-0.007	-0.030	-0.013	0.021***	0.014*	0.005	0.005
	(0.007)	(0.022)	(0.008)	(0.007)	(0.007)	(0.006)	(0.004)
Fourth	-0.005	-0.019	-0.015	0.019**	0.013	0.009	0.007
	(0.010)	(0.030)	(0.011)	(0.010)	(0.010)	(0.008)	(0.006)
Fifth+	0.002	-0.037	-0.023	0.016	0.011	0.017	0.014*
	(0.014)	(0.041)	(0.015)	(0.013)	(0.013)	(0.012)	(0.008)
Observations	206560	125720	138576	197143	179366	206560	206560
R^2	0.546	0.751	0.653	0.629	0.613	0.647	0.606
Adjusted R^2	0.167	0.347	0.179	0.292	0.263	0.352	0.276
Dep Var Mean	0.680	2.992	3.317	0.297	0.259	0.274	0.089

Note: This table reports parameter estimates and standard errors (in parentheses) for regressions of the listed dependent variable on birth order dummies. All regressions include family fixed effects and dummy variables for year of birth. Standard errors are clustered at the family level. * p<0.1 ** p<0.05 *** p<0.01

Table 3: Birth Order and Other Long-Run Outcomes: Native-Born Blacks

(a) Non-South

	Non-South					
	(1) Any Earnings	(2) Log Earnings	(3) Log Occ Score	(4) Farm Occ	(5) Same Occ as Father	(6) Migrant
Second	0.011	0.000	0.004	-0.015	-0.009	-0.011
	(0.023)	(0.054)	(0.029)	(0.021)	(0.023)	(0.024)
Third	0.024	0.025	-0.006	-0.046	-0.034	0.003
	(0.033)	(0.075)	(0.042)	(0.030)	(0.034)	(0.035)
Fourth	0.044	0.024	-0.036	-0.057	-0.043	-0.010
	(0.044)	(0.100)	(0.056)	(0.040)	(0.045)	(0.048)
Fifth+	0.074	-0.004	-0.038	-0.088	-0.058	-0.014
	(0.058)	(0.130)	(0.076)	(0.054)	(0.061)	(0.062)
Observations	8112	5987	6465	7536	6648	8112
R^2	0.506	0.680	0.591	0.617	0.598	0.622
Adjusted R^2	0.068	0.225	0.081	0.229	0.192	0.286
Dep Var Mean	0.822	2.876	3.111	0.142	0.147	0.328

(b) South

	South						
	(1) Any Earnings	(2) Log Earnings	(3) Log Occ Score	(4) Farm Occ	(5) Same Occ as Father	(6) Migrant	(7) Northern Migrant
Second	0.006	0.033	-0.018	0.004	0.012	0.004	0.005
	(0.014)	(0.043)	(0.021)	(0.015)	(0.015)	(0.014)	(0.012)
Third	0.009	0.006	-0.016	0.011	0.016	-0.014	-0.004
	(0.019)	(0.058)	(0.028)	(0.021)	(0.020)	(0.019)	(0.016)
Fourth	0.008	0.017	-0.019	0.004	0.022	-0.013	0.000
	(0.024)	(0.076)	(0.035)	(0.027)	(0.025)	(0.024)	(0.020)
Fifth+	0.023	0.028	-0.021	-0.010	0.003	-0.014	-0.006
	(0.033)	(0.103)	(0.047)	(0.036)	(0.035)	(0.033)	(0.028)
Observations	31124	19841	19154	29298	27368	31124	31124
R^2	0.537	0.710	0.668	0.604	0.600	0.592	0.598
Adjusted R^2	0.122	0.212	0.096	0.207	0.199	0.226	0.239
Dep Var Mean	0.719	2.534	3.084	0.346	0.258	0.359	0.211

Note: This table reports parameter estimates and standard errors (in parentheses) for regressions of the listed dependent variable on birth order dummies. All regressions include family fixed effects and dummy variables for year of birth. Standard errors are clustered at the family level. * p<0.1 ** p<0.05 *** p<0.01

Table 4: Birth Order, Compulsory Schooling, and Educational Attainment: Native-born Whites

(a) Non-South

	(1) Yrs Ed	(2) 8th Grade	(3) HS	(4) Any Coll
8 Years Required x Later Born	0.404**	0.034*	0.052**	0.026***
	(0.179)	(0.019)	(0.022)	(0.008)
9+ Years Required x Later Born	0.548	0.051*	0.069	0.035
	(0.498)	(0.028)	(0.046)	(0.021)
Observations	728993	728993	728993	728993

(b) South

	(1) Yrs Ed	(2) 8th Grade	(3) HS	(4) Any Coll
1-5 Years Required x Later Born	0.528	0.061	0.050	0.029
	(0.513)	(0.060)	(0.039)	(0.031)
6 Years Required x Later Born	0.732	0.106	0.129	0.035
	(2.296)	(0.516)	(0.192)	(0.169)
Observations	151132	151132	151132	151132

Note: This table reports parameter estimates and standard errors (in parentheses) for two-stage difference-in-differences estimation in equation (3) with dummies for laws requiring 8 or 9 or more years of schooling in Panel (a) and laws requiring 1-5 or 6 years of schooling in panel (b). Standard errors are clustered at the state-by-cohort level. Standard errors are obtained using bootstrap (100 iterations) as recommended for large datasets with many fixed effects by Butts (2022). * p<0.1 ** p<0.05 *** p<0.01

Table 5: Birth Order, Compulsory Schooling, and Educational Attainment: Native-born Blacks

(a) Non-South

	(1) Yrs Ed	(2) 8th Grade	(3) HS	(4) Any Coll
8 Years Required x Later Born	-0.389	-0.036	-0.056	-0.077
	(0.680)	(0.108)	(0.091)	(0.067)
9+ Years Required x Later Born	-0.165	-0.169	0.107	0.011
	(1.455)	(0.325)	(0.229)	(0.118)
Observations	5951	5951	5951	5951

