Comment

Reevaluating the Long-Term Impact of In Utero Exposure to the 1918 Influenza Pandemic

Brian Beach

Vanderbilt University and National Bureau of Economic Research

Ryan Brown

University of Colorado Denver

Joseph Ferrie

Northwestern University and National Bureau of Economic Research

Martin Saavedra

Oberlin College

Duncan Thomas

Duke University and National Bureau of Economic Research

Almond (2006) argues that in utero exposure to the 1918 influenza pandemic reduced the 1919 birth cohort's adult socioeconomic status (SES). We show that this cohort came from lower-SES families, which is

At the request of the editor, this paper integrates results from Brown and Thomas (2022), first circulated in 2013, with those from Beach, Ferrie, and Saavedra (2021), first circulated in 2017. Comments from Douglas Almond, Mark Anderson, Alan Barreca, Michael Carter, Eileen Crimmins, Janet Currie, William Evans, Erica Field, Walker Hanlon, James Heckman,

Electronically published May 25, 2022

Journal of Political Economy, volume 130, number 7, July 2022.

https://doi.org/10.1086/719757

incompatible with Almond's cohort-comparison identification strategy. The adult SES deficit is reduced after background characteristics are controlled for; it is small and statistically insignificant in models that include household fixed effects. Replicating Almond's state-level dose-response analysis, we find no evidence in census data that influenza exposure reduced adult SES. Evidence from a city-level dose-response analysis on educational attainment using WWII enlistees from 287 cities is mixed.

I. Introduction

In a seminal paper, Almond (2006) provides important evidence on the fetal-origins hypothesis (Barker 1990) by leveraging the 1918 influenza pandemic to identify the causal effects of shocks to the in utero environment on SES in adulthood. Arguing that the pandemic was severe, unexpected, and temporary, he used two identification strategies. The first, a cross-cohort approach, compares outcomes of the 1919 birth cohort, who were in utero at the time of the 1918 pandemic, with outcomes of comparison cohorts (1912–22). The second, a dose-response approach, uses maternal mortality rates as a proxy for intensity of exposure to make within-cohort comparisons. Drawing on results from both approaches, Almond concludes that SES in adulthood was reduced by in utero exposure to the pandemic.

Almond (2006) is widely cited and is considered definitive evidence on the lasting effects on SES of in utero influenza exposure, in particular, and fetal health shocks in general. Our reevaluation of the evidence is motivated by three facts. First, the results have powerful implications for science and policy, particularly during the SARS-CoV-2 pandemic, which has brought increased focus and salience to understanding the long-term and persistent economic impact of a health shock experienced while in utero. The global interest in the evolving understanding of the impacts of SARS-CoV-2 highlights the critical importance of getting the science right. Second, there is a substantial body of rigorous evidence establishing that, unlike in the case of certain long-term health issues, early-life disadvantage is not immutable in considerations of adult economic outcomes. This provides a theoretical foundation for why one might not expect to find adult SES deficits for children exposed to an in utero health shock (Heckman 2006). Third, the coincidence of the pandemic

Dan LaFave, Jeremy Lebow, Grant Miller, Andrew Noymer, John Parman, Daniel Rees, Nick Sanders, Seth Sanders, T. Paul Schultz, James P. Smith, Tony Sun, Alessandro Tarozzi, Yinhong Zhao, and Nic Ziebarth have been very helpful. Thomas acknowledges funding from the National Institute for Child Health and Human Development (R01HD052762) and the National Institute on Aging (R01AG031266). Replication files are provided as supplementary material online. This paper was edited by James J. Heckman.

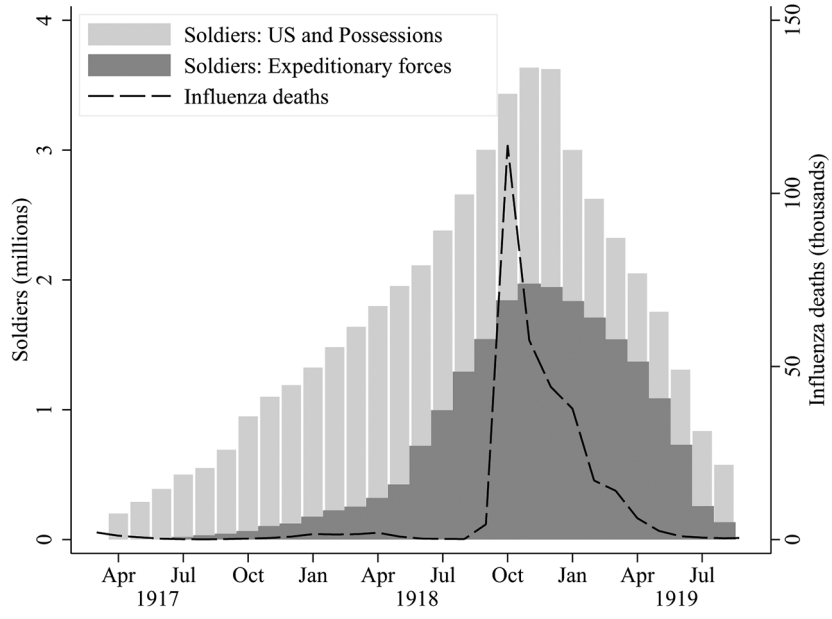

FIG. 1.—Monthly overseas troop deployments and influenza deaths in the registration area, April 1917–October 1919. Troop data from Ayers (1919); influenza deaths from Bureau of the Census (1919, 1920, 1921).

and World War I (WWI) raises legitimate concerns regarding potential omitted factors that confound Almond's conclusions.

The United States declared war on Germany in April 1917 and started deploying large numbers of troops to Europe in the summer of 1918, as shown in figure 1. Military deployments (dark gray bars) rose dramatically in the last half of 1918, peaked at the end of the year, and declined slowly during the first three months of 1919. The dashed black line in figure 1 plots influenza deaths recorded in vital statistics registration areas. The overlap in timing is striking. Further, men who served in WWI tended to be positively selected on SES relative to the general population, so that men who remained in the United States were negatively selected and more likely to be the fathers of those who were in utero during the 1918 pandemic.[1] Moreover, there is evidence that fertility patterns were affected by both the war and the pandemic (Mamelund 2004; Vandenbroucke 2014; Chandra et al. 2018; Kitchens and Rodgers 2020). These findings suggest

[1] This selection is driven by several factors. First, WWI was the first war in which a US citizen was not allowed to hire a proxy to serve in his place, ruling out the possibility of the upper class buying their way out of service. Second, men were placed in a lower-priority draft group if their family had little financial support apart from himself, because the family would have "insufficient" income to sustain itself if he were drafted (Nudd 2004). Finally, deferments were awarded for health reasons, and so the less healthy were less likely to be drafted.

that parents of children born during or after the pandemic may have been different from parents of children born before the pandemic. This raises questions about the exchangeability of the 1919 and surrounding cohorts and, therefore, the validity of Almond's cohort-comparison identification strategy.

We document that these are more than theoretical concerns. Parents of the 1919 birth cohort had lower SES than parents of surrounding cohorts. The 1919 birth cohort was more likely to be nonwhite and to be born in the South, and their fathers were more likely to be illiterate and work in lower-SES occupations. We find evidence of this selection in three data sets: the 1960 census (analyzed by Almond), the full-count 1920 census, and WWII enlistment records linked to the 1930 census. Taken together, these data do not support the assumption that the 1919 and comparison cohorts are exchangeable, which is necessary for the validity of the cross-cohort identification strategy.

Four approaches are used to assess the degree to which parental selection affects conclusions regarding the 1919 birth cohort deficit in adult SES. First, models of adult SES reported by Almond are extended by drawing on a set of background characteristics in the 1960 census. Second, aggregate proxies for parental SES are constructed from the 1920 census and included in Almond's models. Third, using WWII enlistment records linked to the 1930 census, we control for individual-specific parental characteristics. Fourth, household fixed effects are included for a subsample of links to absorb all shared time-invariant parental characteristics of brothers. Regardless of the approach, estimates of the 1919 deficit are attenuated when we take into account observed background characteristics. When unobserved differences in parental characteristics are taken into account, the magnitude of the estimates becomes even less economically meaningful, and none is statistically significant. We conclude that failure to take into account parental selection is critical.

