

Depressed in School: The Great Depression, Educational Attainment, and Productivity Growth*

Jason Dunn

Department of Economics, Boston University

October 9, 2026

Abstract

While the Great Depression caused severe immediate declines in employment and output, this paper studies a positive additional consequence: increased human capital accumulation. To identify this, I construct a granular shift-share instrument for county exposure to youth unemployment, using 8 sectors, 47 manufacturing industries, and 8 demographic subgroups. Counties more exposed to the Depression saw greater increases in high school attainment rates. By 1940, this shock-induced education sorted these cohorts into higher-skilled occupations. Harder-hit counties experienced greater increases in the education of their workforce and labor productivity. These findings suggest that although highly exposed counties were negatively affected from the initial shock of the Depression, it induced educational attainment and put them on the trajectory for higher growth in the long run.

*Correspondence address: jtdunn@bu.edu. I am grateful to my advisor James Feigenbaum for guidance and support through this project. I also thank Adam Guren, Josh Goodman, Martin Fiszbein, seminar participants at the Boston University Applied Micro Workshop, and Harvard Economic History Workshop for helpful comments and suggestions. First version: October 2026

1 Introduction

The Great Depression was a significantly negative economic event, with considerable declines in real economic indicators. US GDP bottomed out at 30% below trend, and total (nonfarm) unemployment peaked at 23% (32%).¹ Despite such a negative economic downturn, the Depression also had some additional consequences that altered the long run development of the United States. One such example was the New Deal, an extensive federal program of relief, recovery, and reform whose legacy remains in the reach of the modern welfare state.

This paper explores another consequence of the Great Depression with lasting effects: educational attainment. With the Depression severely reducing employment opportunities, adolescents in high school faced lower incentives to drop out, and may have been more likely to complete high school. I find this opportunity cost channel to exist across counties, particularly for those harder hit by the Depression. To reduce measurement error and address local confounding shocks, I construct a Bartik-style shift-share instrument of county-level exposure to unemployment shocks. Local employment shares are measured for 8 demographic subgroups, 8 sectors, and 47 manufacturing industries, while national employment shifts are measured for the same sectors and industries.² I find that counties more exposed to the Depression saw larger increases in the high school attainment rates of their 18 year old cohorts. Specifically, a 1 percentage point higher unemployment rate raised the cohort graduation share by an additional 0.39 percentage point from 1929 to 1933. This was an effect that grew over the course of the decade.

By 1940, this employment shock induced a series of long run structural effects. First, I find that young individuals in harder-hit counties were more likely to sort into higher-skill/status occupations, and particularly less likely to enter agriculture. Second, I find that the education level in the workforce increased in counties harder hit by the Depression, where a 1 percentage point higher unemployment rate raised the high school share of the workforce by an additional 0.2 percentage point from 1930 to 1940. I decompose this effect and find the change is almost entirely accounted for by labor-force entry, the margin where high school graduates factor in. Third, I find that counties more exposed to the Depression saw greater increases in labor productivity (value-added per worker), where a 1 percentage point higher unemployment rate raised 1929–1947 productivity growth by 2 percentage points. Taken

1. These national estimates are taken from Tables Ca9-19 and Ba470-477 of the *Historical Statistics of the United States* Database.

2. I measure high school completion at the county level for 8 demographic subgroups: sex (male v. female) $\times$ race (white v. non-white) $\times$ father SES (strictly above, or at or below, the within-county median among ages 16–25 children).

together, these findings suggest that while the Great Depression had immediate negative effects, the channel through which it increased educational attainment had positive long-term consequences on the skill composition of the workforce and the productivity of the local economy.

This paper relates most directly to a large literature on education in the early 20th Century US, and during the Great Depression in particular. (Goldin 1998) and (Goldin and Katz 2008) document the movement toward significant high school enrollment and completion across the US between 1910 and 1940. More recent work has documented the drivers of this movement along various margins, including correspondence schools (Vidart 2026), high school openings (Doxey, Karger, and Nencka 2026), agricultural booms (Jaworski, Kitchens, and Rodgers 2026), and junior colleges (Bleemer and Quincy 2026). This paper focuses on the opportunity cost mechanism driving additional high school attainment during the Depression, tracing the attainment itself while also studying the later skill composition and productivity of local labor markets.

More specific to the Depression, (Goldin and Katz 2008) also document an associational relationship between unemployment and high school graduation, exploiting variation across states. (Janas 2026a) expands on this work by considering local-level variation in cities with populations over 25,000 and applying a DiD framework for causal interpretation, finding effects particularly for boys from high income families. This work most closely follows this strand of the literature, and I expand upon it along three dimensions. First, I study outcomes at the county level, a finer geographic scope than the former study but offering more coverage than the latter. Second, I identify local exposure to the Depression by constructing a granular shift-share instrument using 8 demographic subgroups, 8 sectors, and 47 manufacturing industries. To my knowledge, this is the first application of a Bartik instrument of employment during the Great Depression. Third, I extend beyond the short-run schooling response of the Depression and consider the long-run structural changes that take place in the form of human capital accumulation, skill sorting, and labor productivity growth.

This paper also speaks to work on the long-run consequences of the Great Depression. Most of these studies find negative effects of the Depression in the long run, whether it be on prime-age earnings and health for early-life exposed individuals (Duque and Schmitz 2025), wages for less-educated labor market entrants (Moulton 2017), or employment in general (Mathy 2018). The focus of this paper runs counter to these other channels, studying how the immediate negative shock of the Depression generated more educational attainment and favorable outcomes for those high school graduates in the long run.

Relatedly, this paper also connects to work studying structural change and the labor market during the Depression. Such works consider the interaction between migration and local labor markets (Boustan, Fishback, and Kantor 2010), regional employment effects based upon industry mix (Rosenbloom and Sundstrom 1999), state-sector specific productivity shocks (Cheng, Crucini, and Kim 2026), and the interaction between financial frictions (Benmelech, Frydman, and Papanikolaou 2019) and unemployment (Margo 1993). (Boyd 2002) and (Boone and Wilse-Samson 2023) are the most closely related papers here, which find a form of “reverse” structural transformation taking place during this time, whereby many workers returned to farm labor as a form of migratory insurance. I complement these studies by finding that although many incumbent workers returned to agricultural endeavors, the younger labor force entrants were actually more likely to sort into higher skilled jobs and industries. Only considering the overall response covers this distinct margin for young individuals.

Lastly, this paper interacts with a more modern literature on the relationship between the labor market and educational investment. Economic booms have been documented reducing educational enrollment (Black, McKinnish, and Sanders 2005; Charles, Hurst, and Notowidigdo 2018), while more recent downturns have increased it (Hershbein 2012; Barr and Turner 2015; Goodman and Winkelmann 2025). The time of the Depression in this paper offers an especially informative setting relative to these studies, where both youth employment and high school dropout rates were notably high, and the Depression itself a significantly negative shock. The channel these papers study become highly relevant in this context, not only at the individual level, but also from an aggregate standpoint.

The remainder of the paper is organized as follows: Section 2 provides a historical background on the period and describes the theoretical motivation for the studied mechanisms. Section 3 describes the empirical framework and strategy. Section 4 describes the various datasets and variables used. Section 5 presents and discusses the short-run and long-run results. Section 6 concludes.

2 Background/Context

2.1 Early 20th Century US Education

The early 20th century in the US saw a significant increase in the prevalence of secondary schooling, with local governments making an increasingly concerted effort to provide adequate schooling at this level (Goldin and Katz 2008). Panel A of Figure 1 shows a steady

increase in both the rate of high school attainment and total average years of schooling for 18 year old cohorts during this time.³

There is, however, no discernible increase in schooling following the Depression's outset in late 1929. At the national level, the high school movement of this era dominates the trend in high school attainment. Only once differentiating across space can one observe a relationship between the Depression and educational attainment. Panel B of Figure 1 shows the high school attainment rate of 18 year old cohorts over time for two sets of counties: those that were in the top quartile of unemployment in 1930 and those in the bottom quartile. Multiple items are of note from this comparison. First, the level of high school attainment is higher even before the Great Depression in harder hit counties. This likely has to do with correlates to the shock, such as the level of industrialization and urbanization in a county, which both tended to relate to higher levels of educational attainment.

Second, one can observe that although their pre-levels are different, these two county groups appear to exhibit similar pre-trends in the high school completion share through 1929. Similar pre-trends are required before the later divergence can be attributed to the shock.

Finally, Figure 1 shows that following the outset of the Great Depression, the shares of 18 year olds completing high school starts increasing considerably beyond the projected pre-trend for hard hit counties, while a minimal increase beyond the pretrend occurs for low unemployment counties. This is suggestive that some dynamics during the Great Depression contributed to a structural change in the attainment of education differentially across counties. I will describe the exact theoretical mechanism supporting this in Section 2.2, and rigorously test this in Section 5.1 further.

2.2 The Opportunity Cost of Schooling

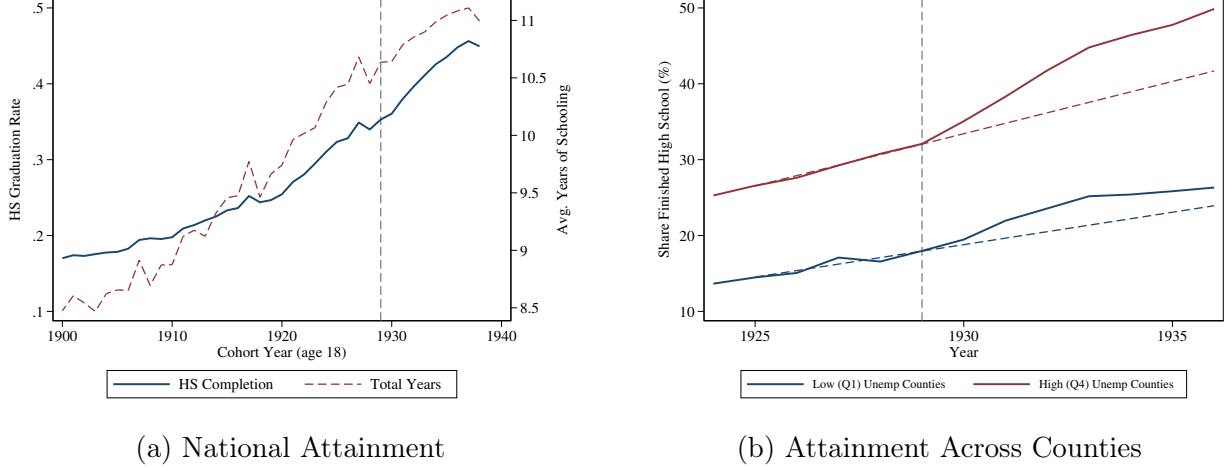
To first fix ideas regarding the channel through which the Great Depression may have affected educational attainment, I consider a simple discrete choice model where teenagers choose between education and work.

There are two time periods: $T = \{1, 2\}$, where the first period is during the Great Depression and the second period is after.⁴ Consider a continuum of measure one agents aged between 16 and 18 years old during period 1. They face the discrete choice set C of

3. These measures of education are taken from the 1940 full count census, so are retrospective from that point.

4. The logic of this model could also extend to infinite periods with the same conclusions being drawn.

Figure 1: US Educational Attainment



Note: Panel A plots the high school completion share and average years of schooling by the year a cohort turned 18, using respondents aged 20–58 in the 1940 full-count census. The vertical line marks 1929. Panel B plots the high school completion share, in percentage points, for the ages 16–18 county aggregate. Counties are split into the bottom and top quartiles of that cell’s 1930 unemployment rate, and county shares are weighted by the ages 16–18 population. The series ends with the 1936 cohort. Dashed lines are linear trends fit on cohorts who turned 18 in 1929 or earlier and are drawn from 1925.

either completing high school or dropping out and entering the labor market.

$$C = \{s, l\}$$

More specifically, the choice s entails attending school in $t = 1$ and working in $t = 2$, whereas the choice l entails working in $t = 1, 2$.

Agents choose the option that maximizes their total utility, which is determined by the following subjective expected utility function:

$$c_i^* = \max_{c \in C} SEU_i(c)$$

and the utility from the two choices are the following:

$$\begin{aligned} c_{i,s} &= p_2(L_1^d)w_{i,s} \\ c_{i,l} &= p_1(L_1^d)w_{i,l} + \beta p_2(L_2^d)w_{i,l} \end{aligned}$$

where β is the time discount factor and $p_t(L_t^d)$ is the probability of finding a job in period t , which I assume is increasing in that period’s labor demand. I also assume that the wage earned by a high school graduate is greater than a non-highschool graduate: $w_{i,s} > w_{i,l}$. The indifference point for an agent would be at the given job finding probabilities and wages such that the expected payoffs of the two choices are equal:

$$p_2(L_2^d)w_{i,s} = p_1(L_1^d)w_{i,l} + \beta p_2(L_2^d)w_{i,l}$$

This fully contextualizes the problem faced by the agent. To further compare the payoffs of the two choices, note the opportunity cost of the agent choosing to go to school is equivalent to the foregone wages from working in period 1: $p_1(L_1^d)w_l$.

