Abstract: This paper investigates the role of unions in the Great Compression of U.S. wage inequality after 1939. Specifically, it tests whether places with higher exposure to unionization during the 1940s, due to their pre-existing industrial composition, tended to have larger declines in inequality, conditional on local economic and demographic observables and regional trends. We find a strong negative correlation between exposure to unionization and changes in local inequality from 1940-50 and 1940-60. This does not appear to be underpinned by skill-specific sorting of workers or by firms leaving places with high exposure to unionization. We also find that the correlation between exposure to unionization in the 1940s and the change in inequality after 1940 persists in long-difference regressions to the end of the twentieth century.

Contacts: william.collins@vanderbilt.edu; niemesgt@miamioh.edu.

Acknowledgements: The authors gratefully acknowledge the assistance of IPUMS-USA staff members in using the new 1960 5-percent sample, as well as helpful discussions with Ran Abramitzky, Leah Boustan, Brant Callaway, James Feigenbaum, Andrew Goodman-Bacon, Taylor Jaworski, Ilyana Kuziemko, Robert Margo, Suresh Naidu, Marianne Wanamaker, Ariell Zimran, and seminar participants at Queen’s University (Canada), the National Bureau of Economic Research (NBER), and the World Congress of Cliometrics. Collins is the Terence E. Adderley Jr. Professor of Economics at Vanderbilt University and a Research Associate of the NBER. Niemesh is an Assistant Professor of Economics at Miami University and Faculty Research Fellow at the NBER.
The widening and polarization of the United States’ wage structure since the 1970s has motivated an extensive literature that studies inequality’s changing patterns and root causes (inter alia, Katz and Murphy 1992, DiNardo, Fortin, and Lemieux 1996, Autor, Katz, and Kearney 2008, Acemoglu and Autor 2011). In contrast, earlier in the twentieth century, the U.S. experienced a large decline in wage inequality, especially during the 1940s, in what Goldin and Margo (1992) termed “the Great Compression.” Scholars in the 1950s remarked on this event (Kuznets and Jenks 1953, Goldsmith 1954, Miller 1958), and several have considered it since that time (e.g., Juhn 1999, Piketty and Saez 2003, Goldin and Katz 2008). Yet the Great Compression has received far less scrutiny than the late twentieth-century rise in inequality, despite its magnitude and persistence well into the 1960s and despite the salience of concerns about long-run inequality trends.

In this paper, we investigate whether changes in labor market institutions—specifically the rise of organized labor—may have contributed to the mid-century narrowing of inequality in the U.S. and, if so, whether that contribution left an imprint on labor markets beyond the 1940s. The Great Compression coincided with a sharp rise in union membership. Indeed, a twentieth-century time-series measure of income inequality at the national level resembles almost a mirror image of a time-series of union density, as shown in Figure 1. Causal inference from this pattern would be tenuous, but several strands of research suggest that a connection between rising unions and falling inequality is plausible. First, labor economists studying the post-1970 period, when data are more abundant, have found that unions tend to compress wage structures and reduce inequality (Freeman 1980, Freeman and Medoff 1984, DiNardo, Fortin, and Lemieux 1996, Card, Lemieux, Riddell 2004). Second, although research that specifically addresses unions and wage compression in the U.S. in the 1940s and 1950s is scarce, Miller (1958) speculated that union bargaining disproportionately raised wages at the bottom of the income distribution; Goldin and Margo (1992) suggested that unions might have helped sustain the relatively narrow dispersion of wartime wages into the post-war period; Frydman and Molloy (2012) found that executive compensation relative to production workers’ pay was negatively correlated with union presence in 1949; and Callaway and Collins (2017) observed that the distribution of union members’ wages was compressed relative to that of observationally similar non-members in 1950. In short, the combination of time-series patterns in Figure 1, micro-based evidence from the modern labor economics literature, and historical evidence from the 1940s motivates a closer examination of the connection between unions and wage inequality during the Great Compression.

It is notable that in the early postwar period, prominent scholars suggested that unions may have increased inequality by introducing a wedge between the wages of union and similar non-union
workers, by raising wages for workers who might have been well paid regardless of unions, and by reducing employment opportunities in unionized firms (Friedman 1962, p. 123-25; Rees 1962, pp. 98-99). But the mid-century evidence was fragmentary and difficult to evaluate (Lewis 1963), reflecting the scarcity of detailed and representative data on workers’ wages and union status. The best nationally representative data on wages for this period are from the federal census of population, but the census has never inquired about workers’ union status, and the Current Population Survey (CPS) first reported union status in 1973. To make headway on the question of whether the unions significantly influenced U.S. wage structures during the Great Compression, researchers must confront this fundamental data challenge.

To do so, we develop evidence based on changes in inequality and union density at the level of local labor markets. In essence, we test whether places exogenously exposed to larger increases in union density tended to have larger declines in wage inequality, conditional on many other factors that may have simultaneously affected local economies. To measure cross-place differences in exposure to unionization, we combine local-level information on the distribution of workers over industries in 1940 with national-level industry-specific measures of changes in unionization during the 1940s. Variation in exposure to unionization then forms the basis of the identification strategy outlined below. This emphasis on variation in ex ante exposure is akin to recent literature that measures the labor market effects of changes in international trade or technology (Autor, Dorn, and Hanson 2013, Acemoglu and Restrepo 2017).

To measure changes in inequality, we use new and large micro-level datasets for the 1940 and 1960 censuses in combination with the 1-percent public use sample for the 1950 census (Ruggles et al. 2015). The measures of inequality, because they include all local wage and salary workers, capture both the direct and indirect effects of exposure to unions on the local wage structure, including the net effect of spillovers to the wages of non-union workers (if they exist). In theory, such spillovers could positively or negatively influence local inequality, depending on the relative importance of “crowding effects” (whereby unions reduce employment in the union sector and workers crowd the non-union sector) versus “threat effects” (whereby the threat of unionization leads firms to proactively change wage structures and conditions of employment) (Lewis 1963, Farber 2005).

The 1940s were a tumultuous time for the U.S. economy due to the Second World War.

---

1 Card, Lemieux, and Riddell (2004) provide a useful encapsulation of the literature on unions and inequality and its evolution since the 1950s, as well as microdata analyses of the U.S., Canada, and United Kingdom from the 1980s to 2001.
Therefore, isolating the role of unions apart from the many other influences on local wage structures poses a challenge, and some caution must be taken in interpreting the regression results. Nonetheless, we find evidence that places that were more exposed to increases in unionization due to their pre-existing industrial structure experienced sharper declines in wage inequality after 1940, controlling for potentially confounding factors such as the value of wartime contracts to local producers, the shift in the relative demand for skilled versus unskilled workers, the pre-existing skill level of the local workforce, and state or regional trends. Further examination does not suggest that this pattern was driven by the sorting of skilled workers or firms out of locations that experienced sharp increases in union density. Finally, even though rising wage inequality and falling unionization were prominent features of the last two decades of the twentieth century, we find that the imprint of mid-century unionization on wage structures was persistent at the local level. That is, long-difference regressions show that places intensively exposed to changes in unionization in the 1940s tended to have smaller changes in inequality between 1940 and 2000 than others.

1. Brief background on the rise of U.S. unions and the Great Compression

From the late 1930s to the early 1950s, union membership in the U.S. increased from approximately 11 to 30 percent of nonfarm employment, as shown in Figure 1. The rise was due, at least in part, to changes in federal policy during the 1930s and 1940s that facilitated union organization and protected unions once established (Seidman 1953, Reder 1988, Freeman 1998). At a deeper level, the changes in policy reflected workers’ demand for representation in the wake of the Great Depression (Ashenfelter and Pencavel 1969), policymakers’ desire to reduce the frequency and severity of violent strikes (Wachter 2012), and the exigencies of the Second World War. Changes in union membership were highly uneven across industries and across space (Troy 1957), a fact that we exploit later in the paper.

