

NBER WORKING PAPER SERIES

FETAL SHOCK OR SELECTION? THE 1918 INFLUENZA PANDEMIC AND HUMAN
CAPITAL DEVELOPMENT

Brian Beach
Joseph P. Ferrie
Martin H. Saavedra

Working Paper 24725
<http://www.nber.org/papers/w24725>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2018, Revised April 2021

We are grateful for feedback from Douglas Almond, Alan Barreca, Ryan Brown, Walker Hanlon, John Parman, Nick Sanders, Duncan Thomas, and Nic Ziebarth. We are also thank participants at the 2017 NBER Summer Institute Development of the American Economy meeting and the 2018 World Economic History Congress. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Brian Beach, Joseph P. Ferrie, and Martin H. Saavedra. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Fetal Shock or Selection? The 1918 Influenza Pandemic and Human Capital Development
Brian Beach, Joseph P. Ferrie, and Martin H. Saavedra
NBER Working Paper No. 24725
June 2018, Revised April 2021
JEL No. I1,J0,N12

ABSTRACT

Almond (2006) argues that in-utero exposure to the 1918 influenza pandemic lowered socioeconomic status in adulthood, whereas subsequent work has argued that exposed cohorts may have been selected. We bring new evidence on the lasting impact of in-utero exposure to the pandemic. Linking census microdata to WWII enlistment records and city-level influenza data allows us to adopt an empirical approach that exploits pandemic intensity as a source of identifying variation. We show that pandemic intensity is less related to parental characteristics, suggesting this approach can more credibly be interpreted as causal. Our results indicate that in-utero exposure to the pandemic lowered high school graduation rates.

Brian Beach
Department of Economics
Vanderbilt University
VU Station B, Box #351819
2301 Vanderbilt Place
Nashville, TN 37235
and NBER
brian.beach@vanderbilt.edu

Martin H. Saavedra
Department of Economics
Oberlin College
Department of Economics, Rice Hall 233
10 N. Professor St
Oberlin, OH 44074
Martin.Saavedra@oberlin.edu

Joseph P. Ferrie
Department of Economics
Northwestern University
2001 Sheridan Road
Evanston, IL 60208-2600
and NBER
ferrie@northwestern.edu

1 Introduction

In a seminal paper, [Almond \(2006\)](#) argued that the 1918 influenza pandemic offered a unique opportunity to examine whether shocks to the in-utero environment have lasting consequences for health and human capital accumulation. The pandemic was severe, unexpected, but ultimately temporary. [Almond](#)'s cross-cohort approach tries to isolate the in-utero effect by comparing outcomes from the 1919 birth cohort, whose mothers had the highest probability of contracting influenza during pregnancy, to outcomes from adjacent cohorts (1912-1922). [Almond](#)'s dose-response approach uses state-by-year maternal mortality rates as a proxy for intensity of exposure to the pandemic to make within-cohort comparisons. Results from both approaches indicate that males from more exposed birth cohorts completed fewer years of education and earned less as adults.

[Almond \(2006\)](#) is often cited as definitive and causal evidence on the lasting effects of in-utero exposure to the pandemic, but a re-evaluation of the evidence is motivated by two factors. First, the 1919 birth cohort may have had systematically different parents because the pandemic corresponded with the height of US WWI deployment. To this point, [Brown & Thomas \(2021\)](#) apply the cross-cohort methodology to the 1920 census (taking background characteristics as the variable of interest) and find evidence that the 1919 birth cohort was born into families that were measurably different. Second, there is evidence that fertility patterns were affected by both the war and the pandemic ([Chandra *et al.*, 2018](#); [Kitchens & Rodgers, 2020](#); [Mamelund, 2004](#); [Vandenbroucke, 2014](#)). These findings raise the possibility that parents that conceived children during or after the pandemic may have been different from parents who conceived before the pandemic.¹ While the dose-response analysis in [Almond \(2006\)](#) addresses these concerns to some extent, they are based on data from 19 states and imprecisely estimated for some years. [Almond](#) finds statistically significant negative effects for males in the 1960 census, but the “1980 estimates are about half as large as the corresponding 1960 effects” and the 1970 census estimates “appear different” and “do not approach statistical significance” ([Almond, 2006](#), p. 706).

This paper uses linked data to provide new evidence on the consequences of in-utero exposure to the 1918 influenza pandemic. Our efforts are made possible by the digitization of the entire 1920 and 1930 U.S. censuses, which allow us to construct an individual-level panel data set by linking World War II enlistment records back to the census.

¹See [Beach *et al.* \(2020\)](#) for a review of the economic effects of the 1918 influenza pandemic.

Linking makes a trade-off between population representativeness and precision. Relative to [Almond \(2006\)](#), our sample only contains males, is slightly more likely to be white and slightly more likely to have foreign born parents (see [Table 1](#)). The main advantage of our linked approach is that it allows us to infer the in-utero environment based on where the individual resided at the time of census enumeration. This allows us to improve upon [Almond's](#) dose-response framework by constructing a measure of pandemic exposure for nearly 300 cities. The second advantage of using linked data is that we observe individuals with their parents and are able to control for a comprehensive set of background characteristics.

The cross-cohort section of our paper has two main results. First, as in [Brown & Thomas \(2021\)](#), we find evidence that the 1919 birth cohort was born into families that were measurably different. We show these patterns in [Almond's](#) 1960 sample, where we find that the 1919 birth cohort was more likely to be nonwhite and more likely to be born in the South, but they were no more or less likely to have foreign born parents. In our linked data set, nearly all of our 15 background characteristics are measurably different for the 1919 birth cohort. These findings support [Brown & Thomas \(2021\)](#) by providing evidence of parental selection that persists after conditioning on surviving to adulthood. Our second finding is that failing to control for these differences leads one to overestimate the impact of in-utero exposure on educational attainment. Estimates of the 1919 effect are 40-60% smaller after controlling for background differences, but the results remain negative and statistically significant. For a subsample of our data we are able to make comparisons between brothers by employing household fixed effects. Those results again indicate that the 1919 birth cohort was less educated, although the estimates are less precisely estimated.

The dose-response section of our paper also has two main results. First, we show that the correlation between background characteristics and regional intensity of the pandemic is weak. This suggests that a within-cohort or dose-response framework is more likely to recover the causal effect of in-utero exposure to the pandemic. Second, we consistently estimate a negative relationship between pandemic intensity and educational attainment. This effect appears to operate on the dimension of high school completion. These patterns hold when we restrict comparisons to just the 1918 versus 1919 birth cohorts as well as when we examine our sample of brothers and employ household fixed effects. Our estimates suggest that those facing the average level of pandemic intensity were 1-1.5 percentage points less likely to graduate high school.

We interpret the collection of results as follows. Our cross-cohort results suggest that the

qualitative claim that men from the 1919 birth cohort were less educated is robust to accounting for background differences but the assumptions needed to interpret those estimates as causal are stronger than previously thought. We show that a dose-response framework offers a more credible identification strategy. Our dose-response results indicate that exposure to the pandemic lowered educational attainment. All of these results were generated from a relatively selected sample (male WWII enlistees that could be successfully linked back to the census). Whether the effects we document extend to women or those born in areas without a record of pandemic intensity remains an open question. Once the 1960, 1970, and 1980 censuses become publicly available, future researchers should adopt a linked-data approach that exploits the intensity of the pandemic to better assess the lasting consequences of early-life exposure to the pandemic.

2 Constructing a Linked Data Set

The last publicly available data set with identifiable information about individuals from the 1919 and adjacent birth cohorts is the WWII enlistment records.² The WWII enlistment records contain information about the individual’s year and place of birth as well as their educational attainment, height (a common proxy for health (Costa, 2015, Sec. 3.2)), and occupation at the time of enlistment. We link men from these records that were born between 1912 and 1922 back to the 1920 or 1930 censuses.³ This creates a panel data set that allows us to examine parental selection and adult outcomes at the individual level.

Unique identifiers are not available in the census and so we overcome this issue by focusing on men who can be uniquely identified by their place of birth, first name, last name, and age. Our linking procedure builds upon earlier work (Long & Ferrie, 2013; Beach *et al.*, 2016) and follows the best practices discussed by Bailey *et al.* (Forthcoming) and Abramitzky *et al.* (Forthcoming). We first standardize all given names (e.g., recoding “Ed” and “Eddie” as “Edward”) in both data sets.⁴ After this standardization, we take each enlistment record and identify any and all census

²IPUMS.org and Ancestry.com have digitized and disseminated full count census data from 1940 and earlier. Deanonymized records from 1950 and later are not yet publicly available. WWII enlistment records were digitized and made available by the National Archives and Records Administration.

³Linking to the 1930 census allows us to maintain comparability with Almond (2006) by studying the 1912-1922 birth cohorts. In our dose-response analysis, however, we assume that the place of enumeration and the in-utero environment were the same. Since the 1930 census was enumerated when the relevant cohorts were between the ages of 7 and 18, this seems like a strong assumption. The 1920 census, however, was enumerated when the 1919 birth cohort was less than 1 year old, which makes this assumption much more believable.

⁴Standardization reduces the likelihood that we discard any individual because they used a nickname in one record and their legal name in another. Standardization also accounts for incorrect spelling variants (e.g., “Elliott” and “Elliot”) that appear because 1920 and 1930 enumerators recorded information that was spoken to them.

records where the individual is of the same race, born in the same state, born around the same time (plus or minus three years), and has a reasonably close name. Names are classified as reasonably close if: 1) the standardized first name initial matches, 2) the last name initial matches, and 3) the Jaro-Winkler string distance between the raw first names is between 0.8 and 1, and 4) the Jaro-Winkler string distance between the raw last names is between 0.8 and 1.⁵ The use of the Jaro-Winkler string distance allows us to relax the “exact name match” criteria and accommodate spelling variants among last names (e.g., Andersen and Anderson) and any minor transcription errors that may have occurred during the process of digitizing the original handwritten records. A successful enlistment-to-census link is one where only one census record satisfies the above criteria.⁶

Among the successful enlistment-to-census links, we restrict our attention to the subset of those links whose age is consistently reported across the two sources. While the enlistment records asked individuals to report their birth year, birth year in the 1920 and 1930 censuses has to be inferred from the reported age. The 1920 census was enumerated as of January 1, 1920, and so when linking to the 1920 census we require the inferred birth year in 1920 to match the reported birth year in the enlistment records. The 1930 census was enumerated as of April 1, 1930, and so when linking to the 1930 census the inferred birth year must match the enlistment records or be off by 1 year. Requiring consistency helps alleviate concerns about intentional or unintentional misreporting.⁷

Our sample may not be representative for two reasons. First, WWII enlistees were a non-random subset of the population. One direct selection mechanism is that registrants could be rejected for failing to meet the minimum education or physical standards. Second, those who consistently report their names and birth years are more likely to be linked to their childhood record, although this will be true for both the 1919 birth cohort and adjacent birth cohorts. Appendix Table

⁵The Jaro-Winkler string distance imposes a penalty for the number of character changes to move from one string to the other. A Jaro-Winkler string distance of 1 implies that the two strings are an exact match.