Even after controlling for heterogeneity in parental composition, it is important to note that other potential sources of bias potentially remain in the cross-cohort approach that are particularly salient for pregnant women. For example, elevated stress in utero has been linked to worse birth outcomes (Mansour and Rees 2012; Brown 2020). In addition, as a result of war-induced increases in food prices and the initiation of the national food conservation campaign, the mothers of the 1919 cohort may have experienced reduced food intake while pregnant, which has been linked to chronic health deficits in adulthood (Rotwein 1945; Roseboom, de Rooij, and Painter 2006).[2] Moreover, the mobilization effort

[2] Famine in Allied countries triggered the "Food Will Win the War" campaign in the United States, which urged citizens to restrict their consumption of meat, wheat, fats, and sugars.

may have caused elevated stress among pregnant women. With those confounders in mind, we exploit information on quarter of birth of the 1919 birth cohort and document that quarter-by-quarter SES deficits in adulthood do not line up with the precise timing implied by in utero exposure during the fall 1918 wave of the pandemic. These results provide further evidence against interpreting estimates based on the cross-cohort strategy as causal.

Almond's second strategy is designed to identify a dose-response effect and potentially addresses these concerns. He reports statistically significant negative dose-response effects for males in the 1960 census but not in the 1970 or 1980 censuses. Our replication of his approach documents two data errors. After correcting the data, we find that of the 15 estimates for males, only two are negative and statistically significant at the 5% level (in the 1960 census) and two are positive and significant (in the 1970 census). Further, of the 30 estimates for females and nonwhites, the only statistically significant estimate is positive. Overall, this approach does not provide consistent evidence of a negative dose-response effect.

The final part of this paper extends the dose-response framework to provide new evidence on the lasting consequences of in utero exposure to the pandemic. Digitization of the entire 1920 US census allows us to construct an individual-level data set linking World War II enlistment records to the censuses, so the enlistee's residence is known as of the census enumeration date. That information is used to construct city-specific measures of pandemic exposure for 287 cities. Linking makes a trade-off between population representativeness and precision. Specifically, compared with the 1960 census, the linked sample contains only males, and they are more likely to be white and have foreign-born parents.

An advantage of our city-level dose-response analysis is that the correlation between background characteristics and city-level pandemic intensity is weak, suggesting that parental selection is unlikely to be an important source of contamination when this approach is used. Our estimates from this model indicate that exposure to greater pandemic intensity lowered educational attainment when the 1912–18 birth cohorts are used as comparisons, even when restricted to differences between brothers in a household fixed effects model. On the other hand, when we restrict the analyses to the 1918–19 birth cohorts, to help rule out unobserved differences across cohorts, none of the estimates is statistically significant.

Section II of this paper evaluates the cross-cohort identification strategy. After we document negative selection of the 1919 birth cohort's parents, estimates of cohort effects are reported that adjust for background in four different ways, and estimates that exploit the quarter of birth are provided. Section III reevaluates the dose-response identification strategy. We first replicate Almond's results using his state-level maternal mortality measure and then use city-level variation with the linked data. The evidence

presented here integrates results reported in Beach, Ferrie, and Saavedra (2021) and Brown and Thomas (2022).

II. Assessment of Cross-Cohort Evidence

A. *Adult SES of the 1919 Birth Cohort*

Using the 1960, 1970, and 1980 censuses, Almond (2006) contrasts indicators of SES in adulthood, y_i, for the 1919 birth cohort against the 1912–22 cohorts by estimating

$$y_i = \beta_0 + \beta_1 \text{YOB} + \beta_2 \text{YOB}^2 + \beta_3 \mathbf{1}[\text{YOB} = 1919] + \epsilon_i, \quad (1)$$

where YOB is year of birth and $\mathbf{1}[\text{YOB} = 1919]$ is an indicator for the 1919 birth cohort.[3]

Table 1 presents estimates of the deviation from trend for the 1919 birth cohort, β_3, for males in 1960 from Almond (2006) in column 1 and our replication in column 2.[4] Relative to those born in surrounding cohorts, males born in 1919 are significantly less likely to have graduated from high school, complete fewer years of education, have lower wage income, are more likely to be poor, and have lower Duncan's Socioeconomic Index (SEI) scores, an indicator of SES that is based on the occupation of the individual.

As shown in columns 2 and 3, the conclusions are generally robust to narrowing the comparison cohorts. Column 3 excludes the 1920–22 birth cohorts, to address concerns that conceptions after October 1918 may be related to the pandemic (Boberg-Fazlic et al. 2021). Estimates of β_3 are similar to those using the 1912–22 cohorts, although the standard errors are larger. For years of education, high school graduation, and the SEI indicator, we continue to find statistically significant deficits. For total income, the point estimate is similar but no longer statistically significant. For wage income and whether the individual is below 1.5 times the poverty level, the estimates are smaller and statistically insignificant. Column 4 further restricts the comparisons to the 1915–18 birth cohorts, all of whom were under the age of 5 when the pandemic struck and were thus unlikely to have had their schooling disrupted by the pandemic. This restriction is

[3] The samples analyzed are a nationally representative 1-in-100 random sample of the population for 1960, a 3-in-100 sample for 1970, and a 5-in-100 sample for 1980. All samples are available at ipums.org, although the 1970 sample is constructed by combining the three (independent) 1% samples of the Form 1 data (Ruggles et al. 2021b). IPUMS took six independent 1% samples of the 1970 census (three from the Form 1 data and three from the Form 2 data). The three "within-form" samples are classified as "State," "Metro," and "Neighborhood," on the basis of identifiable geographic information made available in the sample. The Form 1 vs. Form 2 designation corresponds to which census form the individual received. Almond (2006) uses Form 1 data because respondents were asked about their disability status.

[4] The differences likely reflect differences in the public release versions of the IPUMS samples.

TABLE 1
DIFFERENCES IN ADULT SES OF 1919 BIRTH COHORT RELATIVE TO SURROUNDING COHORTS: MALES IN 1960 CENSUS

	Born in 1919			
	Relative to 1912–22 Cohorts		Relative to Cohorts:	
	Almond (2006)	Replication	1912–18	1915–18
Adult SES Indicator	(1)	(2)	(3)	(4)
High school graduate	−.021*	−.021*	−.022*	−.035*
	(.005)	(.005)	(.009)	(.014)
Years of education (completed)	−.150*	−.150*	−.188*	−.211*
	(.038)	(.038)	(.063)	(.100)
Total income ($/month)	−573	−551	−531	−1,073
	(295)	(288)	(491)	(784)
Wage income ($/month)	−812*	−791*	−543	−1,435*
	(261)	(254)	(445)	(717)
Poor (<1.5 × poverty level)	.010*	.010*	.001	−.003
	(.005)	(.005)	(.008)	(.013)
Duncan's SEI	−.640*	−.631*	−.884*	−.592
	(.259)	(.260)	(.436)	(.694)
Observations	114,031	114,032	80,695	51,462

NOTE.—Estimates of β_3 from eq. (1) reported for each dependent variable listed in the adult SES indicator column and for each comparison listed in cols. 1–4. All income values in 2005 dollars. Robust standard errors in parentheses.

* Statistically significant at 5% size of test.

also useful because age is reported in years and months for these cohorts, which allows us to corroborate our cohort size with data from the vital statistics to rule out the possibility that results are driven by misreporting age (including heaping on preferred digits). In this sample, most 1919 birth cohort deficits are even larger in magnitude relative to the 1912–22 comparisons, although the standard errors are larger because of the smaller sample sizes.

B. Parental SES of Males in the 1919 Birth Cohort

A necessary condition for these estimates to be interpreted as causal is that the 1919 and comparison cohorts are statistically exchangeable. This condition is rejected if, for example, the parental SES of the 1919 birth cohort is different from that of the comparison cohorts. Figure 2 speaks to this question. Figure 2*A* replicates Almond's (2006) figure 1, showing that the 1919 birth cohort completed less education than predicted by trend. Figures 2*B*–2*D* display paternal characteristics for males by birth year from 1912 through 1919, drawing on the full count of the 1920 census (Ruggles et al 2021a).

