Crises and Educational Attainment*

Pawel Janas†

This Draft: February, 2024

Abstract

Between 1910 and 1940, the high school graduation rate in the United States increased five-fold, setting the stage for human capital-led economic growth throughout the 20th century. In this paper, I study the contribution of the Great Depression to this structural shift, and investigate the impact of the crisis on schooling choices for households across the socioeconomic spectrum. I use novel local labor market data to study the effect on secondary education, occupation choice, and wages. Using brothers, I attempt to explain within-household variation in schooling across Depression severity. I find that the crisis created two sets of winners: those from blue-collar households who were making their secondary-school choice at the very start of the Depression, and those from white-collared households who were making that same choice but after the Depression was well under way. These men entered, and stayed in, secondary schools when their labor market opportunity cost shrank but not when their household constaints were binding. The return to education for these groups was high: by 1940, they attained more schooling, worked in more prestigious occupations, and obtained higher wages.

JEL Classification: N3, I2, H4
Keywords: Great Depression, youth unemployment, education attainment, crises

*A previous version of this paper circulated under the title "Crises and Intergenerational Mobility." I would like to thank Mike Andrews, Scott Baker, Carola Frydman, Menaka Hampole, Phil Hoffman, Adam Jørring, Joel Mokyr, Jean-Laurent Rosenthal, Andrew Sinclar, Paola Sapienza, Sarah Quincy, Adriana Troiano, Nic Ziebarth, as well as seminar participants at the Northwestern Economic History Seminar, the Kellogg Finance Brown Bag, and the Western Economic Association International annual meeting. This research has been supported by Kellogg Dean's Office Research Grant, the Dissertation Fellowship from the Economic History Association, and the Center for Economic History at Northwestern.

†Division of the Humanities and Social Sciences, Caltech, email: pawel@caltech.edu, site: paweljanas.com

1 Introduction

Investment in education varies with macroeconomic conditions that alter the available household resources and opportunity cost of schooling. For example, an economic crisis can lower opportunity cost of education and push youth into schools (e.g., fewer low-skill jobs), while also tightening the household budget constraint and push youth into labor (e.g., household work). The relative strength of these forces naturally varies heterogeneously across regions, time, and socioeconomic strata. When a crisis hits, do local differences in opportunities affect the levels of educational attainment? If so, who are the winners and losers of crises, and what are the aggregate implications for economic growth?[1]

This paper studies the schooling behavior of young, urban males in the U.S. during the Great Depression. I find that the Depression created two sets of winners: men from blue-collar households who were making their secondary-school choice at the very start of the Depression, and men from white-collared households who were making that same choice but after the Depression was well under way. These men entered, and stayed in, secondary schools when their labor market opportunity cost shrank but not when their household constaints were binding, consistent with basic prediction of human capital investment theory under credit constraints (Becker (2009)). The return to education for these two groups of men was high: by 1940, they attained more schooling (0.11 and 0.24 years), worked in more prestigious occupations, and obtained higher wages (2.5 and 7 percent).

In the aggregate, the results show that the average Depression-era youth stayed in school for 0.1 - 0.2 more years due to the increase in youth unemployment, with significantly higher achievement rates for those in the Rust Belt, the Northeast, and the West North Central regions of the country. Empirically, I combine full-count Census records with novel local youth unemployment rates to quantify the effect of the change in local youth labor markets during the Great Depression on the educational attainment of males in the 1908-

[1] Consider a two-generational, low-educated, and poor household in financial distress. The household head may decide that the opportunity cost of education is too high for their child: parents may pull youth out of school and push them into the labor market or household production, solidifying the parent-child link in educational attainment at the lower end of the socioeconomic spectrum. At the same time, wealthy households have no such constraints: during a crisis, children of highly educated parents continue going to school, reinforcing the intergenerational link at the upper levels of schooling. In this example, macroeconomic busts reduce educational mobility. More broadly, the changes in intergenerational mobility depend on the labor market opportunities for youth and household constraints.

1918 birth cohorts. I compare outcomes of brothers in 1940 for cohorts that turned 18 years old right before the Depression started (1926 – 1928) with those who turned 18 during it (1931-1937).[2]

The historical context provides several empirical advantages of studying the effect of macroeconomic crises on schooling. First, the effect of the Depression was unevenly felt across the country, creating considerable variation in youth opportunities and economic deprivation across local labor markets. Unlike a regular recession, these shocks were particularly large and forced households into difficult decisions about youth schooling.[3] Second, most states in the U.S. today have laws preventing youth from entering the formal labor market, making studying this relationship impractical with modern survey data. In contrast, youth labor was much more common in the first half of the 20th century. Lastly, the availability of microeconomic Census records of the whole population during and after the high-school movement provides measurable short- and long-run outcomes and permits a holistic analysis of heterogeneous effects.

I create my dataset by merging multiple archival sources. The outcome variables come from the 1940 Census, the first federal Census to ask about each respondent's education level.[4] I create my primary sample of linked (1930 – 1940) urban males using 100 percent count U.S. Census records and linking crosswalks provided by Abramitzky et al. (2020). I combine my sample of young males with newly digtitized unemployment-by-occupation-by-age data from the Special Unemployment Census of 1931. Since the Unemployment Census canvassed only 18 regionally dispersed cities and three boroughs of New York City, I estimate youth unemployment for all other cities by taking a weighted average of regional youth unemployment-by-occupation rates, using 1930 occupation-by-city shares aggregated from the 1930 Census full count census as weights. To the best of my knowledge, mine is the first attempt to quantify locally disaggregated and age-specific unemployment rates during

[2]State-level evidence that the Depression shut off employment opportunities for youth and drove them back into schooling was first introduced by Claudia Goldin and Lawrence Katz (e.g., Goldin and Katz (1999)).

[3]See, for example, Elder (2018)

[4]One well-known issue in intergenerational studies using U.S. Census data is that parent-child links are only identified when parents and children cohabit. Because the proportion of children living with their parents drops from 80 percent to 60 percent once the children reach their 20s (Figure A.4), I link the younger cohorts in my sample to their 1930 households and the older cohorts to their 1920 households to obtain parent and household characteristics.

the Great Depression in the U.S. context.[5]

My empirical strategy attempts to explain within-household variation in education attainment across cohorts using a brothers-only difference-in-differences design with across-city variation in unemployment. Notably, youth unemployment is not systematically related to changes in attainment before 1929, providing evidence that the underlying parallel-trends assumption is not prohibitively strong. The experiment thus compares the education outcomes of younger brothers on the cusp of making secondary schooling decisions during the Great Depression with their older brothers who graduated before the Great Depression, conditional on state-level dynamics, national trends, various local time-varying confounders such as banking resources and manufacturing production, and static city determinants of schooling. After quantifying the average effect for each separate level of education, I conduct heterogeneity analyses based on household characteristics based on parental occupation. I conclude my analysis by aggregating the reduced form estimates under two baseline counterfactuals at the state-level.

This paper contributes to two strands of literature. My work chiefly builds on the economic history of the consequences of U.S. educational investments in the first half of the 20th century, specifically the high school movement and the Great Depression (Goldin and Katz (1997); Schmick and Shertzer (2019); Card et al. (2018); Kisswani (2008); Yamashita (2008)). For example, using state-level data, Goldin and Katz (1997) find that graduation rates increased in states with the largest increases in unemployment during the Depression, and Shanahan et al. (1997) finds that Depression-era cohorts in the Stanford-Terman Study of Gifted Children also obtained more schooling. On the other hand, both Yamashita (2008) and Kisswani (2008) find null or small impacts of the Depression on the average attainment using 1960 Census records.[6] My contributions are threefold. First, I improve on the measurement of the opportunity costs channel by bringing in new data on occupation-age-specific unemployment rates of urban males, and extrapolate this measure to the city (not state) level. To the best of my knowledge, mine is the first attempt at quantifying local

[5]Numerous efforts have been made to compute accurate unemployment rates at a higher level of aggregation, notably Sundstrom (1992), Darby (1975), Margo (1991) and Wallis (1989).

[6]Other papers that study the determinants of educational attainment around the same time period are Baker et al. (2020) (boll weevil), Baran et al. (2020) (Great Migration), and Karger (2021) (public libraries), and Stephens Jr and Yang (2014) (compulsory schooling laws).

labor market shocks for youth, which is the key determinant in schooling choices. Second, I utilize an identification strategy that accounts for unobserved selection into schooling. This is an important empirical contribution as the the estimates in the unidentified regression are downward biased towards zero. And third, unlike the existing literature, I consider heterogeneous effects based on the households socioeconomic status, and find that the average effects mask important timing and effect-size differences across households.

This paper also constributes to the literature that broadly studies the elasticity of schooling choices with respect to changes in labor markets. Researchers have shown that local labor market conditions affect education attainment in both developed (e.g., Betts and McFarland (1995); Charles et al. (2015)) and developing economies (e.g., Shah and Steinberg (2017); Bau et al. (2020); Atkin (2016)). Most of this body of work uses trade or industry-specific labor-demand shocks (e.g., natural resources as in Black et al. (2005) and Cascio and Narayan (2015)) and finds that youth discontinue schooling when opportunities *increase* and the skill premium is low. I extend this literature by studying the elasticity during macroeconomic *downturns* from a non-trade perspective. The closest paper is the one by Feigenbaum (2015), who focuses on the effect of the Great Depression on intergenerational mobility in the U.S. using approximately 5,000 males from the Bureau of Labor Statistics Cost of Living Survey. While his focus is on income, not education, mobility, he does not find a causal effect of the Great Depression on education attainment in his sample father-son pairs. The differences in our findings is not surprising given that we use different samples (BLS sample vs. linked population records) as well as measures of Depression severity (change in retail sales vs. youth unemployment).

The rest of the paper is organized as follows. Section 2 lays out the conceptual framework that guides the empirics. Then, Section 3 provides motivating macroeconomic facts about education attainment in the United States during the Great Depression. Section 4 provides an overview of the linked sample of urban males and youth unemployment in 1931. Section 5 discusses the brothers-only difference-in-differences empirical design while Section 6 presents the results and explores heterogeneous effects. Finally, Section 7 provides concluding remarks.

2 Conceptual Framework

I consider two distinct channels of how business cycles can affect human capital decisions in this paper. The first is the youth unemployment channel: as youth unemployment increases during a recession, the likelihood of finding a job and wages decrease, and the opportunity cost of education decreases. The second channel operates through the family budget constraint: as family income decreases (due to *adult* unemployment), the opportunity cost of schooling for a working-age youth increases. We would expect households to consider the trade-offs between these channels, especially during large economic crises when labor markets are in turmoil. Formally, I consider a two-period model of a household's choice between long-term educational benefits and immediate earnings of the child under imperfect credit markets based on human capital theory (Becker (2009)).

A family i consists of a parent who must decide whether to enroll their child in school (binary e_i), aiming to maximize her period 1 consumption (c_i) and her child's period 2 consumption ($\tilde{c}_i$):

$$\log c_i + \log \tilde{c}_i \tag{2.1}$$

The parent's consumption is limited by her income (y_i) and the direct and opportunity costs of schooling (θ_i) if she decides to send her child to school, without the possibility of negative savings:

$$c_i \leq y_i - e_i\theta_i - s_i \tag{2.2}$$

$$s_i \geq 0 \tag{2.3}$$

The child's consumption and earnings in the second period are determined by the schooling choice in the first period and any accumulated (at interest rate r) savings: she receives the skilled higher wages (w_s) if she attended school, otherwise, she earns the lower unskilled wage (w_u):

$$\tilde{c}_i \leq w_u + e_i(w_s - w_u) + (1 + r)s_i \tag{2.4}$$

For the marginal parents who's optimal choice is bound by the non-negative savings constraint ($s_i = 0$), the decision rule of investing in education becomes:

$$\theta_i \leq y_i \frac{w_s - w_u}{w_s} \tag{2.5}$$

This simple framework has clear testable implications: higher adult incomes (y_i) and a larger difference between skilled and unskilled wages ($w_s - w_u$) lead to more educational investments for the child.

Using the 100 percent records of the U.S. Census (Ruggles et al. (2020)), I verify that these (static) implications hold in 1930: if the conceptual model is a good first-order approximation to human capital choice, I should observe children of richer households in places with a large wage premium enrolling in school at higher rates than their poorer counterparts. I measure family resources using the reported house value (the sample excludes renters), total household income as proxied by a the sum of the father's and mother's occupation income score, and whether or not the individual reported being an only-child.[7] I proxy for the city-level wage premium using wage data available in 1940: I compute the difference between average weekly wages for 18-30 year old male high-school graduates and high-school dropouts.[8] I further control for state fixed effects to account for across-state variation in reported house values, incomes, and school-going, and I cluster the standard errors at the state level.

(Table 1 around here)

As of April 1st 1930, roughly 26 percent of 16 and 17 year olds were exclusively

[7]The sample includes males living in a house-owning household (nonzero and non-missing *valueh*) located in a Census identified city. I further restrict attention to single-family households (1 mother (*nmothers*) and 1 father present (*fathers*), (*nfams*) is one) for whom the father's occupation is classified. Occupation scores (*occscore*) for father and mother occupations are the median wages for occupation types as reported in the 1950 Census.