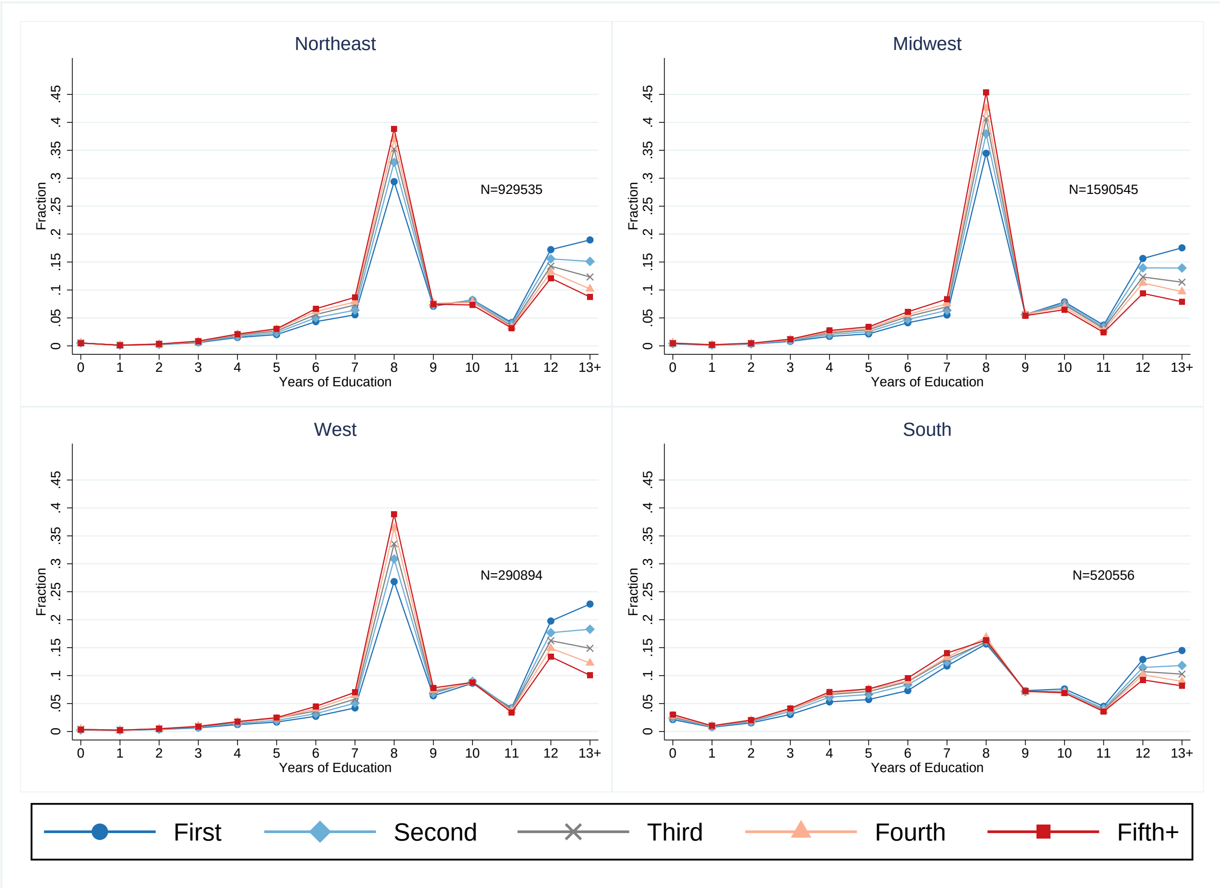
(b) South

	(1) Yrs Ed	(2) 8th Grade	(3) HS	(4) Any Coll
1-5 Years Required x Later Born	0.302	-0.058	0.047	0.016
	(0.575)	(0.103)	(0.061)	(0.040)
6 Years Required x Later Born	-0.417	-0.101	0.007	0.020
	(1.142)	(0.182)	(0.074)	(0.057)
Observations	23217	23217	23217	23217

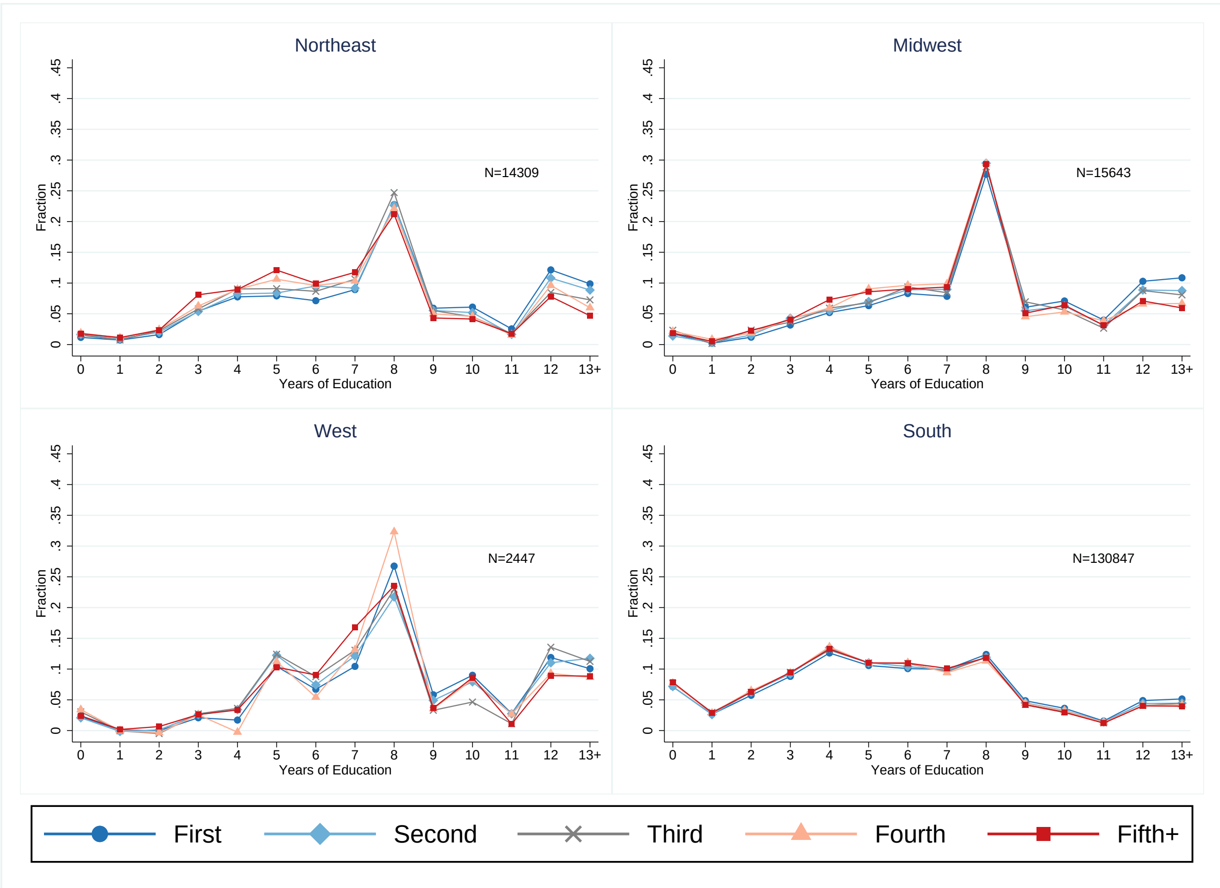
Note: This table reports parameter estimates and standard errors (in parentheses) for two-stage difference-in-differences estimation in equation (3) with dummies for laws requiring 8 or 9 or more years of schooling in Panel (a) and laws requiring 1-5 or 6 years of schooling in panel (b). Standard errors are clustered at the state-by-cohort level. Standard errors are obtained using bootstrap (100 iterations) as recommended for large datasets with many fixed effects by Butts (2022). * p<0.1 ** p<0.05 *** p<0.01

Figure 1: Completed Years of Education by Birth Order

(a) Native-Born Whites



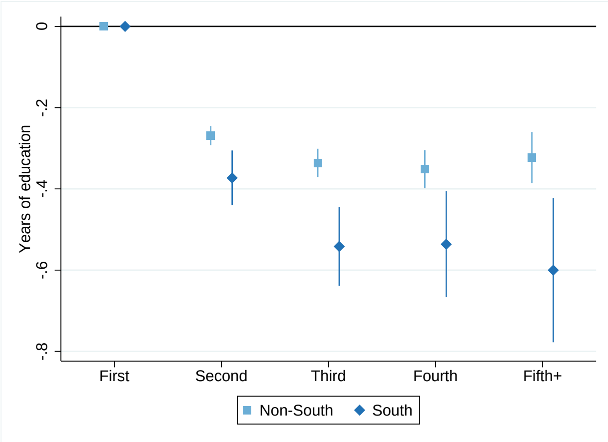
(b) Native-Born Blacks



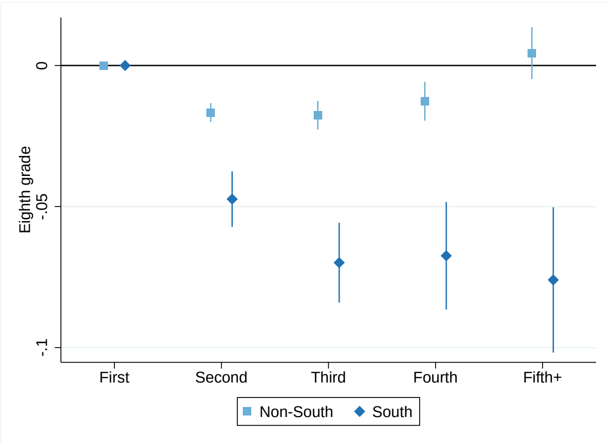
Note: These figures show the age-adjusted fraction of individuals reporting the given level of completed educational attainment in 1940. N represents the number in the sample for each region and race.