⁶We assess the false positive rate of our algorithm by modifying records in the full census sample to incorporate the types of spelling errors, transcription errors, and misreporting of birth years outlined above and in [Goeken *et al.* \(2017\)](#). Using our algorithm to link from the original census to the modified census yields a successful match rate of 36.5 percent and a false positive rate of 1.8 percent.

⁷Some young men intentionally misreported their birth year in order to meet the minimum age requirement for enlistment. This is unlikely to be an issue in our setting, as most enlistment occurred between 1941 and 1944, when the youngest cohort (1922) would have been between 19 and 22 years old. Moreover, while intentional misreporting complicates our ability to obtain a unique link in the enlistment records, those men would not appear in our sample unless their parents misreported the child’s age in 1920 or 1930 in the exact same way. Related, the age heaping phenomenon (i.e., that innumerate parents are more likely to report round or even ages) may affect our ability to link to the census, but is unlikely to affect our sample since the enlistment records asked for birth year. Appendix Figure 3 plots histograms for our linked samples. Younger cohorts are slightly over represented, but there is little evidence to indicate that misreporting is a major issue. The 1921 cohort is slightly larger than adjacent cohorts, but the 1960-1980 censuses reflect this same pattern.

5 compares WWII veterans and non-veterans in the 1950 census. WWII veterans were younger, whiter, and slightly less likely to have U.S.-born parents than non-veterans.

We also make comparisons among brothers. Since the enlistment records do not identify brothers, we infer that two enlistees are brothers if they are observed in the same household when we find each enlistee in the census. This sample is clearly selected relative to the broader population as well as our linked sample. Nevertheless, with this subsample we are able to include household fixed effects, which allows us to rule out whether estimates are explained by parental selection.

3 Revisiting the Cross-Cohort Approach

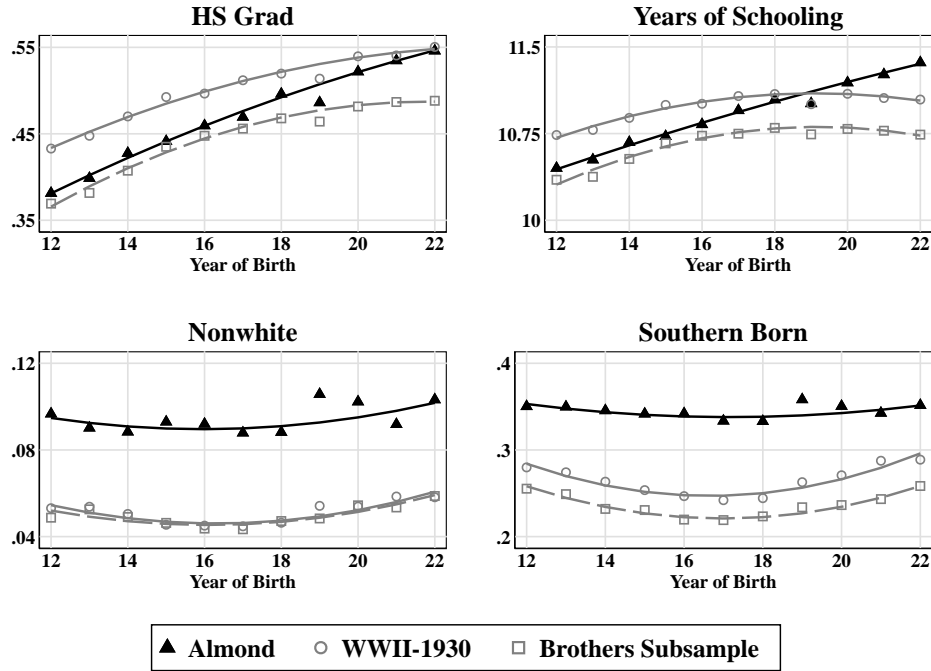
This section adopts variations of Almond’s cross-cohort specification to examine whether parents of the 1919 birth cohort were differentially selected and whether those differences affect estimates of deviation of the 1919 cohort relative to underlying trends. The estimating equation is:

$$y_i = \alpha + \beta \times YOB + \gamma \times YOB^2 + \delta \mathbf{1}[YOB = 1919] + \epsilon_i \quad (1)$$

where y_i is either a background characteristic or adult outcome for individual i . YOB denotes year of birth and δ measures the departure for the 1919 birth cohort from the quadratic trend.

Figure 1 visualizes this approach for three samples by plotting cohort-level averages for the 1912 to 1922 birth cohorts, along with a quadratic cohort trend. The first sample is the 1% 1960 census examined in Almond (2006). The second sample comes from the 1930 census linked to WWII enlistment records. The third sample is the 1930-WWII links restricted to individuals who have a brother in 1930 that was also linked to the enlistment records. The top panels show that for all three samples, the 1919 birth cohort is less likely to graduate from high school and attained fewer years of schooling than the cross-cohort trends would predict. The bottom two panels use the same data but display two background characteristics available in all three samples: whether the individual was nonwhite and whether the individual was born in the South. The 1919 and 1920 birth cohorts are more likely to be nonwhite than trend would predict in the Almond sample. The same is true for the 1919 cohort in the 1930-WWII links, but the effect size is considerably smaller. There is no detectable selection for race in the brothers sample. For all three samples, the 1919 birth cohort was more likely to be born in the South, although the amount of selection is smaller for the 1930 to WWII links than the 1960 census.

Figure 1: Comparing our linked samples to Almond (2006)



3.1 Assessing Family Selection

Table 1 examines whether the 1919 birth cohort appears differentially selected relative to adjacent birth cohorts. Each row corresponds to a parental characteristic.⁸ Column 1 reports the mean and standard deviation for men from the 1912-1922 birth cohorts in the 1% sample of the 1960 census. Column 2 reports the deviation of the 1919 birth cohort from underlying trends, i.e., estimates of δ from equation 1. Columns 3 and 4 are analogous to columns 1 and 2 except that the sample is now men from the 1912-1922 birth cohorts that we link from WWII enlistment records to the 1930 census. Columns 5 and 6 correspond to the sample of men from the 1912-1919 birth cohorts that we can link from WWII enlistment records to the 1920 census.

⁸Note that parental characteristics are observed when the children are at different ages. Appendix 8 asks whether this is an issue by examining patterns of parental selection for a double linked sample, i.e., those that we observe in both 1920 and 1930. Aside from fertility variables, age at observation does not appear to be an issue. The problem with fertility variables (e.g., number of older or younger siblings in the household) is that older siblings may die or leave the household before enumeration while, particularly for younger cohorts, the younger siblings variable is mechanically censored because any younger siblings will be born after enumeration.

Table 1: Assessing background differences for the 1919 birth cohort

Dependent Variable	1912-1922 Birth Cohorts (1960 Census)		1912-1922 Birth Cohorts (WWII-1930 Links)		1912-1919 Birth Cohorts (WWII-1920 Links)	
	Mean [S.D.]	1919 Deviation From Trend	Mean [S.D.]	1919 Deviation From Trend	Mean [S.D.]	1919 Deviation From Trend
	(1)	(2)	(3)	(4)	(5)	(6)
Southern born	0.345 [0.475]	0.018*** (0.005)	0.267 [0.442]	0.006*** (0.002)	0.258 [0.437]	0.015*** (0.003)
Nonwhite	0.095 [0.293]	0.013*** (0.003)	0.052 [0.222]	0.005*** (0.001)	0.038 [0.191]	0.003** (0.001)
Father born outside U.S.	0.223 [0.416]	0.000 (0.004)	0.279 [0.448]	0.006*** (0.002)	0.248 [0.432]	0.010*** (0.003)
Mother born outside U.S.	0.198 [0.399]	0.000 (0.004)	0.241 [0.428]	0.005*** (0.002)	0.213 [0.409]	0.011*** (0.003)
Father is second-gen. immigrant			0.189 [0.392]	-0.005*** (0.001)	0.201 [0.401]	-0.010*** (0.003)
Father moved outside of birthplace			0.464 [0.499]	0.006*** (0.002)	0.458 [0.498]	0.008** (0.003)
Father's age when child was born			32.206 [7.319]	0.366*** (0.026)	32.765 [7.575]	0.489*** (0.051)
Mother's age when child was born			27.843 [6.234]	0.312*** (0.022)	28.312 [6.269]	0.331*** (0.042)
Father's age minus mother's age			4.364 [5.004]	0.054*** (0.018)	4.453 [5.160]	0.158*** (0.035)
Father cannot read or write			0.057 [0.231]	0.004*** (0.001)	0.048 [0.214]	0.003** (0.001)
Mother cannot read or write			0.060 [0.237]	0.003*** (0.001)	0.051 [0.219]	0.003* (0.001)
Father's occupational income score			23.995 [10.450]	-0.194*** (0.040)	23.220 [10.437]	0.020 (0.077)
Family owns home			0.563 [0.496]	-0.000 (0.002)	0.428 [0.495]	0.003 (0.003)
Child's inferred birth order			2.482 [1.495]	0.100*** (0.005)	2.904 [1.909]	0.147*** (0.013)
Mother is in labor force			0.067 [0.249]	0.000 (0.001)	0.035 [0.184]	-0.004*** (0.001)

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Each regression includes the reported "Born in 1919" indicator variable, as well as year of birth and year of birth squared. Robust standard errors are reported in parentheses.

The estimates from Table 1 suggest that the 1919 birth cohort came from families that were measurably different. In the 1960 sample, we see that the 1919 birth cohort was more likely to be nonwhite and born in the South but was no more or less likely to have foreign born parents. Our linked samples also indicate that the 1919 birth cohort was more likely to be nonwhite and born in the South, although the extent of the selection among the 1919 birth cohort is smaller. Our linked samples also provide evidence that the 1919 birth cohort was more likely to have foreign born parents. While this exhausts the background measures available in the 1960 census, we are able to examine a more comprehensive set of background characteristics for our linked samples. 18

of the 22 additional regressions we run are statistically significant at conventional levels.

Table 1 also provides a look at the underlying differences between our linked sample relative to the sample analyzed in Almond (2006). Roughly 25% of our linked sample was born in the South, as compared to 35% in Almond (2006). As for nonwhite share, Almond’s sample was roughly 90% white whereas our samples are roughly 95% white. Our linked sample is also more likely to have foreign born parents; roughly 26% of our links have a foreign born father and 22% of our links have a foreign born mother, while the corresponding numbers for Almond’s sample are 22% and 20%, respectively.