[8]The sample excludes zero wage income (*incwage*) males in the labor force.

working or seeking work with and additional 6 percent seeking work and also attending school . That is, a nontrivial portion of youth were completely or partially out of schooling.[9] Table 1 reports the results of a series of OLS regressions where the outcome variable takes the value of 100 if the individual reported being in school and zero otherwise. The columns use subsamples of 14 through 19 year olds. Starting with 16 year olds, I observe that a 10 percent increase in parental income or housing wealth is associated with a roughly 0.8-1.4 higher probability of an individual enrolling in school. Likewise, youth without siblings report going to school (after age 15) at a 6-10 percent higher rate than those with siblings. Furthermore, the 1940 city-level high school wage premium is also positively associated with 16+ year olds being in school. The correlations in the data show that the basic implications of the human capital model considered here hold in 1930.

During large economic downturns, these dynamics can shift in many ways. A sudden and steep drop in unskilled wages makes education more appealing, due to reduced immediate earning losses when schooling is chosen over work. On the other hand, large crises can also wipe out household savings and lower parental income, which could lead to lower educational investment for the child. While there are other factors that could influence educational decisions during crises, such as increased uncertainty about the future wage premium, my focus is primarily on the effects through contemporaneous youth labor markets.

3 Schooling during the Depression

I begin this section by presenting four macroeconomic facts about trends in high school graduation during the first half of the 20th century - the 1930s in particular - in Panels (A) through (C) of Figure 1. This period of U.S. economic history, typically referred to as the "high school movement," was characterized by a marked increase in the number of youth completing at least 12 years of education (Goldin and Katz (1997)).

(Figure 1 around here)

In the aggregate, trends in high school graduation rates in the U.S. suggest that schooling decisions in the 1930s were, in fact, unique. Using data from the U.S. Department

[9]Figure A.1 plots the proportion of each age cohort in 1930 in these categories.

of Education in Panel A, I plot the ratio of high school graduates to 17-year-olds decennially between 1910 and 1930 and annually after. In 1910, this ratio stood at just 8.8%. While the graduation rate more than tripled to 29% by the eve of the Great Depression in 1929, there was an evident gain relative to the trend beginning in the early 1930s. This uptick did not subside until the U.S. entered the Second World War in the early 1940s.

But did the Depression have anything to do with this acceleration? Panel B offers suggestive evidence that Depression severity at the local level was positively associated with increased educational attainment. I combine a linked sample of males in the 1930 and 1940 Censuses with the change in county-level retail spending between 1929 and 1933 (Fishback et al. (2003)). I restrict attention to persons in this sample from the 1906-1922 birth cohorts who reported living in a city in 1930. I then compute the share of the cohort that reported finishing at least 12 years of education in the 1940 Census, separately by whether or not the individual was living in a county in the top or bottom tercile of the change in retail sales. Panel B shows a slight difference (1 percent) in high school attainment by cohort across the two groups of counties from 1926 until the 1928 graduating cohort. However, beginning with those who turned 14 during and after the Depression started and thus at the point of making high school-going decisions, the graduation rate for those in counties with substantial adverse economic shocks surpassed that of those from the same birth cohort in counties with milder shocks. By 1936, the average rate rose by 12.0% in worse-off counties and only 10% in better-off counties as compared to the 1927 cohort.

This relative gain in worse-off counties differed across households by socioeconomic status. Taking the same sample of linked males as in Panel B, I split the cohorts by local Depression severity and the individual's parent's occupation in 1930 (blue-collar vs. white-collar)[10]. In Panel C, I plot the change in the high school graduation rate separately by local economic condition and family background where the 1930 cohorts serve as the base. The figure reveals that individuals from both backgrounds completed high school at a higher rate in worse-off counties as compared to their peers in better counties. However, the main driver of the aggregate increase was, in fact, individuals from blue-collar families, especially those in places more hit by the Depression.

[10] As I discuss in greater detail in Section 4, linking parents to their kids for older cohorts in 1930 requires linking to the 1920 Census.

These aggregate facts, however, only suggest a heterogeneous impact of the Depression. Only under heroic assumptions could these differential trends be interpreted causally: the use of plausibly noisy estimate of the Depression youth labor markets (retail sales) and omitting several important factors that go into the schooling investment decisions bias these averages. In the remainder of this article, I describe a novel approach of constructing youth unemployment rates and describe my empirical method of comparing the schooling choices of boys who were on the cusp of making high-school going decisions to those of their older brothers.

4 Data Construction

This section describes the construction of the dataset. In section 4.1, I present my method of measuring opportunity cost, which I proxy by the youth unemployment rate. Locally dis-aggregated unemployment data during the Depression for youth (or adults) is not available systematically. Therefore, I use three sources of information to estimate unemployment rates: city-level occupation reports within the state-level publications of the 1930 decennial Census, the Special Unemployment Census of 1931, and the full count records of the 1930 Census publicly available on IPUMS. I describe the linking procedure between the 1920, 1930, and 1940 Census records of young males and their fathers in Section 4.2, the other local economic data that I use to proxy for adult labor market conditions in 4.3.

4.1 Local youth unemployment rates

4.1.1 Unemployment rate (numerator): Special Census of Unemployment 1931

Amid deteriorating labor market conditions in January 1931, the U.S. Congress authorized the Census Bureau to conduct a special Census of Unemployment in 21 urban areas - 18 cities and three boroughs of New York City. The Bureau used the same schedule form and enumerators (to the extent possible) as in the April 1930 Census to make the returns comparable. Since the Bureau's focus was on the characteristics of those unemployed, the reported statistics break down occupational unemployment by sex, age, occupation, marital status, race, and nativity. Enumerators were instructed to visit each household and

Summary of sources to construct youth unmployment rates

	Name	Variable	Level	Scope	Source
(1)	Unemployment Rate - Numerator	Class A and Class B Unemployed - Male	Occupation-city	21 cities*	1931 Special Census of Unemployment (U.S. Census Bureau)
(2)	Unemployment Rate - Denominator	Number of Employed - Male	Occupation-city	same as (1)	1930 Population Census (U.S. Census Bureau)
(3)	Regional occupation-unemployment rates	Average{(1) / ((1) + (2))} across cities within region	Occupation-region	4 regions	Author calculation
(4)	Occupational share	Number of employed in occupation/number in all occupations	Occupation-city	981 cities	1930 Census 100% count records (IPUMS)
(5)	Youth unemployment estimate	$\sum(3) \times (4)$	City	981 cities	Author calculation

ask whether any household member who ordinarily worked at a gainful occupation was unemployed on the preceding day and, if so, ask the specified questions and make detailed entries.

Most of the unemployed fall under two classes, and I collect data on both. Class A contains persons out of a job, able to work, and looking for a job. Across the 1931 census, 20.4 percent of gainful workers from 1930 were classified as Class A unemployed. Class B includes persons having jobs but on lay-off without pay, excluding those sick or voluntarily idle. This class constituted another 3.9 percent of all gainful workers in 1930.

As stated in the introduction to the Special Census statistics, the Bureau published the data in 1931 such that comparisons to 1930 figures would accurately reveal the extent of unemployment in the labor market. Therefore, the age and occupation distribution in these tables was made to conform as closely as possible with the age and occupation distribution of the gainful workers as presented in the 1930 Census. Likewise, the occupations in 1931 were the ones used in 1930.

To obtain the number of unemployed by age group and occupation in each of these cities by 1931, I digitize Table 12 of the Special Unemployment Census of 1931. For example, I observe that 167 deliverymen in the 10-19 age group in Birmingham, Alabama, are Class A unemployed, and 12 are Class B unemployed. In total, I collect data on 21 cities spanning the special enumeration area of the 1931 Census: Boston, Buffalo, New York - Bronx, New York - Brooklyn, New York - Manhattan, Philadelphia, Pittsburgh, Cleveland, Dayton,

Chicago, Detroit, Duluth, Minneapolis, St. Louis, Birmingham, New Orleans, Houston, Denver, Seattle, Los Angeles, and San Francisco.

4.1.2 Unemployment rate (denominator): Census 1930

The total number employed by occupation comes from the April 1930 Census. The employed are "gainful workers" that include everyone ten years old and over who regularly work in an occupation for pay. It does not include women doing housework in their own homes, without wages, and having no other employment, nor does it include children working at home or at odd times on other work. The detailed occupation classification for gainful workers consists of 534 occupations. In the tabulation of the unemployment returns, this list was consolidated to 330.

Employment by occupation for different age groups in 1930 comes from Table 12 in the state-level reports from the 1930 Census. Specifically, this table reports the number of employed persons by occupation in cities of 100,000 or more. Continuing with the example from (1), I observe 458 delivery men enumerated by the Census in the 10-19 age group in Birmingham, AL, in 1930. I collect occupation-city data for the same 21 urban areas enumerated by the special census of unemployment in 1931 in (1)[11].

4.1.3 Constructing regional occupation-unemployment rates

For each occupation in cities reported in both the 1930 and 1931 censuses, I define the youth unemployment rate as:

$$unemp_{ij} = \frac{ClassA_{1931,ij} + ClassB_{1931,ij}}{ClassA_{1931,ij} + ClassB_{1931,ij} + Employed_{1930,_{ij}}} \tag{4.1}$$

where i denotes the occupation, and j denotes the city, and all measures are for the age group 10-19. I then compute the average unemployment rate by occupation for each region by calculating the average occupation unemployment for all cities within the region, weighted by total males aged 10-19 in the labor force as of 1930.

[11]The age brackets are: 10-17, 18-19, 20-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75 and over. The Census reports combined brackets 10-17 and 18-19 in 1931.

4.1.4 Occupational shares

I obtain youth occupational shares for all cities to extrapolate unemployment rates to all Census-enumerated cities by aggregating person-level records from the 100 percent count 1930 Census returns available on IPUMS. My sample includes all 10-19 year-olds reporting an occupation in 1930. The occupation variable in the 100 percent count records was standardized to reflect the 1950 occupational definitions, varying only slightly from those published in Census reports in 1930 and 1931. To merge, I create a crosswalk between 1930/31 and 1950 occupations.

4.1.5 Youth unemployment estimates

Finally, using occupational shares from (4) and the regional rates in (3) I compute average city-level youth unemployment rates:

$$unemp_{j(k)} = \sum_{\forall i} \omega_{i,j} \times unemp_{i,k} \tag{4.2}$$

where $\omega_{i,j}$ denotes the youth occupational share of occupation i in city j and $unemp_{i,k}$ is the unemployment rate of occupation i in region k.

I find significant variation in 1931 unemployment of 10-19 year-olds in the cities enumerated by the Census. Consistent with the literature showing regional patterns of the Depression across the U.S. (Rosenbloom and Sundstrom (1999)), I find that the estimated unemployment was above 40% in industrialized cities specializing in durable goods manufacturing (Buffalo, Detroit, Cleveland) and relatively low (25%) in cities specialized in trade and services (San Francisco, Seattle, Manhattan).[12]

The occupational distribution of youth in these cities drives the variation in total rates. For example, consider the difference between Detroit and San Francisco. The largest share (11.5%) of the youth labor force in Detroit was employed as laborers in the iron and steel industry and experienced a staggering 53% unemployment rate. On the other side of the country, youth in San Francisco primarily worked in low-skill white-collar clerical work, which experienced a much milder Depression of 10.5% youth unemployment.

[12]See Appendix Table A4 for more details.

Not surprisingly, I find strong regional clustering with relatively high rates in the Midwest and Northeast and low rates throughout the South and West.[13] In both the Midwest and Northeast, operatives and laborers in manufacturing constitute much more of the youth labor force - with a higher estimated unemployment rate of between 30 and 40 percent - than in Southern cities. Additionally, the weight placed on these occupations in the computation of the total unemployment rate is considerable, between 20 and 30 percent. On the other hand, the South youth labor force is dominated by servants and retail workers, who saw lower unemployment rates. In only two cities manufacturing laborers make up the largest share in the South, and the weight is below 15 percent.[14]

Within states, I find that youth unemployment was not correlated with the change in per-capita retail sales between 1929 and 1933 at the county level. Table 2 shows the result of five OLS regressions - at the city level - where the dependent variable is mean zero, standard deviation one youth unemployment and the outcomes are analogously standardized measures denoted in the header of the column. I find that one standard deviation in youth unemployment is associated with a positive 0.16 standard deviation in the change of manufacturing output (1929 - 1933), a negative 0.36-0.39 standard deviation in the share of workers employed in wholesale and retail industries (1930), and a positive 0.12 standard deviation in the manufacturing labor force share (1930). As described earlier, these patterns reflect the construction of the estimates based on the types of jobs youth typically held (common laborer in manufacturing) and the relative unemployment rates (high in manufacturing, low in services) in those sectors.

