Unions and the Great Compression of Wage Inequality in the United States at Mid-Century: Evidence from Local Labor Markets

William J. Collins and Gregory T. Niemesh

March 2018

Abstract: This paper tests whether places with higher exposure to unionization during the 1940s, due to their pre-existing industrial composition, tended to have larger declines in wage 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 input from three referees, Ran Abramitzky, Leah Boustan, Brant Callaway, Michael Edelstein, William Even, James Feigenbaum, Alex Field, 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), the World Congress of Cliometrics, and the Economic History Association. 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.[1] 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 termed 'the Great Compression'. Scholars in the 1950s remarked on this event, and several have considered it since that time.[2] 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.[3] Second, although research that specifically addresses unions and wage compression in the U.S. in the 1940s and 1950s is scarce, Miller speculated that union bargaining disproportionately raised wages at the bottom of the income distribution; Goldin and Margo suggested that unions might have helped sustain the relatively narrow dispersion of wartime wages into the post-war period; Frydman and Molloy found that executive compensation relative to production workers' pay was negatively correlated with union presence in 1949; and Callaway and Collins observed that the distribution of union members' wages was compressed relative to that of observationally similar non-members in 1950.[4] 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

---

[1] See inter alia Katz and Murphy, 'Changes in relative wages'; DiNardo, Fortin, and Lemieux, 'Labor market institutions'; Autor, Katz, and Kearney, 'Trends in U.S. wage inequality'; Acemoglu and Autor, 'Skills, tasks and technologies'.
[2] For early work see Kuznets and Jenks, *Shares of upper income*; Goldsmith et al., 'Size distribution'; Miller, 'Changes in the industrial distribution.' For later work see Juhn, 'Wage inequality'; Piketty and Saez, 'Income inequality'; Goldin and Katz, *Race between education and technology*.
[3] Freeman, 'Unionism'; Freeman and Medoff, *What do unions do?*; DiNardo, Fortin, and Lemieux, 'Labor market institutions'; Card, Lemieux, and Riddell, 'Unions'.
[4] Miller, 'Changes in the industrial distribution'; Goldin and Margo, 'Great compression', p. 32; Frydman and Molloy, 'Pay cuts'; Callaway and Collins, 'Unions, workers, and wages'.

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.[5] But the mid-century evidence was fragmentary and difficult to evaluate, reflecting the scarcity of detailed and representative data on workers' wages and union status.[6] 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 exposure to increases in unionization 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 ex ante exposure to unionization then forms the basis of the identification strategy outlined below. This approach attempts to avoid bias from endogenous local changes in unionization, which could occur if unionization responds to local trends in inequality or to local trends that are unobserved but correlated with inequality. The approach may also reflect threat effects associated with potential unionization even when local unions do not form. This emphasis on variation in exposure is akin to recent literature that measures the local labor market effects of changes in international trade or technology.[7]

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.[8] 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

---

[5] Friedman, *Capitalism and freedom*, pp. 123-25; Rees, *Economics of trade unions*, pp. 98-99. Card, Lemieux, and Riddell, 'Unions', provide an encapsulation of the literature on unions and inequality since the 1950s.
[6] Lewis, *Unionism*.
[7] Autor, Dorn, and Hanson, 'China syndrome'; Acemoglu and Restrepo, 'Robots and jobs'.
[8] Ruggles et al., *Integrated public use microdata series*.

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).[9]

The 1940s were a tumultuous time for the U.S. economy due to the Second World War. 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. 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 reveal that this pattern was driven by the sorting of firms or skilled workers out of locations exposed to sharp increases in union density between 1940 and 1960. 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 increases in inequality between 1940 and 2000 than others.

I. Brief background on the rise of US 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.[10] At a deeper level, the changes in federal policy reflected workers' demand for representation in the wake of the Great Depression, policymakers' desire to reduce the frequency and severity of violent strikes, and the exigencies of the Second World War.[11] Subsequent changes in union membership were highly uneven across industries and across space, 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

[9] Lewis, *Unionism*; Farber, 'Nonunion wage rates'.
[10] Seidman, American labor; Reder, 'Rise and fall of unions'; Freeman, 'Spurts in union growth'.
[11] Ashenfelter and Pencavel, 'American trade union'; Wachter, 'Striking success'.

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.[12] 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.[13] The legality of the Wagner Act was challenged, but the Supreme Court's 5-4 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.[14] 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.[15] 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.[16] NWLB decisions also tended to narrow interplant wage differentials within industries and localities.[17] 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 wage compression persisted.