Figure 2: Educational Attainment and Birth Order: Native-Born Whites

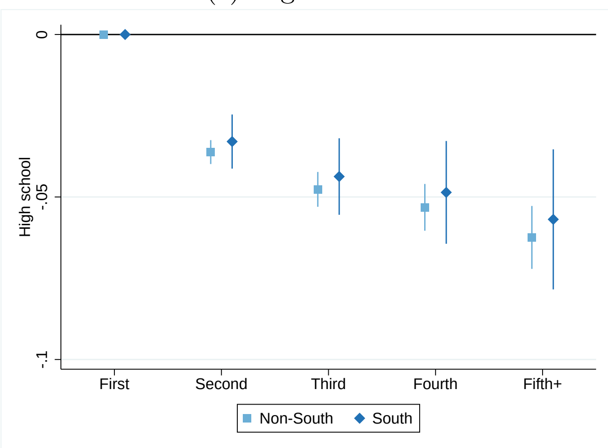
(a) Years of Education



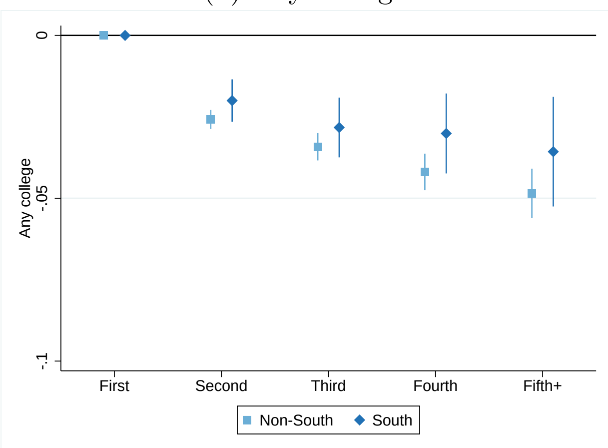
(b) Eighth Grade



(c) High School



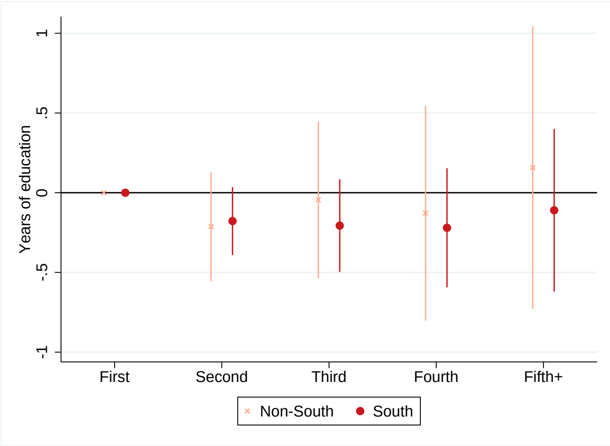
(d) Any College



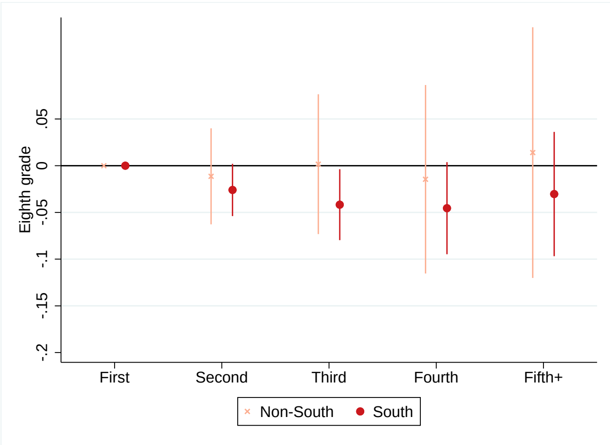
Note: These figures display the point estimates (dots) and 95% confidence intervals (bars) for regressions of the listed dependent variable on birth order dummies. All regressions include family fixed effects and dummy variable for year of birth. Standard errors are clustered at the family level.

Figure 3: Educational Attainment and Birth Order: Native-Born Blacks

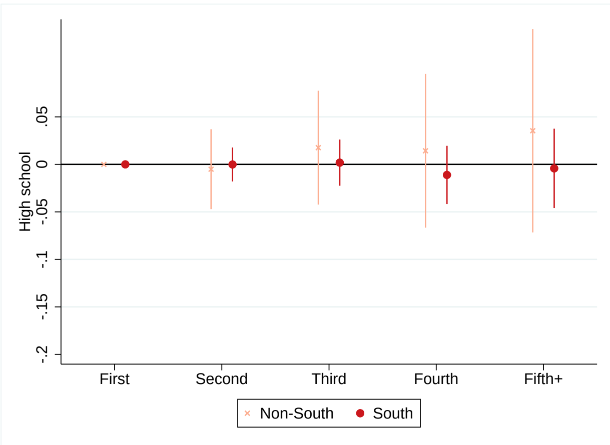
(a) Years of Education



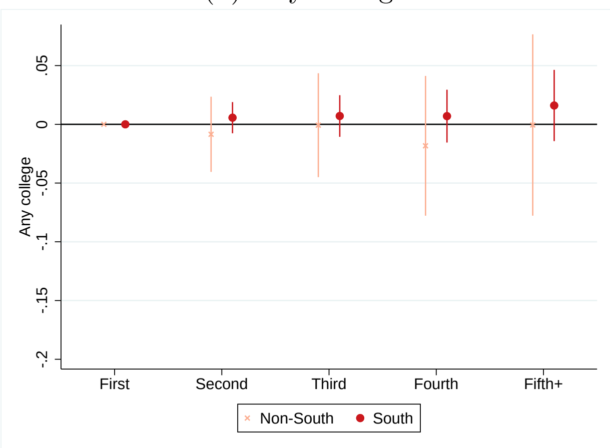
(b) Eighth Grade



(c) High School



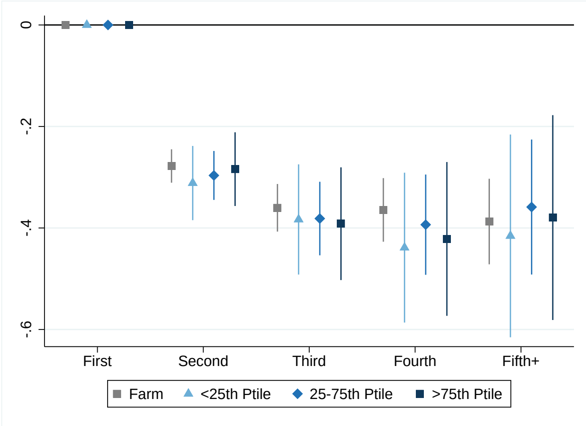
(d) Any College



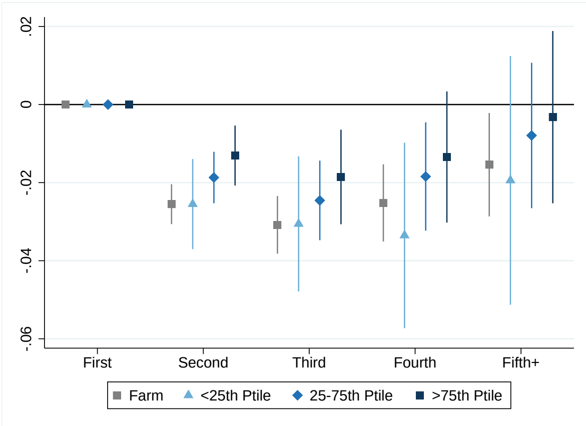
Note: These figures display the point estimates (dots) and 95% confidence intervals (bars) for regressions of the listed dependent variable on birth order dummies. All regressions include family fixed effects and dummy variable for year of birth. Standard errors are clustered at the family level.