This model then motivates the explicit mechanism through which one would expect the Great Depression to influence aggregate educational attainment. During the Depression, a county that was exposed to the economic downturn experienced a decrease in the demand for labor: $L_1^d \downarrow$. With a decrease in labor demand, it becomes more difficult for individuals to find a job: $p_1(L_1^d) \downarrow$. This means that the expected payoff of immediately working (or equivalently, the opportunity cost of going to school) decreases: $p_1(L_1^d)w_l \downarrow$. For the mass of individuals near the previous indifference point that chose to enter the labor market, some of them will now find it optimal to remain in school. In the aggregate, this would mean that a greater share of youths complete high school in these areas. This is the short-run mechanism I test for empirically in Section 5.1.

2.3 Long-Run Theoretical Implications

The discrete choice model outlined above determines more than just the share of high school graduates in the short-run. They also determine the jobs agents hold in period 2: the job of a high school graduate that is paid w_s and the job of a high school dropout that is paid w_l . This lends itself to three testable predictions in the long run.

First, under the assumption $w_s > w_l$, with a lowering L_1^d in Depression-exposed counties, agents near the indifference point will switch from choice l to s . In period 2, there will be more agents holding the higher wage job that pays w_s , raising the average job “quality” in the affected cohorts. I test for this empirically in Section 5.2.1 by looking at the mean occupational standing in young cohorts according to Depression severity. Second, if more agents in Depression-exposed counties make the decision to stay in school, and a proportional amount of them enter the workforce in period 2, one would expect the educational level of the workforces in these counties to grow more from period 1 to 2. I test for this directly in Section 5.2.2, measuring education as the high school attainment share of the workforce. Third, under the assumption that workers are paid their marginal product of labor, counties that were harder hit by the Depression, who gained a greater mass of high school graduates, would exhibit a greater mass of the higher wages w_s , and subsequently contain greater overall labor productivity. I test for this empirically in Section 5.2.3 by looking at both the change

in manufacturing wages and labor productivity (value added per worker) at the county level.

There were certainly other consequences of the same Depression shock that will run counter to these above predictions. A lower L_1^d in an area may lead to worse health outcomes, wage scarring, and unemployment hysteresis (Duque and Schmitz 2025; Moulton 2017; Mathy 2018). These forces would all lead to lower wages and lower productivity growth. The empirical analysis will therefore also test which class of forces, the schooling channel or those described above, ultimately dominates.

3 Methodology

3.1 First Differences Model

To empirically test for the prevalence of an opportunity cost schooling mechanism during the Great Depression, I construct a first differences model of high school attainment change across counties. The baseline specification is the following:

$$\Delta HS_share_{cg} = \alpha + \beta U_{cg} + \mathbf{X}'_{cg}\Gamma + \delta_c + \gamma_g + \epsilon_{cg} \quad (1)$$

This specification is a panel regression at the county level that partitions into 8 subgroups within each county: sex (male, female) $\times$ race (white, non-white) $\times$ father’s occupational score (strictly above, or at or below, the within-county median among ages 16–25 children). ΔHS_share_{cg} measures the change in HS attainment of 18 year old cohorts from 1929 to 1933 in county c subgroup g .⁵ U_{cg} is my primary regressor, which records the unemployment rate for each county-subgroup. $\mathbf{X}'_{cg}$ includes county-subgroup level controls: the initial 1929 cohort share of high school attainment and the change in that share from 1925 to 1929. δ_c and γ_g are county and subgroup fixed effects, respectively. β is the parameter of interest, which captures the relationship between the unemployment rate and the change in high school attainment by county-subgroup. Every county regression is weighted by subgroup-cell population and clusters standard errors by state.⁶

There are several threats to a causal interpretation of this ordinary-least-squares (OLS) framework. The unemployment rate for ages 16–18 is taken from a single week of the 1930 census and is often computed inside a small county-subgroup cell, so classical measurement

5. Looking at the 18 year old cohort in 1933 considers the individuals who were 16 in 1931, the peak national unemployment level of the regression. I show later that this specification is robust to additional cohort end years (1931, 1932, and 1934).

6. Before estimation, I drop county cells that lie strictly outside the pooled 1st or 99th percentile of the 1930 unemployment rate, the 1929–1933 high school change, or the 1929 high school share.

error in U_{cg} attenuates β toward zero. Reverse causality operates through the same measured rate: a collective decision to remain in school reduces employment and mechanically raises unemployment. A more prevalent concern is omitted local shocks that move unemployment and educational attainment together. The failure of Caldwell and Company in 1930 led to many correspondent-bank failures in local economies. In 1930-31, a serious drought hit farmers in the Ohio and Mississippi valleys. Along with school district financing decisions, shocks such as these enter β through the measured unemployment rate.

To address these concerns, I advance the baseline OLS framework with a shift-share instrumental variable (IV) approach (Borusyak, Hull, and Jaravel 2022). This framework consists of a two-stage least squares estimation. The first stage is the following:

$$U_{cg} = \alpha + \gamma Z_{cg} + \mathbf{X}'_{cg}\Gamma + \delta_c + \gamma_g + \eta_{cg} \quad (2)$$

In this equation, the shift-share IV Z_{cg} uses 1929 employment shares by sector-industry for 8 sex-by-race-by-father-SES groups within ages 16–18. The shifts are national sector-industry employment changes, described further in Section 4. This IV regressor predicts the exogenous aspects of the local subgroup unemployment rate, which I use in the second stage:

$$\Delta HS_share_{cg} = \beta_0 + \beta_1 \hat{U}_{cg} + \mathbf{X}'_{cg}\Gamma + \delta_c + \gamma_g + \epsilon_{cg} \quad (3)$$

The second-stage parameter β_1 measures the local average treatment effect (LATE) of unemployment on education, removing concerns regarding reverse causality and omitted local shocks. For a valid causal interpretation of this parameter, I make the standard IV assumptions. First, I assume the exogeneity of national employment shifts by sector-industry. Within the context of this study, this is to say national sector-industry shocks are uncorrelated with unobserved factors affecting local education. The second assumption is exclusion of the instrument, which is to say the predicted Bartik shock affects educational attainment only through changes in local unemployment. Lastly, I assume the IV is relevant, or a strong predictor of actual local unemployment, which I test with the first stage in Section 5.

To assess the credibility of the shift-share design, I follow the shock-level diagnostics in (Borusyak, Hull, and Jaravel 2022). Their representation rewrites the Bartik instrument as an equivalent industry-level IV: national employment changes are the shocks, and the relevant regression weights are products of each county’s population weight and its share of national employment in that industry. The weight on industry k is

$$s_k = \sum_c \omega_c s_{ck},$$

where ω_c is county c ’s share of sample population and s_{ck} is county c ’s share of national

employment in industry k . Identification in the shift-share requires the shocks to be as good as randomly assigned with respect to unobservables that affect local schooling. Variation must also come from a large set of imperfectly correlated industry shocks to average out idiosyncratic industry-level confounders. To test these two identifying assumptions, I report both the concentration of the shock-level weights and a shock-level balance test of pre-Depression county characteristics.

First, no single industry appears to be driving the variation of the instrument. Table A1 of the Appendix reports the shock-level weight concentration for the national employment shifts. Across the 54 sector and manufacturing industries that enter the instrument, the Herfindahl–Hirschman index of shock weights is 0.045, corresponding to an effective number of shocks of about 22. This exceeds the benchmark effective shocks of 20 deemed satisfactory from the Monte Carlo simulations of (Borusyak, Hull, and Jaravel 2022). No single industry accounts for more than about 8 percent of total weight. This dispersion is reassuring for the many-shocks rationale: the first stage is not driven by one or two industries whose national declines might systematically coincide with unobserved local determinants of schooling.

Additionally, shocks are generally balanced along baseline characteristics. Table A2 of the Appendix reports these shock-level balance tests. Exposure-weighted 1925 and 1929 high school shares are unrelated to industry shocks. Urban, white, male, citizen, and high-school-age shares, mean age, the 1925–1929 high school pre-trend, and log population are also balanced. The veteran share is associated with the shocks at the 10 percent level.

Taken together, the sufficiently large effective number of shocks and the balance of pre-Depression schooling support treating national industry shifts as plausibly exogenous for local educational outcomes, and thereby support the exclusion restriction that the Bartik instrument affects attainment only through local labor-market conditions.

Later in the long run analysis, I will use an analogous framework to this short-run IV, but removing the panel subgroup dimension. I regress county unemployment in 1930 as the endogenous regressor, instrument with the county shift-share instrument, and regress on various longer term outcome measures (change in workforce education, change in labor productivity, etc.) The particularities of these specifications are discussed further in Section 5.

3.2 Stacked Event Study

To complement the first differences framework, I also estimate the effect of 1930 unemployment separately for each cohort in a stacked IV regression. The sample is the same

county-by-subgroup panel as in the first differences model: 8 demographic groups within each county, with the unemployment rate and the Bartik shock measured for ages 16–18. Each subgroup-cell is observed once for every cohort that turned 18 between 1924 and 1938. The outcome, HS_{cgt} , is the high school completion share of the county-subgroup-cohort that turned 18 in year t . The unemployment rate U_{cg} and the Bartik shock Z_{cg} do not vary across cohorts within a subgroup-cell, as it is always the 1930 shock.

I interact both the unemployment rate and the instrument with cohort indicators, omitting 1929 for the relative value, and estimate the interactions jointly. The first stage is

$$U_{cg} \times \mathbb{1}\{t = y\} = \pi_y (Z_{cg} \times \mathbb{1}\{t = y\}) + \theta_{cg} + \phi_t + \nu_{cgt} \quad (4)$$

and the second stage is

$$HS_{cgt} = \sum_{y \neq 1929} \beta_y (U_{cg} \times \widehat{\mathbb{1}\{t = y\}}) + \theta_{cg} + \phi_t + \varepsilon_{cgt}. \quad (5)$$

where θ_{cg} denotes the fixed effect for each county-subgroup cell, ϕ_t denotes the cohort fixed effect, and β_y measures the relative causal effect of Depression severity on each cohort year. Regressions are weighted by subgroup-cell population and standard errors are clustered by state. Cohorts that turned 18 in 1939 or 1940 are excluded, because those individuals were still close to high school age when schooling was recorded in the 1940 census.

The IV assumptions are the same as in the first differences model. What this specification adds is the cohort path. Coefficients on cohorts who turned 18 before 1929 are the pretrend check. Coefficients on cohorts who turned 18 after 1929 trace the dynamic effect that the 1929–1933 first difference summarizes in a single parameter.

4 Data

This section details the datasets and data methods used to measure local severity of the Depression, educational attainment, and labor market and manufacturing outcomes in the longer run. To do this, I make use of full count census records, census links, national employment by industry, and manufacturing indicators by county. Where not otherwise indicated, I collect or aggregate this information to the county level.

4.1 Full Count Census of Population

I gather microdata using the full count Censuses of Population from 1920, 1930, and 1940, accessed through IPUMS (Ruggles et al. 2024). I utilize information on individual age, sex, race, state and county of residence, urban status, employment and labor force status, occupation and industry, and a variety of socioeconomic status (SES) scores. The 1940 census is the first to record completed years of schooling for the whole population, which I use to construct educational outcomes. From the 1930 census I use employment status to construct county unemployment rates by demographic subgroup (described further below). I incorporate the 1920 census to interpolate 1929 employment shares and ascertain father SES scores, which I will describe further in the sections on construction of the shift-share instrument and census linking, respectively.

Educational Attainment For the majority of the short run results, I focus on high school completion, which I classify as 1 if the 1940 census respondent indicated having completed at least 12 years of schooling, and 0 otherwise. The national series in Figure 1 and the county summary statistics (Table A3 of the Appendix) also report completed years of schooling from the same census variable. I assign each individual to their respective cohort year, which I define as the year they turned 18 according to the 1940 census.

$$\text{cohort year} = 1940 + 18 - \text{age}_{1940}$$

For the county-level analysis, I aggregate the share of high school graduates by cohort year – demographic subgroup, according to their 1930 county of residence (see Section 4.2 for details on linking). With these data, I calculate changes in high school attainment shares across years, which I outline in more detail in Section 3.

Unemployment I directly measure local labor market exposure to the Depression using employment status from the 1930 census. I classify individuals as in the labor force if they respond as either employed or unemployed. I then aggregate these responses to the county – demographic subgroup level. I focus on the unemployment rate of individuals aged 16–18 in 1930, as this is the potential labor market high school-age students may enter into.

Demographic subgroups To exploit within-county variation in this relationship, I calculate unemployment rates and cohort year high school attainment shares for demographic subgroups within each county. Specifically, I partition the population into 8 subgroups: sex

(male v. female) $\times$ race (white v. non-white) $\times$ father’s socioeconomic status. The score is the co-resident father’s 1930 occupational score, or the 1920 co-resident father score when the 1930 score is missing. High SES means that score is strictly above the within-county median among children aged 16–25. A score equal to the median is low. Individuals with no father score are left out of the 8 subgroup cells.⁷

Occupational Data To study the longer-term effects of shocked individuals, particularly in terms of their job sorting, I use census data on the occupation of workers, as well as the associated SES scores of those occupations (OccScore, EdScor, ErScor, SEI, LIDO).