In 1947, Congress passed the Taft-Hartley Act over President Truman's veto, a significant legislative setback to the labor movement.[18] Among other things, Taft-Hartley circumscribed a

---

[12] Lester, *Economics of labor*, ch. 5.
[13] 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).
[14] Seidman, *American labor*; Freeman, 'Delivering the goods'.
[15] Seidman, *American labor*, pp. 61-62 and 91-108, discusses maintenance or membership clauses in detail. Despite the 'no strike' pledges, there were strikes during the war, most famously the miners' strike of 1943.
[16] Dunlop, 'Appraisal of wage stabilization'; Goldin and Margo, 'Great compression'.
[17] Seidman, *American labor*, p. 115.
[18] Millis and Brown, *From the Wagner Act to Taft-Hartley*.

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.

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.[19] 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—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 motivated 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 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 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.'[20] Jaworksi and Niemesh 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

[19] Goldin and Katz, *Race between education and technology*, ch. 2.
[20] Goldin and Margo, 'Great compression', p. 32.

within-group narrowing of inequality.[21] Piketty and Saez, 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 falling inequality, but without presenting direct evidence on this point.[22] Levy and Temin provide a review and interpretation of changes in U.S. inequality and labor market institutions over the twentieth century; unions feature prominently in this interpretation, but direct econometric evidence regarding the Great Compression is not developed.[23] Lewis reviews and critiques the economics literature on U.S. unions and wages up to the early 1960s. He concludes that unions' effect on overall inequality was likely small, but the evidence at hand was fragmentary and imperfect.[24]

Although integrating international evidence is beyond the scope of this paper, it is worth pointing out that Gazeley and Atkinson and Nolan suggest that institutional changes in the 1940s may have been important in narrowing wage inequality in Britain and Ireland, respectively.[25] 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 that were 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 after the war, are consistent with what one might expect from the rise and persistence 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.

II. Data describing mid-century inequality and unions

Our work benefits from newly available and large micro datasets from the federal census of population in 1940 (complete count) and in 1960 (5-percent sample), which are available from the Integrated Public Use Microdata Series (IPUMS).[26] Their size permits greater precision in measuring local economic characteristics, and importantly, they reveal geographic information at a

[21] Jaworski and Niemesh, 'Revisiting the great compression'.
[22] Piketty and Saez, 'Income inequality', p. 34.
[23] Levy and Temin, 'Inequality'.
[24] Lewis, *Unionism*, p. 295.
[25] Gazeley, 'Leveling of pay'; Atkinson and Nolan, 'Changing distribution of earnings'.
[26] Ruggles et al., *Integrated public use microdata series*.

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'.[27] 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.[28]

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. 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 by Autor and Dorn.[29] 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.[30] The supplementary appendix

---

[27] Bogue, *State economic areas*, p. 1.

[28] 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. Changes outside metropolitan areas were important in the Great Compression. We want to capture that in our analysis.

[29] Autor and Dorn, 'Growth of low-skill service'.

[30] 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.

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 closely in using the census wage data.[31] This ensures consistency with their approach to measuring changes in the mid-century wage structure, and the choices reflect the constraints and imperfections inherent to mid-century U.S. census data. Weekly income is computed by dividing annual income from wage and salary work by the total number of weeks worked.[32] Self-employment income and capital income were not reported in the 1940 census; for consistency, we do not include these income sources in later years. 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. This relies on those who worked most of the year to characterize the local weekly wage structure and omits those whose computed weekly earnings are probably implausibly low. 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. Following Goldin and Margo, we multiply top-coded values by 1.4 in each year.[33] In each case, the relevant census questions refer to the year prior to enumeration; thus, the 1940 census refers to income in 1939. We address sensitivity to sample restrictions at the end of this section.

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.

[31] We follow Goldin and Margo's description of sample restrictions exactly. In turn, they note that their sample restrictions are 'virtually identical' to those used by labor economists studying wage inequality circa 1990. See 'Great compression', p. 6. Goldin and Margo sometimes limit their sample to white men, whereas we do not restrict the sample on race.

[32] 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.

[33] In the 1940 full count dataset, users must 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 top-code the errors at 1.4 times the value of the top-code.