Figure 4: Educational Attainment and Birth Order by Father's Occupation Score: Native-Born Whites

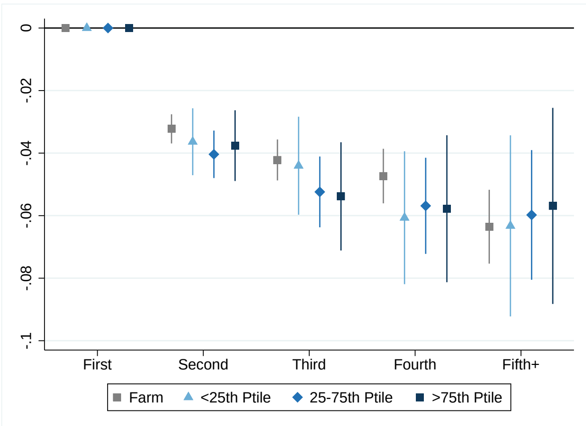
(a) Years of Education



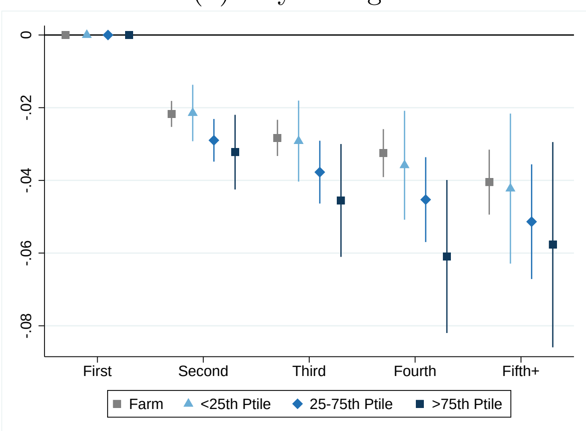
(b) Eighth Grade



(c) High School



(d) Any College



Note: These figures display the point estimates (dots) and 95% confidence intervals (bars) for regressions of the listed dependent variable on birth order dummies. All regressions include family fixed effects and dummy variable for year of birth. Standard errors are clustered at the family level.

Appendix A: Additional Tables and Figures

Table A1: Summary Statistics for Full Matched Sample

(a) Native-Born Whites

	(1) Matched	(2) Std Dev	(3) Diff vs Population	(4) Std Err
1940 Characteristics				
Age	42.764	(8.961)	2.045***	(0.004)
Any Earnings	0.743	(0.428)	-0.020***	(0.000)
Completed Grade 8	0.767	(0.431)	0.020***	(0.000)
Completed High School	0.273	(0.444)	0.004***	(0.000)
Any College	0.135	(0.335)	0.008***	(0.000)
Married	0.837	(0.397)	0.046***	(0.000)
Urban	0.535	(0.497)	-0.028***	(0.000)
Number of Observations	5,648,498		20,643,376	
1900 or 1910 Characteristics				
Number of Children under 18 in Household	3.666	(2.143)	0.026***	(0.001)
First Born	0.374	(0.485)	-0.007***	(0.000)
Second Born	0.262	(0.440)	-0.001***	(0.000)
Third Born	0.165	(0.371)	0.001***	(0.000)
Fourth Born	0.099	(0.296)	0.002***	(0.000)
Fifth Born or Higher	0.100	(0.296)	0.004***	(0.000)
Father is Farm Owner or Worker	0.499	(0.499)	0.032***	(0.000)
Northeast	0.260	(0.453)	-0.035***	(0.000)
Midwest	0.466	(0.493)	0.060***	(0.000)
South	0.186	(0.413)	-0.040***	(0.000)
West	0.088	(0.265)	0.015***	(0.000)
Number of Observations	5,648,498		26,782,211	

(b) Native-Born Blacks

	(1) Matched	(2) Std Dev	(3) Diff vs Population	(4) Std Err
1940 Characteristics				
Age	43.140	(8.731)	2.714***	(0.013)
Any Earnings	0.729	(0.431)	-0.029***	(0.001)
Completed Grade 8	0.336	(0.442)	0.085***	(0.001)
Completed High School	0.105	(0.256)	0.041***	(0.001)
Any College	0.052	(0.181)	0.022***	(0.000)
Married	0.827	(0.408)	0.046***	(0.001)
Urban	0.509	(0.498)	-0.047***	(0.001)
Number of Observations	400,105		2,315,711	
1900 or 1910 Characteristics				
Number of Children under 18 in Household	4.084	(2.342)	0.120***	(0.004)
First Born	0.338	(0.478)	-0.016***	(0.001)
Second Born	0.241	(0.428)	-0.001***	(0.001)
Third Born	0.167	(0.370)	0.004***	(0.001)
Fourth Born	0.109	(0.309)	0.003***	(0.001)
Fifth Born or Higher	0.144	(0.343)	0.010***	(0.001)
Father is Farm Owner or Worker	0.712	(0.444)	-0.021***	(0.001)
Northeast	0.076	(0.236)	0.018***	(0.000)
Midwest	0.086	(0.251)	0.021***	(0.000)
South	0.825	(0.345)	-0.041***	(0.001)
West	0.013	(0.104)	0.002***	(0.000)
Number of Observations	400,105		4,142,369	

Statistics are from the population of native-born males ages 28-59 in 1940 or ages 0-17 in 1910 or 1900. Columns (1) and (2) are the mean and standard deviation, respectively, for the matched sample, which refers to all 1940 men who are matched to any 1910 or 1900 characteristics, including those with missing birth order and educational attainment information. Column (3) displays the coefficient from a regression where the dependent variable is a given characteristic and the independent variable is a dummy that equals 1 for matched sample observations. Column (4) is the standard error of the coefficient. Number of observations is the number in the matched sample in column (1) and the population in column (3). * p<0.1 ** p<0.05 *** p<0.01

Table A2: Summary Statistics for Regression Sample

(a) Native-Born Whites

	(1) Regression Sample	(2) Std Dev	(3) Diff vs Population	(4) Std Err
1940 Characteristics				
Age	41.838	(8.961)	0.595***	(0.007)
Any Earnings	0.737	(0.428)	-0.022***	(0.000)
Years of Education	9.233	(3.404)	0.102***	(0.003)
Completed Grade 8	0.780	(0.431)	0.029***	(0.000)
Completed High School	0.263	(0.444)	-0.008***	(0.000)
Any College	0.127	(0.335)	-0.002***	(0.000)
Married	0.838	(0.397)	0.036***	(0.000)
Urban	0.518	(0.497)	-0.040***	(0.000)
Number of Observations	1,234,067		20,643,376	
1900 or 1910 Characteristics				
Number of Children under 18 in Household	4.712	(2.143)	1.118***	(0.002)
First Born	0.244	(0.485)	-0.147***	(0.000)
Second Born	0.249	(0.440)	-0.015***	(0.000)
Third Born	0.199	(0.371)	0.038***	(0.000)
Fourth Born	0.139	(0.296)	0.046***	(0.000)
Fifth Born or Higher	0.168	(0.296)	0.078***	(0.000)
Father is Farm Owner or Worker	0.520	(0.499)	0.049***	(0.000)
Northeast	0.242	(0.453)	-0.047***	(0.000)
Midwest	0.502	(0.493)	0.087***	(0.000)
South	0.163	(0.413)	-0.057***	(0.000)
West	0.092	(0.265)	0.017***	(0.000)
Number of Observations	1,234,067		26,782,211	