The Norris-LaGuardia Act of 1932 and National Labor Relations Act of 1935, henceforth NLRA or Wagner Act, recast the legal framework under which workers organized and unions bargained in the U.S. Prior to this, employers were relatively unconstrained in the tactics they used to defeat unions and punish employees who promoted unions or went on strike (Lester 1964, ch. 5). Under the NLRA, employees were declared to have a right to organize and bargain collectively; employers in turn were obliged to bargain with unions and prohibited from using a variety of anti-union practices.2 The legality of the Wagner Act was challenged, but the Supreme Court’s 5-4

---

2 The Act did not apply to government employees, agricultural workers, or domestic workers. A separate industry-specific framework covered railroad workers (the Railway Labor Act).
decision in *NLRB v. Jones & Laughlin Steel Corporation* in 1937 upheld it. Thereafter, American Federation of Labor (AFL) and especially Congress of Industrial Organizations (CIO) union membership grew rapidly. Whereas the AFL had traditionally organized workers along skilled craft lines, the CIO aggressively expanded in industrial settings and primarily represented production workers.

In the interest of maintaining maximum production during the Second World War, President Roosevelt created the National War Labor Board (NWLB), which was comprised of representatives from unions, businesses, and the public, to settle labor-management disputes (Seidman 1953, J. Freeman 1978). Unions pledged not to strike during the war, giving up an important source of bargaining power, but when disputes arose the NWLB often granted unions greater security in the form of “maintenance of membership” clauses. In effect, this sustained unions where they were already established and extended unions to firms that had previously resisted them. The NWLB also played an important role in wage setting as part of a broader federal effort to control inflation. Importantly, the NWLB typically allowed larger wage increases for workers earning “substandard” wages, which tended to narrow pay differentials (Dunlop 1950, Goldin and Margo 1992). NWLB decisions also tended to narrow interplant wage differentials within industries and localities (Seidman 1953, p. 115). We do not attempt to separate the influence of unions from the influence of the NWLB in this paper. We see them as complementary and interconnected institutions in the sense that union pressure was often exerted through the NWLB framework, and of course the NWLB included union representatives. After the war, the government disbanded the NWLB and loosened wage controls, but the wartime wage compression persisted.

In 1947, Congress passed the Taft-Hartley Act over President Truman’s veto, a significant legislative setback to the labor movement (Millis and Brown 1950). Among other things, Taft-Hartley circumscribed a variety of union activities (e.g., prohibiting certain kinds of strikes and boycotts), gave employers more scope to voice opposition to unions, banned Communists from leadership in unions, and allowed states to pass “right to work” laws that undermined “union shop” agreements. Even so, union membership as a share of employment held steady, more or less, for about 20 years after Taft-Hartley, as shown in Figure 1. Then, private sector unions began their steady decline through the late twentieth century.

---

3 Seidman (1953, pp. 61-62 and 91-108) discusses maintenance or membership clauses in detail. In sum, “Labor members [of the NWLB] supported membership maintenance, since this form of union security represented a substantial advance to most industrial unions and since craft unions were permitted to keep the union or closed shops that they had previously achieved through collective bargaining” (p. 102). Despite the “no strike” pledges, there were strikes during the war, most famously the miners’ strike of 1943.
The Great Compression in retrospect

Coinciding with the mid-century rise of unions, measures of wage and income inequality declined sharply. It is important to note that earlier in the twentieth century, wage inequality appears to have already declined relative to the late nineteenth century (Goldin and Katz 2008, ch. 2). Yet the 1940s stand out as a period of rapid and sustained compression between and within groups of workers. Figure 2 charts the change in real wages at each percentile of the wage distribution for men in 1939 to 1949 and, for comparison, 1969 to 1989. The series begins in 1939 because the 1940 Census of Population was the first to inquire about the previous year’s wage and salary income. The downward slope of the graph for the 1940s depicts the Great Compression—real wage gains at the bottom of the distribution were much larger than those at the top. Similar plots for the 1950s and 1960s (not shown) are relatively flat but well above zero, reflecting the postwar period of widespread income growth. Finally, a substantial rise in inequality is evident after 1970. It is notable that the compression of wages during the 1940s was similar in magnitude to the widening of the distribution during the 1970s and 1980s (combined), which initiated the modern labor economics literature on rising wage inequality.

As mentioned above, the empirical literature on the Great Compression has rarely addressed the role of unions directly. Goldin and Margo (1992) provide the most detailed national-level investigation of the Great Compression to date. They show that the level of inequality recorded in 1940 was not an anomalous artifact of the Great Depression and that the subsequent compression was not a simple extension of prior trends. Their interpretation emphasizes that the supply of well-educated workers increased between 1940 and 1950, tending to reduce between-group inequality at the national level. They also discuss the wage setting practices of the NWLB, but the role of unions is not in the foreground of their analysis. In their conclusion, however, when discussing the persistence of the Great Compression, Goldin and Margo (1992, p. 32) write, “The American labor movement was never stronger than in the 1950s; and unions, it has been claimed, were strongly in favor of a compressed wage structure, at least for a while.” Niemesh and Jaworski (2016) revisit Goldin and Margo’s analysis with larger public use samples; they replicate Goldin and Margo’s main results but find a larger role for within-group narrowing of inequality. Piketty and Saez (2003), who focus on documenting inequality as reflected in the share of income going to top households, also mention the rise of unions and the importance of institutional factors in contributing to the falling inequality (p. 34), but without presenting direct evidence on this point.

Although integrating international evidence is beyond the scope of this paper, it is worth
pointing out that Gazeley (2006) and Atkinson and Nolan (2010) suggest that institutional changes in the 1940s may have been important in narrowing wage inequality in Britain and Ireland, respectively. They do not develop direct evidence on the role of unions per se, but they do show clear evidence of wage compression and discuss how institutional changes may have contributed to that compression, perhaps reinforcing market forces operating in the same direction.

In sum, although the timing and patterns of wage compression in the 1940s, such as narrowing between- and within-group wage differentials that persisted long after the war, are consistent with what one might expect from the rise of unions, previous scholarship has not developed systematic evidence linking unionization and declining inequality in the 1940s. Given the data constraints and challenges to identification, we turn to empirical patterns observed at the local level to develop new evidence on the question of whether unions may have contributed to the compressed wage structures of the 1940s and beyond.

2. Data describing mid-century inequality and unions

We build a local-level dataset reflecting the mid-century decline in inequality and rise of unions. This work benefits from newly available and large micro datasets from the Census of Population in 1940 (complete count) and in 1960 (5-percent sample), which are available from the Integrated Public Use Microdata Series (IPUMS, Ruggles et al. 2015). Their size permits greater precision in measuring local economic characteristics, and importantly, they reveal geographic information at a finer level of detail than previously available. For 1950, we rely on the standard 1-percent IPUMS sample, which is comparatively small but includes useful geographic identifiers. Ultimately, the baseline dataset includes 467 State Economic Areas (SEAs), covering the entire continental United States.

The Census Bureau created SEAs to identify adjoining counties that were economically similar circa 1940-50. In addition to economic characteristics, the Census Bureau considered a broader set of social, industrial, commercial, demographic, climatic, and cultural factors when delineating SEAs. A report on the procedures for establishing the SEAs explains, “The name ‘state economic areas’ has been given…to convey the implication that each State has been divided into its principal units and that within each unit a distinctive economy prevails, insofar as it is possible to do this using county units” (Bogue 1951, p. 1). We believe SEAs are the best available mid-century
approximation of local labor markets and use them as the paper’s basic unit of analysis.\textsuperscript{4}

The IPUMS 1950 1-percent sample provides SEA identifiers as the smallest geographical unit; it is not possible to measure wage inequality at a lower level of aggregation in 1950 using microdata. We convert all other years to have SEAs as the key unit of geography. The 1940 full count census microdata file includes county identifiers, which are easily aggregated into SEAs. The IPUMS 1960 5-percent sample and those for later years pose a challenge because neither SEA nor county identifiers are provided. The smallest geographic unit in the 1960 sample is the mini public-use microdata area (mini-PUMA), in which census tracts and untracted counties are combined into units with at least 50,000 residents (Ruggles et al. 2016, description of PUMAMINI variable). This is a great improvement in geographic precision relative to the previously available public use sample, but it does not always map seamlessly into SEAs. We construct SEA-level variables for 1960 by following a procedure developed in Autor and Dorn (2013). Each observation in a mini-PUMA is probabilistically allocated to an SEA based on the fraction of the population in a mini-PUMA that maps into an SEA. For example, suppose a mini-PUMA is evenly split over two SEAs. In that case, each observation in the mini-PUMA is weighted by one-half in the first SEA and by one-half in the second SEA, as well.\textsuperscript{5} The supplementary data appendix provides more detailed information.