3.2 Long-Run Estimates After Accounting for Family Selection

Table 2 estimates the deviation in educational attainment for the 1919 birth cohort before and after accounting for parental background differences. Each panel corresponds to a different outcome: total years of schooling, whether the individual completed some high school, and whether the individual graduated high school.⁹ We generate results using the 1960 census sample of men (columns 1 and 2), our sample of male enlistees linked back to the 1930 census (columns 3-4), the brothers subsample of our WWII to 1930 links (columns 5 and 6), and our sample of male enlistees linked back to the 1920 census (columns 7-8). Columns 1, 3, 5, and 7 correspond to baseline estimates, which include our Born-in-1919 indicator variable, as well as a quadratic time trend. Columns 2, 4, 6, and 8 add all observable background characteristics, and in our brothers subsample (column 6) we also include household fixed effects.

Table 2 shows that estimates of the 1919 effect are overstated when background differences are unaccounted for. In general, point estimates fall by 40% or more once background controls are added to the model. While the magnitude of the decline may seem large, it is not uncommon for the new estimate to be contained within the confidence interval of the baseline estimate and vice versa. In terms of qualitative conclusions, 11 of the baseline estimates are negative and significant at the 1% level. Once controls are added, 7 of the 11 remain negative and statistically significant at the 1% level; one remains negative and statistically significant at the 5% level, two remain negative and statistically significant at the 10% level, and one remains negative but is no longer statistically significant (the impact on high school graduation after including household fixed effects). Our

⁹WWII enlistees that did not complete at least one year of high school are coded as having a grammar school education or less (i.e., 8 years of schooling) in the enlistment data. To maintain comparability across samples, we re-code the 1960 census educational attainment variable to adopt this same classification.

analysis points to relatively precise but null effects on height. The estimates are negative and significant for the WWII linked to 1930 sample, but statistically equal to zero for the other two samples. Even when significant, the points estimates are small, suggesting the pandemic lowered enlistee heights by approximately 3 one hundredths of an inch.

Table 2: Differences in adult outcomes of 1919 birth cohort relative to surrounding cohorts

	1960 Census		WWII-1930 Links		WWII-1930 Links Brothers Sample		WWII-1920 Links	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: DV is Years of Schooling								
Born in 1919	-0.109*** (0.030)	-0.070** (0.028)	-0.094*** (0.008)	-0.054*** (0.007)	-0.065*** (0.019)	-0.028* (0.016)	-0.110*** (0.016)	-0.066*** (0.014)
Panel B: DV is Some High School								
Born in 1919	-0.019*** (0.005)	-0.013*** (0.005)	-0.011*** (0.001)	-0.006*** (0.001)	-0.005 (0.004)	-0.002 (0.004)	-0.018*** (0.003)	-0.010*** (0.003)
Panel C: DV is High School Graduation								
Born in 1919	-0.021*** (0.005)	-0.015*** (0.005)	-0.017*** (0.002)	-0.009*** (0.001)	-0.013*** (0.004)	-0.005 (0.004)	-0.015*** (0.003)	-0.006* (0.003)
Panel D: DV is Height at Enlistment								
Born in 1919			-0.035*** (0.010)	-0.023** (0.010)	0.006 (0.029)	0.022 (0.027)	-0.022 (0.020)	-0.011 (0.020)
Limited Controls		Y		Y		Y		Y
Other Controls				Y		Y		Y
Household Fixed Effects						Y		

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Regressions also include a quadratic time trend. Robust standard errors reported in parentheses. Limited controls: race, birth state, father's country of birth, and mother's country of birth fixed effects. Other controls: whether the father is a second generation immigrant, whether the father moved outside their birthplace before the child was born, father's age when the child was born, whether the father is literate, fixed effects for the father's occupation, whether the family owns their home, the number of older siblings observed in the household (a proxy for birth order), the difference between the age of the father and the age of the mother, mother's age when the child was born, whether the mother is literate, and whether the mother is in the labor force.

The results in Table 2 require a nuanced interpretation. One interpretation is that the 1919 birth cohort was less educated and those results are robust to accounting for background differences.¹⁰ Support for this interpretation comes from our linked data, which either accounts for 15 comprehensive background measures or through the inclusion of household fixed effects. A second interpretation is that the selection documented in Table 1 raises concerns about the existence of

¹⁰Since the 1960 results have a more limited set of controls, it is an open question as to what the results in Column 2 would look like if the 1960 census included the same comprehensive background measures. Appendix Table 6 sheds light on this by reweighting our 1930 links to look like the 1960 census, within each year of birth, southern born, father is foreign born, and mother is foreign born cell. Those results indicate that background controls reduce the point estimates by about 50%. The point estimates are statistically differentiable, but the point estimates remain negative and statistically significant at the 1% level.

unobservable differences, which threatens the credibility of the identification strategy.

[Brown & Thomas \(2021\)](#) offers a natural point of comparison for the results presented thus far. [Brown & Thomas](#) document parental selection in the 1920 census. The patterns of selection are similar to what we find in [Table 1](#). Their approach to correcting for background differences is to control for birth-state-by-race-by-birth-year averages, as observed in the 1920 census. This allows them to examine the same set of outcomes and samples analyzed in [Almond \(2006\)](#). By drawing on 1920 data, the authors are unable to include the 1920-1922 birth cohorts in their analysis, which tends to double the size of their standard errors. The controls are also less precisely measured than ours. The inclusion of these controls tends to produce results that are negative but smaller and statistically insignificant.¹¹ While the controlled confidence intervals often contain zero, the confidence intervals also overlap substantially with the uncontrolled specifications, making it difficult to reject whether the attenuation is statistically distinguishable from zero.

4 Revisiting the Dose-Response Approach

A dose-response framework offers a viable path for overcoming cohort-level selection. While the cross-cohort framework asks whether individuals born in 1919 performed worse than individuals from adjacent birth cohorts, the dose-response framework goes further by asking whether the impact was larger for individuals from areas where the pandemic was more widespread. This additional comparison narrows the set of threats to identification, as the identifying assumption is that other key factors of long-run outcomes do not vary with the intensity of pandemic exposure.

[Almond \(2006, Section VI\)](#) adopts a dose-response framework to examine the impact of in-utero pandemic exposure. [Almond](#) relies on state-level variation in pandemic intensity, as all censuses record an individual’s state of birth, and so this is the finest geographic level in which one can infer the in-utero environment. [Almond](#) constructs a measure of pandemic intensity based on annual state-level maternal mortality information for 19 states. He then adopts an empirical framework that includes state-of-birth and year-of-birth fixed effects. These fixed effects help account for average differences in parental characteristics across cohorts. Lastly, [Almond](#) restricts attention to

¹¹The preferred specification in [Brown & Thomas \(2021\)](#) shows that the cross-cohort estimates of the 1919 effect flip sign once background controls were included. After working with the authors to understand why our findings were at odds, it appears that the sign flip is driven by the use of the family size variable “number of father’s children in the household”. This variable reports the number of older and younger siblings in the household at the time of enumeration. Unfortunately, the number of younger siblings component of this variable is censored for young cohorts (see [Appendix 8](#) for more). If one uses a more age-invariant measure of family size (i.e., the number of older children in the household) the sign flips tend to go away. A replication of these findings appears in [Appendix Section 7](#).

the 1918 through 1920 birth cohorts in order to isolate the effect of fetal exposure.

The existing dose-response evidence is mixed. [Almond](#) finds statistically significant and negative effects for males in the 1960 census, but he also notes the effects in 1980 are smaller and the effects in 1970 are not statistically significant. [Brown & Thomas \(2021\)](#) replicate this analysis and present results for males, females, and nonwhites.¹² Only 2 of the 45 estimates they present show a negative and statistically significant impact on long-run outcomes. These results highlight the need for new evidence, in turn motivating the remainder of our paper, which generates new dose-response evidence leveraging variation in pandemic intensity across nearly 300 cities.

4.1 An Improved Measure of Exposure

Our linked data allow us to derive a more localized measure of pandemic exposure. We assume that the city of enumeration in 1920 is the same as the individual’s in-utero environment. We then construct a measure of pandemic intensity based on city-level influenza deaths, which were systematically collected and published in the Census Bureau’s “Mortality Statistics” publications starting in 1900. Although the data only include information for Registration states and cities, this data source allows us to leverage variation from nearly 300 cities.¹³

Influenza mortality data provide a rich set of variation, but one concern is that influenza mortality will capture more than just the severity of the pandemic.¹⁴ [Clay *et al.* \(2018\)](#) show that during the pandemic, mortality rates were higher in places with more coal pollution and worse water quality. These relationships are attributable to the fact that air pollution and poor water quality compromise an individual’s immune system, making them more susceptible to influenza. Thus, observing high influenza mortality rates in 1918 could mean that a city was hit relatively hard by the pandemic, that a city had relatively worse water and air quality, or some interaction of the two. This is concerning since early-life exposure to air pollution ([Sanders, 2012](#); [Isen *et al.*, 2017](#)) or poor water quality ([Beach *et al.*, 2016](#)) also impairs human capital development.

¹²[Brown & Thomas \(2021\)](#) also correct a data transcription error for Virginia and add data for Washington, D.C.

¹³Registration states and cities are those with laws requiring that mortality statistics be collected. In 1900 the Census Bureau worked with those areas to establish uniform reporting standards. The result of this was the adoption of a standardized death certificate and the international classification standard, as well as the distribution of “The Manual of International Classification of Causes of Death,” which cross referenced terms appearing in causes of death from 1890 and 1900 reports with the new uniform classification standard.

¹⁴Appendix Figure 4 shows that influenza fatality rates track the case rates reasonably well on a weekly basis but with a bit of a lag. This lag is unlikely to matter for our analysis, which uses an annual measure of mortality. The underlying data come from the Public Health Service, which went door-to-door collecting morbidity and mortality information for 12 cities in the fall of 1918.

Our solution is to generate a counterfactual estimate of influenza mortality in 1918. To do so, we transcribe all city-level mortality statistics spanning 1900-1930 from the annual Mortality Statistics reports. We then run a series of city-level regressions where we restrict the sample to the 1900-1917 period and regress $\ln(\text{influenza deaths})$ on a city-specific linear time trend.¹⁵ Taking the exponential of the predicted values from this regression yields a prediction of influenza fatalities in the absence of the pandemic for post-1917 years.¹⁶ Subtracting predicted influenza deaths in 1918 from actual influenza deaths in 1918 gives us the unanticipated increase in influenza mortality due to the pandemic. Our options to normalize this measure are to divide by population or to divide by predicted influenza deaths. While the two numbers are correlated, dividing by population ignores the fact that cities of similar sizes may have different underlying disease and pollution environments. Because of this, and also because accurate population data are only available in census years, we use predicted influenza deaths as our denominator. Mechanically this measure is simply the ratio of unexpected influenza deaths occurring in 1918 relative to the number of expected influenza deaths in 1918, where that expectation captures underlying trends in population growth and intrinsic differences in disease and pollution environments.