(b) Native-Born Blacks

	(1) Regression Sample	(2) Std Dev	(3) Diff vs Population	(4) Std Err
1940 Characteristics				
Age	41.774	(8.731)	0.894***	(0.035)
Any Earnings	0.741	(0.431)	-0.013***	(0.002)
Years of Education	6.310	(3.498)	0.866***	(0.019)
Completed Grade 8	0.355	(0.442)	0.090***	(0.002)
Completed High School	0.109	(0.256)	0.039***	(0.002)
Any College	0.054	(0.181)	0.021***	(0.001)
Married	0.831	(0.408)	0.043***	(0.002)
Urban	0.514	(0.498)	-0.034***	(0.003)
Number of Observations	38,129		2,315,711	
1900 or 1910 Characteristics				
Number of Children under 18 in Household	5.462	(2.342)	1.501***	(0.010)
First Born	0.210	(0.478)	-0.146***	(0.002)
Second Born	0.215	(0.428)	-0.028***	(0.002)
Third Born	0.188	(0.370)	0.025***	(0.002)
Fourth Born	0.145	(0.309)	0.039***	(0.002)
Fifth Born or Higher	0.242	(0.343)	0.109***	(0.002)
Father is Farm Owner or Worker	0.709	(0.444)	-0.021***	(0.002)
Northeast	0.091	(0.236)	0.032***	(0.001)
Midwest	0.100	(0.251)	0.033***	(0.002)
South	0.793	(0.345)	-0.070***	(0.002)
West	0.016	(0.104)	0.006***	(0.001)
Number of Observations	38,129		4,142,369	

Statistics are from the population of native-born males ages 28-59 in 1940 or ages 0-17 in 1910 or 1900. Columns (1) and (2) are the mean and standard deviation, respectively, for the regression sample, which refers to all 1940 men who are matched to any 1910 or 1900 characteristics, have non-missing birth order and educational attainment information, and have at least one brother also fulfilling these criteria. Column (3) displays the coefficient from a regression where the dependent variable is a given characteristic and the independent variable is a dummy that equals 1 for matched sample observations. Column (4) is the standard error of the coefficient. Number of observations is the number in the matched sample in column (1) and the population in column (3). * p<0.1 ** p<0.05 *** p<0.01

Table A3: Birth Order and Educational Attainment

(a) Native-Born Whites

	Non-South				South			
	(1) Yrs Ed	(2) 8th Grade	(3) HS	(4) Any Coll	(5) Yrs Ed	(6) 8th Grade	(7) HS	(8) Any Coll
Second	-0.269***	-0.017***	-0.036***	-0.026***	-0.373***	-0.047***	-0.033***	-0.020***
	(0.012)	(0.002)	(0.002)	(0.001)	(0.034)	(0.005)	(0.004)	(0.003)
Third	-0.336***	-0.018***	-0.048***	-0.034***	-0.542***	-0.070***	-0.044***	-0.028***
	(0.018)	(0.003)	(0.003)	(0.002)	(0.049)	(0.007)	(0.006)	(0.005)
Fourth	-0.352***	-0.013***	-0.053***	-0.042***	-0.536***	-0.067***	-0.049***	-0.030***
	(0.024)	(0.004)	(0.004)	(0.003)	(0.067)	(0.010)	(0.008)	(0.006)
Fifth+	-0.323***	0.004	-0.062***	-0.049***	-0.600***	-0.076***	-0.057***	-0.036***
	(0.032)	(0.005)	(0.005)	(0.004)	(0.091)	(0.013)	(0.011)	(0.009)
Observations	1032555	1032555	1032555	1032555	201512	201512	201512	201512
R^2	0.712	0.599	0.666	0.639	0.706	0.648	0.660	0.647
Adjusted R^2	0.470	0.261	0.385	0.336	0.451	0.341	0.364	0.339
Dep Var Mean	9.411	0.821	0.271	0.130	8.318	0.572	0.222	0.111

(b) Native-Born Blacks

	Non-South				South			
	(1) Yrs Ed	(2) 8th Grade	(3) HS	(4) Any Coll	(5) Yrs Ed	(6) 8th Grade	(7) HS	(8) Any Coll
Second	-0.213	-0.011	-0.005	-0.008	-0.178	-0.026*	-0.000	0.006
	(0.175)	(0.026)	(0.021)	(0.016)	(0.109)	(0.014)	(0.009)	(0.007)
Third	-0.045	0.002	0.018	-0.001	-0.206	-0.042**	0.002	0.007
	(0.251)	(0.038)	(0.031)	(0.023)	(0.148)	(0.019)	(0.012)	(0.009)
Fourth	-0.128	-0.015	0.014	-0.018	-0.220	-0.045*	-0.011	0.007
	(0.344)	(0.051)	(0.041)	(0.030)	(0.191)	(0.025)	(0.016)	(0.011)
Fifth+	0.157	0.014	0.035	-0.001	-0.110	-0.030	-0.004	0.016
	(0.452)	(0.068)	(0.055)	(0.039)	(0.260)	(0.034)	(0.021)	(0.015)
Observations	7908	7908	7908	7908	30221	30221	30221	30221
R^2	0.665	0.631	0.611	0.592	0.595	0.567	0.554	0.547
Adjusted R^2	0.355	0.288	0.250	0.214	0.215	0.160	0.134	0.120
Dep Var Mean	8.112	0.610	0.188	0.086	5.839	0.288	0.088	0.046

Note: This table reports parameter estimates and standard errors (in parentheses) for regressions of the listed dependent variable on birth order dummies. All regressions include family fixed effects and dummy variables for year of birth. Standard errors are clustered at the family level. * p<0.1 ** p<0.05 *** p<0.01

Table A4: Birth Order and Educational Attainment: Firstborn vs Later-Born

(a) Native-Born Whites

	Non-South				South			
	(1) Yrs Ed	(2) 8th Grade	(3) HS	(4) Any Coll	(5) Yrs Ed	(6) 8th Grade	(7) HS	(8) Any Coll
Later Born	-0.277***	-0.022***	-0.033***	-0.023***	-0.377***	-0.048***	-0.031***	-0.019***
	(0.011)	(0.002)	(0.002)	(0.001)	(0.031)	(0.005)	(0.004)	(0.003)
Observations	1032555	1032555	1032555	1032555	201512	201512	201512	201512
R^2	0.712	0.599	0.666	0.639	0.706	0.648	0.660	0.646
Adjusted R^2	0.470	0.261	0.385	0.336	0.451	0.341	0.364	0.339
Dep Var Mean	9.411	0.821	0.271	0.130	8.318	0.572	0.222	0.111

(b) Native-Born Blacks

	Non-South				South			
	(1) Yrs Ed	(2) 8th Grade	(3) HS	(4) Any Coll	(5) Yrs Ed	(6) 8th Grade	(7) HS	(8) Any Coll
Later Born	-0.257	-0.014	-0.008	-0.009	-0.213**	-0.033**	0.001	0.003
	(0.159)	(0.024)	(0.019)	(0.015)	(0.100)	(0.013)	(0.008)	(0.006)
Observations	7908	7908	7908	7908	30221	30221	30221	30221
R^2	0.665	0.631	0.611	0.592	0.595	0.567	0.554	0.547
Adjusted R^2	0.355	0.289	0.250	0.214	0.215	0.160	0.134	0.120
Dep Var Mean	8.112	0.610	0.188	0.086	5.839	0.288	0.088	0.046

Note: This table reports parameter estimates and standard errors (in parentheses) for regressions of the listed dependent variable on a dummy variable taking a value of 1 if the individual is not the firstborn child. All regressions include family fixed effects and dummy variables for year of birth. Standard errors are clustered at the family level. * $p<0.1$ ** $p<0.05$ *** $p<0.01$