The occupational income score (OccScore) assigns each occupation the median total income of its workers in 1950, and is standardized into a score between 0 and 80. The Duncan socioeconomic index (SEI) is a prestige-based status score that uses 1947 National Opinion Research Center (NORC) prestige ratings alongside the education and income of workers in 1950 to form an occupation ranking. The occupational earnings score (ErScor) ranks occupations according to median earned income, which includes wages and self-employment income, in 1950. The occupational education score (EdScor) ranks occupations according to the percentage of workers who completed at least one year of college in 1950. The LIDO SES score differentiates beyond the standard SES scores above by allowing scores to differ upon demographic characteristics as well (Saavedra and Twinam 2020).⁸ The summary statistics of these variables on the linked sample are shown in Table A4 of the Appendix.

4.2 Census Linking

Due to the education data existing solely in the 1940 full count census, I must link individuals back to the 1930 census to combine their location during the time of the Depression with their ultimate educational attainment by 1940.⁹ To do this, I make use of the Census Tree 1930 – 1940 crosswalks for both men and women (Price et al. 2023; Price et al. 2021). Due to the imbalanced nature and selectivity of linking for individuals, I estimate a probit model for successful linkage as a function of observable characteristics. The probit and balance

7. Of all linked individuals, 7.5% do not have an observed occupational score from their father in 1920 or 1930.

8. Specifically, the Lasso Industry Demographic and Occupation (LIDO) is built from cross-validated lasso regressions of log income, within industry, on occupation, age, sex, race, and state, using 1950 census earnings.

9. As a more minor note, I also make links to 1920 to determine individuals’ fathers’ SES scores in that year, for where 1930 scores are unavailable.

table use 1940 respondents aged 18–34.¹⁰ The probit includes age indicators interacted with sex, census region, an indicator for living in one’s state of birth, and high school completion interacted with urban status and with race. This probit generates linking probabilities conditional on these characteristics, which I then invert to use as weights in creating a balanced (conditional on weighting) linked sample. Table A5 in the Appendix compares linked and unlinked 1940 individuals with and without these inverse probability weights. Using the weights appears to address much of the disparity between the linked and unlinked populations. With the linked sample and probability weights, I can generate accurate high school attainment shares by county-cohort-subgroup, as described in more detail above.

4.3 Shift-Share Instrument Data

To address the endogenous regressor of local unemployment, I construct an instrument from local industry employment shares and national industry shifts. Local employment shares are measured for eight sex-by-race-by-father-SES groups within ages 16–18. National shifts cover 8 sectors, with manufacturing split into 47 industries. The instrument takes the following form:

$$Z_{cg} = 100 * \sum_i \omega_{cgi,1929} \Delta E_{i,1929 \rightarrow 1931},$$

where $\omega_{cgi,1929}$ is the 1929 employment share of industry (or sector) i in county c for demographic group g , and $\Delta E_{i,1929 \rightarrow 1931}$ is the corresponding national employment change.¹¹

Local industry shares I construct employment shares from full count census industry codes, within county, age group, and sex-by-race-by-father-SES cell. I draw from 8 different sectors: agriculture, mining, construction, finance/insurance/real estate (FIRE), other private services, retail, wholesale, and manufacturing. I further disaggregate manufacturing into 47 different industries, also provided by the census industry classification. Due to the census recordings being in 1930, I linearly interpolate the baseline shares with the 1920 full count census employment shares in that same county, age group, and sex-by-race-by-father-SES cell.

$$Emp_{i,1929} = Emp_{i,1920} + \frac{9}{10} (Emp_{i,1930} - Emp_{i,1920}).$$

10. I also construct an auxiliary linked sample of all individuals aged 19 and above, to calculate workforce education by county in 1930 and 1940.

11. For robustness, I also construct alternative instruments using manufacturing industries only and with the national employment change from 1929 to 1933.

National employment shifts I then assemble national employment shifts from three separate sources. I draw from series in the *Historical Statistics of the United States* (HSUS) for shifts in agriculture, mining, construction, FIRE, and other private services.¹² For the retail and wholesale sectors, I digitize employment indices from the *Survey of Current Business*. For the manufacturing industries, I use national wage earner counts by detailed industry from the 1929, 1931, and 1933 Census of Manufactures (COM). To concord these COM industries to 47 manufacturing industries in the Census of Population (used for the employment shares), I make use of an industry crosswalk between the two sources constructed by (Fiszbein et al. 2020).

4.4 Census of Manufactures

In addition to the national wage earner counts by detailed industry, I also make use of the biennial COMs to study long-run effects on productivity and other aggregate manufacturing outcomes at the county level. The key variables I use are number of establishments, wage earners, total wages, cost of materials, value of product (gross output), and value added. From these indicators I also construct the level and log changes, as well as per-worker and per-establishment measures. My main outcome of interest in this analysis is labor productivity, which I define as value added per worker. I have these data for 1919, 1929, and 1939 from IPUMS NHGIS (Schroeder et al. 2026), 1947 from (Haines and Research 2010), and 1927 from (Janas 2026b).

4.5 Miscellaneous Data

Other auxiliary datasets are implemented in the long run analysis. New Deal spending and lending by county are taken from (Fishback, Horrow, and Kantor 2005) and World War II contract spending is taken from (Brunet, Hilt, and Jaremski 2026), both which are used as controls in the long run productivity growth robustness analysis.

12. For mining, construction, FIRE, and other private services, the series is Table Ba831–839 Employees on nonagricultural payrolls, by industry: 1900-1940 (Lebergott). For agriculture, the series is Table Ba470-477 Labor force, employment, and unemployment: 1890-1990.

5 Results

5.1 Short Run Effects

5.1.1 First Difference Results

To test the extent to which exposure to the Great Depression may have generated additional high school attainment at the county level, I carry out the estimation of the OLS and IV framework of Equations 1 - 3.

There is a strong positive association between local Depression severity and the change in HS attainment rate for 18-year olds during the Depression. Table 1 reports the OLS estimates in Columns (1)–(3). In the full specification, a 1 percentage point increase in the unemployment rate is associated with a 0.139 percentage point increase in the growth of the 18 year old HS attainment rate.¹³

Table 1: **Effect on Δ HS share**

	OLS			IV		
Unemp rate, 1930	0.146*** (0.028)	0.035 (0.026)	0.139*** (0.024)	0.474*** (0.065)	0.681*** (0.118)	0.384*** (0.114)
HS share, 1929	-0.009 (0.015)	-0.075*** (0.015)	-0.591*** (0.018)	-0.038** (0.018)	-0.088*** (0.026)	-0.604*** (0.019)
HS change, 1925-29	-0.411*** (0.011)	-0.403*** (0.010)	-0.169*** (0.011)	-0.395*** (0.012)	-0.398*** (0.015)	-0.164*** (0.012)
Observations	13,975	13,953	13,953	13,972	13,949	13,949
R-squared	0.19	0.43	0.56			
First-Stage F				356.58	275.36	238.27
County FE		✓	✓		✓	✓
Subgroup FE			✓			✓

Note: Each observation is a county by sex-by-race-by-father-SES cell for ages 16–18. The outcome is the percentage-point change, from 1929 to 1933, in the high school completion share of the cohort that turned 18. Unemployment is the 1930 rate, in percentage points, for ages 16–18 in that cell. Every column controls for the cell’s 1929 high school share and the percentage-point change in that share from 1925 to 1929. Columns (1) and (4) include no fixed effects, columns (2) and (5) add county fixed effects, and columns (3) and (6) add subgroup fixed effects. Regressions are weighted by cell population and cluster standard errors by state. The instrument is the 1929–1931 Bartik shock for the same cell. County cells strictly outside the pooled 1st or 99th percentile of the 1930 unemployment rate, the 1929–1933 high school change, or the 1929 high school share are dropped before estimation.

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

Furthermore, the IV estimates state an even larger effect, in Columns (4)–(6). In the full specification, a 1 percentage point higher county-subgroup unemployment rate raised high

13. The left panel in Figure A1 of the Appendix plots a bin scatter of the final OLS relationship. Figure A3 of the Appendix reports the variation of the 1930 unemployment rate across counties.

school attainment by an additional 0.38 percentage point from 1929 to 1933. The first-stage F statistic in that specification is 240.¹⁴

The IV coefficient exceeds the OLS coefficient in all corresponding columns. This is in line with the direction of bias the OLS regression is anticipated to have. The county-subgroup unemployment rate is a noisy measure of the shock facing the cohort, and classical measurement error attenuates the estimate toward 0. The local shocks discussed above, including county bank failures, the 1930–31 drought, idiosyncratic plant shutdowns, and local cuts to school finances, raise measured unemployment and slow high school attainment, so they enter the regression negatively correlated with the schooling residual. The Bartik keeps the part of unemployment that comes from exposure to national industry declines, and along that variation the high school response is larger.

The relationship established here is robust to a wide suite of different specifications, including alternative windows for the high school attainment change, alternative fixed effects, alternative forms of weighting, and alternative forms of the shift-share instrument. The results with these alternatives for the final IV specification are reported in Tables A6–A9 of the Appendix. The estimates remain positive and significant across the alternative attainment windows and fixed effects. The unweighted regression and the manufacturing-only Bartik are positive and insignificant.

This relationship also differs considerably by subgroup, appearing most strongly for white, low-income students. Table 2 reports these results. Within each demographic group, I keep the underlying subgroup cells and absorb county and subgroup fixed effects. A 1 percentage point increase in unemployment raises the high school share by 0.52 percentage points for women and 0.30 for men. The nonwhite estimate is 0.36 and insignificant, while the white estimate is 0.40 and significant. The low-SES estimate is 0.50 and significant, while the high-SES estimate is -0.25 and insignificant, with a weak instrument ($F=3.5$).

5.1.2 Event Study Results

As an additional verification of the short run mechanism, I complement the first differences model with a stacked IV event study that estimates the effect for each individual cohort year. Figure 2 reports these point estimates and standard errors by cohort year.

This exercise reveals a striking pattern on the dynamic of the effect over the course of the decade. While no effect appears in 1930, a small but statistically significant effect emerges for

14. Figure A2 of the Appendix plots a bin scatter of the first-stage relationship. Figure A4 of the Appendix reports the variation of the Bartik shock across counties.

Table 2: **Effect on Δ HS share**

	Gender		Race		SES	
	Female	Male	Non-white	White	Low	High
Unemp rate, 1930	0.524* (0.267)	0.304*** (0.107)	0.358 (0.280)	0.403*** (0.119)	0.495*** (0.103)	-0.246 (0.583)
HS share, 1929	-0.641*** (0.025)	-0.510*** (0.027)	-0.771*** (0.047)	-0.623*** (0.025)	-0.619*** (0.027)	-0.812*** (0.024)
HS change, 1925-29	-0.149*** (0.022)	-0.170*** (0.014)	-0.066* (0.033)	-0.171*** (0.013)	-0.170*** (0.018)	-0.076*** (0.018)
Observations	6,076	7,514	1,751	11,720	7,575	6,033
First-Stage F	126.31	251.95	26.53	229.13	101.38	3.54
County FE	✓	✓	✓	✓	✓	✓
Subgroup FE	✓	✓	✓	✓	✓	✓

Note: IV estimates on ages 16–18 sex-by-race-by-father-SES cells. Each column keeps the cells in that group: women are subgroups 1–4, men 5–8, nonwhite subgroups 1, 2, 5, and 6, white subgroups 3, 4, 7, and 8, low father SES subgroups 1, 3, 5, and 7, and high father SES subgroups 2, 4, 6, and 8. High father SES is a father occupational score strictly above the within-county median among ages 16–25. Every column controls for the cell’s 1929 high school share and the 1925–1929 high school change, absorbs county and subgroup fixed effects, is weighted by cell population, and clusters standard errors by state. The instrument is the cell’s 1929–1931 Bartik shock. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

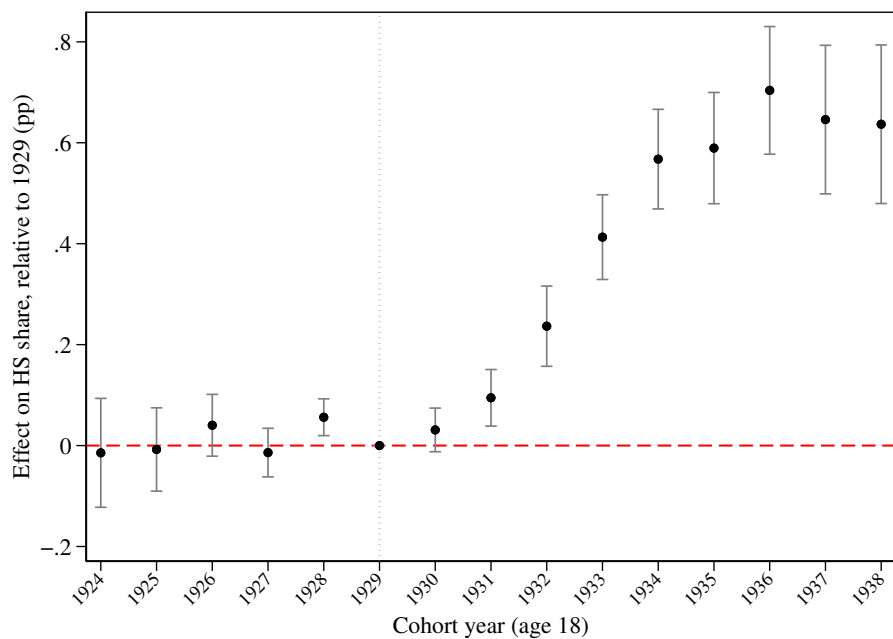
the 1931 cohort, and this effect grows over time, ultimately peaking around a 0.6 percentage point increase for cohort years 1934–1938. Among cohorts who turned 18 before 1929, the 1924, 1925, 1926, and 1927 estimates are insignificant, while the 1928 estimate is 0.06 and significant. This generally supports the notion that pretrends in education are not driving this dynamic post-shock effect.