Appendix Figure 5 plots the average excess influenza ratio by year. The figure spans 1900-1917 (the sample period) as well as 1918 through 1930 (our out-of-sample predictions).¹⁷ There we see the severe and temporary nature of the 1918 pandemic as well as the overall fit of our model. Excess influenza remains close to zero until 1918, during which approximately 35 influenza deaths occurred for every expected influenza death. Influenza deaths are slightly higher than expected during the 1920s, but by 1921 influenza appears to have reached a new steady state and even 13 years after 1917, excess influenza deaths are not far above zero.

4.2 Empirical Approach

Figure 2 displays patterns in the raw data. The top panels correspond to adult outcomes: high school graduation and total years of schooling. The bottom panels correspond to background characteristics: whether the individual is nonwhite and whether the individual was born in the

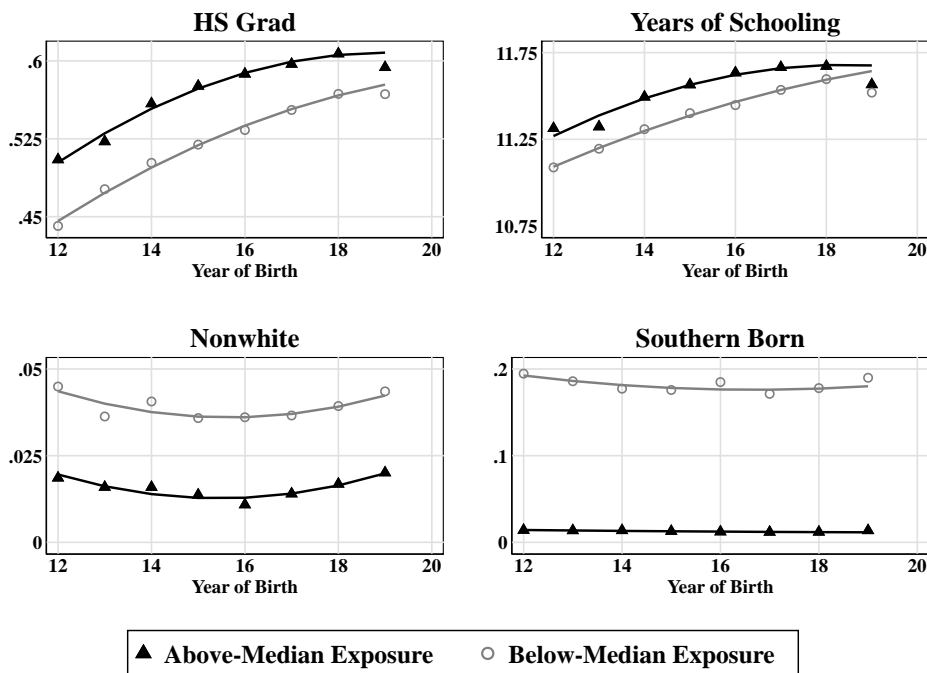
¹⁵We only run these regressions for the 287 cities that appear in every report.

¹⁶The natural logarithm ensures that predicted influenza deaths is always greater than zero.

¹⁷Alternatively, we could have used data from the 1900-1917 and 1920-1930 years, omitting the years during which the pandemic occurred. However, if the pandemic had lingering effects on influenza rates or city population counts, then data from the 1920-1930 period may be endogenous.

South.¹⁸ We plot cohort-level averages as well as underlying trends.¹⁹ This information is displayed separately for those with above-median versus below-median exposure to the pandemic.

Figure 2: Raw patterns in the dose-response data



Notes: High exposure cities had an excess flu ratio greater than 28.2.

Figure 2 conveys two key points. First, while there are compositional differences between cohorts from above-median versus below-median exposure cities, underlying trends are similar. This increases our confidence in using “below-median exposure” cities as a counterfactual. Second, the ability to use members from the same cohort as a counterfactual allows us to difference out any 1919-specific changes (e.g., changes in the nonwhite share), which is useful for identification.

The below estimating equation offers the starting point for the remainder of our analysis:

$$y_{ibc} = \alpha_0 + \beta_b + \gamma_c + \delta \mathbf{1}[yob = 1919] \times \text{Excess Flu}_c + \epsilon_{ibc} \quad (2)$$

where y_{ibc} is outcome y of individual i from birth year b in birth city c (either a background

¹⁸Appendix Figure 6 maps the variation in our sample. While most of our data come from the Northeast, the key takeaway is that there is meaningful sub-state variation in our measure of flu intensity, as there are many situations where neighboring cities had different exposure to the pandemic.

¹⁹Trends are obtained by regressing each outcome on a quadratic cohort trend and a 1919 indicator. We then obtain the predicted values omitting the 1919 estimate.

characteristic or a long-run outcome). The parameters β_b and γ_c are birth year and birth city fixed effects, respectively. The variable Excess Flu_c measures pandemic intensity in city c in 1918. Our main measure is the ratio of total influenza deaths relative to expected influenza deaths, which we normalize by dividing by the sample mean (36.42). This normalization allows coefficients to be interpreted as the average effect of pandemic exposure.

The identifying assumption embedded in Equation 2 is that in the absence of the pandemic, changes in outcomes among cohorts with high exposure would have looked similar to what we observe among cohorts with low exposure. This assumption is not testable. However, it is common to use a generalized difference-in-differences or “event study” design to see if there are meaningful deviations prior to treatment. Those results, which are presented in Appendix Figure 7 show little evidence of a differential pre-trend. The point estimate for the 1918 cohort is negative but less pronounced than the point estimate for the 1919 cohort. However, the 1918 cohort is not a clear placebo cohort as anyone born between October and December of 1918 may have also had in-utero exposure to the pandemic. After dropping potentially treated individuals from the 1918 birth cohort, the 1918 estimate is much closer to zero and still statistically insignificant.

4.3 Assessing Family Selection

Table 3 examines whether background characteristics were measurably different for treated cohorts. As before, we draw on 15 comprehensive background characteristics. Each row corresponds to a different parental characteristic. We present results from two samples. Columns 1 and 2 correspond to a wide set of cohorts (1912-1919), with column 1 reporting the sample mean and standard deviation and column 2 corresponding to the result from estimating equation (2). Columns 3 and 4 mirror this organization but for only the 1918 and 1919 birth cohorts.

Table 3 provides evidence that the dose-response framework offers a more credible identification strategy when considering bias from parental selection. Relative to Table 1, the estimates are largely statistically insignificant and the point estimates are often smaller. In the WWII-1920 columns of Table 1, 13 of the 15 estimates were statistically significant at the 5 or 1 percent level. Among the corresponding results in column 2 of Table 3, only two estimates are significant at the 5 or 1% level. In terms of magnitudes, the cross-cohort estimates indicated that fathers of the 1919 birth cohort were about half a year older (significant at the 1 percent level). Our dose-response estimates indicate that, in the average pandemic city, fathers were 0.12 years older, but the effect is no longer

Table 3: Assessing whether background differences are correlated with pandemic intensity

Dependent Variable	1912-1919 Cohorts		1918-1919 Cohorts	
	Mean [S.D.]	Std. Flu × Born in 1919	Mean [S.D.]	Std. Flu × Born in 1919
Southern born	0.098 [0.297]	-0.000* (0.000)	0.098 [0.297]	0.000*** (0.000)
Nonwhite	0.028 [0.164]	-0.001 (0.001)	0.030 [0.170]	-0.001 (0.001)
Father born outside U.S.	0.440 [0.496]	0.002 (0.004)	0.427 [0.495]	-0.000 (0.005)
Mother born outside U.S.	0.387 [0.487]	0.001 (0.005)	0.368 [0.482]	0.003 (0.005)
Father is second-gen. immigrant	0.247 [0.431]	-0.002 (0.006)	0.250 [0.433]	-0.000 (0.004)
Father moved outside of birthplace	0.602 [0.489]	0.002 (0.003)	0.600 [0.490]	-0.001 (0.004)
Father's age when child was born	32.108 [6.970]	0.137 (0.128)	32.130 [6.926]	0.105 (0.096)
Mother's age when child was born	28.204 [5.961]	0.131 (0.094)	28.287 [5.887]	0.119 (0.086)
Father's age minus mother's age	3.904 [4.694]	0.006 (0.053)	3.843 [4.644]	-0.014 (0.044)
Father cannot read or write	0.052 [0.222]	0.002 (0.002)	0.051 [0.221]	-0.001 (0.003)
Mother cannot read or write	0.074 [0.261]	0.004** (0.002)	0.071 [0.257]	0.003 (0.002)
Father's occupational income score	28.692 [9.172]	-0.169* (0.093)	28.538 [9.039]	-0.206* (0.112)
Family owns home	0.306 [0.461]	-0.009* (0.005)	0.267 [0.442]	-0.007 (0.005)
Child's inferred birth order	2.620 [1.750]	0.049** (0.020)	2.660 [1.790]	0.040** (0.018)
Mother is in labor force	0.035 [0.183]	-0.003 (0.003)	0.026 [0.158]	-0.006* (0.003)

Notes: * p<0.1; ** p<0.05; *** p<0.01. Standard errors (clustered at the city level) in parentheses. Each entry is the coefficient obtained from regressing the indicated outcome variable on our standardized measure of flu intensity as well as cohort fixed effects and birth city fixed effects.

statistically significant. We see consistent evidence that the 1919 birth cohort with more intense exposure had more older siblings and had fathers that were employed in lower paying occupations. In terms of magnitude, these deviations are about 2% of a standard deviation.

4.4 Long-Run Estimates

Table 4 uses our dose-response framework to examine the impact of pandemic intensity on educational attainment. Each row corresponds to a different outcome variable: total years of schooling, an indicator for whether the individual had some high school education, and an indicator for graduating high school. Columns 1 and 2 examine patterns among the 1912-1919 birth cohorts. To strengthen identification, columns 3 and 4 focus on just the 1918 and 1919 birth cohorts, while columns 5 and 6 focus on our brothers subsample. Our baseline estimates are presented in columns 1, 3, and 4. Columns 2 and 4 add the same parental controls that were included in our preferred cross-cohort regressions, while column 6 goes further and includes household fixed effects.

The results in Table 4 point to a negative relationship between in-utero exposure to the pandemic and educational attainment. That effect appears to be operating through high school graduation. We see little evidence that exposure affected high school attendance; while the coefficients are negative the point estimates are small and imprecise relative to high school graduation.