(Table 2 around here)

4.2 Linked Census records of sons and fathers, 1920 - 1940

The primary outcome variable for sons and fathers - education attainment and occupation – comes from the 1940 U.S. Census, the first time questions regarding years of schooling

[13] Appendix Figure A.2 plots $unemp_{j(k)}$ for the full sample of cities.

[14] Online Appendix Table A5 presents the most common youth occupations by region. The column "# Cities" reports the number of cities in which the occupation is the most common, and the "Weight" column reports the share of the youth labor force in that occupation.

and wages appeared in the Decennial Census. Household characteristics and parental information that I use to conduct heterogeneity analysis, however, come from either the 1920 or 1930 Census. This is because of the well-known obstacle in conducting intergenerational studies using U.S. Census data: intergenerational links are identified only if members of both generations live in the same household.[15] This requirement presents an issue to empirical studies in the U.S. because most youth leave the household by their 22nd birthday, and linking an older individual to their parents requires a link to a Census taken during their childhood.[16]

This paper focuses on youth during a short period (the 1930s), and the Census records in 1930 provide good coverage for the sample of 11-17 year-olds in 1930. These are the cohorts making high-school-going decisions during the Depression and graduating high school before 1940, when I can observe education outcomes. However, comparing cohorts with only their older counterparts with intergenerational links as of 1930– those that finished their high-school education before the Depression and still lived with a parent in 1930 - is problematic due to obvious self-selection. Namely, this sample misses all the youth that graduated and established their own households before the Depression. Thus, I complement my sample of older cohorts by linking those who moved out by 1930 to their 1920 households, when they were 8-12 years old.

Starting with the entire population of males between the ages of 11 and 22 in 1930 (16.1 million), I imposed several restrictions to arrive at my primary analysis sample. First, I keep those living in a census enumerated city and drop those living in non-households (e.g., orphanages). Second, for the 10-17-year-olds, I keep only those reporting to be a child of the head of the household, effectively dropping grandsons, nephews, and those living with older siblings as heads of households, as opposed to a parent. Third, for 18-22-year-olds in 1930, I keep only those reporting to be either a child or a head-of-household. For 18-22-year-old head-of-households, I link back to their 1920 records to find their household characteristics and parental information. For 18-22-year-olds reporting to be a child, I use household data

[15] An exception was college students, who were enumerated at their "usual place of abode" and not at their college.

[16] In Figure A.4, I plot the share of each age cohort that lives with a parent in the 1930 Census. While the majority (75 to 95 percent) between 0 to 17 live with a parent, this share drops to 68 percent for 19 and 36 percent for 23-year-olds.

from 1930. The final sample size before linking to 1940 is approximately 4 million.

I use the crosswalks provided by the Census Linking Project (Abramitzky et al. (2020)) and IPUMS publicly available Census data to link records over time using the ABE procedure. In my main analysis, I use records matched using any of the linking methods provided. Overall, I obtain outcome variables in 1940 for 31 – 34 percent of the final sample of sons in each cohort, both in the pre-Depression and Depression groups.[17] The linked sample is not a perfectly representative sample of the urban youth population - children with white-collar fathers, those who lived outside the Southern states, and whites are over-represented. To address this issue, I use inverse probability weighing in my empirical analysis, where the weights are created after predicting the characteristics that are associated with a successful link (see Appendix B.1). Importantly, however, I do not find evidence that the linking probability varies by 1931 youth unemployment across cohorts.[18]

4.3 Other Data

I use several other data sources to control for time-varying local economic conditions in my estimation. First, I use linearly-interpolated Census of Manufactures county-level data on total manufacturing output produced for the years 1927 - 1937, reported biennially (Stewart (1929)). Second, I use total banking deposits at the county level from the Federal Deposit Insurance Corporation for the years 1927 - 1937, reported annually (Federal Deposit Insurance Corporation (1992)). I obtain wholesale and retail shares from the 1930 U.S. Census (ICPSR (2005)).

5 Empirical Framework

I estimate the effect of youth unemployment on education attainment using a pooled difference-in-differences research design that compares individual outcomes across cities be-

[17] Appendix Figure A.3 shows the proportion of each cohort from my final sample linked between 1930 and 1940 (no marker), that lived with a father in 1930 (hallow circle marker), that lived with a father in 1930 *or 1920* and was linked to 1920 (x marker), and finally, those that had either father *or mother* info in either 1920 or 1930 (solid circle marker).

[18] In Appendix Figure A.3, I plot estimated coefficients on cohort fixed effects interacted with youth unemployment when the outcome is a binary taking the value of 1 if a record is linked. These coefficients in my main sample - linked sons with father or mother occ using 1920 or 1930 (yellow circle) – are quantitatively small and statistically insignificant.

fore and after the onset of the Great Depression. Specifically, I estimate regressions of the following form:

$$S_{ijk} = \alpha_j + \beta_k + \sum_{z=1}^{3}(\Delta Unemp_j T_z) \cdot \gamma + C_{jk} \cdot \delta + \Omega_i + \epsilon_{ijk} \tag{5.1}$$

where S_{ijk} is an outcome for person i who reported city of residence j in 1930 and turned 18 (cohort) in year k. $\Delta Unemp_j$ is the standardized (mean zero, standard deviation of one) estimate of youth unemployment in 1931 minus the county-level 1930 unemployment rate, where T_1 - T_3 are dummy variables taking the value of 1 if k = {1926-1928}, {1931 - 1933}, and {1934 - 1937}, respectively. I bin cohorts in this manner to reflect the average early-Depression and late-Depression impacts while excluding the cohorts that were making post-secondary schooling decisions right at the start of the Depression (1929-1930 cohorts). The vector β_k contains cohort fixed effects, α_j includes the city of residence in 1930 fixed effects, C_{jk} is a vector of time-varying city and county control variables, and Ω_i is a vector of person-specific controls. Table 3 reports the summary statistics.

(Table 3 around here)

The primary outcomes are whether the individual finished at least 9-12 years of education (binary) and the number of schooling years completed. The coefficients of interest are contained in the vector γ which measure the differential change in schooling outcomes for cohorts during the Great Depression, holding constant person characteristics and aggregate differences in outcomes across cities and over time. To account for serial correlation and city-specific random shocks, I cluster the standard errors at the city level in all specifications. The 1926 and 1927 cohorts serve as the control group, and all reported coefficients are relative to those cohorts. As discussed above, all regressions are weighted by inverse propensity scores derived after predicting the characteristics that are associated with a successful link (see Appendix B.1).

There are three primary concerns with a causal interpretation of the γ estimates. Since youth unemployment is not randomly distributed across space, omitted *local* variables that are positively correlated with youth unemployment will bias the coefficients. In this

setting, we would expect the bias to be primarily upwards. For example, if places that had higher youth unemployment experienced larger contemporaneous wealth shocks (e.g., via bank closures), stronger anti-child-labor policy changes, or more Federal support from the New Deal, the estimates of γ would overstate the true labor market effect on schooling. To address this bias, I control for several observable and plausibly confounding variables in all specifications. I include state-by-year fixed effects to account for uneven youth schooling dynamics at the state-level driven by state-level policies and regional economic shocks. At the county level, I further include interaction terms between 1929 wholesale share, retail share, and unemployment in 1930 and cohort dummies, annual total banking deposits, and annual manufacturing output.

The second major concern with a causal interpretation of the γ estimates are omitted *person-level* variables that correlate with youth unemployment. For example, if the distribution of latent academic ability within a local area is inversely proportional to the propensity to enter the youth labor force, estimates of γ would be biased downwards. I address this concern by including household fixed effects, effectively comparing the schooling choices of brothers who share similar unobserved biological and household determinants of schooling.[19] The resulting average local estimates of the impact of youth unemployment on educational attainment eliminate the across-household component of selection into schooling during the Depression. However, since this selection is interesting in its own right - as highlighted by the conceptual framework - I conduct heterogeneity analysis by comparing the outcomes of brothers from blue- and white- collar families separately.

The identifying variation that remains in the sample after the aforementioned variables are included in C_{jk} and Ω_i relies on locally-persistent *within* state and *within* household differences, relative to the initial shock in 1931. Said differently, I assume that $\Delta Unemp_j$ is an accurate and quasi-exogeneous shock to the opportunity cost channel, after controlling for state-level dynamics, local confounding variables, and across-household selection into schooling.

The last major identifying assumption is that, in the absence of the Depression, house-

[19] I restrict the sample of older brothers to those who are at least 3 years older then their closest linked younger brother such that all younger brothers starting with the 1931 cohort are compared to 1926-1928 cohorts.

hold choices across cities with different Depression-era youth unemployment would have evolved in parallel. To provide support that this assumption is not prohibitively strong, I show that the 1927-1928 cohorts (pre-Depression) did not exhibit differential trends of schoolgoing in Table 4. This table shows the coefficient estimates of γ in Equation 5.1 when the sample includes only the 1926-1928 cohorts. Columns "9" through "13" use the outcome variable that takes the value of 100 if the individual reported at least 9 through 13 years and zero otherwise, respectively. In the last column, I report the estimates when the number of years is the outcome variable. Across all columns, the estimates are small - only a 0.04 percentage point increase in the likelihood of finishing at least 12 years of education - and statistically insignificant from zero at conventional confidence levels.

(Table 4 around here)

The lack of pre-trends is both reassuring and plausible. In the short period considered in this paper, there is no reason to expect that the trends of school-going should vary significantly across cities unevenly hit by the Depression *before* the 1930s. Indeed, the factors that contribute to different levels of educational attainment in regular times, such as the skill-premium, cultural norms, or the availability and proximity of schools, evolved over the preceding three decades, not years. Conversely, the sharp turn of the economy starting at the end of 1929 was an unexpected and severe shock for households, and choices about a child's investment in education had to be made in the short run.

6 Great Depression and Education Attainment

In this section, I present the paper's main results. I find that young urban males from blue-collar families significantly increased their education investment during the early-part, but not the later-part, of the Depression. Conversely, those from white-collar households increased their school-going at twice the rate of blue-collared males, but only in the second half of the Depression. In particular, more youth started high school (at least 9th grade completion), with some attrition before high school graduation for blue-collared youth and no attrition for white-collared youth. By 1940, the returns to education for the these cohorts was large: those from worse-hit cities had higher earnings and held more prestigious occupations

by 1940. I also find some evidence that the Depression increased college-going. In all, the Depression increased the average schooling by 0.1-0.2 years, with most of these impacts clustered in the Rust Belt, the Northeast, and the West North Central regions of the country.

6.1 Average cohort estimates

Table 5 presents estimates of the pooled difference-in-differences model given in Equation 5.1 using 9 through 13 grade completion as the outcome variable in the first five columns and school years completed in the last column. For ease of interpretation, I scale the binary outcome variables by 100, such that the resulting coefficients represent percentage point changes in attainment, and standardize $\Delta Unemp_j$ to have mean of zero and standard deviation of one. Panel A uses the full sample and Panel B uses only the sample of brothers and controls for household fixed-effects.

(Table 5 around here)

The first row of both panels shows the impact of youth unemployment on cohorts that turned 18 between 1931 and 1933 . Since the vast majority of child-labor laws required school attendance until the age of 16, these cohorts were making education decision in 1929 through 1931 - the early years of the Depression. The second row shows the impact on those who were making these same decisions in 1932 - 1935, once the Depression was well under way.

In the full sample (Panel A), I find a relatively large effects on 9 grade completion for the older cohort (1931 - 1933), with smaller and statistically indistinguishable from zero estimates on all other attainment measures. The effects, however, are quantitatively and statistically larger for the younger (1934 - 1937) cohorts in the full sample. After accounting for across-household selection (Panel B), I find that one standard deviation increase in youth unemployment led to a 1.5 - 2.7 percentage point increase in the likelihood that a young male in a Depression-era cohorts attempted some high school (9 grade completion) and 1.2 percentage point increase in the likelihood that he finished high school. This estimate accounts for roughly 10 percent of the cross-sectional standard deviation of high school graduation rates across cities in 1930. Comparing the full sample with the brother sample,

I observe a significant downward bias due to household selection into schooling, across all outcomes.

These (average) effects are quantitatively significant. To put these numbers in historical perspective, the graduation rate of cohorts in the sample increased from approximately 43 percent in 1927 to 54.5 percent by 1936, or 1.1 percentage points per year - a remarkable increase even in the context of the High School Movement. The 1934-1937 cohorts were, on average, 7.7 – 12 percent more likely to graduate high school than their 1927 counterparts. Therefore, the estimates show that one standard deviation in youth unemployment accounted for 9 – 13 percent of this increase. In terms of the total length of schooling, the 0.11 - 0.13 education years gained for the 1934 - 1937 cohorts is close to the widely-cited impact of compulsory school laws, where men in the 1920-1940 birth cohorts born in the first quarter of the year were found to complete 0.126 fewer years of education (Angrist and Krueger (1991)).