Relative to those of the comparison cohorts, fathers of the 1919 birth cohort have lower SES: they are less likely to be literate (fig. 2*B*), have lower average occupation income scores (fig. 2*C*), and score lower on Duncan's

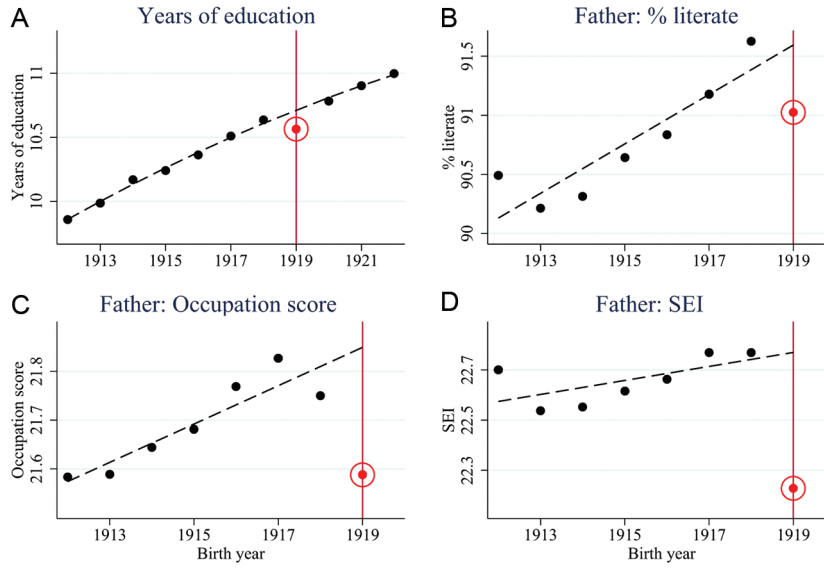

FIG. 2.—Own years of education and paternal characteristics by birth cohort. *A* uses a 1% sample of native-born US males in the 1960 census. *B–D* use fathers of children born in the United States in 1912–19, as reported in the 1920 full-count census. Estimated cohort trends are quadratic in birth year in *A* and linear in *B–D*.

SEI (fig. 2*D*). This evidence of negative paternal selection on characteristics that are known to predict the SES of their children casts doubt on the exchangeability assumption.

To directly test that assumption, model (1) is reestimated, replacing the dependent variable, adult SES, with background characteristics. If the coefficient on the 1919 birth cohort indicator is statistically significant, then the exchangeability assumption is rejected and the estimate of β_3 in model (1) cannot be interpreted as the causal effect of in utero influenza exposure on adult SES. Estimates are reported in column 1 of table 2, using the paternal SES indicators displayed in figure 2 in addition to other background characteristics from the 1920 census. Relative to fathers of the 1912–18 birth cohorts, fathers of the 1919 cohort have significantly lower SES: they are less likely to be literate, have a lower Duncan SEI score, and work in professions with a lower occupation income score. The fathers are also more likely to be born outside the United States and are slightly less likely to be second-generation immigrants. The 1919 birth cohort are less likely to live in a family-owned home and more likely to be nonwhite, to be born in the South, and to have older fathers and more older siblings. All of the estimated differences in column 1 of the table are significantly different from zero.

Age heaping is a legitimate concern in the 1920 census (Myers 1954; Coale 1955; A'Hearn, Baten, and Crayen 2009). If less educated parents

TABLE 2
ESTIMATED DEVIATION OF 1919 BIRTH COHORT'S BACKGROUND CHARACTERISTICS RELATIVE TO SURROUNDING COHORTS

	1920 Census Data		1960 Census Data		1930–WWII Linked Data	
Background Characteristic	1912–19 Cohorts	1915–19 Cohorts	1912–22 Cohorts	1912–19 Cohorts	1912–22 Cohorts	1912–19 Cohorts
	(1)	(2)	(3)	(4)	(5)	(6)
Father cannot read or write	.012*	.012*			.004*	.005*
	(.001)	(.001)			(.001)	(.002)
Father's Duncan SEI	−.719*	−.561*			−.455*	−.517*
	(.037)	(.059)			(.081)	(.151)
Father's occupation income score	−.218*	−.029			−.194*	−.219*
	(.019)	(.031)			(.040)	(.073)
Father born outside United States	.011*	.021*	.000	.005	.006*	.007*
	(.001)	(.001)	(.004)	(.008)	(.002)	(.003)
Father is second-generation immigrant	−.003*	.009*			−.005*	−.012*
	(.001)	(.001)			(.001)	(.003)
Nonwhite	.014*	.015*	.013*	.017*	.005*	.007*
	(.001)	(.001)	(.003)	(.005)	(.001)	(.001)
Southern born	.027*	.017*	.018*	.030*	.006*	.019*
	(.001)	(.001)	(.005)	(.009)	(.002)	(.003)
Family owns home	−.008*	−.008*			−.000	.001
	(.001)	(.001)			(.002)	(.003)
Father's age when child was born	.387*	.341*			.366*	.410*
	(.016)	(.025)			(.026)	(.047)
No. of older siblings in household	.160*	.153*			.100*	.101*
	(.004)	(.007)			(.005)	(.009)

NOTE.—Robust standard errors are in parentheses.
* Statistically significant at 5% size of test.

are more likely to heap on preferred digits, it is possible that heaping could explain the results comparing the 1912–19 cohorts. Birth registration data were first collected in the United States in 1915 and allow us to compare vital statistics with the number of births reported in the census, taking into account age- and state-specific mortality through 1919 as well as changes in the states covered by vital statistics. As described in Brown and Thomas (2022), there is a very high degree of concordance in the number of children alive at the date of the 1920 census, according to the census, and vital statistics data for each of the 1915–19 birth cohorts.[5] Assuming that this

[5] Age is less likely to be heaped on years for these cohorts in the 1920 census because it is reported in both years and months, whereas, for older cohorts, age at last birthday is reported only in years.

concordance applies to cohorts born outside the registration states, we can shield our estimates from age-heaping concerns by restricting attention to the 1915–19 cohorts. Those results appear in column 2. The patterns of negative selection characterizing the 1912–18 comparisons are replicated in 1915–18 comparisons, except that paternal occupation income scores are not significantly different and fathers of the 1919 cohort are more likely to be second-generation immigrants. From these estimates it is clear that age heaping does not explain the result that the 1919 birth cohort is negatively selected.[6]

A second potential concern is that parental differences in early childhood will have disappeared by adulthood if the mortality-SES gradient is large enough. We therefore turn to background indicators reported in the 1960 census. As shown in columns 3 and 4 of table 2, among those who survived to 1960, fathers of the 1919 birth cohort are not more likely to be foreign born, indicating lower survival for those children. This contrast was reported by Almond for the 1912–22 cohorts, and he concluded that this indicated that there was no evidence that the 1919 birth cohort was different from the comparison cohorts. However, for both sets of cohort comparisons, those born in 1919 are more likely to be nonwhite and born in the South, two powerful predictors of adult SES. Moreover, the estimates in columns 1 and 4 (for the 1912–19 cohorts), are very similar, indicating that selective mortality does not contaminate our conclusions based on the 1920 census. Contrasting columns 3 and 4, the estimated gaps are smaller for nonwhite and Southern-born when the 1920–22 cohorts are included in the comparison, indicating fertility selection in the post-1919 birth cohorts.

Third, we use a linked sample of the 1912–22 birth cohorts who are observed first as children with their parents in the 1930 census and again as adults. For the adult observation, we follow Parman (2015) and use World War II (WWII) enlistment records. Since enlistment peaked in 1942, using WWII enlistment records increases the odds that we are observing completed education for the youngest cohorts, who would have been only about 18 at the time of census enumeration in 1940.[7] This sample also avoids concerns about selective mortality before enlistment.

[6] Drawing on the 1930 census, Brown and Thomas (2022) show that extending the comparison cohorts to include 1920–22 does not affect the conclusion that the 1919 birth cohort is negatively selected. The disadvantages of using the 1930 census is that the age measures are not well suited for precise cohort definitions, because age is reported in years as of April 1, 1930, and there is a long hiatus between birth of the child and measurement of paternal characteristics. See Thomas (2010) and Brown and Thomas (2022) for details.

[7] An additional advantage of the WWII enlistment records is that enlistees were asked their year of birth rather than their age. The 1930 and 1940 censuses ask only an individual's age at the time of enumeration (April 1), which complicates one's ability to identify members of the 1919 birth cohort.