The dose-response results are less sensitive to the inclusion of background controls. Our baseline estimates suggest that, relative to exposure between the ages of 0 and 8, an individual born in 1919 with the average level of pandemic exposure would be about 1.7 percentage points less likely to complete high school. Including our background controls reduces the point estimate to 1.3 percentage points, but the effects are statistically indistinguishable. In column 4, we restrict our comparison to just the 1918 and 1919 birth cohorts, and thus try to isolate the in-utero effect by making comparisons with individuals that were exposed between the ages of 0 and 1. This is a strong test, as anyone born between October and December of 1918 may have been exposed as well. Nevertheless, baseline estimates indicate that in-utero exposure lowered high school graduation rates by 1.2 percentage points. Once we adjust for background differences, the point estimate falls to 0.9 percentage points and only remains significant at the 10% level, but the two effects are statistically indistinguishable. In our brothers sample, our baseline estimate is that exposed cohorts were 2.3 percentage points less likely to complete high school, but once we include household fixed effects and other background controls (e.g., birth order and maternal/paternal age when the child

Table 4: Impact of pandemic intensity on adult outcomes for 1919 birth cohort

	1912-1919 Birth Cohorts		1918-1919 Birth Cohorts		Brothers Sample 1912-1919 Birth Cohorts	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: DV is Years of Schooling						
Std. Excess Flu \times Born in 1919	-0.066*** (0.020)	-0.045*** (0.017)	-0.037* (0.020)	-0.023 (0.019)	-0.074 (0.049)	-0.104** (0.049)
Panel B: DV is Some High School						
Std. Excess Flu \times Born in 1919	-0.006* (0.004)	-0.004 (0.003)	-0.001 (0.003)	-0.000 (0.003)	0.008 (0.010)	-0.003 (0.013)
Panel C: DV is High School Graduation						
Std. Excess Flu \times Born in 1919	-0.017*** (0.004)	-0.013*** (0.004)	-0.012** (0.005)	-0.009* (0.005)	-0.023** (0.012)	-0.039*** (0.012)
Panel D: DV is Height at Enlistment						
Std. Excess Flu \times Born in 1919	0.034 (0.021)	0.047** (0.020)	-0.013 (0.030)	-0.014 (0.038)	-0.048 (0.099)	-0.046 (0.099)
City Fixed Effects	Y	Y	Y	Y	Y	Y
Cohort Fixed Effects	Y	Y	Y	Y	Y	Y
Background Controls		Y		Y		Y
Household Fixed Effects						Y

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Robust standard errors, clustered at the city-level, errors are reported in parentheses. Background controls include the limited and other controls defined in Table 2.

was born) the deficit increases to 3.9 percentage points. The height results generally point to null effects. Most estimates are statistically insignificant and all point estimates are smaller than one twentieth of an inch in absolute value.

Our main robustness tests are presented in Appendix Table 7. That table presents results using alternative measures of pandemic intensity (i.e., $\ln(\text{excess flu})$, high vs low, and allowing effects to vary across each quartile of the distribution). The 1912-1919 sample as well as the narrow comparisons (1919 versus 1918) are robust to alternative measures of exposure, while the brothers subsample remains negative but is less precisely estimated. The 1912-1919 results are robust to including city-specific time trends, but the inclusion of birth-state-by-birth-year fixed effects yields negative but imprecise estimates.

5 Conclusion

This paper uses new data and a new identification strategy to re-evaluate the impact of in-utero exposure to the 1918 influenza pandemic on educational attainment in the United States. We linked males in WWII enlistment records to their childhood selves in either the 1920 or 1930 censuses, which allows us to observe parental socioeconomic and demographic characteristics. Like [Brown &](#)

Thomas (2021), we find that the 1919 cohort had measurably different background characteristics. While estimates of the 1919 effect appear to be overstated when these differences are ignored, controlling for these differences does not appear to alter the qualitative conclusion that the 1919 birth cohort in our linked sample was less educated.

Next, we turn to an identification strategy that uses geographic variation in pandemic intensity using data from 287 cities. Since we observe children during their early childhoods, we observe not only their state of birth, but what is likely their city of birth. By using geographic variation in influenza, we are able to include birth-year fixed effects, which will capture any national effect of WWI. We find less evidence that city-level pandemic intensity is related to observable paternal characteristics. Using this approach, we find that the pandemic lowered educational attainment.

We now take stock of the evidence and try to separate facts supported by strong evidence from interpretations of those facts. Our reading of the data is that there is strong evidence for the following claims: The 1919 birth cohort did have lower educational attainment than surrounding birth cohorts, but also had parents who were systematically different, and that selection tends to be negative. After controlling for the observed background differences, the 1919 birth cohort was still worse off, but the effects are smaller. Similarly, those born in cities with high influenza exposure had less educational attainment. There is less selection when using geographic variation in influenza intensity, but some selection is still detected. After controlling for background characteristics, the evidence still suggests that those born in high-exposure cities were less likely to graduate from high school, but the estimates do shrink in size (although not as much as in the cross-cohort analysis).

One interpretation of those facts is that these differences in adult outcomes were caused by the 1918 influenza pandemic, and the estimates of the pandemic are robust to controlling for what parental differences did exist. Another possible interpretation is that the observed parental differences may have been only part of the selection. If we could observe additional parental characteristics, then the estimates of the effect of the pandemic would shrink further, possibly to zero. We see no reason why either interpretation is unreasonable.

Our paper is not without its limitations. Our sample is selected relative to the broader population. Further, due to privacy restrictions, we are not able to observe these cohorts after WWII, and thus we observe these cohorts before some enlistees have completed their years of schooling. Furthermore, income is not available in the WWII records, and even if it were, it would likely not

reflect peak mid-career earnings. Whether these effects continue to later censuses is an open question, but once those censuses become publicly available, future researchers can and should answer that question by using our record-linking approach and dose-response framework.

References

- Abramitzky, Ran, Boustan, Leah Platt, Eriksson, Katherine, Feigenbaum, James J, & Pérez, Santiago. Forthcoming. Automated linking of historical data. *Journal of Economic Literature*.
- Almond, Douglas. 2006. Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population. *Journal of Political Economy*, **114**(4), 672–712.
- Bailey, Martha, Cole, Connor, Henderson, Morgan, & Massey, Catherine. Forthcoming. How Well Do Automated Methods Perform in Historical Samples? Evidence from New Ground Truth.
- Beach, Brian, Ferrie, Joseph, Saavedra, Martin, & Troesken, Werner. 2016. Typhoid fever, water quality, and human capital formation. *Journal of Economic History*, **76**(01), 41–75.
- Beach, Brian, Clay, Karen, & Saavedra, Martin H. 2020. The 1918 influenza pandemic and its lessons for COVID-19. *National Bureau of Economic Research Working Paper Series*.
- Brown, Ryan, & Thomas, Duncan. 2021. On the long term effects of the 1918 US influenza pandemic. *Unpublished Manuscript*.
- Chandra, Siddharth, Christensen, Julia, Mamelund, Sverre-Erik, & Paneth, Nigel. 2018. Short-term birth sequelae of the 1918–1920 influenza pandemic in the United States: state-level analysis. *American journal of epidemiology*, **187**(12), 2585–2595.
- Clay, Karen, Lewis, Joshua, & Severnini, Edson. 2018. Pollution, infectious disease, and mortality: Evidence from the 1918 Spanish influenza pandemic. *The Journal of Economic History*, **78**(4), 1179–1209.
- Costa, Dora L. 2015. Health and the Economy in the United States from 1750 to the Present. *Journal of economic literature*, **53**(3), 503–70.
- Frost, Wade Hampton. 1920. Statistics of influenza morbidity: with special reference to certain factors in case incidence and case fatality. *Public Health Reports (1896-1970)*, 584–597.
- Goeken, Ronald, Lee, Yu Na, Lynch, Tom, & Magnuson, Diana. 2017. Evaluating the Accuracy of Linked U. S. Census Data: A Household Linking Approach. *Unpublished Manuscript*.
- Isen, Adam, Rossin-Slater, Maya, & Walker, W Reed. 2017. Every breath you take—every dollar you’ll make: The long-term consequences of the clean air act of 1970. *Journal of Political Economy*, **125**(3), 848–902.
- Kitchens, Carl T, & Rodgers, Luke P. 2020. *The Impact of the WWI Agricultural Boom and Bust on Female Opportunity Cost and Fertility*. Tech. rept. National Bureau of Economic Research.
- Long, Jason, & Ferrie, Joseph. 2013. Intergenerational occupational mobility in Great Britain and the United States since 1850. *American Economic Review*, **103**(4), 1109–37.
- Mamelund, Sverre-Erik. 2004. Can the Spanish influenza pandemic of 1918 explain the baby boom of 1920 in neutral Norway? *Population*, **59**(2), 229–260.
- Sanders, Nicholas J. 2012. What doesn’t kill you makes you weaker prenatal pollution exposure and educational outcomes. *Journal of Human Resources*, **47**(3), 826–850.
- Vandenbroucke, Guillaume. 2014. Fertility and Wars: the case of world war I in France. *American Economic Journal: Macroeconomics*, **6**(2), 108–36.

6 Appendix A: Additional Results and Figures

6.1 Examining sample selection

Figure 3 shows the distribution of our links by birth year for each linked sample. The samples are somewhat younger on average than the 1912-1922 birth cohorts, as observed in census data. The main takeaway, however, is that there is little evidence of stacking on younger cohorts or on attractive (i.e. round) birth years. This suggests that age heaping or selective misreporting are not major issues for our sample. This is perhaps not surprising, as our linking requirement that age and birth year be consistently reported works to limit contamination of our sample from intentional or unintentional misreporting. Note that the 1921 cohort is somewhat larger than the 1920 and 1922 cohorts. This pattern is also reflected in census data for 1940, 1960, 1970, and 1980.

Figure 3: Distribution of links by census and birth year

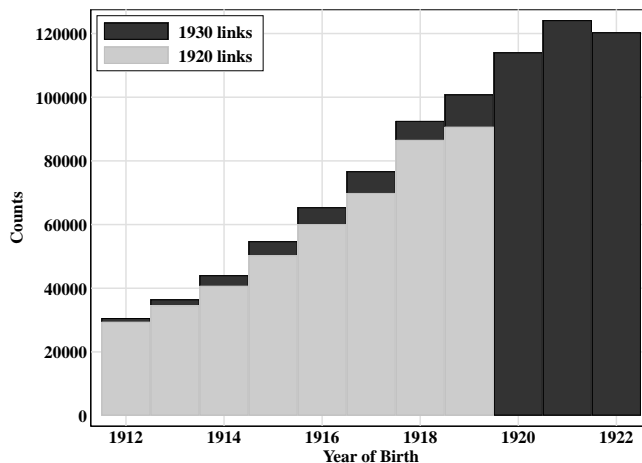


Table 5 uses the IPUMS 1% sample of the 1950 census to examine the margins on which WWII veterans and non-veterans from the 1912-1922 birth cohorts differ. We see that veterans are about 2 years younger, more likely to be white, and slightly less likely to have foreign born parents.