Our measures of SEA-level wage inequality include 90-10, 90-50, and 50-10 percentile differences of log weekly wages, as well as the Gini coefficient and the variance of log weekly wages. The 90-50 and 50-10 perspectives are particularly useful for seeing whether compression occurs at the bottom or top of the income distribution, and we emphasize these measures in our discussion. Many previous studies of inequality have featured the Gini or variance measure of inequality, and so we provide them as well for comparison and completeness.

We follow Goldin and Margo (1992) closely in using the census wage data. We limit the microdata samples to males aged 18 to 64, employed for at least 40 weeks in the prior year, who were wage and salary workers earning more than half of the implied weekly minimum wage. Self-employment and capital income were not reported in the 1940 census; for consistency, we do not

\textsuperscript{4} SEAs are similar to modern commuting zones or metropolitan statistical areas in that they delineate economically meaningful areas below the state level, but they do not follow the same boundaries. Commuting zones were first based on the 1990 census and, thus, reflect the more recent geographic distribution of labor markets. Standard Metropolitan Areas (SMAs) were first defined in the 1950 census, but a relatively small number existed at the time (144 in census microdata). Changes in inequality outside metropolitan areas were important in the Great Compression, and we want to capture that in our analysis.

\textsuperscript{5} Information on how mini-PUMA populations were split over counties can be found at https://usa.ipums.org/usa/volii/1960geotools.shtml. Out of 2,765 mini-PUMAs in the 1960 sample, 1,873 (68 percent) are contained within a single SEA, 654 (24 percent) are split over two SEAs, 198 (7 percent) are split over three SEAs, and 40 (1 percent) are split over four SEAs.
include these income sources in later years. Top-coding of annual wage and salary income affects the top 1.1, 1.2, and 0.4 percent of men in the 1940, 1950, and 1960 samples, respectively. Again, following Goldin and Margo (1992), we multiply top-coded values by 1.4 in each year.\footnote{Observations in the 1940 full count require the user to top-code wage and salary income values at $5,000. The instructions for the census enumerators state, “For amounts above $5,000, enter ‘5,000+'.” This means that you are not to report the actual amount of money wages and salary for persons who have received more than $5,000.” A number of enumerators recorded dollar amounts above $5,000 in error. We choose to top-code the errors at 1.4 times the value of the top-code, as well as top-coded values in other years. Out of observations reporting $5,000 or above in the restricted sample, 14 percent are errors.} Weekly income is computed by dividing annual income from wage and salary work by the total number of weeks worked.\footnote{Weeks worked is recorded as a continuous value in 1940 and 1950 (wkswork1). In 1960 and 1970, weeks worked is recorded as an interval value (wkswork2), which we convert to the midpoint of the interval.} In each case, the relevant census questions refer to the year prior to enumeration; thus, the 1940 census refers to income in 1939.

Figure 3 maps the geographic variation in the compression of the 90-10 wage differential across SEAs for 1939-1949. This information is novel in that no previous work, to our knowledge, has documented or examined cross-SEA variation in wage compression during the 1940s. The post-1939 decline in inequality was nearly ubiquitous, but there was a great deal of cross-place variation. While some regions had more compression than others, the primary visual impression is that there was considerable within-region variation across the U.S. For instance, the change in the 90-10 differential from 1940 to 1950 ranged from -0.372 log points at the 25th percentile of the distribution in changes to -0.107 at the 75th percentile. In addition, it is notable that in 1939 some areas had fairly low levels of inequality whereas others had substantially higher levels; the 90-10 difference in log weekly wages was 1.29 at the 25th percentile and 1.57 at the 75th percentile. We control for these pre-existing differences in inequality levels in the analyses below.

Despite unions’ rising prevalence and power at mid-century, data on union membership are rarely available at the micro- or local-level during the period under study, let alone for the entire United States. As mentioned above, the census has never inquired about union membership, and the CPS began inquiring about union membership in 1973. The Bureau of Labor Statistics (1951) and Troy (1957 and 1965) compiled the best information they could from union records and personal correspondence with union leaders, but consistent local-level information is simply not available.\footnote{Henry Farber, Dan Herbst, Ilyana Kuziemko, and Suresh Naidu are compiling union membership data from mid-century Gallup Polls, which when complete will provide new estimates of state-level union density.} This is a major challenge for quantitative research on mid-century labor market institutions. Troy’s estimates are careful and well informed but imperfect, as he points out: “These records are not equally accessible or reliable and can be explored only with the cooperation and assistance of many
union officials” (1957, p. 1). He further explains that “Whenever possible membership was computed from the financial reports of the union. Where data on dues received from locals…were unavailable, figures were obtained from reports of officers, by correspondence with unions, or were estimated on the basis of voting representation at conventions” (1957, p. 28).

Our analysis emphasizes a measure of local exposure to changes in unionization by combining information from nationwide industry-level changes in union density between 1939 and 1953, available from Troy (1957), with local measures of employment across industries calculated from the full-count 1940 census data. To our knowledge, Troy’s 1939 and 1953 benchmarks are the best estimates of industry-level union density for the period under study. He reports that metals manufacturing and transportation (including railroads) incurred the largest increases in union density in this period, whereas private services (e.g., retail and wholesale trade; finance, insurance, real estate), public services, and mining (which was highly organized before 1939), had the smallest changes in union density.

We calculate a local exposure-to-unions variable ($\Delta U_i$) that is a weighted average of national industry-level changes in unionization as a percent of industry employment ($\Delta U_j$), where the weights ($\omega_{ij}$) correspond to the local mix of employment in 1940 as observed in the full-count census microdata. $\Delta U_i$ is similar to a Bartik-style instrumental variable (Bartik 1991). Such instruments are often used to avoid endogeneity bias, such as (in this setting) the potential response of local unionization to local inequality.

We cannot estimate a “first-stage” SEA-level regression of actual change-in-unionization on the variable $\Delta U_i$ because no direct measures of local unionization exist. But to verify that there is an empirical connection, we can aggregate $\Delta U_i$ to the state level for comparison with separate state-level estimates of unionization from Troy (1957). The simple unweighted correlation across states is 0.77 with a $p$-value of <0.001 (and 0.79 when weighted by 1940 population). A bivariate regression of the change in Troy’s state-level estimates of union density on the $\Delta U_i$ proxy (aggregated to the state

---

9 Unfortunately, Troy’s papers were not deposited in the Rutgers library system to the knowledge of Troy’s colleagues and Rutgers librarians (personal correspondence February 2016). Troy provides industry-level and state-level estimates, but not industry-by-state level estimates. He was aware that industry-state estimates would be useful but concludes that, “Such figures are simply not available” (Troy 1957, p. 3).

10 The weights are based on the sample that enters into the wage inequality calculation.
level) yields a coefficient of 1.61 (s.e.=0.199). Thus, as far as can be told, there is a strong empirical relationship between $\Delta U_i$ and actual changes in unionization. That said, our interpretation of $\Delta U_i$ is generally in terms of its reflection of local exposure to unionization, as opposed to a direct measure of changes in local unionization rates.

Figure 4 maps the SEA-level variation in $\Delta U_i$. Since unionization rarely fell within industries over this period, $\Delta U_i$ is always above zero. There is clearly some geographic concentration in exposure to unions through the industrial belt of the Northeast and Midwest, but there is also considerable variation across SEAs within regions and states. Table 1 reports unweighted summary statistics for $\Delta U_i$. Overall, the average $\Delta U_i$ is 8.2 p.p. with a standard deviation of 3.1 p.p. For the SEA at the 75th percentile, $\Delta U_i$ is almost double that for the SEA at the 25th percentile, so there is considerable cross-place variation.