We refer to our sample as linked because there is not a unique identifier that maps individuals across censuses and to the WWII enlistment records. We focus 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 and Ferrie 2013; Beach et al. 2016) and follows the best practices discussed by Bailey et al. (2020) and Abramitzky et al. (2021). After all given names are standardized (e.g., recoding "Ed" and "Eddie" as "Edward") in both data sets, each enlistment record is matched to every census record where the individual is of the same race, born in the same state, born within 3 years, and has a reasonably close name.[8] A successful enlistment-to-census link is one where only one census record satisfies the above criteria.[9] We restrict 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 1930 census has to be inferred from the reported age as of April 1, 1930. Thus, the inferred birth year must match the enlistment records or be off by 1 year. Requiring consistency mitigates concerns about misreported ages and birth years.[10]

This linked sample of males is not representative of the US population for males for two reasons. First, WWII enlistees were a nonrandom subset of the male population.[11] 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.

[8] 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 name and raw last name is between 0.8 and 1. The Jaro-Winkler string distance imposes a penalty for the number of character changes from one string to the other. A Jaro-Winkler string distance of 1 implies that the two strings are an exact match. 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.

[9] We 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% and a false-positive rate of 1.8%.

[10] Some 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. Relatedly, the age-heaping phenomenon may affect our ability to link to the census but is unlikely to affect our sample, since the enlistment records asked for birth year.

[11] A comparison of WWII veterans and nonveterans in the 1950 census indicates that veterans were younger, whiter, and slightly less likely to have US-born parents than nonveterans (Beach, Ferrie, and Saavedra 2021, 21).

Nonetheless in columns 5 and 6 of table 2, we find, when using the linked enlistee data, that the 1919 birth cohort is significantly different from the comparison cohorts for all indicators other than home ownership. Contrasting the estimates for the 1912–19 cohorts (in cols. 1, 4, and 6), we see that the loss of representativeness in this selected sample of enlistees leads to smaller estimates of β_3 than in the 1920 and 1960 censuses.

All the evidence in table 2 points in one direction: relative to the comparison birth cohorts, the 1919 cohort came from lower-SES environments. The pattern of negative selection does not depend on the choice of comparison cohorts or the data source. The result is important: by rejecting the exchangeability assumption, the 1919 birth cohort deficits in table 1 cannot be interpreted as identifying the causal effect of in utero influenza exposure on adult SES.

C. *Adult SES of Males in the 1919 Birth Cohort Conditional on Background*

This subsection documents that the estimates of the 1919 birth cohort deficit in adult SES reported in Almond (2006) are overstated when background differences are not taken into account. We use four complementary approaches that essentially modify equation (1) by including controls for background differences, P:

$$y_i = \gamma_0 + \gamma_1 \text{YOB}_i + \gamma_2 \text{YOB}_i^2 + \gamma_3 \mathbf{1}[\text{YOB} = 1919] + \gamma_4 P + \nu_i. \quad (2)$$

Our first approach uses the 1960 census, adjusting for four background characteristics in those data: race, birthplace, mother's birth country, and father's birth country.

Our second approach incorporates information from the 1920 census. This method includes race and state of birth fixed effects, but since it is not yet possible to attach individual-level data from the 1920 census to the 1960, 1970, or 1980 censuses, proxies must be constructed for paternal characteristics. For each characteristic, the proxy P_{bsr} is the average over all children born in each state, s, and year of birth, b, calculated separately by race, r, distinguishing whites and nonwhites.

Our third set of estimates draws on the linked enlistee data set. With these data we can include all of the same background characteristics used in our 1920 census proxies model but measure them at the individual level. This approach should increase the precision of our estimates, albeit at the cost of a selected sample that understates the level of parental selection in the 1919 birth cohort. Using the linked enlistee data also allows us to include the 1920–22 birth cohorts and rule out selective mortality as a potential confounder.

All of the previous methods control for differences in background, using observed characteristics or proxies. Our final approach exploits two

features of the enlistee data: many brothers enlisted in WWII, and we are able to identify brothers if we observe them in the same household in 1930. Using this information, we estimate models using the linked enlistee data that additionally include household fixed effects. These estimates have the key advantage that they control for both individual-specific observed characteristics and the unmeasured background characteristics that are shared by brothers.

For the sake of brevity, results reported in table 3 focus on two human capital indicators of males: high school graduation (cols. 1, 2) and years of education (cols. 3, 4). For each outcome we present unadjusted estimates

TABLE 3
CROSS-COHORT ESTIMATES OF 1919 BIRTH COHORT DIFFERENCE RELATIVE TO COMPARISON COHORTS BEFORE AND AFTER ADJUSTMENT FOR BACKGROUND CHARACTERISTICS

		High School Graduation		Years of Education	
Data Source	Cohorts	Unadjusted (1)	Adjusted (2)	Unadjusted (3)	Adjusted (4)
	A. Measured Paternal Characteristics from 1960 Census				
1960 census	1912–22	−.021*	−.015*	−.150*	−.093*
		(.005)	(.005)	(.038)	(.036)
	B. Proxies from 1920 Census for Paternal Characteristics				
1960 census	1912–19	−.022*	−.013	−.188*	−.099
		(.009)	(.009)	(.063)	(.065)
1970 census	1912–19	−.018*	−.008	−.169*	−.060
		(.005)	(.006)	(.038)	(.040)
1980 census	1912–19	−.012*	−.003	−.165*	−.046
		(.004)	(.005)	(.031)	(.032)
1960 census	1915–19	−.035*	−.024	−.211*	−.117
		(.014)	(.015)	(.100)	(.103)
	C. Measured Paternal Characteristics from 1930 Census				
WWII enlistees	1912–22	−.017*	−.009*	−.094*	−.054*
		(.002)	(.001)	(.008)	(.007)
	D. Household Fixed Effects				
WWII enlistees	1912–22	−.013*	−.005	−.065*	−.028
		(.004)	(.004)	(.019)	(.016)

NOTE.—Background characteristics included in panel A are fixed effects for own race, own state of birth, father's country of birth, and mother's country of birth. Panel B background characteristics are race and birthplace fixed effects and a series of proxies. The proxies are averages by birth state, race, and birth year for the following variables from the 1920 census: indicators for an immigrant father, an immigrant mother, a second-generation immigrant father, whether the father moved outside of birthplace before child was born, an illiterate father, an illiterate mother, whether the mother was in the labor force, and whether the family owned the home; father's age when child was born; mother's age when child was born; the difference between father's and mother's ages; father's occupation income score; and birth order. Panel C includes the same set of controls as in panel B, except that the proxies are replaced with individual-level measured parental characteristics from the 1930 census. Panel D replaces parental characteristics in panel C with household fixed effects. Robust standard errors are in parentheses.

* Statistically significant at 5% size of test.

of the 1919 birth cohort gap, $\hat{\beta}_3$ (cols. 1, 3) and adjusted estimates, $\hat{\gamma}_3$ (cols. 2, 4). As shown in panel A, adjusting for the limited characteristics measured in the 1960 census reduces the estimates, but they remain significantly negative. However, as shown in panel B, broadening the set of background controls, albeit with proxies, produces estimates that are smaller in magnitude, and none of the estimates is statistically significant at the 5% level.

Panel C presents results from the linked enlistee sample. Inclusion of background characteristics reduces the point estimates by more than 40%, but the estimates remain significant at the 5% level. In panel D, however, we take into account both observed and unobserved background characteristics by including household fixed effects. Neither of the estimates of the adjusted 1919 birth cohort gap is statistically significant or economically meaningful. For example, the adjusted gap for years of education is 0.028.[12]

When either individual-specific background characteristics or proxies are used, the point estimates of the 1919 birth cohort deficits in education are reduced by between 25% and 80%. There are, almost surely, also unobserved differences between the 1919 and comparison birth cohorts, and when we control for both observed and unobserved background characteristics, neither of the 1919 deficits is statistically significant or economically meaningful.