Table A5: Birth Order by Father's Occupation Subgroups: Native-born Whites

(a) Years of education

	Non-South				South			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Farm	0-25th	25th-75th	75th+	Farm	0-25th	25th-75th	75th+
Second	-0.257***	-0.281***	-0.288***	-0.276***	-0.347***	-0.635***	-0.383***	-0.329***
	(0.018)	(0.038)	(0.025)	(0.039)	(0.042)	(0.155)	(0.102)	(0.124)
Third	-0.311***	-0.348***	-0.363***	-0.364***	-0.518***	-0.756***	-0.566***	-0.552***
	(0.026)	(0.057)	(0.038)	(0.059)	(0.059)	(0.238)	(0.154)	(0.184)
Fourth	-0.315***	-0.404***	-0.380***	-0.396***	-0.504***	-0.802**	-0.524**	-0.534**
	(0.034)	(0.077)	(0.051)	(0.081)	(0.079)	(0.326)	(0.214)	(0.254)
Fifth+	-0.321***	-0.353***	-0.337***	-0.339***	-0.559***	-1.106**	-0.565*	-0.562
	(0.046)	(0.104)	(0.069)	(0.107)	(0.108)	(0.442)	(0.289)	(0.342)
Observations	429841	106150	248227	113046	131689	10057	24027	17446
R^2	0.684	0.673	0.689	0.723	0.675	0.705	0.693	0.693
Adjusted R^2	0.427	0.392	0.422	0.482	0.394	0.438	0.421	0.422
Dep Var Mean	8.804	8.845	9.613	11.041	7.791	7.637	9.015	10.802

(b) Eighth grade completion

	Non-South				South			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Farm	0-25th	25th-75th	75th+	Farm	0-25th	25th-75th	75th+
Second	-0.018***	-0.020***	-0.017***	-0.012***	-0.050***	-0.085***	-0.032**	-0.021
	(0.003)	(0.006)	(0.003)	(0.004)	(0.006)	(0.023)	(0.014)	(0.014)
Third	-0.017***	-0.023**	-0.022***	-0.016**	-0.076***	-0.121***	-0.049**	-0.035
	(0.004)	(0.009)	(0.005)	(0.006)	(0.009)	(0.035)	(0.021)	(0.021)
Fourth	-0.010*	-0.025**	-0.017**	-0.012	-0.073***	-0.131***	-0.030	-0.022
	(0.006)	(0.012)	(0.007)	(0.009)	(0.012)	(0.048)	(0.029)	(0.030)
Fifth+	0.006	-0.006	-0.005	0.002	-0.081***	-0.176***	-0.040	-0.036
	(0.007)	(0.017)	(0.010)	(0.012)	(0.016)	(0.065)	(0.040)	(0.040)
Observations	429841	106150	248227	113046	131689	10057	24027	17446
R^2	0.606	0.592	0.579	0.566	0.634	0.649	0.634	0.607
Adjusted R^2	0.285	0.241	0.216	0.188	0.318	0.332	0.309	0.259
Dep Var Mean	0.787	0.768	0.839	0.911	0.518	0.492	0.657	0.807

(c) High school completion

	Non-South				South			
	(1) Farm	(2) 0-25th	(3) 25th-75th	(4) 75th+	(5) Farm	(6) 0-25th	(7) 25th-75th	(8) 75th+
Second	-0.034***	-0.036***	-0.041***	-0.036***	-0.027***	-0.040**	-0.036***	-0.046***
	(0.003)	(0.006)	(0.004)	(0.006)	(0.005)	(0.018)	(0.014)	(0.017)
Third	-0.043***	-0.045***	-0.052***	-0.051***	-0.038***	-0.030	-0.052**	-0.069***
	(0.004)	(0.008)	(0.006)	(0.009)	(0.007)	(0.027)	(0.020)	(0.025)
Fourth	-0.048***	-0.060***	-0.057***	-0.053***	-0.041***	-0.051	-0.052*	-0.078**
	(0.005)	(0.011)	(0.008)	(0.013)	(0.009)	(0.037)	(0.028)	(0.035)
Fifth+	-0.064***	-0.063***	-0.059***	-0.053***	-0.051***	-0.055	-0.058	-0.069
	(0.007)	(0.015)	(0.011)	(0.017)	(0.013)	(0.050)	(0.038)	(0.046)
Observations	429841	106150	248227	113046	131689	10057	24027	17446
R^2	0.635	0.628	0.646	0.680	0.616	0.652	0.651	0.668
Adjusted R^2	0.338	0.308	0.341	0.402	0.285	0.339	0.341	0.374
Dep Var Mean	0.196	0.203	0.291	0.482	0.162	0.183	0.294	0.496

(d) Any college

	Non-South				South			
	(1) Farm	(2) 0-25th	(3) 25th-75th	(4) 75th+	(5) Farm	(6) 0-25th	(7) 25th-75th	(8) 75th+
Second	-0.024***	-0.021***	-0.028***	-0.031***	-0.014***	-0.022*	-0.035***	-0.036**
	(0.002)	(0.004)	(0.003)	(0.006)	(0.004)	(0.013)	(0.011)	(0.016)
Third	-0.030***	-0.030***	-0.036***	-0.044***	-0.021***	-0.012	-0.051***	-0.053**
	(0.003)	(0.006)	(0.005)	(0.008)	(0.005)	(0.019)	(0.016)	(0.023)
Fourth	-0.035***	-0.036***	-0.044***	-0.060***	-0.022***	-0.026	-0.057***	-0.065**
	(0.004)	(0.008)	(0.006)	(0.011)	(0.007)	(0.025)	(0.022)	(0.031)
Fifth+	-0.044***	-0.043***	-0.049***	-0.055***	-0.024***	-0.020	-0.069**	-0.069*
	(0.005)	(0.011)	(0.008)	(0.015)	(0.009)	(0.035)	(0.030)	(0.042)
Observations	429841	106150	248227	113046	131689	10057	24027	17446
R^2	0.586	0.596	0.626	0.676	0.594	0.621	0.641	0.679
Adjusted R^2	0.249	0.248	0.303	0.393	0.243	0.279	0.322	0.394
Dep Var Mean	0.090	0.081	0.130	0.272	0.075	0.078	0.136	0.303

Note: This table reports parameter estimates and standard errors (in parentheses) for regressions of the listed dependent variable on birth order dummies where the birth order measure only accounts for female siblings. All regressions include family fixed effects and dummy variables for year of birth. Standard errors are clustered at the family level. * $p<0.1$ ** $p<0.05$ *** $p<0.01$.

Table A6: Birth Order among Male Sibling and Educational Attainment

(a) Native-Born Whites

	Non-South				South			
	(1) Yrs Ed	(2) 8th Grade	(3) HS	(4) Any Coll	(5) Yrs Ed	(6) 8th Grade	(7) HS	(8) Any Coll
Second	-0.221***	-0.013***	-0.031***	-0.022***	-0.353***	-0.044***	-0.027***	-0.019***
	(0.011)	(0.002)	(0.002)	(0.001)	(0.032)	(0.005)	(0.004)	(0.003)
Third	-0.228***	-0.006**	-0.035***	-0.029***	-0.374***	-0.048***	-0.026***	-0.020***
	(0.019)	(0.003)	(0.003)	(0.002)	(0.052)	(0.008)	(0.006)	(0.005)
Fourth	-0.220***	0.004	-0.043***	-0.033***	-0.404***	-0.054***	-0.031***	-0.025***
	(0.027)	(0.004)	(0.004)	(0.003)	(0.074)	(0.011)	(0.009)	(0.007)
Fifth+	-0.198***	0.018***	-0.047***	-0.037***	-0.469***	-0.057***	-0.039***	-0.029***
	(0.037)	(0.005)	(0.006)	(0.004)	(0.099)	(0.015)	(0.012)	(0.009)
Observations	1032555	1032555	1032555	1032555	201512	201512	201512	201512
R^2	0.712	0.599	0.666	0.639	0.706	0.648	0.660	0.646
Adjusted R^2	0.470	0.261	0.384	0.335	0.451	0.341	0.364	0.339
Dep Var Mean	9.411	0.821	0.271	0.130	8.318	0.572	0.222	0.111