3. Empirical strategy

Our empirical strategy exploits the local variation shown in Figures 3 and 4 to see whether there was an association between the rise of unions and fall of inequality after 1940, conditional on observable local characteristics and regional or state trends. To fix ideas, consider a simple regression of the following form:

$$\Delta I_{ir} = \alpha + \tau \Delta U_{ir} + X_{ir}^t \beta + \gamma_r + e_{ir},$$

where $\Delta I_{ir}$ is a measure of the change in inequality in locality $i$ in region (or state) $r$ from 1939 to 1949 (or in separate regressions, from 1939 to 1959); $\tau$ is the coefficient of main interest associated with (exposure to) changes in union density in locality $i$ ($\Delta U_{ir}$); $X_{ir}$ is a vector of local characteristics (detailed below) that serve as control variables; and $\gamma_r$ is a set of region (or state) indicator variables. First differencing inequality eliminates local fixed effects, and the regional (or state) indicators absorb differential regional trends. Identification of $\tau$ then comes from within-region variation in $\Delta U_{ir}$, conditional on $X_{ir}$. Baseline regressions are weighted by the count of men in each area’s wage data from the 1940 full-count census records; results are similar if unweighted.

In such a regression, the clear threats to a causal interpretation of $\tau$ are the potential endogeneity of unions with respect to local concerns about inequality and, broadly, omitted variables that are correlated with $\Delta U_{ir}$ and affect local inequality through other channels. As described above, the variable $\Delta U_{ir}$ combines national-level changes in union membership rates with pre-existing local employment shares by industry. Because $\Delta U_{ir}$ is driven by changes in industry-level unionization at the national level, it is insulated from endogenous local changes in unionization.
Omitted variable bias remains a concern and motivates an extensive set of control variables that allow for differential post-1939 inequality trends depending on local economic characteristics circa 1940 and economic shocks after 1940. Baseline control variables in \( X_{it} \) include measures of the 1939 wage structure (the median log wage and the 90-10 log wage differential), economic conditions in 1940 (the employment rate for men, the share of workers who completed high school, and the percent of population in urban areas), wartime demand and investment shocks as reflected in war production and facilities contracts per capita, and labor demand shift variables. Summary statistics for these variables are reported in Table 1.

A particularly important concern is that industry-specific shocks during the 1940s could affect local inequality and be correlated with \( \Delta U_{it} \). This motivates the control variable for war-related production and facilities contracts. But we further address this concern by constructing two control variables that are tied directly to the local industry structure. One is a labor demand index that interacts the local distribution of employment over broad industries in 1940 with national-level industry employment growth during the 1940s. Places with local labor markets that were relatively concentrated in fast-growing industries might have experienced especially tight labor markets and, perhaps, differential changes in inequality. The labor demand control variable should help capture this hypothesized channel from industry structure to changing inequality. The second is a more finely tuned index and is meant to capture local shifts in the relative demand for skilled versus unskilled workers. Following Goldin and Margo (1992), it reflects the local distribution of employment over industries and the local skill mix within those industries in 1940. Changes in the relative demand for skill at the local level are then driven by national-level changes in industry employment shares, which affect areas differently depending on their initial skill mix within and across industries. The supplemental appendix discusses the construction of both industry-based control variables in more detail.

Robustness to several additional control variables is discussed below, including pre-1940
trends in inequality as reflected in changes in occupational and industrial employment distributions, the passage of right-to-work laws, and the bite of minimum wage laws.\textsuperscript{12} In addition, we find that moving from within-region to within-state comparisons as the basis for identifying $\tau$ (i.e., adding state fixed effects), which substantially narrows the geographic scope for confounding shocks and trends, has only a small impact on the estimates of $\tau$.

Ultimately, to sustain a causal interpretation of $\tau$, one has to assume that conditional on $X_i$ and $\gamma_i$ there were no unobserved shocks or trends in local inequality that were correlated with $\Delta U_i$. Such unobserved confounders cannot be ruled out in this setting, but to bias estimates of $\tau$ they would have to operate across localities within regions or states, after controlling for the pre-existing local wage structure and supply of skilled workers, pre-existing trends in occupational and industrial employment distributions, contemporaneous shifts in labor demand tied to war contracts and industrial concentration, and several other pre-war economic characteristics (e.g., employment rates, urbanization, and so on).

One of the benefits of the local perspective on inequality is that spillover effects on the wages of non-union workers are captured; that is, both union and non-union workers are in the wage sample. But we cannot separately observe union and non-union workers in the census data, which limits what we can say about within- and between-union-sector wage inequality. In addition, the local perspective will not capture union effects on inequality across places. Empirically, however, within-place compression was far more important than across-place compression in driving the overall decline in wage inequality between 1939 and 1949.\textsuperscript{13} A related concern, common to empirical strategies that rely on cross-place comparisons, is that general equilibrium effects could confound inference. For instance, if skilled workers left areas with large increases in unionization due to a reduced skill premium, they might increase the supply of skilled workers in other areas and reduce the skill premium; this would spread the compression effect of unions across areas and undermine the identification strategy outlined above. Of course, it is possible to imagine other scenarios in which general equilibrium effects confound measurement based on cross-place comparisons. These are difficult to assess, but we return to this issue below when assessing mechanisms, and in general causal interpretations should be qualified accordingly.

\textsuperscript{12} The U.S. census first recorded wage income in 1940, making it impossible to assemble nationally representative trends in wage inequality before 1940, let alone at the local level. We discuss this challenge in more detail later in the paper.

\textsuperscript{13} The variance in log weekly earnings can be decomposed into the between- and within-SEA component for each year. Of the total change in variance from 1940 to 1950 (-0.062), the change in within SEA variance (-0.052) makes up 83 percent.
4. Mid-century results

Estimates of $\tau$ from equation 1 are reported in Table 2. Each cell presents an estimate from a separate regression with different measures of wage inequality (across columns) and with different specifications and robustness checks (across rows). Panel A reports results for the 1939-49 period, whereas Panel B pertains to 1939-59. Looking over the 20-year period in Panel B allows us to use the relatively large 5-percent public use sample for 1960, and of course it provides a longer-term perspective on the conditional correlation between changes in inequality and exposure to unions. Standard errors are clustered by state in each case. Full results for the baseline specifications, with coefficients for all the covariates, are reported in the supplemental appendix.

The first row of Panel A shows that $\Delta U_{ir}$ was associated with declines in inequality as measured by the 90-10 differential, the 50-10 differential, the Gini coefficient, and the variance of the log weekly wage distribution. All those estimates of $\tau$ are statistically significant at the 1 percent level. Much of the inequality reduction associated with unions appears to be concentrated in the lower portion of the wage distribution, as reflected in the results for the 50-10 differential as opposed to those for the 90-50 differential. The baseline estimates of $\tau$ are sizable: a one standard deviation increase in $\Delta U_{ir}$ (3.1 p.p.) is associated with 0.072 decline in 90-10 wage inequality, equivalent to 32 percent of the mean decline in 90-10 inequality across SEAs during the 1940s.

The first row of Panel B reports results from the base specification for the 1939 to 1959 change in inequality. Again, the results indicate that exposure to unionization was associated with the compression of local wage structures, conditional on regional trends and controls for local characteristics, including industry-specific and skill-specific demand shocks. The coefficients are roughly similar in magnitude to those estimated in Panel A, implying that the connection between unions and inequality extended long after the macroeconomic shock of the war and the dismantling of wartime agencies that governed wage and price controls.

The baseline results are consistent with the hypothesis that the rise of unions caused wage structures to narrow during the 1940s, at least from the perspective of local labor markets. We consider several additional regression specifications that may further limit the scope for bias from unobserved shocks to local inequality that are potentially correlated with $\Delta U_{ir}$.