Although 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 of SEAs. 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. The Bureau of Labor Statistics and Leo Troy compiled the best information they could from union records and personal correspondence with union leaders, but consistent local-level information is simply not available. This is a major challenge for quantitative research on mid-century labor market institutions.[34] 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.'[35] 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.'[36]

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,[37] 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 to date of industry-level union density for the period under study.[38] 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

[34] Bureau of Labor Statistics, *Handbook*, and Troy, *Distribution of union membership* and 'Trade union membership'. Henry Farber, Dan Herbst, Ilyana Kuziemko, and Suresh Naidu are compiling union membership data from mid-century Gallup Polls. When complete, this will provide new estimates of state-level union density and can be used to study household income inequality at that level.
[35] Troy, *Distribution of union membership*, p. 1.
[36] Troy, *Distribution of union membership*, p. 28.
[37] Troy, *Distribution of union membership*.
[38] Troy's papers were not deposited in the Rutgers library system to the knowledge of his 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,' *Distribution of union membership*, p. 3.

estate), public services, and mining (which was highly organized before 1939), had the smallest changes in union density.

We calculate a local exposure-to-unionization variable (ΔU_i) that is a weighted average of national industry-level changes in unionization as a percent of industry employment (ΔU_j), where the weights (ω_{ij}) correspond to the *local* mix of employment in 1940 as observed in the complete-count census microdata.[39]

$$\Delta U_i = \sum_{j=1}^{N} \Delta U_j \times \omega_{ij}$$

ΔU_i is similar to a Bartik-style instrumental variable.[40] Such instruments are often used to avoid endogeneity bias, such as (in this setting) the potential response of local unionization to local inequality or other omitted variables that could be correlated with inequality trends. The local weights are fixed at 1940 levels to avoid endogenous changes in the local employment mix after 1940. The national-level changes in industry unionization are assumed to be exogenous to local circumstances. It is in this sense that we refer to the variable as a measure of 'exposure' rather than a direct measure of unionization.

We cannot estimate a 'first-stage' SEA-level regression of actual change-in-unionization on the variable ΔU_i because no direct measures of local unionization exist. But to verify that there is an empirical connection, we can aggregate ΔU_i to the state level for comparison with separate state-level estimates of unionization from Troy.[41] The simple unweighted correlation across states (and the District of Columbia) is 0.75 with a *p*-value of <0.001. A bivariate regression of the change in Troy's state-level estimates of union density on the ΔU_i proxy (aggregated to the state level) yields a coefficient of 1.65 (s.e.=0.211). Thus, as far as can be told, there is a strong empirical relationship between ΔU_i and actual changes in unionization. That said, our interpretation of ΔU_i is in terms of its reflection of local exposure to unionization, as opposed to a direct measure of changes in local unionization rates per se. It is notable that state-level aggregates of ΔU_i tend to be larger in the South than what Troy estimates directly for southern states. As discussed below, all subsequent analyses will include regional or state fixed effects and, therefore, will base estimates on within-region or within-state variation.

---

39 The local weights are based on the sample that enters into the wage inequality calculation. An alternative would be to allow the weights to change after 1940, but this would bring endogenous changes in employment into the key independent variable.

40 Bartik, *Who benefits*.

41 Troy, *Distribution of union membership*.

Figure 4 maps the SEA-level variation in ΔU_i. Since unionization rarely fell within industries over this period, ΔU_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 ΔU_i. Overall, the average ΔU_i is 8.2 p.p. with a standard deviation of 3.1 p.p. For the SEA at the 75th percentile, ΔU_i is almost double that for the SEA at the 25th percentile, so there is considerable cross-place variation.

III. Empirical strategy

Our empirical strategy exploits the local variation shown in Figures 3 and 4. To fix ideas, consider a simple regression of the following form:

$$\Delta I_{ir} = \alpha + \tau \Delta U_{ir} + X'_{ir}\beta + \gamma_r + e_{ir}\ , \qquad (1)$$

where ΔI_{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); τ is the coefficient of main interest associated with exposure to changes in union density in locality *i* (ΔU_{ir}); X_{ir} is a vector of local characteristics (detailed below) that serve as control variables; and γ_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 (or state) trends.[42] Identification of τ then comes from within-region (or state) variation in ΔU_{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; baseline results are similar but slightly smaller if unweighted, as shown below.