Table 5: Differences of WWII veterans and non-veterans in the 1950 census

	(1)	(2)
	WWII veterans	Non-WWII veterans
Year of birth	1917.893	1916.006
White	0.925	0.865
Black	0.072	0.130
Other race	0.001	0.002
U.S.-born mother	0.804	0.822
U.S.-born father	0.779	0.795
Observations	20941	14824

Notes: Data are from males born between 1912 and 1922 in the 1950 census.

6.2 Additional Cross-Cohort Results

Table 6: Differences in adult outcomes of 1919 birth cohort relative to surrounding cohorts

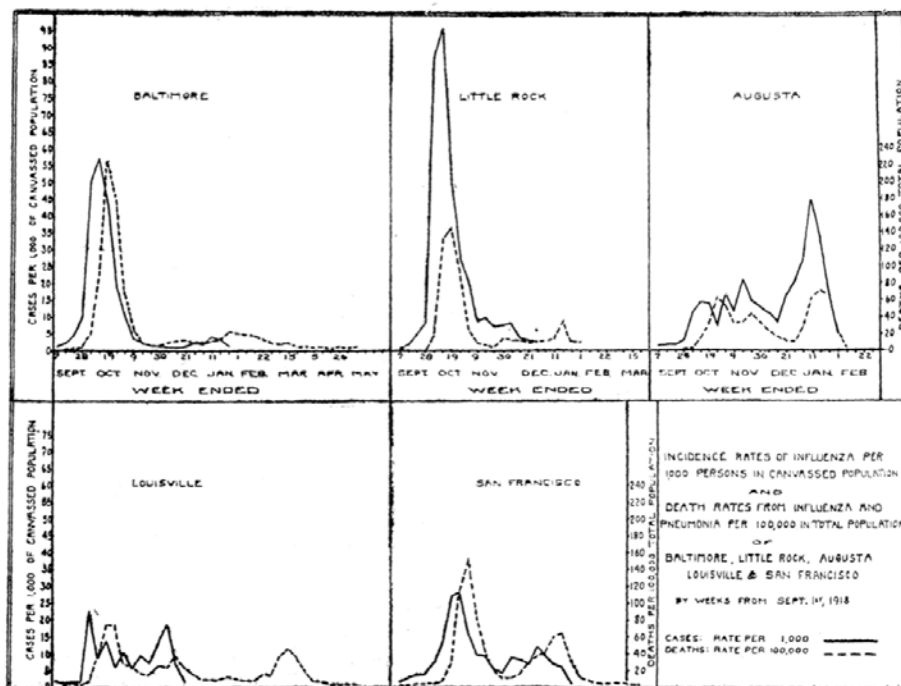
	1960 Census		WWII-1930 Links		WWII-1930 Links Reweighted	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: DV is Years of Schooling						
Born in 1919	-0.109*** (0.030)	-0.070** (0.028)	-0.094*** (0.008)	-0.054*** (0.007)	-0.114*** (0.009)	-0.057*** (0.007)
R-Squared	0.011	0.116	0.002	0.242	0.002	0.252
Panel B: DV is Some High School						
Born in 1919	-0.019*** (0.005)	-0.013*** (0.005)	-0.011*** (0.001)	-0.006*** (0.001)	-0.017*** (0.002)	-0.007*** (0.001)
R-Squared	0.011	0.115	0.004	0.179	0.005	0.198
Panel C: DV is High School Graduation						
Born in 1919	-0.021*** (0.005)	-0.015*** (0.005)	-0.017*** (0.002)	-0.009*** (0.001)	-0.021*** (0.002)	-0.010*** (0.002)
R-Squared	0.011	0.094	0.003	0.174	0.004	0.185
Limited Controls		Y		Y		Y
Other Background Controls				Y		Y

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Following Almond (2006), each regression includes the reported “Born in 1919” indicator variable, as well as year of birth and year of birth squared. Robust standard errors are reported in parentheses. Limited controls include fixed effects for: race, birth state, father’s country of birth, and mother’s country of birth. Other parental controls follow from Table 1 and include: whether the father is a second generation immigrant, whether the father moved outside their birthplace before the child was born, the father’s age when the child was born, whether the father can read or write, fixed effects for the father’s occupational income score, whether the family owns their home, the number of older siblings observed in the household (a proxy for birth order), the difference between the age of the father and the age of the mother, mother’s age when the child was born, whether the mother can read or write, and whether the mother is in the labor force.

6.3 Additional Dose-Response Framework Results

Figure 4 (reproduced from Frost (1920)) illustrates that influenza mortality rates track influenza morbidity rates reasonably well. The slight lag observed in Figure 4 is not a concern for us because we use annual rather than weekly data to construct our measure of excess influenza deaths.

Figure 4: Weekly morbidity and mortality for five cities in the fall of 1918

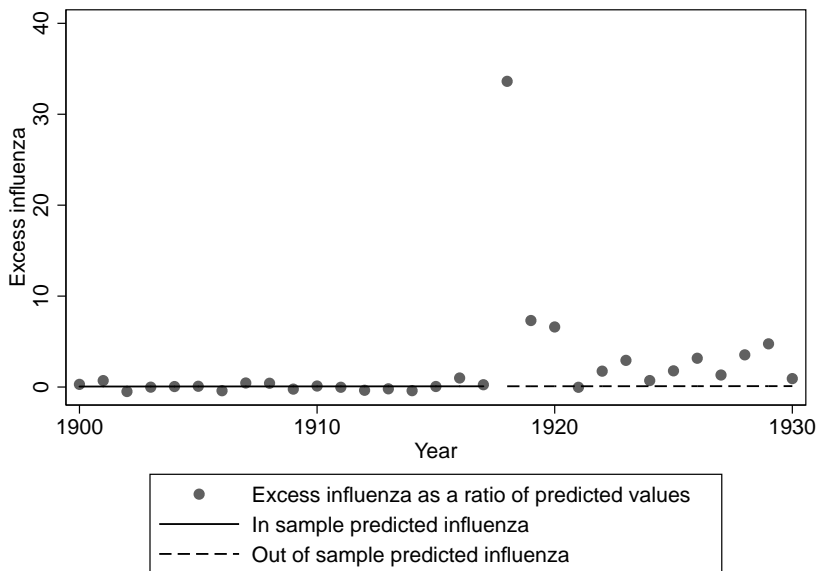


Notes: Figure reprinted from Frost (1920).

Figure 5 plots average excess influenza ratios by year. Note that our sample period is 1900-1917 whereas the 1918-1930 observations are our out-of-sample predictions. Our model fits the data very well in the pre-pandemic period. The excess influenza ratio in the post-pandemic period (1922 and beyond) is noisier, but not far above zero. There is no indication of a post-pandemic trend.

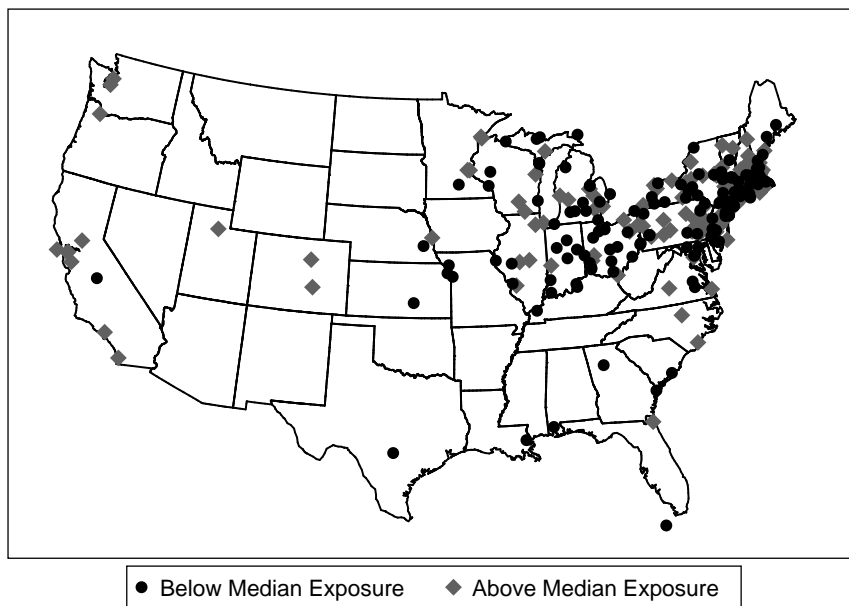
Figure 6 illustrates the spatial variation in our sample. We plot separate markers for cities that had above-median versus below-median exposure to the pandemic. The key takeaway from Figure 6 is that there is meaningful sub-state variation in our measure of flu intensity: there are many situations where neighboring cities had different exposure to the pandemic.

Figure 5: Excess influenza mortality ratio by year



Notes: Excess influenza ratio is calculated by taking actual influenza deaths minus predicted influenza deaths and then dividing by the predicted influenza deaths. City-specific trends in annual $\ln(\text{influenza mortality})$ are estimated over the 1900-1917 period.

Figure 6: Spatial variation in flu intensity



Notes: High exposure cities had an excess flu ratio greater than 28.2.

Table 7 presents the dose-response results with alternative functional forms of influenza intensity. For comparison, Panel A displays our results from Table 4 in which excess influenza enters the regression equation linearly. Panel B replaces the treatment variable with the log of excess influenza. Panel C includes a dummy variable for whether a city was exposed to above-median excess influenza in 1918. Panel D includes dummies for quartile of excess influenza exposure, omitting the first quartile. For all panels we present results examining the 1912-1919 birth cohorts with and without background controls, city time trends, and cohort-by-birth-state fixed effects. We also include specifications that limit the sample to only the 1918 and 1919 birth cohorts (columns 5 and 6). Columns 7 and 8 limit the sample to brothers and include brother fixed effects. For Panels A, B, and C, the coefficients are negative for every specification, however, not all are statistically significant. This is especially true for the brothers-only sample and when including birth state-by-cohort fixed effects. In Panel D, we see that the effect is mostly driven by cities that were in the worst quartile of excess influenza exposure.

Figure 7 presents estimates of the effect of the influenza pandemic over the course of the life cycle using an identification strategy that is analogous to an event study. All cohorts from 1912 to 1919 (which compose our 1920 to WWII links) are potentially affected by the pandemic. While the 1919 cohort would have only been affected through in-utero exposure, the last quarter of the 1918 birth cohort was possibly affected by in-utero exposure as well, and the rest of the 1918 cohort would be have been exposed during infancy. Older cohorts may have also been affected by childhood exposure to the pandemic.