First, we add three control variables to equation (1) to see whether other events or policies that affected labor markets during the 1940s might be driving the baseline inequality results: mobilization of men into the armed services during the war, the local bite of minimum wage
legislation, and the passage of state-level “right to work” laws. Wartime mobilization drew many women into the labor market and varied considerably across states. Acemoglu, Autor, and Lyle (2004) find that mobilization-induced entry by women tended to reduce wage inequality among men. Therefore, we add a control variable for each state’s mobilization rate from Acemoglu, Autor, and Lyle (2004). In addition, Congress instituted a series of minimum wage increases, rising from $0.25 in 1938 to $0.30 in 1939 and $0.40 in 1945 (nominal dollars). We add a control variable for the share of workers in each SEA that earned less than the federal minimum wage for a full-time work week in 1939 ($0.30 x 40 hours = $12). This measures the size of the left-hand tail of the wage distribution that could be strongly affected by minimum wage legislation. Finally, as mentioned earlier, many states passed right-to-work laws after the Taft-Hartley Act. They were intended to reduce unions’ economic power and may reflect cross-place variation in the pro- versus anti-union balance of political power. We include an indicator for whether a state adopted a right-to-work statute or constitutional amendment prior to 1950 (Ellwood and Fine 1987). The results are reported in the second row of Table 2. The strong negative estimates of $\tau$ are robust to the additional institutional control variables and remain statistically significant.

Next, in the third row of Table 2, we replace the region-level fixed effects with state fixed effects. This substantially narrows the geographic scope for confounding shocks and trends by shifting identification to within-state variation rather than within-region. The estimates of $\tau$ are slightly smaller than in the base specification (row 1), but they remain strongly negative and statistically significant at the 1 percent level (with exception of the 90-50 differential).

In the fourth row, we attempt to control for the pre-1940 trend in local labor market inequality. The paucity of nationally representative wage information before the 1940 census is a difficult challenge for this paper’s empirical design. Wage surveys that do exist for earlier in the century tend to focus on urban areas and on middle class families, and they omit large portions of the United States. Instead, for each regression, we construct a 1930-40 pre-trend variable (corresponding to 90-10, 50-10, and so on) that is based on the distribution of industry-by-occupation income scores. The scores, in turn, equal the median income of men in each detailed industry-occupation cell from the full count 1940 census with at least 50 observations. Otherwise, the median income for the detailed occupation is used. This allows us to gauge local inequality pre-trends to the extent that the

---

14 Acemoglu, Autor, and Lyle (2004) define the state-level mobilization rate as the number of men, 18-44, who served in the military during the Second World War divided by the number of men registered for the draft. This is based on data from the Selective Service System.

15 The sample is limited to males aged 18 to 64, employed for at least 40 weeks in the prior year, who were wage and salary workers earning more than half of the implied weekly minimum wage.
evolving mix of local industries and occupations was associated with changing inequality. Obviously, this approach has shortcomings, as it cannot capture trends in within- or across-cell inequality, but it has the advantage of covering the entire U.S. in a consistent fashion. The fourth row shows that regressions augmented with the pre-trend control variable yield results that are similar to the baseline estimates.\textsuperscript{16}

For a much-reduced sample of 83 cities (not SEAs), we can draw on the Bureau of Labor Statistics 1919 Cost of Living Survey to provide an alternative approach to examining pre-trends in inequality. The survey provides micro-level income data for a sample of more than 12,000 families in 99 cities (ICPSR Study 8299, Feigenbaum 2016); some of these cities are not separately identified in the 1940 census. Unfortunately for the purposes of this study, the BLS’s sample frame focused on urban middle-class families. There is, nonetheless, considerable variation across cities in measured inequality in the 1919 data. There is also a statistically significant correlation between the levels of inequality in 1919 (BLS) and 1939 (Census), suggesting that the 1919-based measures are not just noise. Using the implied 1919-1939 change in inequality as the dependent variable in the baseline regression framework yields no significant conditional correlation with the union exposure variable. These results are in the supplemental appendix Table A3, row 3.

Row 5 of Table 2 drops predominantly rural SEAs from the sample to mitigate concern that identification is coming from comparisons of places that are strongly dissimilar in their economic character despite the many control variables included in the analysis.\textsuperscript{17} Again, the results are fairly robust.

Finally, Row 6 reports baseline results from unweighted regressions. The baseline estimates of $\tau$ generally do not depend strongly on the choice of weights, which may reduce concern about model misspecification from omitted heterogeneous effects.

5. Exploring mechanisms of wage compression

Even if one accepts that the conditional correlation between changes in local inequality and exposure to unionization reflects a causal relationship, it is not straightforward to interpret that relationship because the union effect could come through various channels with quite different

\textsuperscript{16} Adding a control for inequality pre-trends measured from 1920 to 1940 provide similar results as the 1930 to 1940 pre-trends displayed in Table 2.

\textsuperscript{17} A related issue concerns agricultural laborers, who are in the sample if they meet the criteria set in Goldin and Margo (1992). We believe they should be part of the analysis because they are clearly part of a locality’s labor force. But we note that excluding agricultural laborers from the calculations of inequality leads to smaller point estimates in the baseline specification (by one-quarter to one-half); results that were statistically significant in the baseline remain so.
implications. For instance, if rapidly rising unionization led firms or workers to sort out of (or into) some locations, then the observed union effect on wage structure might have worked through sample composition rather than by compressing wages for a given set of workers (or a given set of worker characteristics). These mechanisms might be expected in a Rosen-Roback framework if firms viewed heavily unionized areas as having a “productive disamenity” or in a Roy model framework if skilled workers perceived local declines in the return to skill (Rosen 1979, Roback 1980, Roy 1951). To our knowledge, no nationally representative micro-level data reveal skill-specific migration patterns across SEAs in the 1940s. However, it is possible to see whether observed population characteristics exhibit patterns of change that are consistent with an endogenous response to the rise of unions, and if so, whether those responses can account for the empirical link between $\Delta U_{ir}$ and $\Delta I_{ir}$.

For brevity, we summarize results from regressions that are similar in form to the baseline specification described in equation 1 but that use changes in population characteristics as the dependent variable. The results appear in Table 3, columns 1 and 2. First, we do not find consistent evidence that changes in the skill mix, measured as the change in the share of the adult population with a high school degree or more, were correlated with $\Delta U_{ir}$ in baseline regressions. In the 1940-50 period, there is a negative correlation, but over the 1940-60 period, which allows more time for endogenous sorting, the correlation is positive but statistically insignificant. Second, we do not find evidence that total population growth was negatively correlated with local exposure to unionization, which would result if firms fled areas that were highly exposed to unionization. In fact, we see the opposite, a positive correlation over both the 1940-50 and 1940-60 periods, conditional on the baseline controls.

Next, in columns 3-7 of Table 3, we include the potentially endogenous variables (changing skill mix and population growth) as controls in our baseline regressions of inequality on union exposure to see whether estimates of $\tau$ are sensitive to their inclusion. Because the skill mix and population growth may be endogenous to unionization, we would not consider estimates of $\tau$ in such regressions to be preferable to those offered in Table 2. Rather, the point is to see whether changes in those potentially endogenous variables underpinned the baseline results. In short, the estimates of $\tau$ change little relative to those reported in Table 2, consistent with the effects of unions working through channels other than sample composition and firm relocation.

---

18 Callaway and Collins (2017) find relatively low returns to education for union workers relative to non-union workers in a small sample of urban workers in 1950. In a different setting, see Abramitzky (2009) for evidence of highly skilled workers sorting out of Israeli kibbutzim.
Finally, we examine changes in the distribution of the residuals from wage equations that control for education, experience, their squares, and their interaction. The goal is to see whether local exposure to unions was associated with declines in wage variation within observable education and experience categories, as reflected in the distribution of residuals. This would be consistent with findings from more recent decades (e.g., Card, Lemieux, and Riddell 2004), where unions reduce overall inequality by standardizing wages across workers with similar observables. Indeed, the evidence is consistent with this hypothesis. The details of this analysis are reported in the supplemental appendix.