In such a framework, the clear threats to a causal interpretation of τ are the potential endogeneity of unions with respect to local concerns about inequality and, broadly, omitted variables that are correlated with ΔU_{ir} and affect local inequality through other channels. As described above, the variable ΔU_{ir} combines national-level changes in union membership rates with pre-existing local employment shares by industry. Because ΔU_{ir} is driven by changes in industry-level unionization at the national level, it is insulated from endogenous local changes in unionization. However, omitted

[42] Baseline regressions reported in Table 2 use the four large census-defined regions for fixed effects. Estimates of τ are similar when using nine smaller census divisions. They are also similar if Troy's estimates of state-level changes in union density are added as a control variable to the baseline specification. And they are similar when regressions include the interaction of ΔU_{ir} and the South's regional dummy. In those regressions, the coefficient on the interaction tends to be negative but statistically insignificant, but for the Gini regressions, the coefficient on the interaction is significant at the 10 percent level and the estimate of τ declines somewhat (e.g., from -0.0041 in the 1940-50 baseline to -0.0033).

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_{ir} 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 ΔU_{ir}. This motivates the control variable for the value of war-related production and facilities contracts. 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 employment that was 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.[43] Following Goldin and Margo, 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

---

[43] $D_{s,t} = \frac{D_{(skilled)st}}{D_{(unskilled)st}} = \frac{\sum_j \left(\frac{E_{jt}}{E_t}\right)\left(\frac{E_{(skilled)js40}}{E_{s40}}\right)}{\sum_j \left(\frac{E_{jt}}{E_t}\right)\left(\frac{E_{(unskilled)js40}}{E_{s40}}\right)}$, where *s* denotes SEA, *t* denotes year, and *j* denotes industry. The first term, $\left(\frac{E_{jt}}{E_t}\right)$, captures the employment in each industry as a share of national employment in year *t*. All time variation in the index comes from this term. The second term, $\left(\frac{E_{(skilled)js40}}{E_{s40}}\right)$ and $\left(\frac{E_{(unskilled)js40}}{E_{s40}}\right)$, is fixed at 1940 base levels and captures the SEA-specific skill mix within each industry. Regressions control for the decadal difference in the index (i.e., $D_{s,1950} - D_{s,1940}$). We treat high school graduates and above as the skilled group. In 1940, approximately 30 percent of the labor force had completed high school, and so it seems appropriate to consider them a relatively skilled group.

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.[44] In addition, we find that moving from within-region to within-state comparisons as the basis for identifying τ (i.e., adding state fixed effects), which substantially narrows the geographic scope for confounding shocks and trends, has only a small impact on the point estimates of τ.

Ultimately, to sustain a causal interpretation of τ, one would have to assume that conditional on X_i and γ_r there were no unobserved shocks or trends in local inequality that were correlated with ΔU_i. Such unobserved confounders cannot be ruled out in this setting, but to bias estimates of τ 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.[45] 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 resulting in downward pressure on the skill premium in those locations; 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, and in general causal interpretations should be qualified accordingly; we return to this issue below when assessing mechanisms.

[44] 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.
[45] 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.

IV. Mid-century results

Estimates of τ 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 supplemental appendix Table A2.

The first row of Panel A shows that ΔU_{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 τ 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.[46] The baseline estimates of τ are sizable: a one standard deviation increase in ΔU_{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.

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 ΔU_{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 results: mobilization of men into the armed

[46] Supplemental appendix Figure A1 plots the coefficients and confidence intervals from regressions that estimate τ for each decile separately for 1940-50. This provides a finer decomposition of the overall change in 90-10 inequality. Roughly 75 percent of the union exposure coefficient on the 90-10 differential comes from compression between the 20th to 40th percentiles; another 20 percent appears between the 60th and 70th.

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 find that mobilization-induced entry by women tended to reduce wage inequality among men.[47] Therefore, we add a control variable for each state's mobilization rate.[48] 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.[49] 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 or 1960, depending on the timeframe examined.[50] The results are reported in the second row of Table 2. The strong negative estimates of τ are robust to the additional 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 basing identification on within-state variation rather than within-region. The estimates of τ are slightly smaller than in the base specification (row 1), but they remain strongly negative and statistically significant, 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 difference, 50-10 difference, and so on) that is based on the distribution of industry-by-

---

[47] Acemoglu, Autor, and Lyle, 'Women, war, and wages.'

[48] Acemoglu, Autor, and Lyle, 'Women, war, and wages.' They 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.

[49] State minimum wage laws in this period tended to focus on women and children, but some states did extend coverage to men who were not covered by the federal law. Including a dummy variable equal to one for those states and the interaction of the dummy variable with the share of workers below the federal minimum in 1939 does not alter key results in the second row of Table 2. See Women's Bureau, 'State minimum-wage laws'.