We have focused on two indicators of education of males. However, in support of his conclusion, Almond (2006) reports results for a broader set of indicators of adult SES for males, females, and nonwhites. These include four additional income- and education-related indicators of SES in the 1960 census as well as the same six measures plus four more indicators (two that are income related and two that are health related) reported in the 1970 and 1980 censuses. Relative to using the enlistee data, an important advantage of the proxy approach to measuring background

[12] Since proxies for background are noisy indicators of own background, those estimates are likely to overstate the magnitude of the 1919 birth cohort deficit. We have checked this potential bias, using the enlistee data. Relative to the uncontrolled deficits for high school graduation and years of education, controlling with proxies reduces the deficit by about 10%. Controlling own parental characteristics further reduces the estimated deficits by 35%–50%. While these magnitudes may reflect less variation in the background of enlistees relative to the population, they do suggest that the estimates in panel B of table 3 are upward biased.

None of the results in table 3 include a control for the total number of children in the household. The total number of children in the household is a function of the number of older siblings and the number of younger siblings in the household. Our birth-order control captures the older-siblings component of that variable. Brown and Thomas (2022) present results controlling for the total number of children (as observed in 1920) and find evidence that the 1919 birth cohort was of higher adult SES than the comparison cohorts. For example, in the 1960, 1970, and 1980 censuses, males have completed 0.29 (SE = 0.06), 0.28 (SE = 0.04), and 0.27 (SE = 0.03) more years of education than males in the 1912–18 cohorts, respectively. It remains an open question whether this pattern tells us something about parental selection and investments in the 1919 birth cohort or whether the results are driven by incomplete fertility of parents of the later cohorts, particularly the 1919 birth cohort.

is that the same SES indicators can be examined for males, females, and nonwhites in all three censuses (Brown and Thomas 2022).

Considering all 26 estimates of the 1919 birth cohort gap for males, relative to the 1912–18 cohorts, eight are negative and statistically significant, while one is significantly positive in models that do not adjust for background. Adjusting for background with proxies reduces the magnitude of the estimated deficit for all but two estimates. Furthermore, only two of the 26 estimates are significant: one is negative, indicating a deficit, whereas the other is significantly positive, indicating an advantage relative to the comparison cohorts.

The patterns for females are similar. Four of the unadjusted estimates indicate a significant deficit, while two indicate a significant advantage; with adjustment for background, two indicate a significant deficit, and three indicate a significant advantage. Among nonwhites, none of the unadjusted estimates is statistically significant, and the only adjusted estimate that is significant indicates that the 1919 birth cohort had higher adult SES than the comparison cohorts. These results do not support the conclusion that the 1919 birth cohort had a significant deficit in adult SES.

D. Using Quarter of Birth to Evaluate the Cross-Cohort Identification Strategy

This section examines an alternative approach that can provide supplementary evidence concerning the validity of the cross-cohort identification strategy. This analysis exploits information on timing of births and examines the pattern of deficits at the quarterly level.

The majority of influenza cases in the United States during the pandemic occurred over the last four months of 1918. While the United States suffered a subsequent influenza outbreak in the spring of 1919 and the virus lingered in some areas for considerably longer, the scale of those incidents is generally below what occurred during the fall of 1918. If the 1919 deficit were driven primarily by in utero exposure during the fall wave, then we would expect the deleterious effects on adult outcomes to be greatest for those born in the first two quarters of 1919 and smallest for those born in the fourth quarter of 1919.

Table 4 explores whether the cohorts exposed in utero during the fall wave drive the 1919 birth cohort's worse socioeconomic standing as adults. Panel A displays the deviation from trend in completed years of education for the 1919 birth cohort relative to the 1915–18 birth cohorts. Then, in panel B, quarter-of-birth fixed effects and four 1919 birth-quarter indicators, which replace the single 1919 birth cohort indicator, are added to the model.

The results of this exercise do not align with the temporal composition of effects that would be expected if the reported deficits to adult

TABLE 4
DIFFERENCES IN COMPLETED YEARS OF EDUCATION OF MALES IN 1919 BIRTH COHORT BY YEAR OF BIRTH AND BY QUARTER OF BIRTH RELATIVE TO 1915–18 COHORTS

	A. Differences by Year of Birth	B. Differences by Quarter of Birth			
Data Source	Born in 1919 (1)	1919Q1 (2)	1919Q2 (3)	1919Q3 (4)	1919Q4 (5)
1960 census	−.211*	−.166	−.226	−.133	−.315*
	(.100)	(.121)	(.121)	(.122)	(.120)
1970 census	−.182*	−.098	−.151*	−.174*	−.320*
	(.061)	(.073)	(.073)	(.073)	(.072)
1980 census	−.088	−.075	−.083	−.029	−.169*
	(.049)	(.059)	(.059)	(.059)	(.058)

NOTE.—Sample sizes are 51,462, 140,082, and 207,318 in the 1960, 1970, and 1980 censuses, respectively. Birth-quarter models in cols. 2–5 replace 1919 year-of-birth cohort indicator in eq. (1) with four 1919 birth-quarter indicators. Models in panel B include birth-quarter fixed effects. Robust standard errors are in parentheses.

* Statistically significant at 5% size of test.

SES were purely a result of in utero influenza exposure to the fall wave of the pandemic. All of the quarterly estimates are statistically equal, and the results of table 4 indicate that, if anything, the deficit was most pronounced among those born in the fourth quarter 1919, that is, those conceived after the fall wave.

E. Summarizing the Cross-Cohort Identification Strategy

We have established that the 1919 birth cohort is negatively selected on parental characteristics. This fact invalidates the assumptions necessary to interpret the cross-cohort results presented in Almond (2006) as causal. Moreover, we have shown that taking this selection into account is important. The magnitude of the estimated adult SES deficit of the 1919 birth cohort is reduced in all models that include observed background controls, and when estimates take into account both observed and unobserved characteristics, the deficit is not statistically different from zero. Fine-grained analyses that focus on quarter of birth provide little evidence that the deficit in adult SES found for the 1919 birth cohort fits the temporal pattern expected if it was solely capturing in utero pandemic exposure during the fall wave.

Our results cast doubt on the identifying assumptions necessary to successfully implement a cross-cohort strategy to recover the causal effect of in utero exposure to the 1918 pandemic on long-run outcomes. Moreover, these issues are not relevant for the US context only, as similar concerns arise with other studies that use the cross-cohort strategy in other countries (see Vollmer and Wójcik 2017 and Brown and Thomas 2022). Overall, our conclusions imply that in order to make progress on understanding the persistent impact of in utero exposure to the influenza

pandemic, we need an alternative estimation strategy that relies on variation that can plausibly be considered quasi-random.

III. Reassessment of Dose-Response Approach

A. *Replication of Almond's Results*

The second approach used in Almond (2006) to identify the causal effect of in utero exposure to the pandemic exploits spatial and temporal variation in virulence of the influenza pandemic. Restricting attention to the 1918–20 birth cohorts in order to isolate the effect of fetal exposure, he investigated how each adult outcome, y_i, varies with a function of year- and state-specific maternal mortality rates in the year before birth, $t - 1$ that he used to estimate the maternal infection rate (MIR):

$$y_i = \alpha_0 + \alpha_1 \text{MIR}_{s,t-1} + \mu_s + \mu_t + \epsilon_{ist}, \tag{3}$$

where y_i is adult SES, μ_s and μ_t are state and birth-year fixed effects, respectively.

Using the 1960 census, Almond reports statistically significant negative dose-response effects for three of five SES indicators for males (high school graduation, year of education, and log of total income). He comments that the 1970 census estimates for males "do not approach statistical significance" and that the "1980 estimates are about half as large as the corresponding 1960 effects" (Almond 2006, 705–6).

Brown and Thomas (2022) replicate these results after correcting two data issues. Almond assigns a maternal mortality rate of 6.3 for Virginia in 1919, whereas the rate recorded in US PHS (1947) is 8.3. Brown and Thomas also use data from Washington, DC, that is recorded in the same source but was not used in Almond's analysis. After correction of those errors, only two of the five SES indicators are significantly related to MIR (table 5, col. 1). Results for males in the 1970 and 1980 censuses are reported in columns 2 and 3, respectively, of table 5. Whereas in 1960, males who were born in states with higher levels of excess maternal mortality were significantly less likely to have graduated from high school and completed significantly fewer years of education, by 1970, as shown in column 2 of table 5, the dose-response estimates indicate that these same males were no less likely to have graduated from high school and reported having completed significantly more years of education. The reason for the reversal of the results is unclear. The difference between the 1960 and 1970 census estimates is more than a year of education, and it is very unlikely that these men completed more years of education in their forties. Moreover, in 1970, the dose-response estimates indicate that males are also significantly less likely to be poor. By 1980, none of the dose-response estimates is statistically significant (table 5, col. 3).