6.2 Heterogeneity with respect to father occupation

In this section, I investigate the schooling behavior of youth based on their household's socioeconomic status, as proxied by their father's occupation. Even though around 68 percent of the youth in my sample come from a family with a father working in a blue-collar occupation, only 54 percent of high school graduates came from a blue-collar family in the 1930 cohort.[20] This gap suggests that students from lower socio-economic households drop out of school earlier than their peers from higher socioeconomic status households. A deep economic recession, on the one hand, can negatively impact poorer students disproportionately – they may be less likely to stay in school and enter the labor force full-time to support the family, either through the formal labor market or through household production. On the other hand, if a recession lowers the opportunity cost for youth to stay in school, and if poor students are the ones for whom this cost is salient, then a recession may entice them to stay in school. Did the Depression widen or narrow this education gap?

Table 6 reports the results of Table 5 separately for brothers of blue- and white-

[20]Consistent with the literature, I use a broad definition of blue-collar: individuals working as craftsmen, farmers, operatives, service, laborers (*occ1950* between 100 and 200, and between 500 and 978) and white-collar: professionals, managers, clerks, sales (*occ1950* between 0 and 100, and between 200 and 500)

collar households in Panels A and B. To obtain standard errors of the estimate of the difference between the estimates, Panel C shows the fully interacted triple-difference model. Not surprisingly, I find that the factors leading to higher education attainment during the Depression impacted students differently across socio-economic strata. The timing of the impact for each strata, however, is noteworthy: blue-collar youth who were making high-school going decisions at the start (but not at the end) of the Depression invested more in schooling. A one standard deviation in youth unemployment for those turning 18 in 1931 through 1933 led to 0.11 more years of education for the earlier blue-collar cohorts of brothers. I observe a 1.1 percentage point increase in the likelihood of them finishing high school and a 0.92 percentage point increase in entering post-secondary education. These impacts are sizable, considering that the mean high school graduation and college-entry rates for the 1931 cohort in this sample was only 38 and 12 percent, respectively.

Unlike their older peers, I do not find any significant impact on blue-collar youth turning 18 in 1934 to 1937 besides 9-grade completion. Conversely, I find no impact on white-collar youth at the start of the Depression, but I do find pronounced effects for the later cohorts. White-collared youth turning 18 between 1934 and 1937 increased their schooling investment (at all levels) at roughly twice the rate as for those found in 1931-1933 blue-collar cohorts: a one standard deviation in youth unemployment led to 0.24 more years of schooling, 3.60 percentage point increase in their likelihood of finishing high school, and a 2.16 percentage point increase in their likelihood of entering college.

(Table 6 around here)

What explains these timing differences? While I lack authoritative data to discern between the various explanations, the evidence so far does not reject the simple budget-constraint human capital conceptual framework of Section 2. Faced with an lack of labor market opportunities and declining family income, sons of blue-collar household reacted swifty to the falling opportunity cost of education. On the other hand, sons of white-collared households - for whom the opportunity costs of education were lower to begin with and family incomes and savings were plausibly higher - did not. But once the Depression was well under way, blue-collar families realized that the recession in 1929 and 1930 was

turning into a full-blown Depression: sons were called upon to contribute monetarily to the family budget as a matter of survival. The white-collar sons, upon seeing poor labor market opportunities for high-school graduates and fewer family obligations, decided to extend their schooling in hopes of obtained larger returns of education in adulthood.

6.3 Sensitivity Analysis

The main results found above are robust to various sample selection and alternative variable definition concerns. First, instead of using the full linked sample of cohorts, I limit my analysis to those linked using a more conservative matching procedure, eliminating those youth that did not match on exact age from the sample. The estimates change only slightly for years of education completed: for the full sample, the 1931-1933 estimate decreases to 0.04 from 0.06 and the 1934-1937 estimate decreases to 0.10 from 0.11. Likewise, in the brother sample, the estimates remain in the same order of magnitude (close to 0.1). Second, instead of defining blue- and white-collar households based on the father's occupation, I re-estimate the heterogeneity analyses based on the father's occupational income score and occupational education score. The results are qualitatively and quantitatively similar: urban youth from blue-collar (below median in either score) households who faced the choice of dropping out or continuing their high school education in the first years of the Depression increased their schooling investment while their younger counter-parts did not. The point estimate is very close to the one reported in Table 6 across both alternative definitions of blue and white collar (0.10 vs. 0.11). The converse was true for white-collar youth, and again the point estimates changed only slightly (0.20 vs. 0.24). The results for these analyses can be found in Appendix Tables A1, A2, and A3.

6.4 Youth outcomes in young adulthood (1940)

I now turn to the discussion of post-Depression outcomes for Depression-era cohorts by using the 1940 Census wage and occupation data: did schooling improve outcomes in young adulthood for youth? And, importantly, what was their return to education? For the sample of youth that were in the labor force in 1940 and reported a wage - thus ignoring the impact of any post-secondary education for the youngest of these cohorts still in school

- I find that blue-collar individuals in high youth unemployment areas significantly outperformed their peers in low youth unemployment areas, relative to the outcomes of their older brothers. Table 7 reports the regression estimates when log weekly earnings ("wage"), the Duncam Socioeconomic Index ("SEI"), Siegal prestige score ("presgl"), and the standardized median earned income by occupation ("erscor") are the outcome variables. I augment the main specification to account for labor market experience differences across individuals, where *experience* is defined as (age minus six) in 1940 minus the number of schooling years completed.

(Table 7 around here)

I find significant returns to education for blue-collared youth that turned 18 during the initial phase of the Depression. On average, the Depression increased average earnings of the 1931-1933 blue-collar and the 1934-1937 white-collar cohorts by approximately 2.5 and 7 percent, respectively. It did so because these individuals entered, by 1940, into more prestigious, higher-wage occupations. For example, the 3 log point change in the Duncan Socioeconomic Index (SEI) is approximately 60 percent of the difference between a painter (except construction) (Log SEI of 2.89) and a carpenter (Log SEI of 2.94).

To a first-order approximation, the return to education for the 1931-1933 blue-collared and the 1934-1937 white-collared cohorts was large. A back of the envelope return to education in this sample of blue-collared urban youth is approximately 22 percent: 0.024 (increase in log earning)/0.11 (increase in school years completed) and 29 percent for white-collared youth (0.069/0.24). For the younger cohorts (1934 - 1937) and for the white-collared older youth, I do not find any statistically significant impact of youth unemployment during the Depression on their young adulthood outcomes by 1940.

6.5 Spatial Aggregate Distribution of Schooling Gains during the Depression

The impact of Depression-era youth unemployment was unevenly distributed across the U.S. states. In this section, I assess the quantitative significance of my findings at the state-level by aggregating the reduced form estimates in Table 6 for cohorts in each state in my sample. Recall that the standardized unemployment measure that appears in the

specification is the difference between actual estimated youth unemployment in 1931 and the total county unemployment as reported by the Census in 1930, expressed as standard deviations away from its mean. In order to find the total quantitative impact, however, an assumption about the baseline counterfactual evolution of youth unemployment in the absense of the Depression must be made. Here, I consider two counterfactual baselines: (1) no change between 1930 total and 1931 youth unemployment ($\Delta Unemp_j$ is standard deviations away from zero), and (2) youth unemployment at 5 percentage point greater than the 1930 county average ($\Delta Unemp_j$ is standard deviations away from 5 percent plus the the county unemployment).[21] Under each scenario, I find the population weighted average of $\Delta Unemp_j$ ($E[\Delta Unemp_{j(s)}]$) across cities in each state and compute the following change in average years of schooling completed:

$$\Delta_{sk} = (\hat{\gamma}_{blue,k} \times share_{blue,sk} + \hat{\gamma}_{white,k} \times share_{white,sk}) \times E[\Delta Unemp_{j(s)}] \tag{6.1}$$

where s denotes the state, k denotes the cohort, γ's are the reduced form estimates from Table 6, and *share* is the total share of the cohort that is from a white- or blue- collared household. Panels A and B of Figure 2 plot the the average Δ_{sk} across cohorts for each state.

(Figure 2 around here)

Under both counterfactuals, the states with the largest increases in average school attainment in the 1931-1937 cohorts were those of Rust Belt (Ohio, Indiana, Pennsylvania, Michigan), those in the West North Central (Iowa, Nebraska, Kansas, and Missouri), and those in the Northeast (Maine, New Hampshire, Vermont). The more conservative estimates (Panel B) show gains of 0.16 to 0.26 years of school on average for Depression-era cohorts in these states. Conversely, there was little to no impact on urban youth in the South and West regions of the country. This spatial clustering is not entirely surprising given the distribution of youth labor-intensive industries and Depression employment shocks in those industries around the country, as discussed in Section 4.

[21] The difference between peak youth unemployment and adult unemployment during the peak of the 1949 recession (October) was 5 percent according to Federal Reserve Economic Data.

7 Summary

This paper used new spatially disaggregated data on youth occupation and unemployment during the Depression and found immediate but fleeting effects on secondary school completion of blue-collared youths and delayed but relatively larger effects of white-collared youth. I derived the estimates by comparing the outcomes of urban men who made their their secondary schooling choices before the Depression to those of their brothers who were making choices during the Depression. The evidence presented here shows that, on average, the opportunity cost channel was responsible for a 0.1 to 0.2 more years of schooling for these cohorts, suggesting not only an essential role of labor market opportunity costs on human capital investment decisions but also an important contribution of the Depression to the broader schooling gains during High School Movement. Because these impacts varied across cohorts, socioeconomic strata, and the intensity of the Depression in local areas, the aggregate reduced-form gains varied widely across U.S. states. In all, the heterogeneous impacts uncovered do not point to a unilateral improvement of post-schooling outcomes for poorer households, as both white and blue-collared individuals realized high returns to education in the short-run (by 1940). While this project focused on short-term outcomes *immediately after* the crisis itself, it opens up interesting and important follow-up questions regarding how, and for whom, the Depression changed *lifetime* earnings and outcomes in adulthood, and its overall impact on the achievement gap and inequality.

References

Abramitzky, Ran, Leah Boustan, and Myera Rashid, "Census Linking Project: Version 1.0 [dataset]," 2020. 3, 16, 33, 41

Angrist, Joshua D and AlanB Krueger, "Does Compulsory School Atttendace Affect Schooling and E," *The Quarterly Journal of Economics*, 1991. 21

Atkin, David, "Endogenous skill acquisition and export manufacturing in Mexico," *American Economic Review*, 2016, *106* (8), 2046–85. 5

Bailey, Martha J, Connor Cole, Morgan AC Henderson, and Catherine G Massey, *How Well Do Automated Methods Perform in Historical Samples?: Evidence from New Ground Truth*, National Bureau of Economic Research, 2017. 48

Baker, Richard B, John Blanchette, and Katherine Eriksson, "Long-run impacts of agricultural shocks on educational attainment: Evidence from the boll weevil," *The Journal of Economic History*, 2020, *80* (1), 136–174. 4

Baran, Cavit, Eric Chyn, and Bryan A Stuart, "The Great Migration and Educational Opportunity," Technical Report, Working Paper 2020. 4

Bau, Natalie, Martin Rotemberg, Manisha Shah, and Bryce Steinberg, "Human Capital Investment in the Presence of Child Labor," Technical Report, National Bureau of Economic Research 2020. 5

Becker, Gary S, *Human capital: A theoretical and empirical analysis, with special reference to education*, University of Chicago press, 2009. 2, 6

Betts, Julian R and Laurel L McFarland, "Safe port in a storm: The impact of labor market conditions on community college enrollments," *Journal of Human resources*, 1995, pp. 741–765. 5

Black, Dan, Terra McKinnish, and Seth Sanders, "The economic impact of the coal boom and bust," *The Economic Journal*, 2005, *115* (503), 449–476. 5

Card, David, Ciprian Domnisoru, and Lowell Taylor, "The intergenerational transmission of human capital: Evidence from the golden age of upward mobility," Technical Report, National Bureau of Economic Research 2018. 4

Cascio, Elizabeth U and Ayushi Narayan, "Who Needs a Fracking Education? The Educational Response to Low-Skill-Biased Technological Change," *ILR Review*, 2015, p. 0019793920947422. 5

Charles, Kerwin Kofi, Erik Hurst, and Matthew J Notowidigdo, "Housing booms and busts, labor market opportunities, and college attendance," Technical Report, National Bureau of Economic Research 2015. 5

Darby, Michael R, "Three-And-A-Half Million US Employees Have Been Mislaid: Or, An Explanation of Unemployment, 1934-1941," Technical Report, National Bureau of Economic Research 1975. 4

Elder, Glen H, *Children of the Great Depression: Social change in life experience*, Routledge, 2018. 3

Federal Deposit Insurance Corporation, "Federal Deposit Insurance Corporation Data on Banks in the United States, 1920 - 1936," Inter-university Consortium for Political and Social Research 1992. DOI: https://doi.org/10.3886/ICPSR0007.v1. 16

Feigenbaum, James J, "Intergenerational mobility during the great depression," 2015. 5

Fishback, Price V, Shawn Kantor, and John Joseph Wallis, "Can the New Deals three Rs be rehabilitated? A program-by-program, county-by-county analysis," *Explorations in Economic History*, 2003, *40* (3), 278–307. 9