Our conclusion is that the mid-century conditional correlation between local union exposure and changing inequality was probably causal and unlikely to have been driven by the endogenous sorting of workers and firms, at least in the short to medium run. Based on the historical record, it is plausible that the mechanisms entailed both wartime wage-setting practices by the NWLB and longer-standing reinforcement of wage compression by unions, which remained powerful after the war. It is also plausible that in practice unions reduced inequality by standardizing wages and compressing wage distributions directly, as suggested by results in Callaway and Collins (2017), and perhaps through threat effects on non-union firms. Although more difficult to pin down, it is also possible that unions influenced local norms about wage setting (Western and Rosenfeld 2011). Deeper and more direct insight into the mechanisms linking unions and wage inequality at mid-century would be valuable and, therefore, a promising topic for future research.

6. Ramifications beyond 1960

As described above, labor market and institutional conditions in the United States in the 1940s were particularly favorable for the growth of unions. At a national level, union density remained fairly constant at its mid-century peak for about two decades until a strong secular decline began in the 1970s. Whether the sharp increase in unions during the 1940s had effects on local wage structures that endured beyond the 1960s is an open question. As private sector unions waned in importance at the national level, reflecting a combination of less favorable laws, more effective management resistance, increased product market competition, and a shift toward service-based output (Freeman 1988, Reder 1988), it is possible that unions’ imprint on local wage structures faded away. On the other hand, it also possible that places with greater exposure to unions and wage compression in the 1940s retained relatively compressed wage structures into the future, at least in comparison with other localities.
Here, we examine whether exposure to unionization in the 1940s had a persistent correlation with local inequality in long-difference regressions that span from 1940 to the end of the 20th century (1940-60, 1940-70, 1940-80, etc.). In essence, this simply extends our basic empirical approach through the end of the century; all the regressions include the complete set of controls from the base specification of equation 1. Each entry in Table 4 is from a separate regression, where \( \tau \) is still the coefficient on \( \Delta U_{ir} \) and where the dependent variable is the change in inequality measured over varying periods of length. To be clear, \( \Delta U_{ir} \) is still keyed to industry structure in 1940 and industry-specific growth in unions at the national level between 1939 and 1953, corresponding to the relatively short but dramatic burst of union growth in the United States. Thus, estimates of \( \tau \) correspond to the long-run association between union growth in the 1940s and changes in local inequality measured over longer and longer timeframes.

Results in Table 4 show that exposure to unionization circa 1940 was negatively associated with changes in local inequality even 50 years after the Second World War, with the most prominent manifestation in the 50-10 differential. Even though there is some evidence that the empirical connection faded between 1970 and 1990, it did not disappear entirely, and the point estimate rebounded somewhat by 2000. Although somewhat speculative, these correlations may reflect persistent variation in union density and influence across localities. At a more aggregate level, there is evidence that states with high levels of unionization circa 1964 continued to have relatively high levels circa 2000, even though unionization was declining everywhere.\(^{19}\) We cannot, however, rule out other channels connecting the experience of the 1940s to the wage distribution circa 2000, such as persistent norms or long-run selection of workers and firms across localities in a way that corresponded to wage patterns established in the 1940s and 1950s, even as union membership waned.

7. Conclusions

Unions rose to prominence in the United States from the ashes of the Great Depression and reached their peak in the early 1950s. At the same time, wage inequality declined sharply and remained relatively compressed for decades. This national-level pattern is readily apparent in time series data, but making a causal inference from that pattern is tenuous, as many things changed at the same time. In this paper, we exploit cross-place variation to shed light on the connection between

\(^{19}\) It is evident that states with relatively high union density in the 1960s also had relatively high density circa 2000, even though the levels had declined; the correlation between 1964 and 2000’s state-level density is 0.82 (Hirsch, Macpherson, and Vroman 2001). This correlation is calculated using union membership density data available at: http://unionstats.gsu.edu/MonthlyLaborReviewArticle.htm, accessed July 19, 2017. See also Cohen, Malloy, and Nguyen (2016).
rising unions and falling inequality. While we benefit from new datasets, especially the full count 1940 census data, there are still significant data challenges to this investigation. First and foremost, systematic local-level measures of unionization do not exist for the United States; therefore, we rely on a measure of exposure to post-1939 unionization based on pre-existing industrial characteristics. Second, because the census did not inquire about wages before 1939, it is difficult to account for pre-trends in local inequality, though it is possible to control for the local wage structure at the start of the period and local pre-trends in an inequality measure derived from occupational and industrial composition. The baseline regressions allow for differential trends by region and a number of economic characteristics in 1940; they also control for post-1940 shocks to labor demand through war contracts and local industrial concentration interacted with national industry-level employment growth trends.

We find a robust negative correlation between exposure to unionization and changes in local inequality. Taking the baseline estimate of $\tau$ for the 1939-49 period at face value, moving from the 25th to 75th percentile of the exposure-to-unions variable is associated with a 0.10 log point reduction in 90-10 wage inequality, ceteris paribus. This is a relatively large change compared to the median decline of 0.25. Unions may have influenced local inequality by standardizing wages within and across firms and by influencing wage levels of non-union firms. But union pressures could also have caused endogenous relocation of firms and skilled workers. This too might influence observed local income inequality, but through channels that are quite different from those typically emphasized in studies of unions. Our analysis, however, finds no evidence that such endogenous sorting underpins the baseline results. Finally, the conditional correlation between exposure to unions circa 1940 and changes in local inequality persisted for many decades. Thus, even as unions faded from the private sector after the 1970s in the United States, traces of the differential compression of the 1940s remained visible at the end of the century.

Our findings leave open many questions about the Great Compression and the role that unions played in the mid-twentieth century economy. In addition to suggesting more research into the mechanisms and persistence of compression, we conclude by pointing out several related topics that demand attention. Due to data limitations, we have not assessed non-wage dimensions of compensation, such as health insurance, paid vacations, and pensions, or the role of unions in advancing them. But these became important components of compensation in the second half of the century and certainly had implications for the wellbeing of workers and their families. Also, we have not assessed the potential unintended consequences of unions and wage compression for the
investment and location decisions of firms or young workers. Such responses may have been second-order in the short run, but also may have accumulated with time. Finally, we have not developed international comparisons in this paper. Labor laws and movements differed across developed economies. Such variation might help us better understand the connections between labor market institutions and outcomes, though of course, the destruction and disarray of the Second World War complicate international comparisons for the mid-century period. All of the above represent promising avenues for new research on economic history of inequality and labor market institutions.
References


Table 1: U.S. SEA-level summary statistics

Panel A: Inequality and union exposure

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St. Dev</th>
<th>25\textsuperscript{th} pctile</th>
<th>Median</th>
<th>75\textsuperscript{th} pctile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940-1950</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta 90-10) wage differential</td>
<td>-0.224</td>
<td>0.229</td>
<td>-0.372</td>
<td>-0.248</td>
<td>-0.107</td>
</tr>
<tr>
<td>(\Delta 90-50) wage differential</td>
<td>-0.188</td>
<td>0.129</td>
<td>-0.266</td>
<td>-0.182</td>
<td>-0.105</td>
</tr>
<tr>
<td>(\Delta 50-10) wage differential</td>
<td>-0.036</td>
<td>0.234</td>
<td>-0.194</td>
<td>-0.073</td>
<td>0.089</td>
</tr>
<tr>
<td>(\Delta) Variance of log wages</td>
<td>-0.048</td>
<td>0.075</td>
<td>-0.094</td>
<td>-0.055</td>
<td>-0.010</td>
</tr>
<tr>
<td>(\Delta) Gini</td>
<td>-0.057</td>
<td>0.036</td>
<td>-0.080</td>
<td>-0.057</td>
<td>-0.033</td>
</tr>
<tr>
<td>1940-1960</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta 90-10) wage differential</td>
<td>-0.249</td>
<td>0.153</td>
<td>-0.356</td>
<td>-0.244</td>
<td>-0.158</td>
</tr>
<tr>
<td>(\Delta 90-50) wage differential</td>
<td>-0.199</td>
<td>0.104</td>
<td>-0.274</td>
<td>-0.199</td>
<td>-0.129</td>
</tr>
<tr>
<td>(\Delta 50-10) wage differential</td>
<td>-0.050</td>
<td>0.160</td>
<td>-0.160</td>
<td>-0.686</td>
<td>0.056</td>
</tr>
<tr>
<td>(\Delta) Variance of log wages</td>
<td>-0.069</td>
<td>0.047</td>
<td>-0.103</td>
<td>-0.068</td>
<td>-0.038</td>
</tr>
<tr>
<td>(\Delta) Gini</td>
<td>-0.052</td>
<td>0.026</td>
<td>-0.071</td>
<td>-0.053</td>
<td>-0.035</td>
</tr>
<tr>
<td>1939-1953</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Delta U_{ir})</td>
<td>8.2</td>
<td>3.1</td>
<td>5.7</td>
<td>7.5</td>
<td>10.1</td>
</tr>
</tbody>
</table>