Both the county first difference and stacked event study results shown support the notion that increased severity of the Great Depression in a local economic area generated increased high school educational attainment. With this short run mechanism firmly established, I move next to analyzing the longer-run implications of increased educational attainment.

5.2 Long Run Effects

In this section, I empirically test for the proposed predictions laid out in Section 2.3: First, the individuals who gained additional education may tend to sort into higher skilled jobs given their expanded potential skill set. This would be influencing growth along the extensive margin, where higher productive jobs and industries are able to expand because more individuals are capable of operating in them. I test for this in Section 5.2.1. Second, I consider the intensive margin, where additional schooling may raise a worker’s productivity

Figure 2: **Effect of 1930 Unemployment on HS Attainment Share**



Note: Stacked IV estimates. Each point is the effect of the ages 16–18 county-subgroup unemployment rate on that cohort’s high school completion share, in percentage points, relative to the 1929 cohort. The sample stacks cohorts who turned 18 in 1924–1938 in the same county-subgroup cells as Table 1, keeping cells with a nonmissing 1929 share. Unemployment and the 1929–1931 Bartik shock do not vary across cohorts. Both are interacted with cohort indicators, omitting 1929, which is plotted at zero. The regression absorbs county-by-subgroup and cohort fixed effects, is weighted by cell population, and clusters standard errors by state. Bars are 95 percent confidence intervals.

in whatever job or industry they hold. I test for this in Section 5.2.2 by gauging the change in the educational attainment of the overall workforce during this time. Last, one might expect an increase in skills and capabilities, both at the individual and occupational level, to manifest in higher degrees of labor productivity overall, which I check for in Section 5.2.3.

5.2.1 Skill sorting

Young adults who finish high school as a result of the Depression may sort into higher-skilled jobs. I test that prediction for the cohorts who turn 18 between 1931 and 1939, using the Census Tree links between 1930 and 1940. Each person is attached to the 1930 unemployment rate of adults aged 16–18 in the same county, sex, race, and father-SES cell. The instrument is the Bartik shock for that same cell. Regressions use the inverse-probability linking weights and cluster standard errors by state. The outcome is the person’s occupation status in 1940, which need not be in the 1930 county.

To measure the status of one’s occupation, I use 5 different socioeconomic indicators: occupational income scores, the Duncan SEI, ErScor, EdScor, and the LIDO score.¹⁵ I standardize each score to mean zero and standard deviation one in the estimation sample, using the linking weights, so coefficients are in standard deviations. For person i in 1930 county c , subgroup g , and cohort t , I estimate the following regression:

$$SES_i = \beta U_{cg} + \gamma HS_{cg}^{1929} + \theta \Delta HS_{cg}^{1925-29} + \delta \overline{SES}_c^{1930} + \alpha_c + \alpha_g + \alpha_t + \varepsilon_i, \quad (6)$$

with U_{cg} instrumented by the cell Bartik shock. HS_{cg}^{1929} is the 1929 high school share in the cell, $\Delta HS_{cg}^{1925-29}$ is that cell’s high school share change from 1925 to 1929, and $\overline{SES}_c^{1930}$ is the 1930 county workforce average of the same score variable. County, subgroup, and cohort fixed effects are also included.

Relative to other youths in less exposed areas, Depression-exposed youths tended to sort into higher-status occupations by 1940. Table 3 reports the IV estimates. A 1 percentage point higher cell unemployment rate raises 1940 occupational status by 0.033 standard deviations on the occupational income score, 0.054 on ErScor, 0.007 on EdScor, and 0.012 on LIDO. The Duncan SEI coefficient is 0.002 and insignificant.

A common concern with SES measures is the lack of variation in measuring agricultural occupations, which were the dominant form of employment for the US population at this

15. The LIDO score is taken from (Saavedra and Twinam 2020), which incorporates 1940 full count income data into the occupational ranking. For more information on all of these SES measures, see the discussion in Section 4.

Table 3: **Effect on 1940 Job SES, Youth Cohorts**

	OccScore	SEI	ErScor	EdScor	LIDO
Unemp rate, 1930	0.033*** (0.004)	0.002 (0.002)	0.054*** (0.003)	0.007** (0.003)	0.012*** (0.004)
County HS share, 1929	0.002*** (0.000)	0.003*** (0.000)	0.003*** (0.001)	0.006*** (0.001)	0.001*** (0.000)
HS change, 1925-29	-0.001*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.000 (0.000)
Workforce score, 1930	0.029*** (0.003)	0.019*** (0.001)	0.000*** (0.000)	0.001*** (0.000)	0.059*** (0.007)
Observations	5,720,577	5,621,292	5,680,687	5,680,687	5,244,545
First-Stage F	142.968	263.297	341.302	315.475	161.945
County+Subgroup+Cohort FE	✓	✓	✓	✓	✓

Note: IV estimates for linked individuals who turn 18 in 1931–1939. Outcomes are 1940 occupational scores, standardized to mean zero and standard deviation one in the estimation sample using the linking weights. The unemployment rate is the ages 16–18 rate in the person’s 1930 county and sex-by-race-by-father-SES cell, instrumented by that cell’s 1929–1931 Bartik shock. Regressions control for that cell’s 1929 high school share, its 1925–1929 high school change, and the 1930 mean of the same score among workers aged 19 and older in that county and subgroup. They absorb county, subgroup, and cohort fixed effects, use inverse-probability linking weights, and cluster by state. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

time. If the specification of Table 3 is rerun after dropping farmers, farm managers, and farm laborers, all SES scores are indistinguishable from zero.

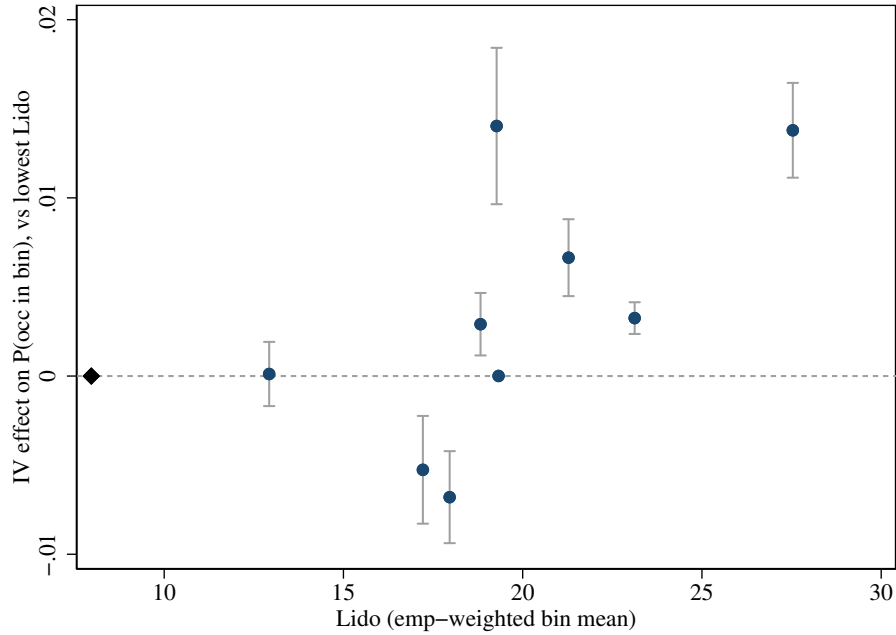
Much of this skill sorting is related to Depression-exposed youths being less likely to sort into agriculture. Table A11 of the Appendix estimates the probability of entering 1940 industries directly. A 1 percentage point higher unemployment rate lowers the probability of being in the agricultural industry by 0.5 percentage points, while it increases the probability of being in the services industry by 0.3 percentage points.

The skill sorting found here is generally consistent across the distribution of the score. Figure 3 further decomposes the effect on the LIDO score into ten bins of approximately equal total linking weight. I run a stacked IV regression that estimates each individual’s probability of sorting into each bin, relative to the lowest LIDO score bin:

$$\begin{aligned}
\mathbf{1}\{bin_i = k\} = & \sum_{m=2}^{10} \beta_m U_{cg} \times \mathbf{1}\{k = m\} + \sum_{m=2}^{10} \gamma_m HS_{cg}^{1929} \times \mathbf{1}\{k = m\} \\
& + \sum_{m=2}^{10} \theta_m \Delta HS_{cg}^{1925-29} \times \mathbf{1}\{k = m\} + \alpha_{ck} + \alpha_g + \alpha_t + \varepsilon_{ik},
\end{aligned} \tag{7}$$

I instrument each unemployment interaction with the corresponding Bartik interaction.

Figure 3: **Effect of 1930 Unemployment on 1940 Job SES, Youths**



Note: Stacked individual IV estimates for linked people who turn 18 in 1931–1939. Occupations with a nonmissing LIDO score, excluding OCC1950 codes 995, 997, and 999, are ordered by mean LIDO and packed into ten bins of approximately equal linking weight. The regression expands each person by bin. The outcome is an indicator for the person’s own bin. Each point is the effect of a 1 percentage point higher 1930 unemployment rate, in the ages 16–18 sex-by-race-by-father-SES cell of the 1930 county, on the probability of that bin relative to the lowest-LIDO bin. Unemployment, the 1929 high school share, and the 1925–1929 high school change are interacted with bin, omitting the lowest bin, and each unemployment interaction is instrumented by the corresponding Bartik interaction. The regression absorbs county-by-bin, subgroup, and cohort fixed effects, uses linking weights, and clusters by state. The Kleibergen-Paap first-stage F statistic is 33. The omitted bin is plotted at zero. The horizontal axis is the linking-weight-weighted mean LIDO of the bin. Bars are 95 percent confidence intervals.

County fixed effects are interacted with the bin to control for any baseline differences. The coefficient β_k is the effect of a 1 percentage point higher cell unemployment rate on the probability of sorting into bin k , relative to the lowest-LIDO bin. The joint first-stage F statistic is 33.

Figure 3 shows a consistent story with Table 3. Depression-exposed youths are more likely to sort into the higher-binned LIDO score occupations. While the shift is not completely monotone, the higher status occupations appear to be driving this higher average.

5.2.2 Workforce Education

One important consideration for increased educational attainment is its effect on the workforce. If more youths are completing high school and entering the labor market, this could generate a nontrivial increase in the human capital stock of the workforce, which can have

subsequent effects on output. To test this, I apply the same framework as in the county baseline of Equation 3, but with the outcome variable as the change in a county’s high school share of the workforce from 1930 to 1940. I also consider only aggregate effects here, so there is no differentiation by demographic subgroup. The unemployment rate and Bartik shock are the ages 16–18 county aggregate. The outcome is the 1930–1940 change in the high school share for the labor force aged 19 and older. By dropping the subgroup dimension, I can only use state fixed effects, so the fuller columns also include the urban share and the 1930 high school share of the workforce.

Increased depression severity did indeed increase the growth in the educational attainment of a county’s workforce. Table 4 reports these results. In the full specification, a 1 percentage point higher unemployment rate raised the high school share of the workforce by an additional 0.21 percentage point from 1930 to 1940. As in the short-run attainment results, the unemployment rate is a noisy measure of the shock, and the Bartik keeps the part of that variation that comes from exposure to national industry declines.