Goldin, Claudia and Lawrence F Katz, "Why the United States led in education: Lessons from secondary school expansion, 1910 to 1940," Technical Report, National Bureau of Economic Research 1997. 4, 8

_ and _ , "Human Capital and Social Capital: The Rise of Secondary Schooling in America, 1910–1940," *Journal of Interdisciplinary History*, 1999, *29* (4), 683–723. 3

ICPSR, "Historical, Demographic, Economic, and Social Data: The United States, 1790-1970," 2005. https://doi.org/10.3886/ICPSR00003.v1. 16

Jr, Melvin Stephens and Dou-Yan Yang, "Compulsory education and the benefits of schooling," *American Economic Review*, 2014, *104* (6), 1777–92. 4

Karger, Ezra, "The Long-Run Effect of Public Libraries on Children: Evidence from the Early 1900s," 2021. 4

Kisswani, Khalid M, "Did the Great Depression Affect Educational Attainment in the US?," *Economics Bulletin*, 2008, *9* (30), 1–10. 4

Margo, Robert A, "The microeconomics of depression unemployment," *The Journal of Economic History*, 1991, *51* (2), 333–341. 4

Rosenbloom, Joshua L and William A Sundstrom, "The sources of regional variation in the severity of the Great Depression: evidence from US manufacturing, 1919–1937," *The Journal of Economic History*, 1999, *59* (3), 714–747. 13

Ruggles, Steven, Flood Sarah, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, "IPUMS USA: Version 10.0 [dataset], Minneapolis, MN," `https://doi.org/10.18128/D010.V10.0` 2020. Accessed: 2020-06-01. 7

Schmick, Ethan J and Allison Shertzer, "The Impact of Early Investments in Urban School Systems in the United States," Technical Report, National Bureau of Economic Research 2019. 4

Shah, Manisha and Bryce Millett Steinberg, "Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital," *Journal of Political Economy*, 2017, *125* (2), 527–561. 5

Shanahan, Michael J, Glen H Elder Jr, and Richard A Miech, "History and agency in men's lives: Pathways to achievement in cohort perspective," *Sociology of Education*, 1997, pp. 54–67. 4

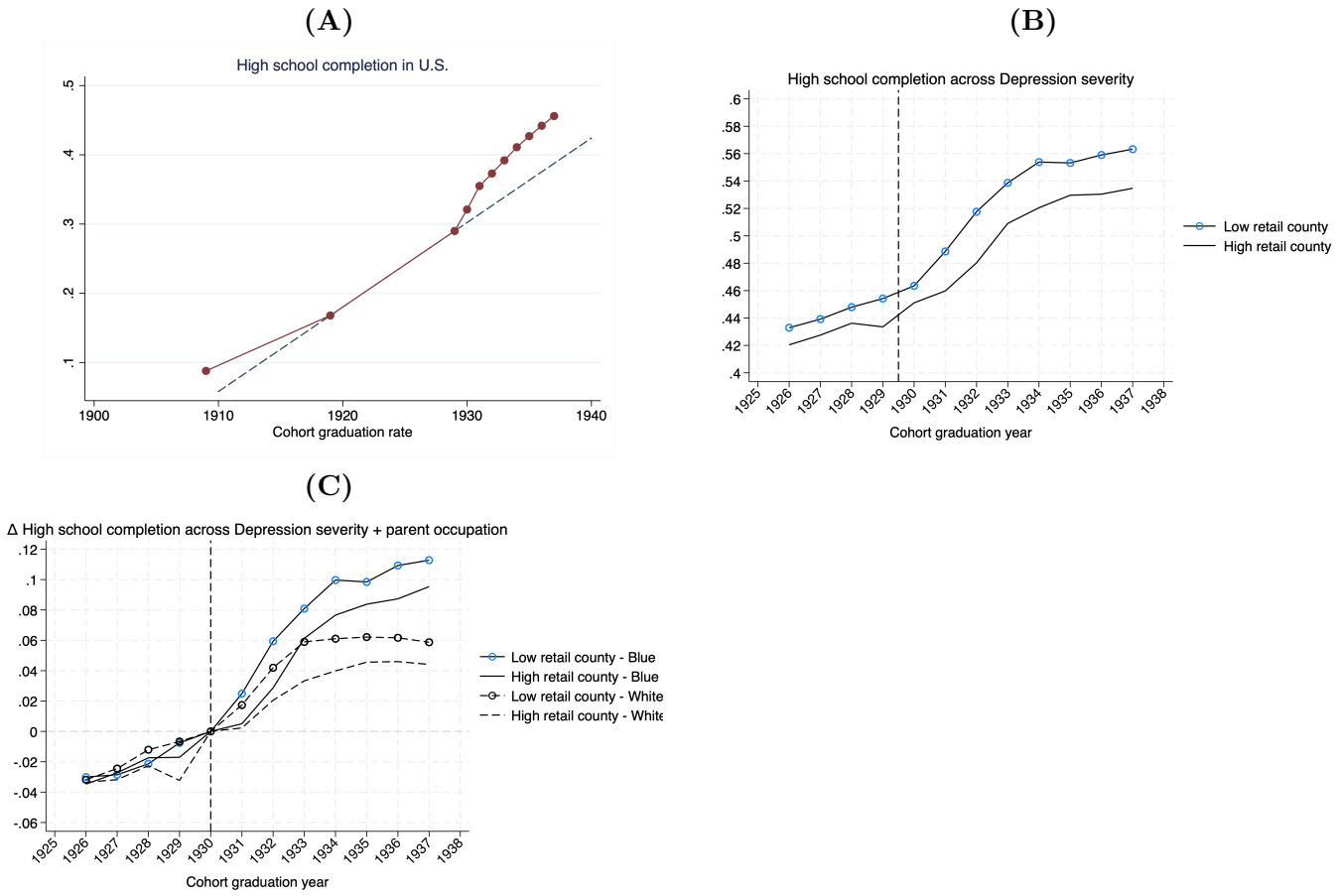
Stewart, Paul William, *Market Data Handbook of the United States*, Washington: U. S. Govt. Print. Offs, 1929. 16

Sundstrom, William A, "Last hired, first fired? Unemployment and urban black workers during the Great Depression," *The Journal of Economic History*, 1992, *52* (2), 415–429. 4

Wallis, John Joseph, "Employment in the Great Depression: New data and hypotheses," *Explorations in Economic History*, 1989, *26* (1), 45–72. 4

Yamashita, Takashi, "The effects of the Great Depression on educational attainment," Technical Report, Citeseer 2008. 4

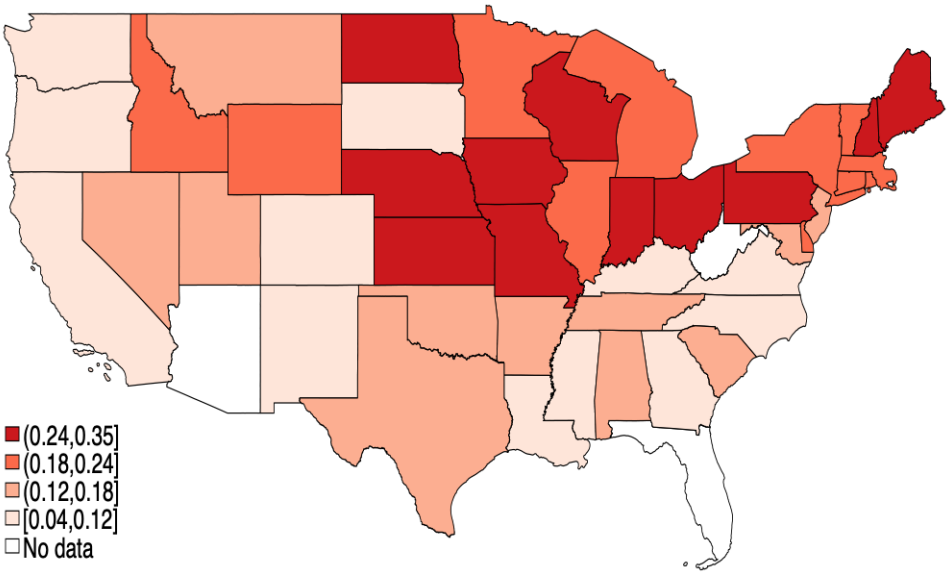
Figure 1: High School Movement and Great Depression: U.S. Males



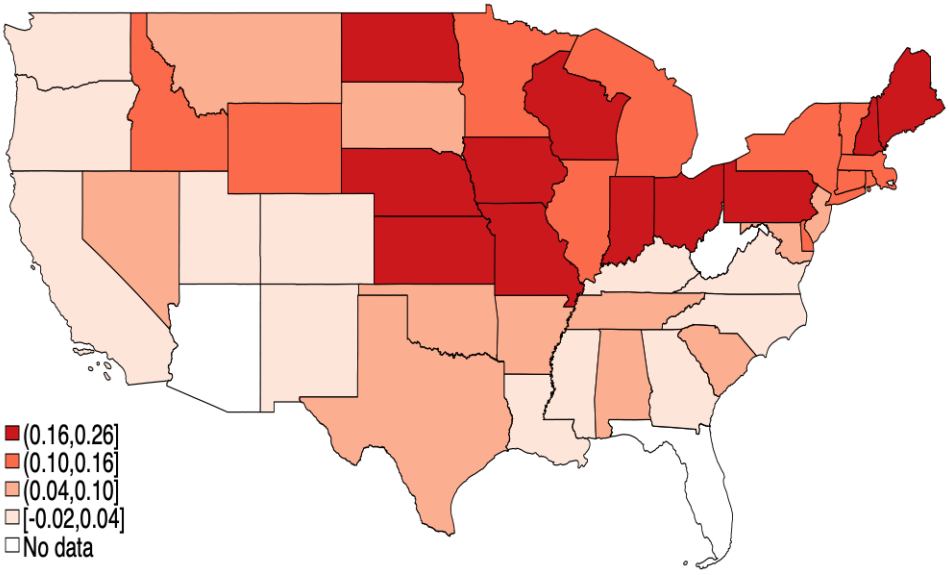
Notes: Panel (A) plots the number of high school graduates as a propotion of 17-year-olds in the United States for the years 1910, 1920, and 1930-1938. The dashed line denotes the average 1920-1930 growth rate extrapolated to earlier and later decades. The source of the data is U.S. Department of Commerce, Bureau of the Census, *Historical Statistics of the United States, Colonial Times to 1970* as reproduced in Table 19 of *120 Years of American Education: A statistical portrait* published by the National Center for Education Statistics. Panel (B) plots the proportion of high school graduates in the main sample of Census-linked young males used in this paper. Cohort graduation year is when the respondent turns 18 years old. "Low retail county" denotes all counties in the lowest tercile of retail sales growth between 1929 and 1933 and "high retail county" denotes all cohorts in the top tercile. In Panel (C), I plot the change from 1930 in the average high school graduation rates for 4 types of cohorts: those in "low retail" counties with a blue-collar (solid line, circle) and white-collar (dashed line, circle) parent and those in "high retail" counties with a blue-collar (solid line, no circle) and white-collar (dashed line, no circle) parent. The 1930 graduation rates (%) for those groups are: 37, 68, 36, 67, respectively. All education outcomes come from the 1940 Census. See Section 4.1 for details.

Figure 2: State Level Aggregate Impacts

Panel A. Using County Unemployment in 1930 as Baseline



Panel B. Using 5% + County Unemployment in 1930 as Baseline



Notes: These map show the average change in years of schooling for cohorts that turned 18 between 1931 - 1937 aggregated to the state-level as in Equation 6.1. Baselines used to compute $E[\Delta Unemp_{j(s)}]$ denoted in the title of each panel. Data for Florida, Arizona, and West Virginia is missing.

Table 1: School choice by age in 1930

Sample: Youth in Home-Owning Households Only

	Age:					
	14	15	16	17	18	19
Ln(House Value)	0.5***	2.5***	8.1***	12.1***	13.1***	12.3***
	(0.2)	(0.4)	(1.0)	(0.8)	(0.8)	(0.6)
Ln(Household Income)	0.6**	3.6***	14.0***	22.0***	23.2***	20.0***
	(0.3)	(0.5)	(1.3)	(1.4)	(1.0)	(0.7)
I(Only)=1	-0.8***	0.5	6.1***	10.1***	11.6***	9.0***
	(0.2)	(0.4)	(1.0)	(1.1)	(0.9)	(0.7)
log(wage premium adult)	0.1	0.1	3.3***	5.8***	6.3***	6.4***
	(0.2)	(0.4)	(1.0)	(1.3)	(1.0)	(0.8)
State FE	✓	✓	✓	✓	✓	✓
E[y]	92.96	83.15	61.45	43.89	30.42	22.31
R-sq	0.00	0.03	0.08	0.11	0.10	0.08
N	124,736	122,382	121,510	116,358	113,229	104,070

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the estimate results of an OLS regression of a school attendance binary variable (scaled by 100) on various household and city characteristics using 1930 Census 100 percent count records. Each column presents the results for a different age group denoted in the header. More details can be found in Section 2.