(b) Native-Born Blacks

	Non-South				South			
	(1) Yrs Ed	(2) 8th Grade	(3) HS	(4) Any Coll	(5) Yrs Ed	(6) 8th Grade	(7) HS	(8) Any Coll
Second	-0.077	0.004	-0.005	0.006	-0.225**	-0.031**	-0.003	0.002
	(0.164)	(0.024)	(0.020)	(0.015)	(0.101)	(0.013)	(0.009)	(0.006)
Third	-0.035	0.009	-0.009	0.004	-0.212	-0.033	-0.011	-0.004
	(0.271)	(0.040)	(0.033)	(0.024)	(0.160)	(0.021)	(0.013)	(0.010)
Fourth	-0.075	-0.012	-0.012	0.014	-0.291	-0.037	-0.019	-0.011
	(0.381)	(0.055)	(0.046)	(0.033)	(0.225)	(0.029)	(0.019)	(0.014)
Fifth+	0.388	0.091	0.013	0.022	-0.204	-0.023	-0.002	0.002
	(0.520)	(0.075)	(0.063)	(0.046)	(0.298)	(0.039)	(0.025)	(0.019)
Observations	7908	7908	7908	7908	30221	30221	30221	30221
R^2	0.665	0.631	0.611	0.592	0.596	0.567	0.554	0.547
Adjusted R^2	0.354	0.289	0.250	0.213	0.215	0.160	0.135	0.120
Dep Var Mean	8.112	0.610	0.188	0.086	5.839	0.288	0.088	0.046

Note: This table reports parameter estimates and standard errors (in parentheses) for regressions of the listed dependent variable on birth order dummies where the birth order measure only accounts for male siblings. All regressions include family fixed effects and dummy variables for year of birth. Standard errors are clustered at the family level. * $p<0.1$ ** $p<0.05$ *** $p<0.01$

Table A7: Birth Order relative to Female Siblings and Educational Attainment

(a) Native-Born Whites

	Non-South				South			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Yrs Ed	8th Grade	HS	Any Coll	Yrs Ed	8th Grade	HS	Any Coll
Second	-0.114***	-0.010***	-0.013***	-0.008***	-0.157***	-0.019***	-0.014***	-0.008**
	(0.014)	(0.002)	(0.002)	(0.002)	(0.037)	(0.006)	(0.005)	(0.003)
Third	-0.105***	-0.009***	-0.013***	-0.009***	-0.140***	-0.018**	-0.018***	-0.006
	(0.021)	(0.003)	(0.003)	(0.002)	(0.053)	(0.008)	(0.006)	(0.005)
Fourth	-0.153***	-0.011*	-0.021***	-0.018***	-0.184**	-0.018	-0.021*	-0.013
	(0.035)	(0.006)	(0.005)	(0.004)	(0.092)	(0.014)	(0.011)	(0.008)
Fifth+	0.035	-0.015	0.010	0.005	-0.458	-0.084**	-0.020	0.010
	(0.108)	(0.017)	(0.016)	(0.013)	(0.303)	(0.041)	(0.038)	(0.030)
Observations	1032555	1032555	1032555	1032555	201512	201512	201512	201512
R^2	0.712	0.599	0.666	0.639	0.706	0.647	0.660	0.646
Adjusted R^2	0.469	0.261	0.384	0.335	0.449	0.340	0.364	0.339
Dep Var Mean	9.411	0.821	0.271	0.130	8.318	0.572	0.222	0.111

(b) Native-Born Blacks

	Non-South				South			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Yrs Ed	8th Grade	HS	Any Coll	Yrs Ed	8th Grade	HS	Any Coll
Second	-0.057	-0.027	0.002	-0.014	-0.057	-0.015	0.000	0.003
	(0.204)	(0.029)	(0.024)	(0.018)	(0.112)	(0.015)	(0.009)	(0.007)
Third	-0.149	-0.022	0.019	-0.023	0.026	-0.009	-0.002	0.007
	(0.285)	(0.045)	(0.035)	(0.025)	(0.156)	(0.020)	(0.013)	(0.009)
Fourth	0.206	-0.010	0.038	-0.004	0.063	-0.021	0.014	0.026
	(0.511)	(0.077)	(0.059)	(0.042)	(0.265)	(0.035)	(0.022)	(0.016)
Fifth+	-0.061	0.027	0.198	-0.005	-0.388	-0.015	-0.051	0.010
	(1.713)	(0.305)	(0.143)	(0.031)	(0.725)	(0.093)	(0.065)	(0.054)
Observations	7908	7908	7908	7908	30221	30221	30221	30221
R^2	0.665	0.631	0.611	0.592	0.595	0.567	0.554	0.547
Adjusted R^2	0.354	0.288	0.250	0.213	0.214	0.160	0.134	0.120
Dep Var Mean	8.112	0.610	0.188	0.086	5.839	0.288	0.088	0.046

Note: This table reports parameter estimates and standard errors (in parentheses) for regressions of the listed dependent variable on birth order dummies where the birth order measure only accounts for female siblings. All regressions include family fixed effects and dummy variables for year of birth. Standard errors are clustered at the family level. * p<0.1 ** p<0.05 *** p<0.01

Appendix B: Effects of Compulsory Schooling Laws with Alternative Difference-in-Differences Estimators

A recent literature has shown that, in the presence of differential timing and heterogenous treatment effects, two way fixed effects (TWFE) regressions can produce biased estimates of treatment effects (see De Chaisemartin and d'Haultfoeuille (2020); Callaway and Sant'Anna (2021); Goodman-Bacon (2021); Sun and Abraham (2021), among others). While a variety of solutions have been proposed in the case of binary treatment, few estimators have examined solutions in the case of multilevel treatments such as compulsory schooling laws. One exception is Gardner (2021), who proposes a two-stage difference-in-differences estimator that can be used in the case of multilevel treatments. In the section below, we replicate the TWFE results of Clay et al. (2021) showing the relationship between compulsory schooling laws and completed years of schooling for cohorts born between 1883 and 1910. Next, focusing on the set of laws most relevant in each region, we present the results using Gardner's two-stage difference-in-differences estimator.

We begin with native-born white and black men ages 28-59 in the 1940 Census. Following Clay et al. (2021), we calculate each individual's year of birth as $yob = 1940 - age - 1$ since the 1940 Census took place in April. Based on state and year of birth, we link each individual to the number of years of schooling legally required, $SCHYRS$. Developed by Clay et al. (2021), this measure captures the effects of compulsory attendance laws, taking into account exemptions for child labor and years of schooling as well as continuation schools. We adjust the measure by Clay et al. (2021) to prohibit treatment reversibility: if a state enacts a law requiring 9 years of schooling, we consider all subsequent cohorts to have at least 9 years required. If switching from a low education requirement to a higher education requirement puts the treated cohorts on a differential trend, the initial switching point is the relevant treatment point (Goodman-Bacon, 2021). In this case, changes in compulsory schooling may influence social norms around schooling. One wouldn't expect these norms shifts to completely dissipate in response to legislative changes reducing years of schooling.

We then replicate Clay et al. (2021)'s two-way fixed effects regression:

$$y_{isc} = \alpha + \beta_1 SCHYRS^{1-5}_{isc} + \beta_2 SCHYRS^{6}_{isc} + \beta_3 SCHYRS^{7}_{isc} + \\ \beta_4 SCHYRS^{8}_{isc} + \beta_5 SCHYRS^{9}_{isc} + \gamma_s + \eta_{rc} + \epsilon_{isc} \quad \text{(B.4)}$$

where y_{isc} is completed years of schooling reported in 1940 and $SCHYRS^{1-5}_{isc}$, $SCHYRS^{6}_{isc}$, $SCHYRS^{7}_{isc}$, $SCHYRS^{8}_{isc}$, and $SCHYRS^{9}_{isc}$ are dummies representing a state requirement of 1-5, 6, 7, 8, and 9 or more years of schooling, respectively. γ_s is a set of state fixed effects and η_{rc} is a set of region-by-year of birth fixed effects. We run separate regressions by region (non-South vs South) and race (white and black).