Table 4: **Effect on Workforce Education**

	OLS			IV		
Unemp rate, 1930	0.293*** (0.041)	0.133*** (0.044)	0.010 (0.023)	0.430*** (0.053)	0.417*** (0.138)	0.208* (0.114)
% Urban		0.005 (0.008)	-0.009 (0.006)		-0.020 (0.016)	-0.022* (0.011)
Workforce HS share, 1930		0.218*** (0.033)	0.198*** (0.026)		0.197*** (0.036)	0.191*** (0.030)
Observations	3,094	3,094	3,093	3,094	3,094	3,093
R-squared	0.202	0.408	0.661			
First-Stage F				434.78	179.66	71.78
State FE			✓			✓

Note: The outcome is the 1930–1940 change, in percentage points, in the high school completion share of the labor force aged 19 and older, counting the employed and the unemployed and using 1940 schooling in both years. The sample is one observation per county, the ages 16–18 aggregate cell. Columns (1) and (4) include only the 1930 unemployment rate. Columns (2) and (5) add the urban share and the 1930 workforce high school share. Columns (3) and (6) add state fixed effects. Regressions are weighted by the ages 16–18 population and cluster standard errors by state. The instrument is the ages 16–18 Bartik shock using 1929–1931 national employment changes. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To understand the key drivers behind this increase in workforce education, I decompose the change into five key margins. First, there are stayers: those who were in the workforce and same county in 1930 and 1940. Due to the retrospective nature of education being collected in the 1940 census, this margin will have no effect on the change in the education of the workforce, because it is static throughout. Then, there are two other classes of margins I consider: through migration, with workers in 1930 who either entered the county

Table 5: **Decomposition of Workforce Δ HS Attainment**

	Aggregate	Stayer	Migration		Labor Force	
			In-migrant	Out-migrant	Enterer	Exit
Unemp rate, 1930	0.208*	0.000	0.032*	-0.006	0.212*	-0.030*
	(0.114)	(.)	(0.018)	(0.029)	(0.126)	(0.015)
Workforce HS share, 1930	0.191***	0.000	0.052***	-0.015	0.145***	0.008*
	(0.030)	(.)	(0.013)	(0.012)	(0.024)	(0.004)
% Urban	-0.022*	0.000	-0.006***	-0.003	-0.023*	0.011***
	(0.011)	(.)	(0.002)	(0.002)	(0.012)	(0.001)
Observations	3,093	3,093	3,093	3,093	3,093	3,093
First-Stage F	71.78	.	71.78	71.78	71.78	71.78
State FE	✓	✓	✓	✓	✓	✓

Note: Each column is that margin's contribution, in percentage points, to the 1930–1940 change in the county high school share in Equation 8. The labor force is the employed and the unemployed aged 19 and older, and schooling is the 1940 report in both years. Stayers are in the labor force in both years and the same county; their contribution is zero because both means use 1940 schooling for the same people. In- and out-migrants are in the labor force in both years and a different county in 1940. Entrants are in the 1940 labor force only, and exiters are in the 1930 labor force only. Each non-stayer term is the group's linking-weight share times its high school gap relative to stayers. The specification matches column (6) of Table 4: the ages 16–18 county aggregate, state fixed effects, the 1930 workforce high school share, the urban share, population weights, and the 1929–1931 Bartik shock. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

from another area in 1940 or exited the county after 1930, and labor force participation, with individuals both deciding to enter and exit the labor force between 1930 and 1940. The channel I have been studying, with increased high school graduates, would enter along this labor force entry margin, as they would be entering the labor force (if at all) sometime between 1930 and 1940. Equation 8 summarizes this decomposition. All margins are measured relative to the stayer group and multiplied by the group's person-weighted share of that year's workforce, s , so that all coefficients will sum to the aggregate change and can be interpreted as a contribution share.

$$\begin{aligned}
\Delta H_c = & (\bar{H}_{c,40}^{\text{stay}} - \bar{H}_{c,30}^{\text{stay}}) \\
& + s_{c,40}^{\text{in}} (\bar{H}_{c,40}^{\text{in}} - \bar{H}_{c,40}^{\text{stay}}) + s_{c,30}^{\text{out}} (\bar{H}_{c,30}^{\text{stay}} - \bar{H}_{c,30}^{\text{out}}) \\
& + s_{c,40}^{\text{ent}} (\bar{H}_{c,40}^{\text{ent}} - \bar{H}_{c,40}^{\text{stay}}) + s_{c,30}^{\text{exit}} (\bar{H}_{c,30}^{\text{stay}} - \bar{H}_{c,30}^{\text{exit}}).
\end{aligned} \tag{8}$$

Labor-force entry, the channel which contains recent high school graduates from the short-run mechanism, accounts for the majority of the aggregate workforce educational response. Table 5 reports the changes within each margin. As mentioned, the contribution of the stayer group is mechanically 0 by construction. Of the 0.21 aggregate effect, in-migration contributes 0.03, out-migration contributes -0.01 (and is insignificant), labor-force entry contributes 0.21, and labor-force exit contributes -0.03 . As a share of total effects, labor market entry accounts for essentially all of the increase.

To further understand the source of this workforce educational increase, I decompose in an alternative way: changes in schooling inside industries and changes in the industrial mix. In this setup, industries contribute three different terms to the aggregate workforce educational change: the within term weights each industry’s change in its high-school share by its 1930 employment share. The between term weights the change in its employment share by its 1930 high-school share. The interaction is the product of the two changes. Intuitively, this is comparing the workforce educational increase that comes from all industries becoming more educated against the increase coming from ex-ante higher educated industries becoming larger.

From this decomposition, Table A10 of the Appendix shows that of the 0.208 aggregate effect, a 1 percentage point higher unemployment rate raises the aggregate high-school share by 0.147 percentage points through the within industry change, -0.026 (and insignificantly) through the between industry change, and 0.086 through the interaction term. This finding suggests that much of the educational increase in the workforce comes from a general increase in the education of workforces across all industries, rather than higher-educated industries becoming larger.

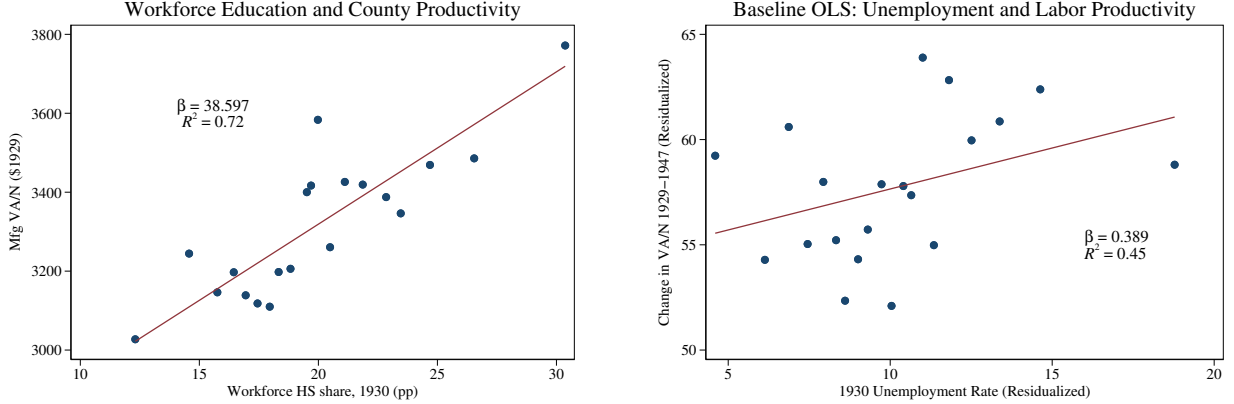
5.2.3 Productivity growth

As a result of individuals sorting into higher skilled jobs and more productive industries, together with a more educated workforce, one might expect downstream effects upon the efficiency of production in a shocked area. Panel A of Figure 4 shows that counties with a higher 1930 workforce high school share had higher 1939 manufacturing value added per wage earner. The idea is that with a greater human capital stock, these counties’ production functions have a more efficient labor input in their production of goods.

I move beyond the static relationship between education and productivity to gauge whether the severity of the Depression *induced* higher changes in labor productivity (value added per worker). To test this, I apply the same framework as in Equation 3, but with the outcome variable as the change in the county’s labor productivity from 1929 to 1947. The sample is the ages 16–18 county aggregate. In the full specification, I control for the level of productivity in 1929, the pretrend in productivity from 1919–1929, the share of the workforce who graduated high school in 1930, and the county urban share. I also include state fixed effects, weight by the age 16–18 population, and cluster errors at the state level.

Increased severity of the Depression led to greater growth in labor productivity. Panel B of Figure 4 plots the OLS relationship of the full specification, and Table 6 reports the whole

Figure 4: Education and Manufacturing Labor Productivity



(a) 1930 High School Share and 1929 Productivity

(b) Unemployment and Productivity Growth

Note: Both panels are binscatters on the ages 16–18 county aggregate, weighted by that population, with state fixed effects. Panel A plots 1929 manufacturing value added per wage earner, in 1929 dollars, against the 1930 workforce high school share, in percentage points, partialling out the urban share and 1929 materials per wage earner. Panel B plots 100 times the 1929–1947 log change in value added per wage earner against the 1930 unemployment rate, partialling out 1929 value added per wage earner, the 1919–1929 log change, the 1930 workforce high school share, and the urban share.

set of specifications. For the full specification, the OLS coefficient is 0.39 and significant at the 10 percent level. The IV coefficient is 2.04, with a first-stage F of 57. The IV once again generates larger coefficients than the OLS. Weaker local demand and firm exit in depressed counties, for example, can pull the OLS estimates toward zero, while the Bartik keeps the variation that comes from exposure to national industry declines.

I test the robustness of this relationship across several specifications. The added controls are 1929 materials per wage earner and their 1929–1939 change, 1929 wage earners and their 1929–1947 change, and the male and white population shares. I also drop counties at or above the 95th percentile of manufacturing employment per population, measure productivity growth from 1927 to 1947 and from 1929 to 1939, replace the 1919–1929 pre-trend with the 1927–1929 pre-trend, and add per capita New Deal spending and World War II contract spending. Table A12 in the Appendix reports all of these robustness specifications. The coefficient remains positive in every column, and statistically significant except when productivity growth is measured from 1929 to 1939.¹⁶ This further bolsters the notion that Depression severity in some way influenced the growth in labor productivity.

As an additional exploration, Table A13 in the Appendix reports the identical framework

16. As shown in Figure A5 of the Appendix, most counties had still not recovered even to their pre-Depression levels of productivity by 1939, so the 1929–1939 time span does not measure growth, per se.

Table 6: **Effect on $\Delta VA/N$**

	OLS			IV		
Unemp rate, 1930	0.695*** (0.249)	0.415* (0.226)	0.389* (0.210)	1.309*** (0.421)	2.234*** (0.646)	2.039*** (0.679)
Level, 1929	-0.014*** (0.001)	-0.014*** (0.001)	-0.015*** (0.001)	-0.015*** (0.001)	-0.014*** (0.001)	-0.015*** (0.001)
Pretrend		-0.211*** (0.029)	-0.225*** (0.028)		-0.210*** (0.031)	-0.223*** (0.030)
% Urban		0.197*** (0.043)	0.139*** (0.036)		0.075* (0.042)	0.032 (0.038)
Workforce HS share, 1930			0.689*** (0.162)			0.650*** (0.158)
Observations	2,035	2,034	2,034	2,035	2,034	2,034
R-squared	0.324	0.438	0.448			
First-Stage F				285.34	54.18	57.10
State FE		✓	✓		✓	✓

Note: The outcome is 100 times the log change in Census of Manufactures value added per wage earner from 1929 to 1947, so coefficients are approximate percent changes. The sample is the ages 16–18 county aggregate. Within OLS and within IV, the first column controls for 1929 value added per wage earner and includes no fixed effects. The second adds the 1919–1929 log change, the urban share, and state fixed effects. The third adds the 1930 workforce high school share. Regressions are weighted by the ages 16–18 population and cluster standard errors by state. The instrument is the ages 16–18 Bartik shock using 1929–1931 national employment changes. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

as above, but using changes in various manufacturing variables as the outcome instead of labor productivity. These results confirm that the changes observed appear to be particular to a story about labor productivity, consistent with the theory that increased educational attainment is driving this change. In addition to labor productivity, there is a positive and statistically significant effect on the change in wages per worker, consistent with the theory that workers are paid according to their marginal product, and the workforce is becoming more educated (productive) in these areas. The effect on the change in total value added is negative and significant at the 10 percent level. I use non-wage surplus per worker ($\Delta \ln((VA - W)/N)$) as a proxy for capital, and a statistically significant effect of Depression severity on that variable is supportive of the notion that capital deepening also played some role in generating more labor productivity growth.¹⁷ The change in the number of workers is negative and significant, though as shown in Table A12, it does not largely account for the growth in value added per worker.

Taken together, these results suggest that the Depression brought about increased growth in labor productivity in the long run. Increased education of the workforce is a theoretical channel for this growth, along with the expansion of more productive occupations and

17. There are no widely available alternatives for capital/investment by county over this period. The COM discontinued its measures of investment and horsepower by county in 1919 and 1929, respectively.

industries.

6 Conclusion

This paper studies increased human capital accumulation as a positive consequence of the Great Depression. Through lowered opportunity costs, hard hit counties saw increases in the high school completion of their 18 year old cohorts. I identify this short-run relationship with a shift-share instrument, covering 8 demographic subgroups, 8 sectors, and 47 manufacturing industries. I find this increase in attainment to be an effect that grew over the course of the decade.

These short-run schooling responses had long-run implications for the local economies they initially affected. In aggregate, harder-hit counties saw larger increases in the high school share of their workforce, accounted for mostly by labor-force entry. Young adults in higher-exposed counties sorted into higher-status occupations. Downstream of these effects, higher shocked counties experienced greater growth in the labor productivity of their manufacturing sectors. Taken together, these results suggest that Depression-induced schooling reshaped not only individual attainment but also the skill composition and productive efficiency of local economies.

These findings complement existing work on the short-run educational attainment channel in cities, as well as reverse structural transformation, during the Great Depression. With respect to structural change, I find a distinct trend within the subset of shocked young individuals, who tended to sort into higher SES jobs. More broadly, these results illustrate how large collapses in labor demand can accelerate human capital accumulation and leave lasting imprints on local growth.