Table 2: Relationship between 1931 Unemployment and other city-level characteristics

Share	Δ Retail Sales	Δ Manu. Output	Wholesale Share	Retail Share	Manu Share
	(1)	(2)	(3)	(4)	(5)
Δ Unemployment - Youth	0.056	0.157***	-0.362***	-0.385***	0.122*
	(0.070)	(0.038)	(0.069)	(0.074)	(0.061)
State FE	✓	✓	✓	✓	✓
E[y]	-0.43	-0.55	0.01	0.03	0.09
R-sq	0.45	0.14	0.24	0.26	0.45
N	566	446	575	575	570

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the estimation results of an OLS regression of 1931 Youth Unemployment on the variable denoted in the header and state fixed effects. All variables are standardized to have mean zero and standard deviation of one. The unit of observation is a city. See text for sources and descriptions of the variables. Standard errors shown in parentheses and are clustered at the state-level.

Table 3: Summary Statistics

Panel A: Census data

	N	Mean	SD	Median	25 pct	75 pct
School years completed (1940)	1,304,302	10.886	2.86	11.00	9.00	12.00
I(Finish 9)	1,304,302	75.204	43.18	100.00	100.00	100.00
I(Finish 10)	1,304,302	67.568	46.81	100.00	0.00	100.00
I(Finish 11)	1,304,302	56.335	49.60	100.00	0.00	100.00
I(Finish 12)	1,304,302	48.634	49.98	0.00	0.00	100.00
I(Finish 13)	1,304,302	19.232	39.41	0.00	0.00	0.00
I(blue collar occ)	1,304,302	0.596	0.49	1.00	0.00	1.00
I(white collar occ)	1,304,302	0.278	0.45	0.00	0.00	1.00
Log(weekly wage)	898,049	3.159	0.72	3.26	2.86	3.61
County Unemployment	1,304,302	0.075	0.02	0.08	0.06	0.09
Unemployment - Youth - City Average	1,304,302	0.200	0.04	0.21	0.18	0.23
Δ Unemployment - Youth	1,304,302	0.124	0.05	0.12	0.09	0.15
Manufacturing share	1,300,529	0.104	0.05	0.11	0.06	0.14
Retail share	1,304,302	0.042	0.02	0.04	0.03	0.05
Bank Deposits pc	1,180,669	0.538	0.83	0.39	0.21	0.58
Manufacturing Output pc	1,261,618	641.232	398.03	552.19	368.27	842.05
Age (1930)	1,304,302	16.176	3.38	16.00	13.00	19.00

Panel B: City-Occupations in 1930 and 1931

	N	Mean	SD	Median	25 pct	75 pct
Total under 20 workers [city, 1930]	925	1537.0	8271.5	459.0	288.0	948.0
Youth occupation categories [city, 1930]	925	67.1	25.1	59.0	49.0	79.0
Youth occupation categories w/rates [city, 1931]	925	35.8	13.3	33.0	26.0	43.0
%Youth covered by occupation categories w/rates [city, 1931]	925	73.4	10.5	74.0	65.6	82.3
%Weight per occupation [city x occ, 1930]	33,071	2.8	5.4	0.9	0.4	2.5
%Regional unemployment rate [city x occ, 1931]	33,071	25.5	11.7	24.2	16.6	33.9

Notes: Panel A presents the summary statistics of U.S. Decennial Census variables of 1930 - 1940 linked males between between the ages of 11 and 22 in 1930. Sample includes only males living in Census enumerated cities for which youth unemployment rates could be obtained (534 cities). For 18-22 year olds, household and parent characteristics come from the 1920 Census (when available) if the individual reported as a head-of-household in in the 1930 Census. Otherwise, they come from the 1930 Census, Youth living in any non-households (institutions) were dropped from the sample, as were youth with parents in unclassified occupations. Census records were linked using crosswalks obtained from Abramitzky et al. (2020). Panel B shows the summary statistics of the main variables used in the construction of the youth unemployment estimate from the 1931 Special Census of Unemployment. Total under 20 workers reports the size of the under-20 labor force. Regional unemployment rate denotes the 1931 unemployment estimates for each city-occupation. Youth occupation categories is the number of occupations reported in the 1931 Special Census of Unemployment.

Table 4: Pre-Depression dynamics of schooling behavior and 1931 youth unemployment

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1927 - 1928 × Δ Unemployment - Youth	0.17	0.18	-0.05	0.04	0.02	0.02
	(0.33)	(0.34)	(0.37)	(0.37)	(0.27)	(0.02)
City FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓
E[y]	73.24	65.16	53.95	46.80	19.99	10.85
R-sq	0.05	0.04	0.04	0.04	0.02	0.05
N	267,280	267,280	267,280	267,280	267,280	267,280

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the estimation of the pooled difference-in-differences specification of Equation 5.1 using the 1926-1928 full cohorts of males. Those that turned 18 in 1926 serve as the omitted reference group. The outcome variable across the the first five columns is a binary variable taking the value of 100 if the individual reported finishing (at least) the amount of years of school denoted in the header and zero otherwise. The outcome variable in the last column is the number of school years completed. Δ Unemployment - Youth is the standardized (mean zero, standard deviation one) measure of the youth unemployment rate at the city level in 1931 estimated using the 1931 Unemployment Census minus the 1930 county total unemployment rate. Regressions are weighted for representativeness. Standard errors shown in parentheses and are clustered at the city level.

Table 5: Pooled difference-in-differences estimates

Panel A: Full Sample

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1931 - 1933 × Δ Unemployment - Youth	1.07***	0.38	0.15	0.27	0.18	0.06***
	(0.30)	(0.29)	(0.30)	(0.27)	(0.23)	(0.02)
1934 - 1937 × Δ Unemployment - Youth	1.79***	1.04***	0.62*	0.48	0.57**	0.11***
	(0.34)	(0.31)	(0.33)	(0.32)	(0.22)	(0.02)
City FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓
E[y]	73.24	65.16	53.95	46.80	19.99	10.85
R-sq	0.06	0.06	0.06	0.05	0.02	0.06
N	946,239	946,239	946,239	946,239	946,239	946,239

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel B: Brother Sample

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1931 - 1933 × Δ Unemployment - Youth	1.46***	1.17**	0.14	1.14**	0.46	0.07**
	(0.56)	(0.54)	(0.55)	(0.56)	(0.37)	(0.03)
1934 - 1937 × Δ Unemployment - Youth	2.65***	1.87***	1.02	1.21	0.93**	0.13***
	(0.69)	(0.69)	(0.70)	(0.76)	(0.44)	(0.04)
Household FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓
E[y]	68.79	59.83	47.65	40.68	15.13	10.47
R-sq	0.40	0.41	0.42	0.42	0.40	0.51
N	90,268	90,268	90,268	90,268	90,268	90,268

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the estimation of the pooled difference-in-differences specification of Equation 5.1 using the full sample (Panel A) and the linked brother sample (Panel B). Individuals that turned 18 in 1926, 1927, or 1928 serve as the omitted reference group. The outcome variable across the the first five columns is a binary variable taking the value of 100 if the individual reported finishing (at least) the amount of years of school denoted in the header and zero otherwise. The outcome variable in the last column is the number of school years completed. Δ Unemployment - Youth is the standardized (mean zero, standard deviation one) measure of the youth unemployment rate at the city level in 1931 estimated using the 1931 Unemployment Census minus the 1930 county total unemployment rate. Regressions are weighted for representativeness. Standard errors shown in parentheses and are clustered at the city level.

Table 6: Heterogeneous impacts on blue- and white-collar households

Panel A. Blue-Collar Households

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1931 - 1933 × Δ Unemployment - Youth	2.13***	1.02	-0.02	1.71**	0.92**	0.11***
	(0.74)	(0.68)	(0.67)	(0.67)	(0.44)	(0.04)
1934 - 1937 × Δ Unemployment - Youth	2.07**	0.62	-0.01	0.15	0.33	0.08
	(0.92)	(0.95)	(0.91)	(0.87)	(0.50)	(0.05)
Household FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓
E[y]	68.12	58.55	45.63	38.01	12.11	10.26
R-sq	0.37	0.38	0.37	0.37	0.28	0.45
N	52,946	52,946	52,946	52,946	52,946	52,946

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel A. White-Collar Households

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1931 - 1933 × Δ Unemployment - Youth	0.11	0.50	-0.63	0.45	0.10	0.03
	(1.06)	(1.30)	(1.42)	(1.41)	(1.23)	(0.08)
1934 - 1937 × Δ Unemployment - Youth	3.19**	3.74**	2.95*	3.60**	2.16	0.24**
	(1.54)	(1.81)	(1.68)	(1.78)	(1.46)	(0.10)
Household FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓
E[y]	86.40	81.80	74.31	67.98	37.83	12.27
R-sq	0.36	0.37	0.40	0.42	0.44	0.50
N	20,726	20,726	20,726	20,726	20,726	20,726

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel C. Fully Interacted Model

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1931 - 1933 x White x Δ Unemployment - Youth	-2.02*	-0.51	-0.61	-1.26	-0.83	-0.07
	(0.083)	(0.709)	(0.698)	(0.407)	(0.530)	(0.367)
1934 - 1937 x White x Δ Unemployment - Youth	1.12	3.12	2.96	3.45*	1.82	0.16
	(0.559)	(0.139)	(0.132)	(0.079)	(0.254)	(0.185)

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Panels A and B present the estimation of the pooled difference-in-differences specification of Equation 5.1 using subsamples of linked brothers that reported living in a blue- or white-collar household in 1930, respectively. See text for the definition of blue and white collar. Older brothers that turned 18 in 1926, 1927, or 1929 serve as the omitted reference group. The outcome variables come from the 1940 Census. Panel C presents the estimation results of a fully-interacted triple difference model. Δ Unemployment - Youth is the standardized (mean zero, standard deviation one) measure of the youth unemployment rate at the city level in 1931 estimated using the 1931 Unemployment Census minus the 1930 county total unemployment rate. Regressions are weighted for representativeness. Standard errors shown in parentheses and are clustered at the city level.

Table 7: Young adulthood outcomes across siblings in 1940

Panel A. Sub-samples

	White				Blue			
	wage	SEI	presgl	erscor	wage	SEI	presgl	erscor
1931 - 1933 × Δ Unemployment - Youth	-0.039	0.001	-0.025*	-0.030	0.024*	0.032**	0.018***	0.024
	(0.034)	(0.031)	(0.015)	(0.029)	(0.014)	(0.014)	(0.007)	(0.015)
1934 - 1937 × Δ Unemployment - Youth	0.069**	-0.056*	-0.012	-0.033	0.025	0.004	-0.009	0.013
	(0.034)	(0.033)	(0.016)	(0.041)	(0.020)	(0.019)	(0.009)	(0.015)
Household FE	✓	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓	✓	✓
Experience	✓	✓	✓	✓	✓	✓	✓	✓
E[y]	3.30	3.54	3.58	4.01	3.19	3.18	3.44	3.93
R-sq	0.24	0.34	0.29	0.14	0.23	0.32	0.26	0.12
N	10,341	9,912	10,221	10,341	29,048	28,046	28,709	29,048

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel B. Fully Interacted Model

	wage	SEI	presgl	erscor
1931 - 1933 x Blue x Δ Unemployment - Youth	0.06*	0.03	0.04***	0.05
	(0.082)	(0.392)	(0.009)	(0.118)
1934 - 1937 x Blue x Δ Unemployment - Youth	-0.04	0.06	0.00	0.04
	(0.234)	(0.125)	(0.963)	(0.316)

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

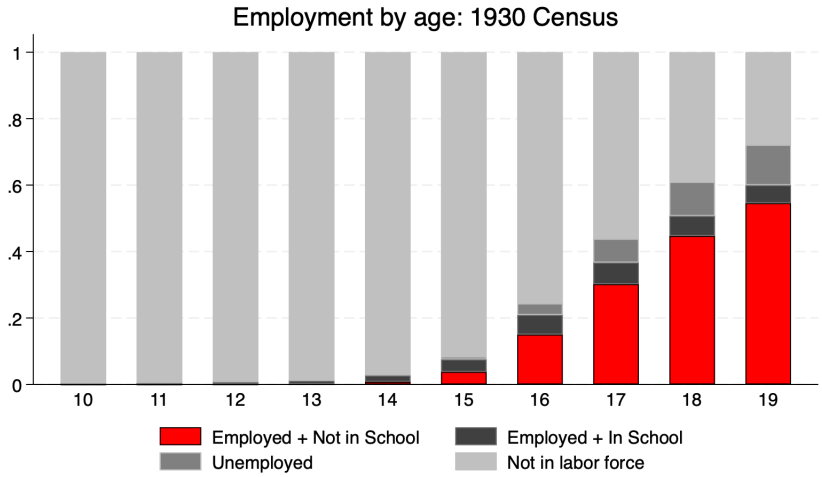
Notes: Panel A presents the estimation of the pooled difference-in-differences specification of Equation 5.1 using subsamples of linked brothers. Older brothers that turned 18 in 1926, 1927, or 1929 serve as the omitted reference group. The outcome variables come from the 1940 Census. Panel B presents the estimation results of a fully-interacted triple difference model. Δ Unemployment - Youth is the standardized (mean zero, standard deviation one) measure of the youth unemployment rate at the city level in 1931 estimated using the 1931 Unemployment Census minus the 1930 county total unemployment rate. Regressions are weighted for representativeness. Standard errors shown in parentheses and are clustered at the city level.