The coefficients estimated by equation B.4 are shown in Table B1. Focusing first on whites outside the South (column (1) of panel (a)), our results are similar to those in Clay et al. (2021) and indicate no significant relationship between completed education and laws requiring 1-5 or 6 years of compulsory schooling, and a small relationship with laws requiring 7 years. This is expected because, as shown in Figure 1, the vast majority of students outside the South were attending school through eighth grade during this period. Furthermore, few non-Southern states required 1-5 or 6 years only over the period, so the estimates rely on only a few states. For example, only 3 non-Southern states (Delaware, Missouri, and Oklahoma) had no compulsory schooling laws for cohorts born in 1890. By the 1900 cohort, all states outside the South had required at least 6 years of schooling. Since laws requiring 8 or 9+ years of schooling have the greatest potential to increase schooling outside the South, in column (2) we re-run equation B.4 using only dummies for 8 and 9+ years of schooling required. We again see a strong relationship: relative to having fewer years of schooling required, laws requiring 8 or 9+ years are associated with an increase of 0.04 and 0.2 years of schooling, respectively.

As discussed above, the estimates in columns (1) and (2) may be subject to bias given the differential timing of laws across states. Goodman-Bacon (2021) shows that TWFE models make use of various comparison groups in estimating coefficients: treated versus never treated, earlier treated vs later treated, later treated vs earlier treated, and treated vs always treated. In the presence of dynamic treatment effects, the latter two comparisons introduce bias into the TWFE estimates. As a first step at removing bad comparisons, in column (3) we re-run column (2) excluding groups that are "always treated" (that is, have 8 years of required schooling at the beginning of the sample period) or birth years that are "always treated." The coefficients are similar in magnitude to those in column (2), indicating a limited bias from the inclusion of "always treated" groups and periods. However, comparisons between later and earlier treated groups are still included.

Finally, in column (4), we run the two-stage DiD model proposed by Gardner (2021). In the first stage, we perform the following regression for the *untreated* groups (i.e. those with less than 8 years of schooling required) :

$$y_{isc} = \alpha + \gamma_s + \eta_{rc} + \epsilon_{isc} \tag{B.5}$$

The generates uncontaminated estimates of the group and time fixed effects. Using these estimates, we then obtain residualized years of education for both the treated and untreated units in the sample: $\tilde{y}_{isc} = y_{isc} - \hat{\gamma}_s - \hat{\eta}_{rc}$. Note that since fixed effects are based on untreated units, we are not able to obtain residuals for time periods in which all units in the region are treated or for always treated states. These are omitted from the analysis. We then regress these residuals on treatment dummies:

$$\tilde{y}_{isc} = \omega + \phi_1 SCHYRS^8_{isc} + \phi_2 SCHYRS^9_{isc} + \mu_{isc} \tag{B.6}$$

Since the dependent variable of equation B.6 is estimated in the first stage, we bootstrap to obtain appropriate standard errors. The coefficients (treatment effects) using the Gardner method are shown in column (4). The coefficients are substantially larger than those obtained by TWFE: laws requiring 8 and 9+ years increase completed education by 0.14 and 0.24 years, respectively. In column (1) of Table B2, we perform Gardner's 2SDiD estimate for the matched sample only. The effects of compulsory schooling laws are similar in the matched sample used for the birth order regressions. In columns (1)-(4) of Panel (b) in Table B1, we replicate the results for native-born blacks. We see no significant effects of compulsory schooling laws in most specifications, including the preferred 2SDiD specification (column (4)). Given the well-documented barriers to education for black individuals in this time period, this is not surprising. Additionally, the smaller sample size for native-born blacks living outside of the South reduces the precision of the estimates.

In columns (5)-(9), we replicate the results for Southern whites (panel (a)) and blacks (panel (b)). In the South, all states are untreated before cohorts born in 1898. Over the first few years of the twentieth century, all Southern states enact some compulsory schooling laws. By the time the first Southern state requires 7 years (1908 year of birth), all states have been treated and there are no valid comparison groups remaining. As a result, the analysis for Southern states focuses on laws requiring 1-5 and 6 years of compulsory schooling. The preferred 2SDiD estimates for Southern whites and blacks indicate no significant effect of compulsory schooling laws on completed educational attainment.

Table B1: Compulsory Schooling Laws and Completed Years of Education: Full 1940 Census Sample

(a) Native-Born Whites

	Non-South				South			
	(1) TWFE	(2) TWFE	(3) TWFE	(4) 2SDiD	(5) TWFE	(6) TWFE	(7) TWFE	(8) 2SDiD
1-5 Years Required	0.017				0.091***	0.013	0.107***	0.137
	(0.027)				(0.015)	(0.020)	(0.015)	(0.162)
6 Years Required	0.011				0.161***	0.006	0.098*	0.138
	(0.027)				(0.032)	(0.032)	(0.053)	(0.310)
7 Years Required	0.058**				0.212***			
	(0.028)				(0.052)			
8 Years Required	0.081***	0.041***	0.066***	0.136**	0.309***			
	(0.031)	(0.011)	(0.013)	(0.058)	(0.085)			
9+ Years Required	0.198***	0.151***	0.127***	0.240***	0.000			
	(0.035)	(0.017)	(0.021)	(0.079)	(.)			
Observations	14311278	14311278	11371304	11371304	3972004	3972004	3004343	3004343

(b) Native-Born Blacks

	Non-South				South			
	(1) TWFE	(2) TWFE	(3) TWFE	(4) 2SDiD	(5) TWFE	(6) TWFE	(7) TWFE	(8) 2SDiD
1-5 Years Required	0.021				0.093***	0.019	0.142***	0.131
	(0.042)				(0.022)	(0.026)	(0.028)	(0.137)
6 Years Required	0.020				0.215***	0.066	0.167***	0.118
	(0.043)				(0.052)	(0.040)	(0.051)	(0.357)
7 Years Required	-0.002				0.162***			
	(0.044)				(0.060)			
8 Years Required	0.025	0.017	-0.007	0.024	0.296***			
	(0.052)	(0.026)	(0.028)	(0.152)	(0.082)			
9+ Years Required	0.088	0.083	0.094	0.079	0.000			
	(0.071)	(0.050)	(0.074)	(0.221)	(.)			
Observations	301360	301360	252744	252744	1734823	1734823	1316657	1316657

Note: This table displays parameter estimates and standard errors (in parentheses) for the specification in equation (B.4) for columns (1)-(3) and (5)-(7) and in equation (B.6) for columns (4) and (8). Columns (3), (4), (7), and (8) exclude "always treated" states and years. The dependent variable is completed years of schooling as reported in the 1940 Census. The sample consists of all native-born white and black men ages 28-59 in the 1940 Census. Standard errors are clustered at the state-by-cohort level. In columns (4) and (8), standard errors are obtained using bootstrap (50 iterations) as recommended for large datasets with many fixed effects by Butts (2022). * p<0.1 ** p<0.05 *** p<0.01

Table B2: Compulsory Schooling Laws and Completed Years of Education: Matched Regression Sample, 2SDiD

	(1) White Non-South	(2) White South	(3) Black Non-South	(4) Black South
8 Years Required	0.117**		-0.112	
	(0.058)		(0.262)	
9+ Years Required	0.244***		0.112	
	(0.072)		(0.411)	
1-5 Years Required		0.187		0.197
		(0.124)		(0.191)
6 Years Required		0.233		0.164
		(0.280)		(0.481)
Observations	798321	174027	6549	26361

Note: This table displays regression coefficients based on the specification in equation (B.6). The dependent variable is completed years of schooling as reported in the 1940 Census. The sample consists of all native-born white and black men ages 28-59 in the 1940 Census who are matched to childhood records in 1900 or 1910 and have at least one brother who is also matched. Standard errors are clustered at the state-by-cohort level. Standard errors are obtained using bootstrap (100 iterations) as recommended for large datasets with many fixed effects by Butts (2022). * $p<0.1$ ** $p<0.05$ *** $p<0.01$