References

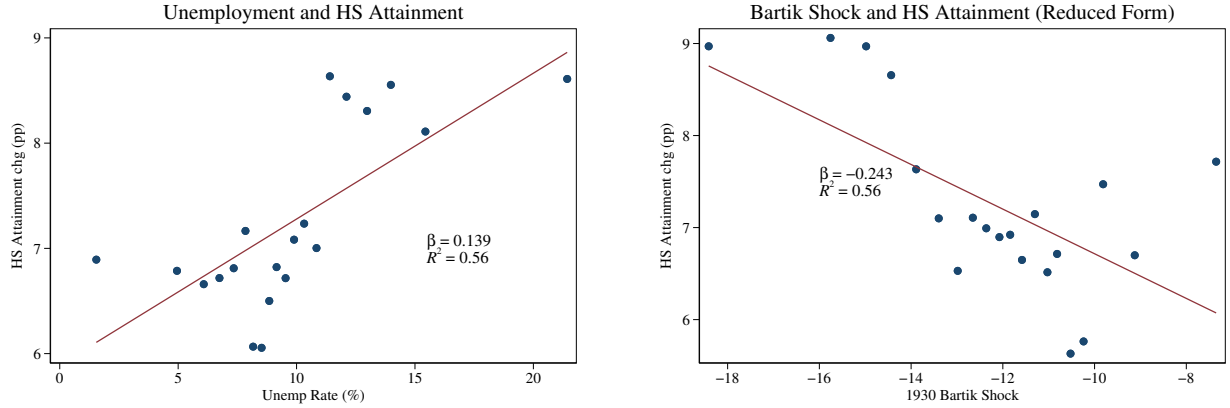
- Barr, Andrew, and Sarah Turner. 2015. “Out of work and into school: Labor market policies and college enrollment during the Great Recession.” *Journal of Public Economics* 124 (April): 63–73. <https://doi.org/10.1016/j.jpubeco.2014.12.009>.
- Benmelech, Efraim, Carola Frydman, and Dimitris Papanikolaou. 2019. “Financial frictions and employment during the Great Depression.” *Journal of Financial Economics*, JFE Special Issue on Labor and Finance, 133, no. 3 (September): 541–563. <https://doi.org/10.1016/j.jfineco.2019.02.005>.
- Black, Dan A., Terra G. McKinnish, and Seth G. Sanders. 2005. “Tight Labor Markets and the Demand for Education: Evidence from the Coal Boom and Bust.” *ILR Review* 59, no. 1 (October): 3–16. <https://doi.org/10.1177/001979390505900101>.
- Bleemer, Zachary, and Sarah Quincy. 2026. “Junior Colleges and Local Development.” *Working Paper*.
- Boone, Christopher D. A., and Laurence Wilse-Samson. 2023. “Structural Change and Internal Labor Migration: Evidence from the Great Depression.” *The Review of Economics and Statistics* 105, no. 4 (July): 962–981. https://doi.org/10.1162/rest_a_01116.
- Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. “Quasi-Experimental Shift-Share Research Designs.” *The Review of Economic Studies* 89 (1 (324)): 181–213. <https://doi.org/https://doi.org/10.1093/restud/rdab030>.
- Boustan, Leah Platt, Price V. Fishback, and Shawn Kantor. 2010. “The Effect of Internal Migration on Local Labor Markets: American Cities during the Great Depression.” *Journal of Labor Economics* 28, no. 4 (October): 719–746. <https://doi.org/10.1086/653488>.
- Boyd, Robert L. 2002. “A “Migration of Despair”: Unemployment, the Search for Work, and Migration to Farms During the Great Depression.” *Social Science Quarterly* 83 (2): 554–567. <https://doi.org/10.1111/1540-6237.00100>.
- Brunet, Gillian, Eric Hilt, and Matthew Jaremski. 2026. “Inflation, War Bonds, and Voter Backlash in the 1950s.” *The Review of Economics and Statistics* (May): 1–45. <https://doi.org/10.1162/REST.a.1783>.
- Charles, Kerwin Kofi, Erik Hurst, and Matthew J. Notowidigdo. 2018. “Housing Booms and Busts, Labor Market Opportunities, and College Attendance.” *American Economic Review* 108, no. 10 (October): 2947–2994. <https://doi.org/10.1257/aer.20151604>.
- Cheng, Dong, Mario J. Crucini, and Hanjo T. Kim. 2026. “Which U.S. States Suffered a Greater Great Depression and Why?,” Working Paper Series (April). <https://doi.org/10.3386/w35028>.
- Doxey, Alison, Ezra Karger, and Peter Nencka. 2026. “High Schools and the Uneven Rise in American Opportunity,” Working Paper Series (April). <https://doi.org/10.3386/w35068>.
- Duque, Valentina, and Lauren L. Schmitz. 2025. “Early-Life Exposure to the Great Depression and Long-Term Health and Economic Outcomes.” *Journal of Human Resources* 60, no. 6 (November): 1937–1970. <https://doi.org/10.3368/jhr.0421-11584R1>.

- Fishback, Price, William C Horrace, and Shawn Kantor. 2005. “Did New Deal Grant Programs Stimulate Local Economies? A Study of Federal Grants and Retail Sales During the Great Depression.” *Journal of Economic History* 65, no. 1 (March). <https://doi.org/https://doi.org/10.1017/S0022050705050023>.
- Fiszbein, Martin, Jeanne Lafortune, Ethan G. Lewis, and José Tessada. 2020. “Powering Up Productivity: The Effects of Electrification on U.S. Manufacturing” (November). <https://doi.org/https://doi.org/10.3386/w28076>.
- Goldin, Claudia. 1998. “America’s Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century.” *The Journal of Economic History* 58, no. 2 (June): 345–374. <https://doi.org/10.1017/S0022050700020544>.
- Goldin, Claudia Dale, and Lawrence F. Katz. 2008. *The Race Between Education and Technology*. Harvard University Press, June. <https://doi.org/https://doi.org/10.2307/j.ctvjf9x5x>.
- Goodman, Joshua, and Joseph Winkelmann. 2025. “Labor Market Strength and Declining Community College Enrollment.” *NBER Working Paper Series*, <https://doi.org/https://doi.org/10.3386/w34498>.
- Haines, Michael R., and Inter-university Consortium for Political and Social Research. 2010. *Historical, Demographic, Economic, and Social Data: The United States, 1790-2002*, May. <https://doi.org/10.3386/ICPSR02896.v3>.
- Hershbein, Brad J. 2012. “Graduating High School in a Recession: Work, Education, and Home Production.” *The B.E. journal of economic analysis & policy* 12, no. 1 (January): Article 3. <https://doi.org/10.1515/1935-1682.2599>.
- Janas, Pawel. 2026a. “Crises and Educational Attainment.” *The Journal of Economic History* 86, no. 2 (June): 501–536. <https://doi.org/10.1017/S002205072510106X>.
- Janas, Pawel. 2026b. *U.S. County Manufacturing, 1927 - 1937 | ICPSR*, August. <https://doi.org/https://doi.org/10.3386/E207242V1>.
- Jaworski, Taylor, Carl T. Kitchens, and Luke P. Rodgers. 2026. “Skill Formation, Child Labor, and the Schooling Consequences of the World War I Agricultural Boom,” Working Paper Series (April). <https://doi.org/10.3386/w35032>.
- Margo, Robert A. 1993. “Employment and Unemployment in the 1930s.” *Journal of Economic Perspectives* 7, no. 2 (June): 41–59. <https://doi.org/10.1257/jep.7.2.41>.
- Mathy, Gabriel P. 2018. “Hysteresis and persistent long-term unemployment: the American Beveridge Curve of the Great Depression and World War II.” *Econometrica* 12, no. 1 (January): 127–152. <https://doi.org/10.1007/s11698-016-0158-1>.
- Moulton, Jeremy Grant. 2017. “The Great Depression of Income: Historical Estimates of the Longer-Run Impact of Entering the Labor Market during a Recession.” *The B.E. Journal of Economic Analysis & Policy* 17, no. 4 (October). <https://doi.org/10.1515/bejeap-2017-0119>.
- Price, Joseph, Casey Buckles, Adrian Haws, and Haley E.B. Wilbert. 2023. *The Census Tree*.

- Price, Joseph, Kasey Buckles, Jacob Van Leeuwen, and Isaac Riley. 2021. “Combining family history and machine learning to link historical records: The Census Tree data set.” *Explorations in Economic History* 80 (April): 101391. <https://doi.org/10.1016/j.eeh.2021.101391>.
- Rosenbloom, Joshua L., and William A. Sundstrom. 1999. “The Sources of Regional Variation in the Severity of the Great Depression: Evidence from U.S. Manufacturing, 1919–1937.” *The Journal of Economic History* 59, no. 3 (September): 714–747. <https://doi.org/10.1017/S0022050700023548>.
- Ruggles, Steven, Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, David Hacker, Evan Roberts, and Robert Warren. 2024. *IPUMS Ancestry Full County Data*. Minneapolis, MN. <https://doi.org/https://doi.org/10.18128/D014.V4.0>.
- Saavedra, Martin, and Tate Twinam. 2020. “A machine learning approach to improving occupational income scores.” *Explorations in Economic History* 75 (January): 101304. <https://doi.org/10.1016/j.eeh.2019.101304>.
- Schroeder, Jonathan P, David Van Riper, Steven M Manson, Grace Cooper, Zachary Krause, Tracy A Kugler, Tsu Zhu, and Steven Ruggles. 2026. *IPUMS National Historical Geographic Information System*. Minneapolis, MN. <https://doi.org/http://doi.org/10.18128/D050.V21.0>.
- Vidart, Daniela. 2026. “Learning by Mail: The Impact of Correspondence Schools in Early 20th Century America,” Working Paper Series (April). <https://doi.org/10.3386/w35147>.

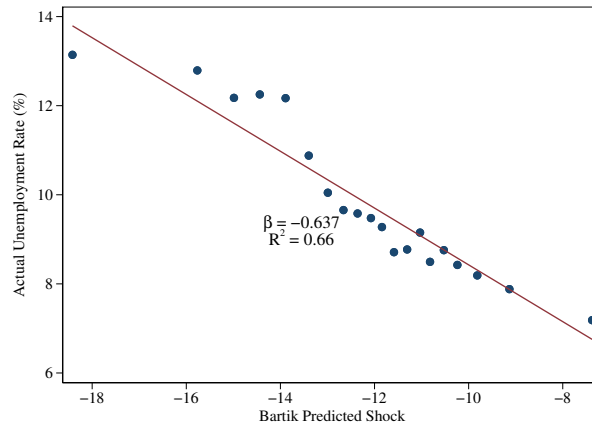
Appendix

Figure A1: OLS and Reduced Form Results



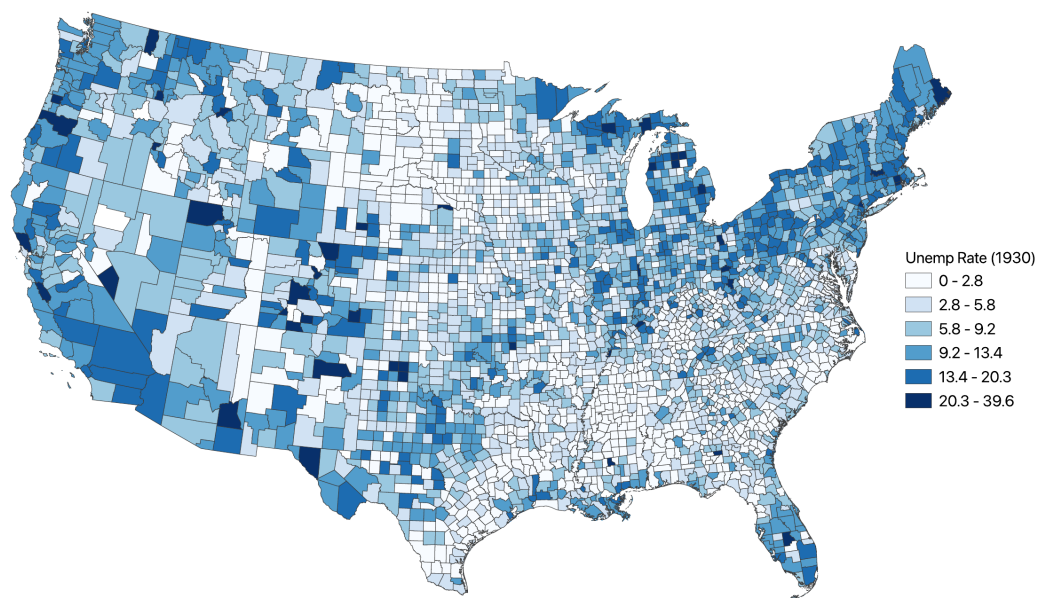
Note: Population-weighted binscatters of county-subgroup cells for ages 16–18. Both panels partial out county fixed effects, subgroup indicators, the 1929 high school share, and the 1925–1929 high school change. The left panel plots the 1929–1933 percentage-point change in the high school share against the 1930 unemployment rate. The right panel plots the same outcome against the 1929–1931 Bartik shock.