TABLE 5

Estimates of Relationships between Maternal Influenza Infection Rates and Adult SES for Males, Females, and Nonwhites in 1960, 1970, and 1980 Censuses

	A. Males			B. Females			C. Nonwhites		
Adult SES Indicator	1960 Census (1)	1970 Census (2)	1980 Census (3)	1960 Census (4)	1970 Census (5)	1980 Census (6)	1960 Census (7)	1970 Census (8)	1980 Census (9)
High school graduate	−.085*	.015	.006	.025	.030	−.021	−.325	−.052	.132
	(.040)	(.022)	(.019)	(.044)	(.024)	(.017)	(.193)	(.105)	(.081)
Years completed schooling	−.694*	.407*	.108	.049	.189	−.078	−.973	.535	.317
	(.310)	(.170)	(.103)	(.263)	(.127)	(.107)	(1.336)	(.444)	(.565)
Log(total income)	−.162	.060	−.065	−.113	.051	.085	.396	.011	.330
	(.088)	(.051)	(.041)	(.118)	(.040)	(.052)	(.320)	(.222)	(.238)
Poor (<1.5 × poverty level)	.031	−.064*	.013	.003	−.004	−.022	−.245	−.136	.027
	(.032)	(.017)	(.011)	(.034)	(.011)	(.017)	(.200)	(.079)	(.087)
Duncan SEI	−2.336	2.633	−1.386	−2.833	.729	.686	−.865	12.163*	−1.768
	(1.984)	(1.671)	(1.004)	(1.609)	(1.033)	(.993)	(7.652)	(3.378)	(3.588)
Observations	16,659	46,238	71,048	17,164	49,440	80,916	1,866	5,319	8,232

Note.—Standard errors, in parentheses, are clustered at the levels of state and year of birth.

* Statistically significant at 5% size of test.

Estimates for females and nonwhites are displayed in panels B and C of table 5, respectively. The only significant estimate (of 30 estimates) indicates that nonwhites born in states with higher levels of excess maternal mortality had higher SES in 1970.

This evidence does not support the conclusion that there is a significant negative dose-response effect using the approach taken by Almond. Specifically, of 45 estimated coefficients, only two indicate a statistically significant negative link between adult SES and excess maternal mortality, and those estimates are not consistent over time. In contrast, three of the estimates indicate that the link is significantly positive. Given the number of comparisons, it is appropriate to adopt a testing procedure that takes into account the multiple comparisons in these analyses (Hochberg 1988). In that case, the evidence indicates that variation in the intensity of exposure to the 1918 influenza pandemic in utero has no statistically significant impacts on SES in adulthood.

B. Extending the Dose-Response Approach

This section extends the dose-response framework to assess whether within-cohort comparisons are a viable path to identifying the causal effect of in utero exposure to the pandemic. While the cross-cohort framework asks whether individuals born in 1919 performed worse than individuals from adjacent birth cohorts, the dose-response framework goes farther, 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. The results in section III highlight the need for new dose-response evidence, in turn motivating the remainder of our paper.

C. An Improved Measure of Exposure

Our linked data allow us to derive a more localized measure of pandemic exposure. We apply our same linking algorithm to link male WWII enlistees to the 1920 census. Next, 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, published in the Census Bureau's *Mortality Statistics* publications starting in 1900. Although the data include only information for registration states and cities, this data source allows us to leverage variation from nearly 300 cities.[13]

[13] Registration 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

One concern is that influenza mortality captures more than the severity of the pandemic.[14] Clay, Lewis, and Severnini (2018) show that during the pandemic, mortality rates were higher in places with more coal pollution and worse water quality. These relationships may be 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, Rossin-Slater, and Walker 2017) or poor water quality (Beach et al. 2016) also impairs human capital development.

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(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 available only 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 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.

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.

[14] When the 1918 influenza strain killed, it tended to kill quickly. Accordingly, influenza fatality rates track the case rates reasonably well on a weekly basis, but with a 2–3-week lag (see app. fig. 4 in Beach, Ferrie, and Saavedra 2021). This lag is unlikely to matter for our analysis, which uses an annual measure of mortality.

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

[16] The natural logarithm ensures that predicted influenza deaths are always greater than zero.

D. Empirical Approach

The starting point for our analysis is

$$y_{ibc} = \alpha_0 + \beta_b + \gamma_c + \delta \mathbf{1}[\text{YOB} = 1919] \times \text{Flu}_c + \epsilon_{ibc}, \quad (4)$$

where y_{ibc} is either a background characteristic or a long-run outcome y of individual i from birth year b in birth city c. The parameters β_b and γ_c are birth-year and birth-city fixed effects, respectively. The variable Flu_c measures pandemic intensity in city in c in 1918. Our main measure is the ratio of total influenza deaths to expected influenza deaths, which we normalize by dividing by the sample mean (36.42).[17] This normalization allows coefficients to be interpreted as the average effect of pandemic exposure.

The identifying assumption 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 whether there are meaningful deviations before treatment.

Figure 3 presents estimates with 1917 as the omitted period. Relative to this omitted group, the only coefficient that is statistically different is the coefficient for the 1919 birth cohort. That coefficient is also the most negative. The second-most-negative coefficient corresponds to the 1918 birth cohort, and, importantly, the 1919 and 1918 estimates are statistically equal. However, the 1918 cohort is not a clear placebo cohort, as anyone born between October and December of 1918 may have been exposed to the pandemic during the final trimester.

E. Assessing Family Selection

Table 6 examines whether background characteristics were measurably different for treated cohorts in the linked data. We draw on the same comprehensive background characteristics used in the previous section. Each row corresponds to a different characteristic. We present results from two samples. Columns 1 and 2 correspond to a wide set of cohorts (1912–19), 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. When we turn to our long-run estimates, this restriction

[17] Flu_c is an annual measure. A more flexible estimation strategy would involve obtaining monthly or quarterly data and applying a variation of this strategy that more precisely leverages the timing of exposure. This type of approach would also require month-of-birth information, which is available for some of the cohorts that we observe in 1920 but is often missing and contains heaping on six months, which raises questions about the quality of that variable.

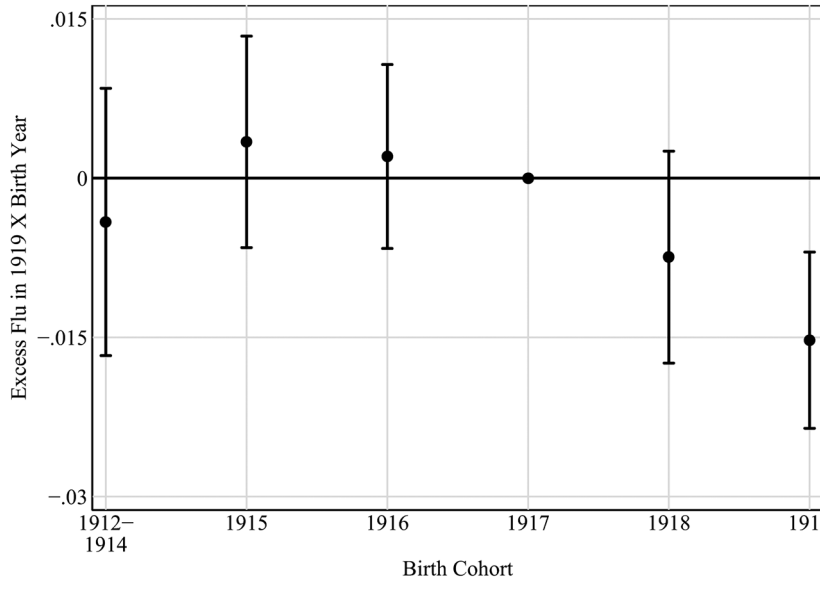

FIG. 3.—Event-study estimates of the impact of pandemic exposure on high school completion: regression estimates and 95% confidence intervals from models that include city of enumeration and birth-year fixed effects as well as background controls listed in table 3. Standard errors are clustered at the city level.

limits the set of observable and unobservable differences that may confound our estimates, but it is somewhat conservative in its ability to isolate the effect of in utero exposure, since individuals born in the final months of 1918 were also exposed to the pandemic while in utero.