Panel B: Control variables

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St. Dev</th>
<th>25\textsuperscript{th} pctile</th>
<th>Median</th>
<th>75\textsuperscript{th} pctile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median log weekly wage 1939</td>
<td>3.006</td>
<td>0.248</td>
<td>2.795</td>
<td>3.033</td>
<td>3.219</td>
</tr>
<tr>
<td>Initial log 90-10 gap in 1939</td>
<td>1.418</td>
<td>0.179</td>
<td>1.287</td>
<td>1.435</td>
<td>1.566</td>
</tr>
<tr>
<td>Per capita WWII contracts (1940$)</td>
<td>1.172</td>
<td>1.805</td>
<td>0.192</td>
<td>0.467</td>
<td>1.347</td>
</tr>
<tr>
<td>Male employment rate in 1940</td>
<td>0.674</td>
<td>0.054</td>
<td>0.642</td>
<td>0.679</td>
<td>0.707</td>
</tr>
<tr>
<td>Percent urban 1940</td>
<td>0.417</td>
<td>0.245</td>
<td>0.214</td>
<td>0.364</td>
<td>0.622</td>
</tr>
<tr>
<td>Percent males HS grads</td>
<td>0.248</td>
<td>0.054</td>
<td>0.208</td>
<td>0.249</td>
<td>0.285</td>
</tr>
<tr>
<td>Local demand shock index</td>
<td>1.896</td>
<td>0.262</td>
<td>1.725</td>
<td>1.865</td>
<td>2.051</td>
</tr>
<tr>
<td>Local skill-specific demand shock index</td>
<td>-0.002</td>
<td>0.019</td>
<td>-0.013</td>
<td>-0.004</td>
<td>0.007</td>
</tr>
<tr>
<td>State WWII mobilization rate</td>
<td>0.469</td>
<td>0.032</td>
<td>0.449</td>
<td>0.468</td>
<td>0.49</td>
</tr>
<tr>
<td>Percent of males earning below minimum wage in 1939</td>
<td>0.313</td>
<td>0.178</td>
<td>0.161</td>
<td>0.279</td>
<td>0.453</td>
</tr>
<tr>
<td>State passed right to work law by 1950</td>
<td>0.259</td>
<td>0.439</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Notes: Unweighted summary statistics over State Economic Areas.
Sources: Inequality measures in Panel A and wage measures, demand indices, high school educational attainment, and the share earning below the minimum wage in Panel B are derived from the 1940, 1950, or 1960 IPUMS samples (Ruggles et al. 2015), as described in the text. Exposure to unions (\(\Delta U_{ir}\)) is calculated by interacting industrial employment data from the 1940 full count census and industry-level changes in union density from Troy (1957). The employment rate, share of population in urban areas, and per capita war expenditure variables are calculated with data from Haines (2010). War mobilization is from Acemoglu, Autor, and Lyle (2004). Right to work laws are from Ellwood and Fine (1987). More detailed information is provided in the supplemental appendix.
Table 2: Regressions results, changes in wage inequality and unionization

<table>
<thead>
<tr>
<th>Panel A: 1940-50</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Δ90-10</td>
<td>Δ90-50</td>
<td>Δ50-10</td>
<td>ΔGini</td>
<td>ΔVar</td>
</tr>
<tr>
<td>(1) Base specification</td>
<td>-0.0232</td>
<td>-0.00305</td>
<td>-0.0201</td>
<td>-0.0041</td>
<td>-0.0066</td>
</tr>
<tr>
<td></td>
<td>(0.00619)</td>
<td>(0.00326)</td>
<td>(0.00435)</td>
<td>(0.0011)</td>
<td>(0.0019)</td>
</tr>
<tr>
<td>(2) Base with controls for mobilization and policy</td>
<td>-0.0190</td>
<td>0.00017</td>
<td>-0.0192</td>
<td>-0.0031</td>
<td>-0.0049</td>
</tr>
<tr>
<td></td>
<td>(0.00626)</td>
<td>(0.00309)</td>
<td>(0.00439)</td>
<td>(0.0010)</td>
<td>(0.0016)</td>
</tr>
<tr>
<td>(3) Base with state f.e.</td>
<td>-0.0202</td>
<td>-0.00415</td>
<td>-0.0160</td>
<td>-0.0035</td>
<td>-0.0057</td>
</tr>
<tr>
<td></td>
<td>(0.00649)</td>
<td>(0.00328)</td>
<td>(0.00508)</td>
<td>(0.0010)</td>
<td>(0.0019)</td>
</tr>
<tr>
<td>(4) Base w/ pre-trend control</td>
<td>-0.0233</td>
<td>-0.0033</td>
<td>-0.0199</td>
<td>-0.0040</td>
<td>-0.0058</td>
</tr>
<tr>
<td></td>
<td>(0.0062)</td>
<td>(0.0035)</td>
<td>(0.0044)</td>
<td>(0.0011)</td>
<td>(0.0018)</td>
</tr>
<tr>
<td>(5) Base, restrict to urban SEAs (&gt;=50%)</td>
<td>-0.0208</td>
<td>0.00019</td>
<td>-0.0210</td>
<td>-0.0030</td>
<td>-0.0034</td>
</tr>
<tr>
<td></td>
<td>(0.0102)</td>
<td>(0.0056)</td>
<td>(0.0070)</td>
<td>(0.0018)</td>
<td>(0.0028)</td>
</tr>
<tr>
<td>(6) Base, unweighted</td>
<td>-0.0199</td>
<td>-0.00319</td>
<td>-0.0167</td>
<td>-0.0040</td>
<td>-0.0074</td>
</tr>
<tr>
<td></td>
<td>(0.00520)</td>
<td>(0.00254)</td>
<td>(0.00458)</td>
<td>(0.0009)</td>
<td>(0.0019)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: 1940-60</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Δ90-10</td>
<td>Δ90-50</td>
<td>Δ50-10</td>
<td>ΔGini</td>
<td>ΔVar</td>
</tr>
<tr>
<td>(1) Base specification</td>
<td>-0.0214</td>
<td>-0.00482</td>
<td>-0.0165</td>
<td>-0.0032</td>
<td>-0.0056</td>
</tr>
<tr>
<td></td>
<td>(0.00474)</td>
<td>(0.00248)</td>
<td>(0.00459)</td>
<td>(0.00080)</td>
<td>(0.00150)</td>
</tr>
<tr>
<td>(2) Base with controls for mobilization and policy</td>
<td>-0.0212</td>
<td>-0.00218</td>
<td>-0.0189</td>
<td>-0.00271</td>
<td>-0.00562</td>
</tr>
<tr>
<td></td>
<td>(0.00515)</td>
<td>(0.00260)</td>
<td>(0.00484)</td>
<td>(0.00083)</td>
<td>(0.00161)</td>
</tr>
<tr>
<td>(3) Base with state f.e.</td>
<td>-0.0260</td>
<td>-0.00684</td>
<td>-0.0191</td>
<td>-0.00371</td>
<td>-0.00695</td>
</tr>
<tr>
<td></td>
<td>(0.00488)</td>
<td>(0.00266)</td>
<td>(0.00485)</td>
<td>(0.00081)</td>
<td>(0.00160)</td>
</tr>
<tr>
<td>(4) Base w/ pre-trend control</td>
<td>-0.0220</td>
<td>-0.0048</td>
<td>-0.0162</td>
<td>-0.0032</td>
<td>-0.0058</td>
</tr>
<tr>
<td></td>
<td>(0.0047)</td>
<td>(0.0026)</td>
<td>(0.0044)</td>
<td>(0.00082)</td>
<td>(0.0015)</td>
</tr>
<tr>
<td>(5) Base, restrict to urban SEAs (&gt;=50%)</td>
<td>-0.0166</td>
<td>-0.0009</td>
<td>-0.0156</td>
<td>-0.0029</td>
<td>-0.0035</td>
</tr>
<tr>
<td></td>
<td>(0.0087)</td>
<td>(0.0035)</td>
<td>(0.0080)</td>
<td>(0.0013)</td>
<td>(0.0024)</td>
</tr>
<tr>
<td>(6) Base, unweighted</td>
<td>-0.0181</td>
<td>-0.00543</td>
<td>-0.0126</td>
<td>-0.0025</td>
<td>-0.0053</td>
</tr>
<tr>
<td></td>
<td>(0.00527)</td>
<td>(0.00230)</td>
<td>(0.00539)</td>
<td>(0.0007)</td>
<td>(0.0015)</td>
</tr>
</tbody>
</table>