Figure A2: First Stage: Predicted v. Actual Unemployment



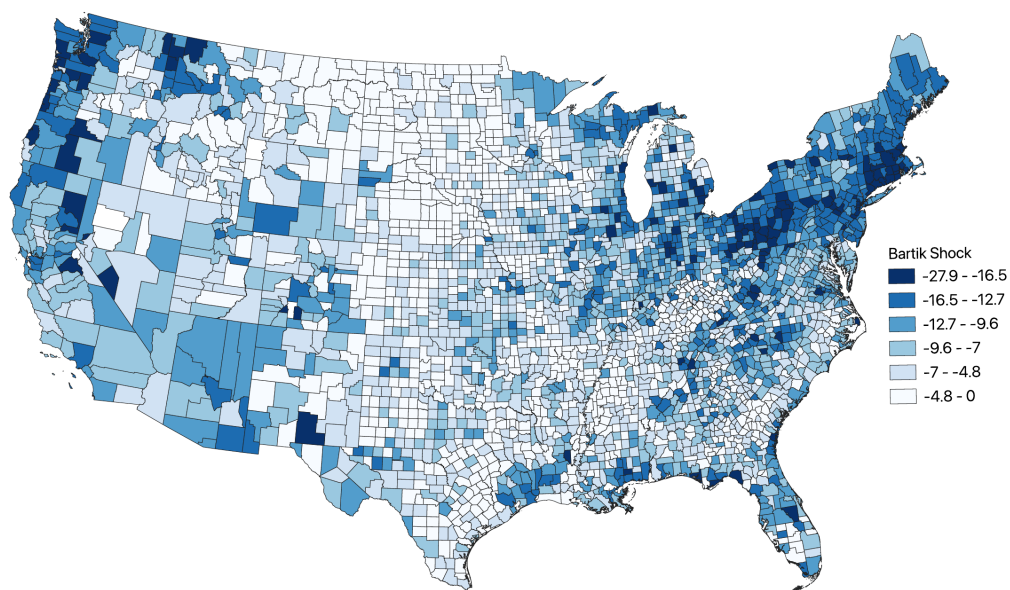
Note: Binscatter of the 1930 unemployment rate against the 1929–1931 Bartik shock. County fixed effects, subgroup indicators, the 1929 high school share, and the 1925–1929 high school change are partialled out. Observations are population-weighted county-subgroup cells for ages 16–18.

Figure A3: 1930 Unemployment Rate Across Counties



Note: The mapped value is the 1930 unemployment rate, in percentage points, for the ages 16–18 county aggregate. The labor force is the employed and the unemployed.

Figure A4: Bartik Shock Across Counties



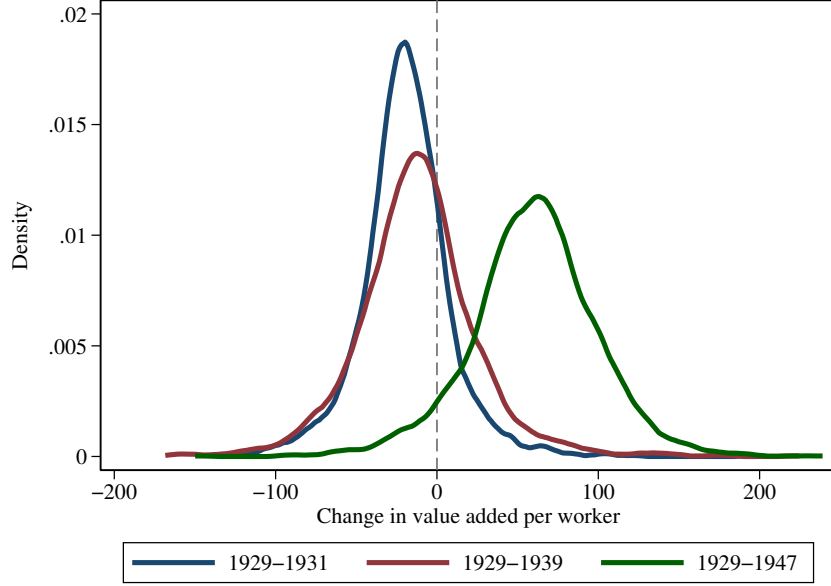
Note: The mapped value is the ages 16–18 county-aggregate Bartik shock, equal to 100 times the sum of 1929 local sector-industry employment shares times national sector-industry employment growth from 1929 to 1931.

Table A1: Shift Share Instrument Decomposition

Sector	Shock weight	Emp
Electrical machinery, equipment and supplies	0.0794	4,535,749
Printing, publishing, and allied industries	0.0789	6,865,521
FIRE	0.0780	10,284,768
Services	0.0723	15,928,735
Motor vehicles and motor vehicle equipment	0.0640	5,288,551
Wholesale	0.0508	2,662,446
Meat products	0.0491	663,814
Retail	0.0466	29,871,022
Canning and preserving	0.0453	3,275,560
Misc machinery	0.0359	1,428,039
Apparel and accessories	0.0318	4,715,356
Fabricated steel products	0.0287	1,218,693
Railroad and misc transportation equipment	0.0280	120,772
Misc manufacturing industries	0.0238	1,744,335
Construction	0.0232	3,828,130
Furniture and fixtures	0.0212	546,211
Misc wood products	0.0192	153,657
Misc chemicals and allied products	0.0164	363,780
Knitting mills	0.0162	1,493,369
Other primary iron and steel industries	0.0128	250,898
Dairy products	0.0110	97,845
Leather products, except footwear	0.0096	50,450
Beverage industries	0.0096	44,014
Grain-mill products	0.0094	40,241
Tobacco manufactures	0.0093	530,295
Misc nonmetallic mineral and stone products	0.0093	34,178
Paperboard containers and boxes	0.0093	44,615
Fabricated nonferrous metal products	0.0081	61,989
Ship and boat building and repairing	0.0076	118,319
Office and store machines	0.0075	48,009
Footwear, except rubber	0.0075	1,096,599
Leather: tanned, curried, and finished	0.0066	100,438
Pulp, paper, and paper-board mills	0.0062	126,419
Misc fabricated textile products	0.0057	41,954
Petroleum refining	0.0052	67,218
Professional equipment	0.0051	38,502
Glass and glass products	0.0050	155,368
Mining	0.0049	2,124,454
Misc paper and pulp products	0.0047	13,834
Rubber products	0.0044	243,780
Misc textile mill products	0.0038	39,123
Carpets, rugs, and other floor coverings	0.0038	53,669
Photographic equipment and supplies	0.0038	36,494
Yarn, thread, and fabric	0.0035	2,439,871
Primary nonferrous industries	0.0032	87,468
Cement, concrete, gypsum and plaster products	0.0019	9,406
Sawmills, planing mills, and mill work	0.0019	162,343
Dyeing and finishing textiles, except knit goods	0.0019	58,711
Misc petroleum and coal products	0.0018	2,545
Structural clay products; pottery and related products	0.0018	83,699
Synthetic fibers; drugs and medicines	0.0015	12,383
Watches, clocks, and clockwork-operated devices	0.0013	23,410
Aircraft and parts	0.0012	3,429
Agriculture	0.0009	2,609,327
Effective # shocks (1/HHI)		22.27
HHI		0.0449
# sectors (G)		54

Note: Shock-level weights for the (Borusyak, Hull, and Jaravel 2022) representation of the 1929–1931 Bartik shock. The industry panel is the ages 16–18 county aggregate. A county’s weight is its share of that sample’s population, and an industry’s weight sums, across counties, the product of that population weight and the county’s share of national employment in the industry. Weights are normalized to sum to one. Employment is the sample employment attached to the industry. The effective number of shocks is the inverse of the Herfindahl-Hirschman index of these weights.

Figure A5: Productivity Changes Across Counties



Note: Unweighted kernel densities of 100 times the log change in manufacturing value added per wage earner, on the ages 16–18 county aggregate. The three series end in 1931, 1939, and 1947. The vertical line is zero.

Table A2: Shift Share Instrument Balance Test

County characteristic	$dEmp_{29-31}$	SE	N
Urban share	4.700	(3.365)	54
White share	-0.096	(1.020)	54
Male share	-0.870	(0.546)	54
Citizen share	-2.602	(2.029)	54
HS-age share	0.042	(0.249)	54
Mean age	0.426	(0.523)	54
Veteran share	-0.433*	(0.217)	54
HS share (1925)	-0.011	(0.013)	54
HS share (1929)	-0.003	(0.011)	54
HS pre-trend (1929-25)	0.008	(0.006)	54
ln(population)	0.255	(0.335)	54

Note: Each row regresses an exposure-weighted county characteristic on the industry's 1929–1931 national employment change. County characteristics are collapsed to the industry using the same shock-level weights as Table A1, from the ages 16–18 county aggregate. Regressions are weighted by those shock weights and use robust standard errors. The high school pre-trend is the 1929 cohort share minus the 1925 cohort share. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: County Summary Statistics

	N	Mean	Median	Std. Dev.	Min	Max
Ages 16–18, demographic subgroups						
Unemp rate, 1930	17,983	9.77	9.70	6.96	0.00	50.00
Bartik shock, 1929-31	18,485	-12.09	-13.32	5.49	-45.75	0.00
HS change, 1929-33	14,875	7.23	7.47	10.93	-74.11	87.88
HS share, 1929	15,646	33.33	32.41	18.14	0.00	100.00
HS change, 1925-29	14,381	4.38	3.77	11.23	-100.00	100.00
County aggregate						
Mfg emp / pop	3,096	6.79	5.51	5.64	0.00	75.15
Workforce years of school, change	3,097	0.86	0.85	0.25	-2.51	6.91
Workforce HS share, change	3,097	9.34	9.82	3.39	-18.65	60.96
Workforce years of school, 1930	3,099	9.00	9.30	1.25	3.29	12.19
Workforce HS share, 1930	3,099	19.31	18.97	7.58	0.00	47.04
Workforce years of school, 1940	3,097	9.85	10.22	1.29	1.14	17.90
VA per worker, 1929	2,035	3,316.16	3,362.80	1,203.12	532.85	11,467.89
VA per worker change, 1919-29	2,035	21.14	26.14	28.25	-202.80	150.54
VA per worker change, 1929-39	2,035	-12.21	-12.61	25.97	-168.37	224.60
VA per worker change, 1929-47	2,035	57.69	54.54	28.89	-149.88	239.02
Wages per worker, 1929	2,035	1,203.23	1,262.25	321.14	297.00	2,374.44
Wages per worker change, 1929-47	2,035	-625.76	-626.76	18.18	-736.37	-539.77
Output per worker, 1929	2,035	7,521.44	6,958.17	4,205.00	812.78	49,657.52
Output per worker change, 1929-39	2,035	-7.22	-10.53	29.95	-215.62	287.21
Materials per worker, 1929	2,035	4,461.51	3,921.27	3,443.37	204.70	41,794.48
Materials per worker change, 1929-39	2,035	685.99	684.05	39.08	431.59	1,090.15
(VA-wages) per worker, 1929	2,035	2,112.92	2,105.72	978.74	158.78	10,344.77
(VA-wages) per worker change, 1929-47	2,035	107.47	101.91	39.45	-133.52	366.30
Establishments, 1929	2,035	1,374.72	140.00	3,241.01	3.00	20,797.00
Establishments change, 1929-47	2,035	11.70	9.30	31.35	-169.46	199.24
Wage earners, 1929	2,035	52,437.21	6,985.00	101,475.98	3.00	462,750.00
Wage earners change, 1929-47	2,035	32.84	28.14	47.61	-393.18	449.98
Workers per establishment change, 1929-47	2,035	21.13	18.44	42.94	-283.32	340.12
New Deal index	3,099	0.04	-0.06	0.61	-0.67	8.39
WWII contracts	3,099	963,649.90	20,739.00	2,237,545.22	0.00	11217327.00

Note: Population-weighted statistics from the county file. The first panel is ages 16–18 sex-by-race-by-father-SES cells. Unemployment and the high school share and change are in percentage points. The high school pre-trend is the 1929 cohort share minus the 1925 cohort share. The Bartik shock is 100 times 1929 local industry employment shares times national industry employment growth from 1929 to 1931. The second panel is the ages 16–18 county aggregate. Workforce education covers the employed and unemployed aged 19 and older and uses 1940 schooling; the high school share is in percentage points and years of schooling are in years. Manufacturing changes are 100 times log changes. Per-worker variables divide by wage earners. Cells strictly outside the pooled 1st or 99th percentile of the 1930 unemployment rate, the 1929–1933 high school change, or the 1929 high school share are dropped before these statistics, with all three cutoffs computed on the pooled file.

Table A4: Individual Summary Statistics

	N	Mean	Median	Std. Dev.	Min	Max
Occupational income score	5,759,325	20.67	22.00	8.46	3.00	80.00
Duncan SEI	5,659,521	29.27	18.00	20.90	3.00	96.00
Occupational earnings score	5,719,311	38.99	41.20	23.57	0.60	100.00
Occupational education score	5,719,311	14.06	4.60	19.01	0.00	93.90
Lido score	5,278,162	15.50	14.99	7.32	0.52	89.59

Note: Linking-weighted statistics for linked individuals who turn 18 in 1931–1939. Scores are the raw 1940 values, before the standardization used in Table 3. OccScore is median 1950 total income in hundreds of 1950 dollars. The Duncan SEI is the prestige-weighted index. ErScor is the percentile rank of occupational median earnings, and EdScor is the percent of incumbents with at least one year of college. LIDO is the (Saavedra and Twinam 2020) occupational income score.