Table 6 provides evidence that the dose-response framework offers a more credible identification strategy when we consider bias from parental selection. Relative to table 2, the estimates are largely statistically insignificant, and the point estimates are often smaller. Applying the cross-cohort strategy to our 1920 sample indicates that 12 of the 15 estimated deviations for the 1919 birth cohort were significant at the 5% level or lower (Beach, Ferrie, and Saavedra 2021). In the dose-response framework (col. 2), the only deviation that is significant at the 5% level or lower relates to the number of older siblings in the household. In terms of magnitude, the cross-cohort estimates indicate that the 1919 birth cohort had 0.15 more older siblings. Our dose-response estimates indicate that, in the average pandemic city, members of the 1919 cohort had 0.05 more older siblings.

F. Long-Run Estimates

The first two rows of table 7 use our dose-response framework to examine the impact of pandemic intensity on educational attainment. We

TABLE 6
ASSESSING WHETHER BACKGROUND DIFFERENCES ARE RELATED TO PANDEMIC INTENSITY IN WWII–1920 LINKED SAMPLE

	1912–19 Birth Cohorts		1918–19 Birth Cohorts	
Background Characteristic	Mean [SD] (1)	Standard Flu × Born in 1919 (2)	Mean [SD] (3)	Standard Flu × Born in 1919 (4)
Father cannot read or write	.052	.002	.051	−.001
	[.222]	(.002)	[.221]	(.003)
Father's Duncan SEI	32.992	−.288	32.601	−.339
	[21.684]	(.195)	[21.436]	(.233)
Father's occupation income score	28.692	−.169	28.538	−.206
	[9.172]	(.093)	[9.039]	(.112)
Father born outside United States	.440	.002	.427	−.000
	[.496]	(.004)	[.495]	(.005)
Father is second-generation immigrant	.247	−.002	.250	−.000
	[.431]	(.006)	[.433]	(.004)
Nonwhite	.028	−.001	.030	−.001
	[.164]	(.001)	[.170]	(.001)
Family owns home	.306	−.009	.267	−.007
	[.461]	(.005)	[.442]	(.005)
Father's age when child was born	32.108	.137	32.130	.105
	[6.970]	(.127)	[6.926]	(.096)
Number of older siblings in household	2.620	.049*	2.660	.040*
	[1.750]	(.020)	[1.790]	(.018)

NOTE.—Standard errors, in parentheses, are clustered at the city level.
* Statistically significant at 5% size of test.

consider total years of schooling and an indicator for graduating from high school. The third row of coefficients in the table displays results using height as the dependent variable. Columns 1 and 2 examine patterns among the 1912–19 birth cohorts. Columns 3 and 4 focus on just the 1918 and 1919 birth cohorts, where we expect the set of observed and unobserved differences to be more limited, while columns 5 and 6 focus on our brothers subsample. Our baseline estimates are presented in columns 1, 3, and 5. Columns 2 and 4 add the same parental controls that were included in our preferred cross-cohort regressions, while column 6 goes farther and includes household fixed effects. The results in table 7 point to a negative relationship between in utero exposure to the pandemic and educational attainment.

Relative to the cross-cohort results, there is less evidence that the dose-response results are driven by parental selection. 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

TABLE 7
IMPACT OF PANDEMIC INTENSITY ON ADULT OUTCOMES FOR 1919 BIRTH COHORT IN WWII–1920 LINKED SAMPLE

	ESTIMATED EFFECT OF STANDARD EXCESS FLU × BORN IN 1919					
	1912–19 Birth Cohorts		1918–19 Birth Cohorts		1912–19 Birth Cohorts (Brothers Sample)	
DEPENDENT VARIABLES	Unadjusted (1)	Adjusted (2)	Unadjusted (3)	Adjusted (4)	Unadjusted (5)	Adjusted (6)
1. Years of schooling	−.065*	−.045*	−.036	−.023	−.072	−.104*
	(.020)	(.017)	(.020)	(.019)	(.049)	(.050)
2. Graduated high school	−.017*	−.013*	−.011*	−.009	−.023*	−.039*
	(.004)	(.004)	(.005)	(.005)	(.012)	(.012)
Observations (rows 1, 2)	148,550	148,550	56,756	56,756	12,864	12,864
3. Height	.036	.047*	−.013	−.014	.106	−.050
	(.021)	(.020)	(.032)	(.038)	(.062)	(.073)
Observations (row 3)	113,609	113,609	45,831	45,831	7,991	7,991

NOTE.—Adjusted specifications include the same background controls described in the note to table 3 for panels C and D. Standard errors, in parentheses, are clustered at the city level.

* Statistically significant at 5% size of test.

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 the 1918 and 1919 birth cohorts and thus try to isolate the in utero effect by making comparisons with individuals who were exposed somewhere between the last trimester of the in utero period and age 1. These estimates indicate that in utero exposure lowered high school graduation rates by 1.1 percentage points. Once we adjust for background differences, the point estimate falls to 0.9 percentage points and is significant only 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 was born), the deficit increases to 3.9 percentage points.

There are several important limitations to mention with regard to this analysis. The first relates to the causal mechanism underpinning the negative estimates. We do not observe whether individuals were exposed to influenza. Our best estimate is the intensity of the pandemic, which means that it is impossible for us to separate the impact of influenza exposure from stress, price fluctuations, or other factors related to the pandemic intensity (see Beach, Clay, and Saavedra 2022). Second, our sample is not population representative, as the sample contains no women and relatively few nonwhite individuals, for whom Brown and Thomas (2022)

report no evidence of negative dose-response effects. Third, because of privacy restrictions, we are not able to follow the enlistees after WWII, and some may have resumed their education. Fourth, income is not available in the WWII records, and even if it were, those incomes would likely not reflect peak midcareer earnings. Whether these effects persist in later censuses is an open question that can be investigated when those censuses become publicly available.

IV. Conclusion

Almond (2006) reports that, relative to surrounding birth cohorts, the 1919 birth cohort in the United States attained lower levels of adult SES. Since this birth cohort was in utero during the 1918 influenza pandemic, this result has been interpreted as evidence of the long-term economic effects of in utero exposure to health insults. A key assumption underlying this inference is that the 1919 birth cohort is exchangeable with surrounding birth cohorts.

This paper documents that the 1919 birth cohort was born into lower-SES environments, relative to adjacent birth cohorts. We establish that this pattern exists in Almond's own 1960 sample, in the 1920 full-count census, and in a sample of WWII enlistees linked to the 1930 census. The surrounding cohorts are not exchangeable, which invalidates the assumptions necessary to interpret estimates from the cross-cohort identification strategy as causal.

Furthermore, the cross-cohort estimates of the 1919 birth cohort gap are overstated when background differences are not taken into account. After observed background differences are controlled for, estimates of the impact of in utero influenza exposure on adult SES are between 20% and 80% smaller than those in the unadjusted models. Furthermore, when estimates take into account both observed and unobserved background differences by contrasting outcomes of brothers, the coefficients are attenuated by around 60%, relative to the unadjusted model, and are not statistically significant.

The full impact of WWI remains unaccounted for in these regressions. Over and above troop deployments, the war and the pandemic also introduced greater uncertainty, along with elevated levels of stress and reductions in income and food consumption. Exploiting variation by quarter of birth, we find that the pattern of adult SES deficits contradicts the interpretation that the cross-cohort strategy solely captures the impact of in utero exposure during the fall wave of the pandemic. Our analysis highlights that cross-cohort comparisons of US birth cohorts fail to recover the causal effect of in utero exposure to the 1918 influenza pandemic.

We then interrogate the dose-response strategy. First, we replicate Almond's results using state-level maternal mortality rates. After correcting data errors, we document that 89% of the dose-response estimates are not

statistically significant and that, among those that are significant, more indicate an SES advantage, rather than a disadvantage, of exposure to greater intensity of influenza at the state level. We conclude that there is no evidence in support of a negative dose-response effect when using Almond's original approach.