Appendix

Table of Contents

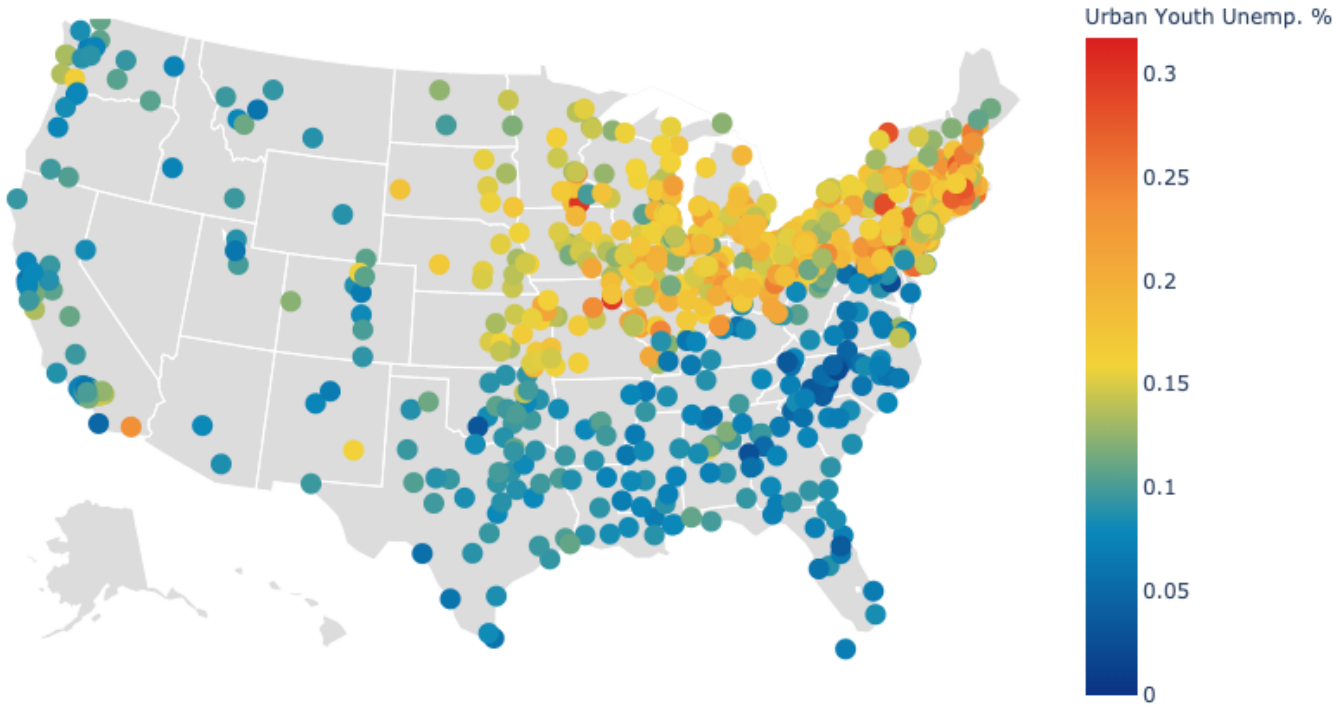
A Supplementary Figures

Figure A.1: Proportion of Employed Males by Age: 1930 U.S. Census



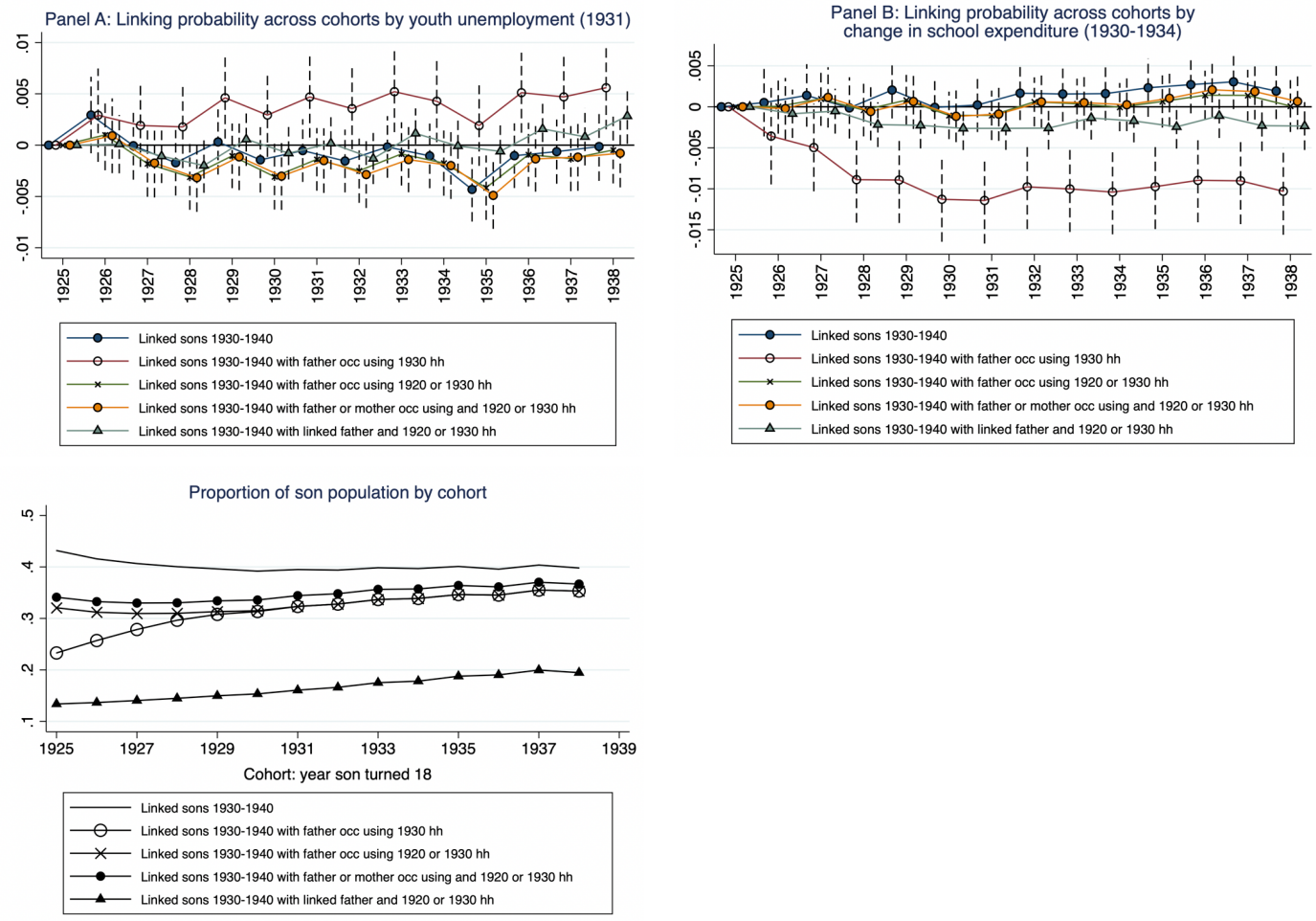
Source: Aggregation of 100 percent count records of the 1930 Decennial Census, available on IPUMS.

Figure A.2: Map of urban youth unemployment estimates



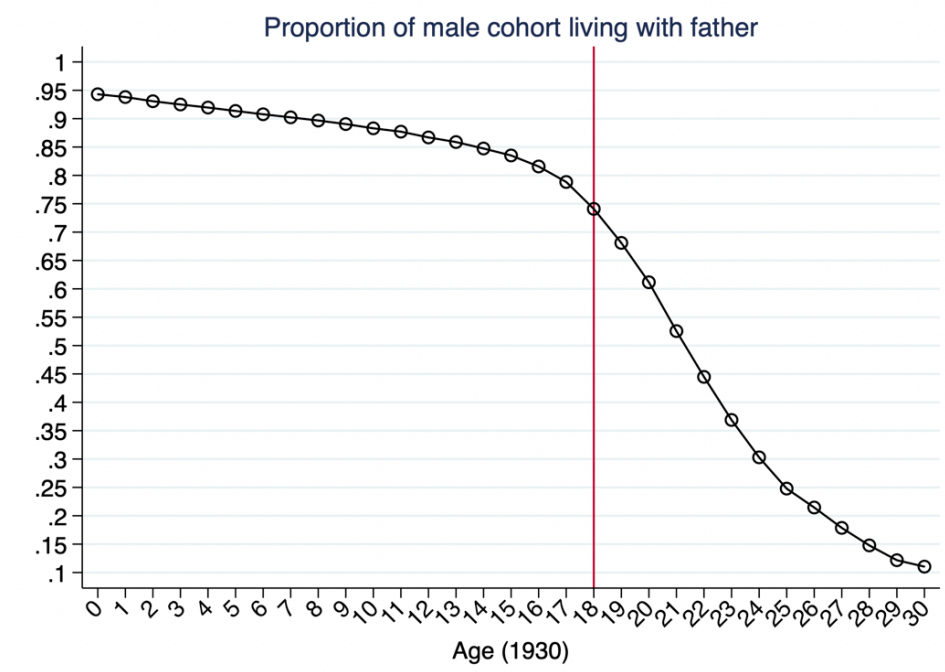
Notes: This figure plots the geographical distribution of youth unemployment in 1931. Youth unemployment is estimated using city-level occupation shares and regional occupational unemployment shares as computed from the Special Unemployment Census of 1931. Section 4.1 describes the construction in more detail.

Figure A.3: Census linking probabilities across cohorts and Depression shocks



Notes: Panels (A) and (B) report the marginal probability of linking a male aged 11-22 between the 1930 and 1940 Censuses across Depression severity - youth unemployment and school expenditure cuts - using crosswalks provided by Abramitzky et al. (2020). Each line denotes the estimated coefficients of an OLS regression of a binary variable taking the value of 1 if a link to 1940 is present on the interaction terms between cohort and Depression shocks. The samples are the following: all males 11-22, all males with an employed father in 1930, all males with an employed father in either 1920 (for 18-22 year-olds) or 1930 (for 11-17 year-olds), all males with either father or mother occupation, and finally all males with non-missing father information in the 1940 Census. Bottom panel shows the proportion of total cohort captured at various steps that lead to my main sample of sons (solid circle) and father-son pairs (solid triangle).

Figure A.4: Proportion of males living with father in the 1930 Census



Source: Aggregation of 100 percent count records of the 1930 Decennial Census, available on IPUMS.

B Supplementary Tables

Table A1: Pooled difference-in-differences estimates with conservative links

Panel A: Full Sample

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1931 - 1933 × Δ Unemployment - Youth	1.01*	0.33	-0.20	0.11	0.17	0.04
	(0.52)	(0.53)	(0.51)	(0.49)	(0.45)	(0.03)
1934 - 1937 × Δ Unemployment - Youth	1.81***	1.09*	0.87*	0.34	0.51	0.10***
	(0.55)	(0.56)	(0.49)	(0.49)	(0.40)	(0.03)
City FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓
s	73.24	65.16	53.95	46.80	19.99	10.85
E[y]	0.06	0.06	0.06	0.05	0.03	0.05
R-sq	222,894	222,894	222,894	222,894	222,894	222,894

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel B: Brother Sample

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1931 - 1933 × Δ Unemployment - Youth	1.97*	2.26**	0.03	2.14**	0.65	0.12**
	(1.10)	(1.08)	(0.98)	(1.06)	(0.77)	(0.06)
1934 - 1937 × Δ Unemployment - Youth	0.70	1.17	0.88	0.32	0.65	0.08
	(1.25)	(1.17)	(1.32)	(1.25)	(0.85)	(0.06)
Household FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓
E[y]	68.79	59.83	47.65	40.68	15.13	10.47
R-sq	0.41	0.42	0.44	0.45	0.44	0.53
N	24,499	24,499	24,499	24,499	24,499	24,499

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents the estimation of the pooled difference-in-differences specification of Equation 5.1 using the full sample (Panel A) and the linked brother sample (Panel B). In contrast with Table 5, this table only uses individuals linked using the nysiis conservative method provided. Individuals that turned 18 in 1926, 1927, or 1928 serve as the omitted reference group. The outcome variable across the the first five columns is a binary variable taking the value of 100 if the individual reported finishing (at least) the amount of years of school denoted in the header and zero otherwise. The outcome variable in the last column is the number of school years completed. Δ Unemployment - Youth is the standardized (mean zero, standard deviation one) measure of the youth unemployment rate at the city level in 1931 estimated using the 1931 Unemployment Census minus the 1930 county total unemployment rate. Regressions are weighted for representativeness. Standard errors shown in parentheses and are clustered at the city level.