Table A5: Differences in Means of Linked v. Unlinked 1940 Samples

	Sex	HS Bin	Age	Race	Urban	Stayer	Northeast	Midwest	South	West
Panel A: Unweighted Balance										
matched	-0.193*** (0.000)	0.088*** (0.000)	-0.570*** (0.002)	-0.118*** (0.000)	-0.083*** (0.000)	0.117*** (0.000)	-0.029*** (0.000)	0.087*** (0.000)	-0.065*** (0.000)	0.006*** (0.000)
Observations	38,041,568	37,228,701	38,041,568	38,041,568	38,041,568	38,041,568	38,041,568	38,041,568	38,041,568	38,041,568
Panel B: Weighted Balance										
matched	0.003*** (0.000)	0.001*** (0.000)	0.021*** (0.002)	0.000 (0.000)	0.003*** (0.000)	-0.003*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	0.003*** (0.000)	0.002*** (0.000)
Observations	37,778,571	37,228,701	37,778,571	37,778,571	37,778,571	37,778,571	37,778,571	37,778,571	37,778,571	37,778,571

Note: Each entry is the difference in means between linked and unlinked 1940 respondents aged 18–34. Panel A is unweighted. Panel B weights by the inverse probability of a successful Census Tree link. The probit is fit on sex interacted with age indicators, census region, an indicator for living in one's state of birth, and high school completion interacted with urban status and with race. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Effect on Δ HS Attainment (Diff HS Windows)

	1929-1933	1929-1931	1929-1932	1929-1934
Unemp rate, 1930	0.384*** (0.114)	0.336*** (0.114)	0.474*** (0.128)	0.408*** (0.104)
HS share, 1929	-0.604*** (0.019)	-0.585*** (0.026)	-0.613*** (0.026)	-0.624*** (0.023)
HS change, 1925-29	-0.164*** (0.012)	-0.171*** (0.014)	-0.165*** (0.013)	-0.167*** (0.013)
Observations	13,949	13,907	13,944	14,044
First-Stage F	238.27	233.99	237.28	235.33
Fixed Effects	County+Subgroup	County+Subgroup	County+Subgroup	County+Subgroup

Note: IV estimates of the specification in column (6) of Table 1. Columns change only the end year of the high school attainment window. The 1929 high school share, the 1925–1929 high school change, county and subgroup fixed effects, population weights, state clustering, and the 1929–1931 Bartik shock are unchanged. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: **Effect on Δ HS Attainment (Diff Fixed Effects)**

	County + Subgroup	County	Subgroup	State
Unemp rate, 1930	0.384*** (0.114)	0.681*** (0.118)	0.326*** (0.068)	0.614*** (0.086)
HS share, 1929	-0.604*** (0.019)	-0.088*** (0.026)	-0.183*** (0.043)	-0.098*** (0.022)
HS change, 1925-29	-0.164*** (0.012)	-0.398*** (0.015)	-0.335*** (0.021)	-0.379*** (0.012)
Observations	13,949	13,949	13,972	13,972
First-Stage F	238.27	275.36	358.67	384.19
Fixed Effects	County+Subgroup	County	Subgroup	State

Note: IV estimates of the baseline county specification for ages 16–18. Columns replace the county and subgroup fixed effects with the set named in the header. The outcome, 1929 high school share, 1925–1929 high school change, population weights, state clustering, and 1929–1931 Bartik shock match column (6) of Table 1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: **Effect on Δ HS Attainment (Diff Weights)**

	Population	Labor Force	Census Links	No weights
Unemp rate, 1930	0.384*** (0.114)	0.299** (0.114)	0.343*** (0.104)	0.167 (0.199)
HS share, 1929	-0.604*** (0.019)	-0.552*** (0.022)	-0.595*** (0.021)	-0.692*** (0.019)
HS change, 1925-29	-0.164*** (0.012)	-0.181*** (0.013)	-0.173*** (0.011)	-0.100*** (0.012)
Observations	13,949	13,949	13,949	13,949
First-Stage F	238.27	330.02	245.96	118.85
Fixed Effects	County+Subgroup	County+Subgroup	County+Subgroup	County+Subgroup

Note: IV estimates of column (6) of Table 1. Columns weight by cell population, the cell labor-force count, the unweighted number of Census Tree links in the county-subgroup, or give every cell equal weight. The 1929 high school share, the 1925–1929 high school change, county and subgroup fixed effects, state clustering, and the 1929–1931 Bartik shock are unchanged. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: **Effect on Δ HS Attainment (Diff Bartiks)**

	1929-1931 Shift	1929-1933 Shift	Mfg Only (1929-1931)	Ages 19–25	Ages 26–64
Unemp rate, 1930	0.384*** (0.114)	0.532*** (0.119)	0.189 (0.152)	0.438*** (0.098)	0.426*** (0.100)
HS share, 1929	-0.604*** (0.019)	-0.612*** (0.020)	-0.594*** (0.019)	-0.607*** (0.019)	-0.607*** (0.019)
HS change, 1925-29	-0.164*** (0.012)	-0.161*** (0.012)	-0.168*** (0.011)	-0.164*** (0.012)	-0.163*** (0.012)
Observations	13,949	13,949	13,949	13,944	13,940
First-Stage F	238.27	157.88	33.51	221.50	179.87
Fixed Effects	County+Subgroup	County+Subgroup	County+Subgroup	County+Subgroup	County+Subgroup

Note: IV estimates of column (6) of Table 1. The outcome, the 1929 high school share, the 1925–1929 high school change, county and subgroup fixed effects, population weights, and state clustering are unchanged. The first three instruments are the ages 16–18 Bartik shock using 1929–1931 national employment changes, 1929–1933 changes, or 1929–1931 changes in manufacturing industries only. The last two columns instrument ages 16–18 unemployment with the 1929–1931 Bartik shock built from employment shares of ages 19–25 and ages 26–64 in the same county and demographic cell. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: **Decomposition of Workforce Δ HS Attainment**

	Aggregate	Within Industry	Between Industry	Interaction
	Aggregate	Within	Between	Interaction
Unemp rate, 1930	0.208* (0.114)	0.147** (0.073)	-0.026 (0.042)	0.086** (0.041)
Workforce HS share, 1930	0.191*** (0.030)	0.161*** (0.018)	0.124*** (0.022)	0.058*** (0.008)
% Urban	-0.022* (0.011)	-0.011 (0.007)	0.001 (0.004)	0.010*** (0.004)
Observations	3,093	3,093	3,093	3,093
First-Stage F	71.78	71.78	71.78	71.78
State FE	✓	✓	✓	✓

Note: IV estimates of the contribution, in percentage points, of each Kitagawa term to the 1930–1940 change in the county workforce high school share. The labor force is the employed and the unemployed aged 19 and older, with 1940 schooling in both years. Industry is IND1950, and a missing industry is its own category. The 1930 workforce is counted in the 1930 county of residence and the 1940 workforce in the 1940 county. For industries present in both years, the within term weights the change in the industry high school share by its 1930 employment share, the between term weights the change in its employment share by its 1930 high school share, and the interaction is the product of the two changes. Industry entry and industry exit close the identity and are not shown. The specification matches column (6) of Table 4: the ages 16–18 county aggregate, state fixed effects, the 1930 workforce high school share, the urban share, population weights, state clustering, and the 1929–1931 Bartik shock. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: **Youth Probability of Entry, by Industry**

	Agriculture		Manufacturing		Services	
	OLS	IV	OLS	IV	OLS	IV
Unemp rate, 1930	-0.002*** (0.000)	-0.005*** (0.002)	0.000 (0.000)	0.002 (0.001)	0.001** (0.000)	0.003*** (0.001)
County HS share, 1929	-0.000* (0.000)	-0.000** (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.001*** (0.000)	0.000*** (0.000)
HS change, 1925-29	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Industry share, 1930	0.007*** (0.000)	0.006*** (0.000)	0.006*** (0.000)	0.006*** (0.000)	0.005*** (0.000)	0.005*** (0.000)
Observations	5,726,396	5,726,396	5,726,396	5,726,396	5,726,396	5,726,396
R-squared	0.350		0.104		0.128	
First-Stage F		64.07		108.50		263.97
County+Subgroup+Cohort FE	✓	✓	✓	✓	✓	✓

Note: OLS and IV linear probability estimates for linked individuals who turn 18 in 1931–1939. Outcomes are indicators for the 1940 industry, so a coefficient of 0.01 is one percentage point. Agriculture is IND1950 code 105, manufacturing is 306–499, and services is 806–899, after dropping industry codes 0, 979, 997, and 999. Services are the 1950 census services major group; trade, finance, transportation, and public administration are not included. Each column controls for that industry’s 1930 employment as a share of the workforce aged 19 and older in the person’s 1930 county and sex-by-race-by-father-SES cell, in percentage points, and for that cell’s 1929 high school share and 1925–1929 high school change. Regressions absorb county, subgroup, and cohort fixed effects, use inverse-probability linking weights, and cluster by state. The unemployment rate is the ages 16–18 rate in that cell, instrumented by the cell’s 1929–1931 Bartik shock. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: **Effect on $\Delta VA/N$**

	Baseline	Inputs	Workers	Demos	No mfg hubs	27 Base	39 End	27-29 Trend	New Deal	WWII
Unemp rate, 1930	2.039*** (0.679)	2.333*** (0.726)	1.691** (0.647)	1.925*** (0.646)	2.620*** (0.635)	1.823** (0.716)	0.659 (0.830)	1.987*** (0.644)	2.405** (0.904)	2.255*** (0.649)
Observations	2,034	2,034	2,034	2,034	1,917	1,586	2,034	1,586	2,029	2,034
R-squared										
First-Stage F	57.10	58.21	56.51	43.95	43.77	41.86	57.10	41.69	68.23	58.32
State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: IV estimates on the ages 16–18 county aggregate. Unless noted, the outcome is 100 times the 1929–1947 log change in value added per wage earner, and the controls are the 1929 level, the 1919–1929 log change, the 1930 workforce high school share, and the urban share. Inputs adds 1929 materials per wage earner and their 1929–1939 log change. Workers adds 1929 wage earners and their 1929–1947 log change. Demos adds the male and white population shares. No mfg hubs drops counties at or above the 95th percentile of manufacturing employment per population. 1927 Base measures productivity growth from 1927 to 1947 and replaces the 1929 level and 1919–1929 change with the 1927 level and the 1919–1927 change. 1939 End measures productivity growth from 1929 to 1939 with the baseline controls. 27–29 Trend keeps 1929–1947 growth and the 1929 level, and replaces the 1919–1929 log change with the 1927–1929 log change. New Deal adds per capita spending on loans, relief, public works, and the Agricultural Adjustment Administration. WWII adds supply-contract spending in thousands of dollars, with counties that have no recorded contracts set to zero. All columns include state fixed effects, weight by the ages 16–18 population, cluster standard errors by state, and instrument unemployment with the 1929–1931 Bartik shock. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: **Effect on ΔY**

	VA/N	VA	Wages/N	Output/N	Inputs/N	(VA-W)/N	Firms	Workers
Unemp rate, 1930	2.039*** (0.679)	-2.503* (1.394)	3.230*** (0.566)	-0.506 (0.818)	-0.368 (1.077)	2.531*** (0.762)	1.424 (2.140)	-5.489*** (1.583)
Level, 1929	-0.015*** (0.001)	-0.000*** (0.000)	-0.054*** (0.008)	-0.003*** (0.000)	-0.004*** (0.000)	-0.031*** (0.002)	0.000 (0.000)	0.000 (0.000)
Pretrend	-0.223*** (0.030)	-0.087*** (0.030)	-0.173*** (0.039)	-0.092** (0.040)	-0.093** (0.037)	-0.161*** (0.025)	-0.160*** (0.056)	-0.119*** (0.043)
Workforce HS share, 1930	0.650*** (0.158)	1.844*** (0.482)	0.272* (0.144)	0.614** (0.238)	0.490 (0.333)	0.707*** (0.185)	1.479*** (0.276)	2.125*** (0.549)
% Urban	0.032 (0.038)	-0.148 (0.111)	0.004 (0.043)	0.101** (0.044)	-0.002 (0.064)	-0.025 (0.053)	-0.135 (0.187)	-0.018 (0.126)
Observations	2,034	2,034	2,034	2,034	2,034	2,033	2,034	2,034
First-Stage F	57.10	53.87	59.71	65.64	63.81	55.57	51.89	56.06
State FE	✓	✓	✓	✓	✓	✓	✓	✓

Note: IV estimates on the ages 16–18 county aggregate. Each outcome is 100 times a log change in the indicated Census of Manufactures variable. Value added per wage earner, value added, wages per wage earner, non-wage surplus per wage earner, establishments, and wage earners run from 1929 to 1947. Output per wage earner and materials per wage earner run from 1929 to 1939 (due to data limitations). Firms are establishments. Per-worker outcomes divide by wage earners. Controls are the 1929 level of the outcome, its 1919–1929 log change, the 1930 workforce high school share, and the urban share. Regressions include state fixed effects, are weighted by the ages 16–18 population, and cluster standard errors by state. The instrument is the 1929–1931 Bartik shock. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.