We estimate the following model:

$$y_{ibc} = \alpha_0 + \beta_b + \gamma_c + \sum_{yob \neq 1917} \delta_{yob} \mathbf{1}[b = yob] \times \text{Excess Flu}_c + \epsilon_{ibc} \quad (3)$$

where δ_{yob} is the effect of the pandemic on a particular birth cohort relative to the effect on the 1917 birth cohort. We let 1917 be the omitted group because it is the closest cohort to 1919 for whom no one in that cohort would have been exposed to the pandemic in utero. To increase precision, we also estimate only one parameter for the 1912 to 1914 birth cohorts. We also include our background controls, as defined in Table 4.

Table 7: Assessing robustness of dose-response results

	Baseline Links 1912-1919				1918-19 Only		Brothers (1912-1919)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Treatment is Standardized Excess Flu								
Exposure \times 1919	-0.017*** (0.004)	-0.013*** (0.004)	-0.013*** (0.003)	-0.011** (0.005)	-0.012** (0.005)	-0.009* (0.005)	-0.023** (0.012)	-0.039** (0.017)
Panel B: Treatment is ln(Excess Flu)								
Exposure \times 1919	-0.019*** (0.006)	-0.014*** (0.005)	-0.012** (0.006)	-0.011 (0.008)	-0.014** (0.006)	-0.011* (0.006)	-0.025 (0.020)	-0.038 (0.030)
Panel C: Treatment is Above Median Exposure								
Exposure \times 1919	-0.023*** (0.008)	-0.017** (0.007)	-0.014** (0.007)	-0.006 (0.008)	-0.014* (0.008)	-0.012 (0.008)	-0.027 (0.021)	-0.035 (0.030)
Panel D: Treatment is Relative to 0-25th Percentile Exposure								
1[flu 25-50th] \times 1919	0.002 (0.012)	0.005 (0.009)	-0.002 (0.008)	-0.008 (0.014)	-0.019* (0.011)	0.005 (0.009)	0.025 (0.032)	0.025 (0.033)
1[flu 50-75th] \times 1919	-0.016 (0.012)	-0.009 (0.011)	-0.012 (0.010)	-0.008 (0.013)	-0.020 (0.014)	-0.009 (0.011)	-0.008 (0.042)	-0.008 (0.044)
1[flu 75+] \times 1919	-0.028*** (0.010)	-0.020** (0.010)	-0.017* (0.010)	-0.013 (0.013)	-0.027** (0.011)	-0.020** (0.010)	-0.045 (0.044)	-0.042 (0.045)
Background Controls		Y	Y	Y		Y		Y
City Time Trends			Y	Y				
Birth State by Cohort Fixed Effect				Y				

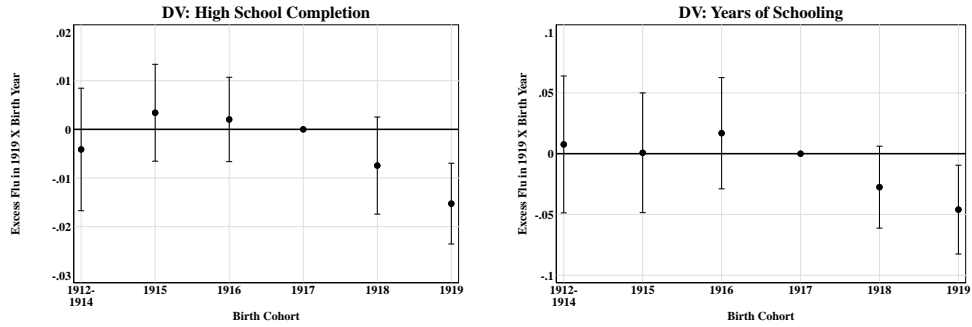
Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Following Almond (2006), each regression includes cohort fixed effects and city fixed effects. Robust standard errors (clustered at the city-level) reported in parentheses. Background controls follow Table 2

The results from estimating Equation 3 are in the top panels of Figure 7. Relative to exposure between 1 and 2 (i.e., the 1917 cohort) there is no statistically significant effect of the pandemic on high school graduation rates for any cohort from 1912 to 1916. There is also no clear trend among these cohorts and the point estimates are close to zero. The point estimate for the 1918 cohort, which could be thought of as partially treated, is negative, suggesting the pandemic lowered high school graduation by approximately three quarters of a year, but is statistically insignificant. The effect of the 1919 cohort is negative with a coefficient around -1.5 percentage points and is highly significant. While these results do not necessarily imply that other cohorts were unaffected by the pandemic (because these estimates are relative to the 1917 birth cohort), they do suggest that the 1919 cohort was disproportionately affected.

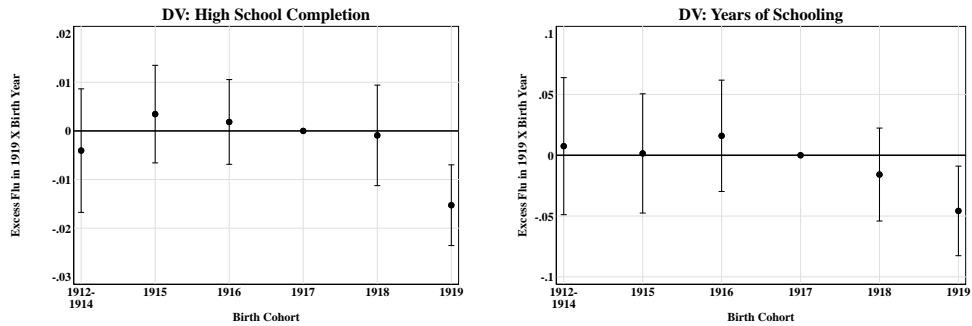
One way to address the partial treatment of 1918 is to use census information about the individual's age in months. Census enumerators were instructed to obtain the age in years and months for any individual under the age of 5 at the time of census enumeration. Unfortunately, not all of the census enumerators recorded this information. For instance, among those aged 1 (i.e., the 1918 birth cohort) this information is missing for 13.9% of our sample. Many of those are likely individuals that are close to 1 year. Only 2% of our 1-year-olds show up as 1 year and 0 months, whereas 6.6% show up as 1 year and 1 month and 8.6% show up as 1 year and 2 months.

Our solution to address the partial treatment is to drop from the sample any individual that appears as 1 year and 0, 1, 2, or 3 months. Those individuals would have been born between October of 1918 and December 31, 1918. We also drop anyone with missing information. The bottom panels of Figure 7 present results after implementing these restrictions. The results suggest that the negative point estimates that we saw in the top panels of Figure 7 for the 1918 birth cohort were, to some extent, driven by the inclusion of individuals that were potentially treated by the pandemic.

Figure 7: Event study estimates of the dose-response effect
 Panel A: Baseline Event Study



Panel B: Dropping potentially exposed individuals from 1918 Cohort



Notes: Regressions include city-of-enumeration and birth year fixed effects as well as the background controls from Table 4. Standard errors are clustered at the city-level.

7 Appendix B: Revisiting Brown & Thomas (2021)

This section examines the sensitivity of the Brown & Thomas (2021) approach to accounting for background differences. Their approach to accounting for selection is to take birth-state-by-birth-year-by-race averages of each background characteristic. The averages are calculated using the 1920 census, where birth year is calculated as $1920 - \text{age} - 1$.²⁰ The authors then include these measures as controls to assess the robustness of Almond (2006).

Table B.1 replicates this approach for various sets of aggregate background controls. The three panels correspond to the 1960, 1970, and 1980 censuses examined in Almond (2006). Each row corresponds to a different outcome. Column 1 corresponds to the results as reported in Brown & Thomas (2021, Table 4). Column 2 is our replication of those results. Column 3 omits the number-of-children control. Column 4 adds inferred birth order (i.e., the number of older siblings in the household). Column 5 includes aggregate versions of our 15 background measures. Finally,

²⁰Individuals were asked to report their age as of January 1, 1920.

Column 6 reports the baseline Almond result, but restricting to the 1912-1919 birth cohorts.

The main takeaway from Table B.1 is that the sign flips documented in Brown & Thomas (2021) are largely driven by the inclusion of the “number of father’s children in the household” variable. When that control is omitted or replaced with birth order, the sign flips are much less pronounced, although the confidence intervals often contain zero. The standard errors in these specifications are so large that it is often difficult to reject whether the results in columns 3, 4, or 5 are different from what Almond would estimate if he omitted the 1920-1922 birth cohorts (Column 6).

The results in Table B.1 raise the question as to whether the number of children in the household variable is trustworthy. We explore this issue further in Appendix Section 8.

Table B.1: Assessing whether selection for the 1919 birth cohort varies over time

	BT (2018) Table 4 (1)	Our Replication (2)	No # child (3)	w/ Birth Order (4)	15 BFS Controls (5)	Almond 1912-19 (6)
Dependent Variable						
Panel A: 1960 Census (IPUMS 1% Sample)						
Years of Schooling	0.266*** (0.062)	0.240*** (0.056)	-0.047 (0.054)	-0.034 (0.054)	-0.046 (0.055)	-0.168*** (0.056)
HS Graduation	0.035*** (0.009)	0.036*** (0.009)	-0.012 (0.009)	-0.010 (0.009)	-0.008 (0.009)	-0.022** (0.009)
Wage Income	2160*** (452)	2110.149*** (453.543)	669.142 (439.918)	721.513 (439.678)	1113.786** (446.067)	-243.665 (447.049)
Total Income	2790*** (504)	2984.563*** (495.889)	231.053 (479.229)	339.557 (478.741)	637.820 (485.353)	-531.312 (491.203)
Panel B: 1970 Census Form 1 (IPUMS 3% Sample)						
Years of Schooling	0.242*** (0.038)	0.248*** (0.034)	-0.038 (0.033)	-0.028 (0.033)	-0.031 (0.034)	-0.136*** (0.034)
HS Graduation	0.041*** (0.006)	0.045*** (0.006)	-0.007 (0.005)	-0.005 (0.005)	0.000 (0.005)	-0.018*** (0.005)
Wage Income	2314*** (397)	2036.630*** (390.393)	126.768 (377.639)	204.379 (377.378)	540.663 (381.500)	-895.703** (381.842)
Total Income	3060*** (433)	3565.502*** (436.612)	289.566 (422.585)	418.613 (421.965)	541.503 (426.579)	-651.284 (427.723)
Panel C: 1980 Census (IPUMS 5% Sample)						
Years of Schooling	0.258*** (0.033)	0.247*** (0.029)	-0.037 (0.028)	-0.024 (0.028)	-0.026 (0.028)	-0.141*** (0.029)
HS Graduation	0.045*** (0.004)	0.054*** (0.004)	-0.000 (0.004)	0.002 (0.004)	0.008* (0.004)	-0.012*** (0.004)
Wage Income	-132 (447)	-15.130 (459.510)	-994.376** (453.289)	-942.474** (453.279)	-1035.327** (454.358)	-1121.454** (455.166)
Total Income	1770*** (489)	1903.330*** (502.173)	-766.472 (495.636)	-665.869 (495.406)	-701.427 (496.630)	-801.548 (502.377)

Notes: * p<0.1; ** p<0.05; *** p<0.01. Following Almond (2006), each regression includes the reported “Born in 1919” indicator variable, as well as year of birth and year of birth squared. Robust standard errors are reported in parentheses.