Table A2: Heterogeneous impacts on father's occupation income score

Panel A. Father below median occupational income score

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1931 - 1933 × Δ Unemployment - Youth	1.29*	1.17	0.42	1.31*	0.76	0.10**
	(0.78)	(0.76)	(0.78)	(0.74)	(0.52)	(0.04)
1934 - 1937 × Δ Unemployment - Youth	2.00**	1.16	0.48	0.59	0.51	0.08*
	(0.90)	(0.91)	(0.89)	(0.91)	(0.55)	(0.04)
Household FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓
E[y]	68.09	59.07	47.15	39.94	14.98	10.39
R-sq	0.38	0.40	0.40	0.40	0.35	0.48
N	45,414	45,414	45,414	45,414	45,414	45,414

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel A. Father above median occupational income score

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1931 - 1933 × Δ Unemployment - Youth	1.73*	1.04	-0.39	1.16	0.35	0.03
	(0.95)	(0.94)	(0.91)	(1.00)	(0.84)	(0.05)
1934 - 1937 × Δ Unemployment - Youth	3.30***	2.20**	1.53	1.88	1.20	0.20***
	(1.06)	(1.10)	(1.24)	(1.39)	(0.95)	(0.06)
Household FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓
E[y]	80.89	74.22	64.09	56.94	27.08	11.53
R-sq	0.38	0.40	0.42	0.42	0.42	0.51
N	34,167	34,167	34,167	34,167	34,167	34,167

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Panels A and B present the estimation of the pooled difference-in-differences specification of Equation 5.1 using subsamples of linked brothers that reported living with a father with a below median (Panel A) or above-median (Panel B) occupation income score (1950). See text for more details. Older brothers that turned 18 in 1926, 1927, or 1929 serve as the omitted reference group. The outcome variables come from the 1940 Census. Panel C presents the estimation results of a fully-interacted triple difference model. Δ Unemployment - Youth is the standardized (mean zero, standard deviation one) measure of the youth unemployment rate at the city level in 1931 estimated using the 1931 Unemployment Census minus the 1930 county total unemployment rate. Regressions are weighted for representativeness. Standard errors shown in parentheses and are clustered at the city level.

Table A3: Heterogeneous impacts on father's occupational education score

Panel A. Father below median educational (occupation) score

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1931 - 1933 × Δ Unemployment - Youth	1.98**	1.15	-0.06	1.43*	0.80	0.11**
	(0.85)	(0.80)	(0.79)	(0.79)	(0.54)	(0.04)
1934 - 1937 × Δ Unemployment - Youth	1.38	-0.14	-1.05	-1.22	0.06	0.02
	(0.97)	(0.91)	(0.82)	(0.88)	(0.52)	(0.04)
Household FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓
E[y]	64.77	54.75	41.61	34.19	10.28	10.02
R-sq	0.36	0.37	0.36	0.35	0.27	0.44
N	40,489	40,489	40,489	40,489	40,489	40,489

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel N. Father above median educational (occupation) score

	Outcome: I(completed X years)					Years of education
	9	10	11	12	13	
1931 - 1933 × Δ Unemployment - Youth	0.87	0.40	-0.39	1.11	0.26	0.02
	(0.96)	(1.04)	(1.15)	(1.21)	(0.87)	(0.06)
1934 - 1937 × Δ Unemployment - Youth	3.43***	3.27**	3.09**	3.98***	1.82*	0.24***
	(1.18)	(1.37)	(1.37)	(1.43)	(1.04)	(0.08)
Household FE	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
State x Year	✓	✓	✓	✓	✓	✓
1930 Industry Shares x Year	✓	✓	✓	✓	✓	✓
Manufacturing	✓	✓	✓	✓	✓	✓
Banking	✓	✓	✓	✓	✓	✓
E[y]	83.33	77.43	68.25	61.27	30.60	11.80
R-sq	0.37	0.39	0.41	0.43	0.42	0.50
N	33,154	33,154	33,154	33,154	33,154	33,154

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Panels A and B present the estimation of the pooled difference-in-differences specification of Equation 5.1 using subsamples of linked brothers that reported living with a father with a below median (Panel A) or above-median (Panel B) occupation education score (1950). See text for more details. Older brothers that turned 18 in 1926, 1927, or 1929 serve as the omitted reference group. The outcome variables come from the 1940 Census. Panel C presents the estimation results of a fully-interacted triple difference model. Δ Unemployment - Youth is the standardized (mean zero, standard deviation one) measure of the youth unemployment rate at the city level in 1931 estimated using the 1931 Unemployment Census minus the 1930 county total unemployment rate. Regressions are weighted for representativeness. Standard errors shown in parentheses and are clustered at the city level.

Table A4: Unemployment rates for youth in 1931 in select occupations and cities

City	Rate	Weight	Occupation
Duluth	68.1%	6.2%	Laborers : Iron and steel industries
Seattle	59.4%	1.1%	Operatives : Iron and steel industries
Chicago	55.9%	2.0%	Operatives : Paper, printing, and allied industries
Chicago	54.6%	1.6%	Operatives : Food and allied industries
Denver	54.1%	1.0%	Operatives : Iron and steel industries
Detroit	53.9%	11.5%	Laborers : Iron and steel industries
Philadelphia	53.7%	1.7%	Operatives : Electrical machinery and supply factories
Buffalo	53.2%	1.1%	Laborers : Food and allied industries
San Francisco	53.0%	1.1%	Laborers : Building construction, laborers, and helpers
Detroit	52.8%	2.0%	Mechanics
Philadelphia	52.3%	1.1%	Operatives : Leather industries
Boston	51.5%	5.0%	Operatives : Leather industries
Duluth	51.4%	3.5%	Laborers : Building construction, laborers, and helpers
Detroit	50.8%	9.0%	Operatives : Iron and steel industries
Boston	50.7%	1.1%	Operatives : Clothing industries
Buffalo	50.4%	1.1%	Operatives : Leather industries
New Orleans	50.0%	1.4%	Porters (except in stores)
Chicago	50.0%	1.0%	Laborers : Building construction, laborers, and helpers
Buffalo	49.5%	3.3%	Laborers : Building construction, laborers, and helpers
Seattle	49.1%	3.3%	Laborers : Building construction, laborers, and helpers
Boston	48.9%	2.4%	Laborers : Building construction, laborers, and helpers
Chicago	48.9%	1.8%	Mechanics
Buffalo	48.8%	1.5%	Operatives : Textile industries
Birmingham	2.8%	1.3%	Engineers (stationary), cranemen, hoistmen, etc
San Francisco	4.3%	7.9%	Servants (except cooks)
Denver	4.9%	3.5%	Bookkeepers, cashiers, and accountants
San Francisco	5.2%	2.3%	Waiters
San Francisco	5.4%	8.8%	Sailors, deck hands, boatmen, etc.
Denver	5.4%	2.1%	Retail dealers
San Francisco	7.1%	3.3%	Bookkeepers, cashiers, and accountants
Minneapolis	7.6%	3.9%	Bookkeepers, cashiers, and accountants
St. Louis	8.0%	2.8%	Bookkeepers, cashiers, and accountants
Seattle	8.1%	2.3%	Bookkeepers, cashiers, and accountants
San Francisco	8.5%	2.2%	Retail dealers
Denver	9.0%	27.6%	Clerks (except "clerks" in stores)
Seattle	9.4%	8.5%	Servants (except cooks)
Los Angeles	9.8%	3.1%	Bookkeepers, cashiers, and accountants
San Francisco	10.5%	31.0%	Clerks (except "clerks" in stores)
Seattle	10.6%	17.5%	Clerks (except "clerks" in stores)
Manhattan	11.1%	1.5%	Retail dealers
San Francisco	11.1%	2.0%	Machinists, millwrights, and toolmakers
Los Angeles	11.2%	19.9%	Clerks (except "clerks" in stores)
Manhattan	11.8%	6.5%	Salesmen and saleswomen
New Orleans	11.8%	2.7%	Bookkeepers, cashiers, and accountants
Denver	11.9%	3.2%	Laborers : Iron and steel industries
Seattle	11.9%	6.8%	Sailors, deck hands, boatmen, etc.

Notes: This table lists the occupations with the largest and smallest youth (10-19 years old) unemployment rates as defined in Section 4.1 across cities in the 1931 Special Census of Unemployment. Column "Weight" refers to share of the youth labor force at the occupation-city level, as of 1930.

Table A5: Most common youth occupations and unemployment rates by region

Rank	Modal Occupation	Unemployment Rate	# Cities	Weight
	Midwest			
1	Retail workers	7%	176	16%
2	Servants (except cooks)	16%	50	24%
3	Operatives: Leather industries	43%	15	24%
4	Clerks (except "clerks" in stores)	15%	15	21%
5	Laborers: Iron and steel industries	50%	6	17%
6	Operatives: Clothing industries	44%	4	19%
7	Laborers: Food and allied industries	45%	3	30%
8	Farm laborers (wageworkers)	33%	3	17%
9	Operatives: Metal industries (except iron and steel)	33%	3	17%
10	Operatives: Iron and steel industries	46%	2	17%
11	Laborers: Metal industries	34%	1	19%
12	Operatives: Clay, glass, and stone industries	31%	1	15%
	Northeast			
1	Retail workers	5%	91	15%
2	Operatives: Textile industries	42%	72	31%
3	Clerks (except "clerks" in stores)	13%	61	20%
4	Servants (except cooks)	12%	40	23%
5	Operatives: Leather industries	47%	29	29%
6	Laborers: Iron and steel industries	49%	20	24%
7	Operatives: Clothing industries	44%	16	20%
8	Stenographers and typists	18%	8	14%
9	Operatives: Cigar and tobacco factories	27%	5	15%
10	Laborers: Clay, glass, and stone industries	38%	4	21%
11	Operatives : Metal industries (except iron and steel)	29%	3	14%
12	Public service - non-laborers	11%	3	38%
13	Operatives: Rubber factories	33%	2	29%
14	Farm laborers (wageworkers)	29%	1	13%
15	Operatives: Electrical machinery and supply factories	29%	1	13%
	South			
1	Retail workers	5%	78	15%
2	Servants (except cooks)	8%	69	18%
3	Clerks (except "clerks" in stores)	11%	2	13%
4	Laborers : Food and allied industries	26%	1	26%
5	Stenographers and typists	15%	1	15%
6	Laborers : Iron and steel industries	42%	1	13%
7	Waiters	15%	1	15%
8	Laborers : Clay, glass, and stone industries	43%	1	15%
	West			
1	Retail workers	4%	56	16%
2	Servants (except cooks)	6%	10	24%
3	Farm laborers (wageworkers)	39%	10	23%
4	Laborers : Lumber and furniture industries	29%	5	24%
5	Clerks (except "clerks" in stores)	9%	5	16%
6	Oil and gas well operatives	29%	1	20%
7	Operatives : Food and allied industries	17%	1	14%
8	Fishermen and oystermen	40%	1	12%
		Total	878	

Notes: This table shows the most common occupations reported by urban 10-19 year olds and their estimated unemployment rates in 1931 from city-level data obtained from the Special Census of Unemployment. The column "# Cities" reports the number of cities in which the occupation listed is the most common occupation within the city. The "Weight" column reports the share of youth that hold the occupation as a proportion of all city youth workers. Midwest includes the states: IA, IL, IN, KS, MI, MN, MO, ND, NE, OH, SD, WI. Northeast includes the states: CT, MA, ME, NH, NJ, NY, PA, RI, VT. South includes the states: AL, AR, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV. West includes the states: AZ, CA, CO, ID, MT, NM, NV, OR, UT, WA, WY.

B.1 Weighing

Following Bailey et al. (2017), I construct inverse propensity weights to adjust for observable differences between linked and linked records in two steps. First, using the population of 11-22 year old males living in an Census-identified city in 1930, I estimate a probit regression of link status (whether an individual is matched) on the following variables: indicator for being white, indicator for father having a white-collar occupation, indicators for each Census region, age and age squared, and a constant. The results are shown in Table A6. I then compute the inverse propensity scores for each person as (1-p)/p times m/(1-m), where (p) is the predicted likelihood of an individual being matched based on the estimated probit coefficients and (m) is the actual match rate (22.1 percent).

Table A6: Predicting characteristics of successful links using a probit regression

	(1) I(in sample)
I(White)	0.453***
	(0.026)
I(White Collar Father)	0.106***
	(0.012)
Middle Atlantic Division	-0.105***
	(0.040)
East North Central Div.	0.086**
	(0.042)
West North Central Div.	0.138**
	(0.056)
South Atlantic Division	-0.489**
	(0.195)
East South Central Div.	-0.223**
	(0.109)
West South Central Div.	-0.137*
	(0.076)
Mountain Division	0.115
	(0.080)
Pacific Division	0.029
	(0.081)
Age	0.106***
	(0.005)
Age Squared	-0.004***
	(0.000)
Constant	-1.801***
	(0.054)
N	5,910,651

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$