Notes: Each figure in the table is the estimated coefficient (τ) on the change-in-unions variable from a separate regression. The base specification’s control variables are described in the text; they include regional fixed effect, measures of the 1939 wage structure (the median log wage and the 90-10 wage gap), 1940 economic conditions (the employment rate for men, the share of workers who completed high school, and the percent of population in urban areas), wartime demand and investment shocks as reflected in war production and facilities contracts per capita, and labor demand shift variables. “Mobilization and policy” adds controls for the share of the 1940 labor force earning less than the minimum wage, state-level indicators for right-to-work laws, and military mobilization rates. The pre-trend control is described in the text and reflects changing occupational distributions.

Sources: These are described in the text, below table 1, and in the supplemental appendix.
Table 3: Population and skill-mix responses to union exposure

<table>
<thead>
<tr>
<th>Dep. variable</th>
<th>Δ share HS+</th>
<th>Δ ln(pop)</th>
<th>Δ90-10</th>
<th>Δ90-50</th>
<th>Δ50-10</th>
<th>ΔGini</th>
<th>ΔVar</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940-50</td>
<td>-0.0029</td>
<td>0.0072</td>
<td>-0.0228</td>
<td>-0.0024</td>
<td>-0.0204</td>
<td>-0.0040</td>
<td>-0.0061</td>
</tr>
<tr>
<td></td>
<td>(0.0017)</td>
<td>(0.0040)</td>
<td>(0.0063)</td>
<td>(0.0032)</td>
<td>(0.0044)</td>
<td>(0.0011)</td>
<td>(0.0019)</td>
</tr>
<tr>
<td>1940-60</td>
<td>0.0020</td>
<td>0.0267</td>
<td>-0.0229</td>
<td>-0.0042</td>
<td>-0.0187</td>
<td>-0.0032</td>
<td>-0.0061</td>
</tr>
<tr>
<td></td>
<td>(0.0017)</td>
<td>(0.0084)</td>
<td>(0.0048)</td>
<td>(0.0026)</td>
<td>(0.0045)</td>
<td>(0.0008)</td>
<td>(0.0015)</td>
</tr>
</tbody>
</table>

Notes: Each figure in the table is the estimated coefficient (τ) on the change-in-unions variable from a separate regression with different dependent variables. Columns (1) and (2) use as the dependent variable, respectively, the change in the share of the sample that is a high school graduate or above and the change in log population. Columns (3) to (7) correspond to the baseline regressions of Table 2 but include control variables for the changing skill mix and population growth. The base specification’s control variables are described in the text and include regional fixed effects, measures of the 1939 wage structure (the median log wage and the 90-10 wage gap), 1940 economic conditions (the employment rate for men, the share of workers who completed high school, and the percent of population in urban areas), wartime demand and investment shocks as reflected in war production and facilities contracts per capita and labor demand shift variables. Sources: See Table 1 and supplemental appendix.
Table 4: Long-difference regression results, changes in wage inequality and mid-century unionization

<table>
<thead>
<tr>
<th></th>
<th>(1) 1940-50</th>
<th>(2) 1940-60</th>
<th>(3) 1940-70</th>
<th>(4) 1940-80</th>
<th>(5) 1940-90</th>
<th>(6) 1940-2000</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Long differences from 1940</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90-10</td>
<td>-0.0232</td>
<td>-0.0214</td>
<td>-0.0254</td>
<td>-0.0167</td>
<td>-0.0121</td>
<td>-0.0186</td>
</tr>
<tr>
<td></td>
<td>(0.00619)</td>
<td>(0.00474)</td>
<td>(0.00370)</td>
<td>(0.00302)</td>
<td>(0.00507)</td>
<td>(0.00618)</td>
</tr>
<tr>
<td>90-50</td>
<td>-0.00305</td>
<td>-0.00482</td>
<td>-0.00458</td>
<td>-0.00411</td>
<td>-0.00604</td>
<td>-0.00403</td>
</tr>
<tr>
<td></td>
<td>(0.00326)</td>
<td>(0.00248)</td>
<td>(0.00272)</td>
<td>(0.00262)</td>
<td>(0.00243)</td>
<td>(0.00272)</td>
</tr>
<tr>
<td>50-10</td>
<td>-0.0201</td>
<td>-0.0165</td>
<td>-0.0209</td>
<td>-0.0126</td>
<td>-0.00610</td>
<td>-0.0146</td>
</tr>
<tr>
<td></td>
<td>(0.00435)</td>
<td>(0.00459)</td>
<td>(0.00368)</td>
<td>(0.00354)</td>
<td>(0.00473)</td>
<td>(0.00578)</td>
</tr>
</tbody>
</table>

Notes: The Gini coefficient and variance of the log weekly wage distribution follow a similar pattern as the 90-10 differential.
Sources: Income inequality is calculated from the census microdata (Ruggles et al. 2015). The $\Delta U_t$ variable is described in the text; it combines information from Troy (1957) and the 1940 complete count U.S. census microdata for 1940 (Ruggles et al. 2015).
Figure 1: Unions and income inequality trends in the 20th-century United States

Sources: The series for income share of the top 10 percent of tax units is from Piketty and Saez (2003, updated to 2014, Table A1, income shares excluding capital gains, accessed August 3, 2015); 1917 is the first year available in the income inequality series. The union membership data up to 1983 are the Troy-Sheflin series, minus Canadian membership of U.S. unions, as reported in Carter et al. (2006, series Ba4785 and Ba4786). The civilian nonfarm employment data up to 1983 are also from Carter et al. (2006, series Ba471 and Ba472). In 1983, the union density series is spliced to the Bureau of Labor Statistics series for union members as a share of employed wage and salary workers (series id: LUU0204899600).
Figure 2: Change in real log weekly wages, by percentile of the distributions

Notes: The earnings sample includes men reporting wage and salary employment in the reference week, aged 18 to 64 years who reported positive wage and salary income in the year prior to the census, worked more than 39 weeks, and earned more than half the minimum wage at a full-time basis (weekly wages of $6 in 1940 and $8 in 1950, $26 in 1970, and $27 in 1990).
Sources: Earnings data from decennial census microdata provided by IPUMS (Integrated Public Use Microdata Series) 1940 100 percent complete count, and 1950, 1970, and 1990 1 percent samples.
Figure 3: SEA-level changes in wage inequality, 1939-1949

Notes: The map displays the geographic variation in the compression of the 90-10 differential in log weekly earnings between 1939 and 1949 for state economics areas (SEAs). Sources: IPUMS (Integrated Public Use Microdata Series) 1940 100% complete count and 1950 1% sample provide wage and salary earnings data.

Figure 4: SEA-level changes in unions, 1939-1953

Notes: The map displays the geographic variation in the $\Delta U_i$ variable for state economics areas (SEAs).