8 Appendix C: Examining Parental Selection with a Double-Linked Sample

This section uses our linked data to examine whether the individual’s age at the time of census enumeration affects our estimates of parental selection. We focus on individuals that we can link between the 1920 census, the 1930 census, and the WWII enlistment records. We further restrict attention to those that are residing with their parents in both censuses. This yields a final sample of 274,640 individuals.

Figure B.1 provides a graphical assessment of the age-at-observation issue. Each panel plots cohort-level averages based on information observed in both the 1920 census and the 1930 census. This graphical assessment gives us a sense of whether the data are exhibiting expected patterns, which speaks to the trustworthiness of our underlying links.

The first 4 panels of Figure B.1 correspond to characteristics that, if perfectly observed, would not change between the 1920 and 1930 censuses. These characteristics are: whether the individual is nonwhite, whether the father was born outside of the United States, the father’s age when the child was born, and the child’s birth order. Each panel corresponds to a different characteristic, with 1920 cohort-level means plotted as black diamonds and 1930 cohort-level means plotted as grey circles.

The first takeaway from Figure B.1 is that our linked data are of high quality. The cohort-level averages for nonwhite and whether the father is foreign born are nearly identical, which is not an artifact of linking because these pieces of information are not part of our linking procedure. The extent to which markers do not perfectly overlap often reflects some combination of transcription errors, errors in enumeration (e.g., relying on information from a neighbor or another member of the household rather than the individual of interest), or errors in the linking. These do not appear to be major issues.

While father’s age at the time of birth and birth order are time invariant, the patterns do differ between the two censuses. Fathers in 1920 are about 0.25 years older on average at the time of the child’s birth relative to 1930. This is likely due to changes in enumeration procedures. Since the 1920 census was enumerated as of January 1 while the 1930 census was enumerated as of April 1, there may be slight discrepancies in age. It is worth noting that if we restrict to the set of records where the child’s age in 1930 is equal to the reported age in 1920 + 10 we see almost complete overlap between the two data series. Birth order, which is inferred based on the number of older

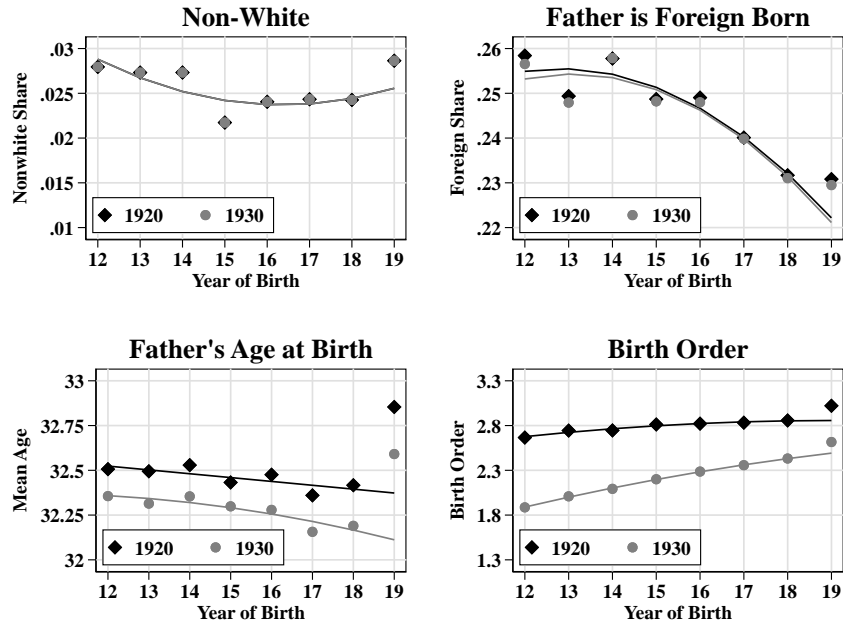
siblings in the household, also changes between 1920 and 1930. On average, all of our cohorts have fewer older siblings in the 1930 census relative to the 1920 census. This pattern reflects the fact that older siblings may leave the household or die by 1930, particularly for older cohorts (the 1912 birth cohort is roughly 18 at the time of enumeration in 1930).

The bottom four panels of Figure B.1 plot characteristics that may vary over time. The characteristics are father’s occupational income score, an indicator for whether the mother is in the labor force, an indicator for whether the family owns their home, and the total number of father’s own children in the household. Between 1920 and 1930 cohort-level averages change in the following ways: father’s occupational standing increases, maternal labor force participation increases, home ownership rates increase, and (aside from the 1912 cohort) total number of children in the household increases. For the first three panels, the change between the two censuses reflects a level shift (but not a change in trend) that is predictable based on life-cycle patterns. The patterns for total number of children in the household are slightly different. In 1920 we see that younger cohorts were from smaller families, but by 1930 there is much less of a distinction. The decline in family size in 1920 likely reflects the inherent truncation of the “number of younger siblings” component of the family size variable. That is, for the 1912 birth cohort, family size will be a function of both the number of older siblings in the household and any younger siblings born between 1913 and 1920. For the 1919 birth cohort, the family size variable will only be a function of the number of older siblings in the household because any younger siblings will be born after enumeration in 1920.

Table B.2 examines whether our estimated deviations of the 1919 birth cohort vary based on the census year. Each row corresponds to a different background characteristic. Columns 1 and 2 correspond to the 1930 census, columns 3 and 4 correspond to the 1920 census, and column 5 tests whether the estimated deviations are statistically equal. The majority of selection patterns that we observe in 1920 are stable over time and are thus unlikely to be an artifact of observing different birth cohorts at different ages. Across the 16 outcomes, only two estimates are statistically distinguishable based on whether we use the 1920 census or the 1930 census. The first is the total number of father’s children (as discussed above). The second variable is inferred birth order. It is worth pointing out that while we can reject equality between the estimates of selection on birth order, the point estimates are not meaningfully different from each other (0.16 in 1920 and 0.12 in 1930). For these reasons, and because family size plays an important role in determining long-run outcomes, we choose to include birth order as one of our background controls.

Figure B.1: Patterns of parental characteristics by birth cohort for those observed with parents in 1920 and 1930

Time Invariant Characteristics



Time Variant Characteristics

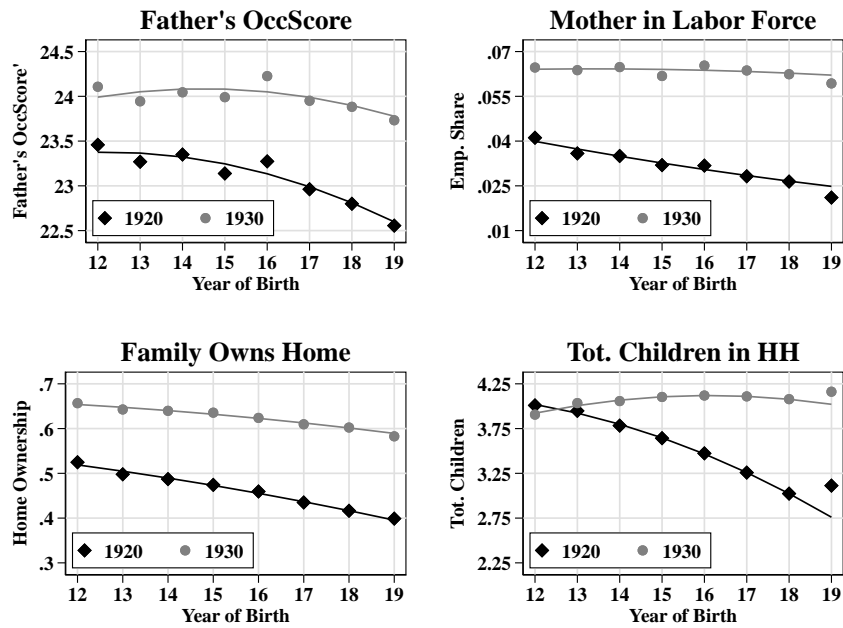


Table B.2: Assessing whether selection for the 1919 birth cohort varies over time

	WWII-1930 Links		WWII-1920 Links		Reject (2)=(4)? (5)
	Mean (1)	1919 Deviation From Trend (2)	Mean (3)	1919 Deviation From Trend (4)	
Dependent Variable					
Southern Born	0.225	0.013*** (0.004)	0.225	0.013*** (0.004)	1
Nonwhite	0.026	0.003** (0.001)	0.026	0.003** (0.001)	1
Father born outside U.S.	0.242	0.009** (0.004)	0.241	0.008** (0.004)	0.958
Mother born outside the U.S.	0.206	0.008** (0.004)	0.205	0.009** (0.004)	0.802
Father is second-gen. immigrant	0.220	-0.009** (0.004)	0.224	-0.010*** (0.004)	0.861
Father moved outside of birthplace	0.435	0.003 (0.004)	0.436	0.003 (0.004)	0.964
Father's age when child was born	32.523	0.481*** (0.066)	32.320	0.479*** (0.065)	0.985
Mother's age when child was born	28.301	0.321*** (0.056)	28.078	0.288*** (0.056)	0.677
Father's age minus mother's age	4.222	0.159*** (0.045)	4.242	0.190*** (0.045)	0.623
Father cannot read or write	0.039	0.002 (0.002)	0.034	0.004** (0.002)	0.488
Mother cannot read or write	0.041	-0.000 (0.002)	0.035	0.002 (0.002)	0.301
Father's occupational income score	23.531	-0.057 (0.106)	24.698	-0.115 (0.107)	0.704
Family owns home	0.447	0.003 (0.005)	0.616	-0.007 (0.004)	0.119
Child's inferred birth order	2.845	0.164*** (0.017)	2.317	0.124*** (0.013)	0.062*
Number of father's children in HH	3.399	0.353*** (0.018)	4.092	0.140*** (0.019)	0.000***
Mother is in labor force	0.029	-0.004*** (0.001)	0.063	-0.003 (0.002)	0.706

Notes: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Following Almond (2006), each regression includes the reported “Born in 1919” indicator variable, as well as year of birth and year of birth squared. Robust standard errors are reported in parentheses.