Finally, we turn to an identification strategy that uses geographic variation in pandemic intensity among 287 cities. This is possible only in our linked data set, which allows us to observe children during their early childhoods. Thus, relative to most censuses, where the finest level of geography is state of birth, we observe what is likely the individual's city of birth (i.e., the place of residence as of January 1, 1920). By using geographic variation in influenza, we are able to include birth-year fixed effects, which 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 exposure to greater pandemic intensity lowered educational attainment among the male enlistees who were successfully linked to the 1920 census when comparisons are drawn either between the 1912–19 birth cohorts or among brothers in a household fixed effects model. On the other hand, when we restrict attention to the 1918–19 birth cohorts, to help rule out unobserved differences across cohorts, none of the estimates is statistically significant.

Whether fetal health shocks do have long-lasting impacts on SES is an extremely important topic for science and policy. In this paper, our conclusions are based on the most thorough evaluation of this question in the context of the 1918 influenza pandemic in the United States that is possible, given the data available. As new data and approaches become available, it behooves the field to advance the science on this important and timely question.

References

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, James Feigenbaum, and Santiago Pérez. 2021. "Automated Linking of Historical Data." *J. Econ. Literature* 59 (3): 865–918.

A'Hearn, Brian, Jörg Baten, and Dorothee Crayen. 2009. "Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital." *J. Econ. Hist.* 69 (3): 783–808.

Almond, Douglas. 2006. "Is the 1918 Influenza Pandemic Over? Long-Term Effects of *In Utero* Influenza Exposure in the Post-1940 U.S. Population." *J.P.E.* 114 (4): 672–712.

Ayers, Leonard P. 1919. *The War with Germany: A Statistical Summary*. 2nd ed. Washington, DC: Government Printing Office. http://digital.library.unt.edu/ark:/67531/metadc276266/.

Bailey, Martha J., Connor Cole, Morgan Henderson, and Catherine Massey. 2020. "How Well Do Automated Methods Perform in Historical Samples? Evidence from New Ground Truth." *J. Econ. Literature* 58 (4): 997–1044.

Barker, David J. P. 1990. "The Fetal and Infant Origins of Adult Disease." *BMJ* 301 (6761): 1111.

Beach, Brian, Karen Clay, and Martin Saavedra. 2022. "The 1918 Influenza Pandemic and Its Lessons for COVID-19." *J. Econ. Literature* 60 (1): 41–84.

Beach, Brian, Joseph Ferrie, and Martin Saavedra. 2021. "Fetal Shock or Selection? The 1918 Influenza Pandemic and Human Capital Development." Working Paper no. 24725 (April), NBER, Cambridge, MA.

Beach, Brian, Joseph Ferrie, Martin Saavedra, and Werner Troesken. 2016. "Typhoid Fever, Water Quality, and Human Capital Formation." *J. Econ. Hist.* 76 (1): 41–75.

Boberg-Fazlic, Nina, Maryna Ivets, Martin Karlsson, and Therese Nilsson. 2021. "Disease and Fertility: Evidence from the 1918 Spanish Flu epidemic in Sweden." *Econ. and Human Biology* 43:101020.

Brown, Ryan. 2020. "The Intergenerational Impact of Terror: Did the 9/11 Tragedy Impact the Initial Human Capital of the Next Generation?" *Demography* 57 (4): 1459–81.

Brown, Ryan, and Duncan Thomas. 2022. "On the Long-Term Effects of the 1918 US Influenza Pandemic." BREAD Working Paper no. 569, Bureau for Research in the Economic Analysis of Development, London.

Bureau of the Census. 1919. *Mortality Statistics 1917: Eighteenth Annual Report.* Washington, DC: Government Printing Office.

———. 1920. *Mortality Statistics 1918: Nineteenth Annual Report.* Washington, DC: Government Printing Office.

———. 1921. *Mortality Statistics 1919: Twentieth Annual Report.* Washington, DC: Government Printing Office.

Chandra, Siddharth, Julia Christensen, Svenn-Erik Mamelund, and Nigel Paneth. 2018. "Short-Term Birth Sequelae of the 1918–1920 Influenza Pandemic in the United States: State-Level Analysis." *American J. Epidemiology* 187 (12): 2585–95.

Clay, Karen, Joshua Lewis, and Edson Severnini. 2018. "Pollution, Infectious Disease, and Mortality: Evidence from the 1918 Spanish Influenza Pandemic." *J. Econ. Hist.* 78 (4): 1179–209.

Coale, Ansley J. 1955. "The Population of the United States in 1950 Classified by Age, Sex, and Color—A Revision of Census Figures." *J. American Statis. Assoc.* 50 (269): 16–54.

Goeken, Ronald, Yu Na Lee, Tom Lynch, and Diana Magnuson. 2017. "Evaluating the Accuracy of Linked U. S. Census Data: A Household Linking Approach." Working Paper no. 2017-1 (December), Minnesota Population Center, Minneapolis.

Heckman, James J. 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science* 312 (5782): 1900–1902.

Hochberg, Yosef. 1988. "A Sharper Bonferroni Procedure for Multiple Tests of Significance." *Biometrika* 75 (4): 800–802.

Isen, Adam, Maya Rossin-Slater, and W. Reed Walker. 2017. "Every Breath You Take—Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970." *J.P.E.* 125 (3): 848–902.

Kitchens, Carl T., and Luke P. Rodgers. 2020. "The Impact of the WWI Agricultural Boom and Bust on Female Opportunity Cost and Fertility." Working Paper no. 27530 (July), NBER, Cambridge, MA.

Long, Jason, and Joseph Ferrie. 2013. "Intergenerational Occupational Mobility in Great Britain and the United States since 1850." *A.E.R.* 103 (4): 1109–37.

Mamelund, Svenn-Erik. 2004. "Can the Spanish Influenza Pandemic of 1918 Explain the Baby Boom of 1920 in Neutral Norway?" *Population* 59 (2): 229–60.

Mansour, Hani, and Daniel I. Rees. 2012. "Armed Conflict and Birth Weight: Evidence from the al-Aqsa Intifada." *J. Development Econ.* 99 (1): 190–99.
Myers, Robert J. 1954. "Accuracy of Age Reporting in the 1950 United States Census." *J. American Statis. Assoc.* 49 (268): 826–31.
Nudd, Jean. 2004. "U.S. World War I Draft Registrations." *Yesterdays* 24 (1): 34–41.
Parman, John. 2015. "Childhood Health and Sibling Outcomes: Nurture Reinforcing Nature during the 1918 Influenza Pandemic." *Explorations Econ. Hist.* 58:22–43.
Roseboom, Tessa, Susanne de Rooij, and Rebecca Painter. 2006. "The Dutch Famine and Its Long-Term Consequences for Adult Health." *Early Human Development* 82 (8): 485–91.
Rotwein, Eugene. 1945. "Post–World War I Price Movements and Price Policy." *J.P.E.* 53 (3): 234–57.
Ruggles, Steven, Catherine Fitch, Ronald Goeken, et al. 2021a. *IPUMS Ancestry Full Count Data: Version 3.0* [data set]. Minneapolis: IPUMS.
Ruggles, Steven, Sarah Flood, Sophia Foster, et al. 2021b. *IPUMS USA: Version 11.0* [data set]. Minneapolis: IPUMS. https://doi.org/10.18128/D010.V11.0
Sanders, Nicholas J. 2012. "What Doesn't Kill You Makes You Weaker: Prenatal Pollution Exposure and Educational Outcomes." *J. Human Resources* 47 (3): 826–50.
Thomas, Duncan. 2010. "Health and Socioeconomic Status: The Importance of Causal Pathways." In *People, Politics, and Globalization: Annual World Bank Conference on Development Economics—Global*, edited by Justin Yifu Lin and Boris Pelskovic, 355–83. Washington, DC: World Bank.
US PHS (Public Health Service). 1947. *Vital Statistics Rates in the United States, 1900–1940*. Washington, DC: Government Printing Office.
Vandenbroucke, Guillaume. 2014. "Fertility and Wars: The Case of World War I in France." *American Econ. J. Macroeconomics* 6 (2): 108–36.
Vollmer, Sebastian, and Juditha Wójcik. 2017. "The Long-Term Consequences of the Global 1918 Influenza Pandemic: A Systematic Analysis of 117 IPUMS International Census Data Sets." Discussion Paper no. 242, Courant Res. Centre, Göttingen.