[50] Ellwood and Fine, 'Impact of right to work laws.'

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.[51] Otherwise, the median income for the detailed occupation is used (without the industry interaction). This allows us to gauge inequality pre-trends to the extent that the 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 based on wages in highly detailed job categories before the Second World War. The fourth row shows that regressions augmented with the pre-trend control variable yield results that are similar to the baseline estimates.[52]

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; some of these cities are not separately identified in the 1940 census.[53] 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. Using the implied 1919-1939 change in inequality as the dependent variable in the baseline regression framework reveals no significant conditional correlation with the union exposure variable. In other words, exposure to unionization after 1939 does not predict the pre-trend in wage inequality. These results are in the supplemental appendix Table A3, row 7.

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.[54] Again, the results are fairly robust.

Finally, Row 6 reports baseline results from unweighted regressions. In some cases, the point estimates of τ are smaller when unweighted, but the basic pattern of results remains. This may reduce concern about model misspecification from omitted heterogeneous effects.

---

[51] 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.

[52] 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.

[53] U.S. Department of Labor, 'Cost of living' (ICPSR Study 8299), Feigenbaum, 'Intergenerational mobility'.

[54] A related issue concerns agricultural laborers, who are in the sample if they meet the criteria set in Goldin and Margo, 'Great compression'. 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.

Supplemental appendix Tables A3 and A4 assess the sensitivity of estimates to a number of sample restriction and modeling choices. Estimates are not very sensitive to using census division fixed effects rather than census region, using weeks worked in the previous year expressed in intervals for all years (even when more precise information is available), or including observations with computed weekly wages that are less than half the minimum wage instead of dropping them (bottom coding them at half the minimum wage). Changing the sample restriction on the minimum number of weeks worked from 40 to 26 or 1 tends to lower the estimate of τ for the 90-10 distribution (from -0.0232 to -0.0176 for the 1940-50 baseline specification) and tends to shift the observed compression to the 90-50 portion of the distribution, though the estimates are still sizable and negative. Estimates of τ for Gini and variance measures of inequality change little.

The supplemental appendix also examines sensitivity to adding a control variable for the change in the black share of local population from 1940-50 or 1940-60 (table A5). The idea is that black migration was high during the 1940s and 1950s, which could influence observed levels of local inequality and might be correlated with local exposure to unionization. Estimates of τ change little, however. The same is true when we add a control variable for the share of the foreign-born population in 1940. Here, the idea is that if the foreign born assimilate rapidly in terms of wages and happen to be located in places with high exposure to unions, their relative gains might confound the estimate of τ, but this does not appear to be the case.

V. Exploring mechanisms of wage compression

Even if one were to accept 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 it could come through various channels with quite different 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.[55] To our knowledge, no nationally

[55] Rosen, 'Wages-based indexes'; Roback, 'Wages, rents, and quality of life'; Roy, 'Some thoughts'; Callaway and Collins, 'Unions, workers, wages', find relatively low returns to education for union workers relative to

representative micro-level data are available that can reveal skill-specific gross 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 ΔU_{ir} and ΔI_{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 ΔU_{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 likely 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 τ are sensitive to their inclusion. Because the skill mix and population growth may be endogenous to unionization, we would not consider estimates of τ 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 τ change little relative to those reported in Table 2, consistent with the effects of unions working primarily through channels other than sample composition and firm relocation.

Finally, we examine changes in the distribution of the residuals from wage equations that control for education, experience (age minus years of schooling minus six), their squares, and their interactions. 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, where unions reduce overall inequality by standardizing wages across workers with similar observables.[56] Indeed, the evidence is consistent with this hypothesis. The details of this analysis are reported in the

non-union workers in a small sample of urban workers in 1950. In a different setting, see Abramitzky, 'Effect of redistribution' for evidence of highly skilled workers sorting out of Israeli kibbutzim.

[56] See Card, Lemieux, and Riddell, 'Unions'.

supplemental appendix.

Our conclusion is that the mid-century conditional correlation between local union exposure and changing inequality is fairly robust, plausibly 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 and also perhaps through threat effects on non-union firms. By extension, it is possible that union threat effects mattered even where unionization did not occur widely. Although more difficult to pin down, it is also possible that unions influenced local norms about wage setting.[57] 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.

VI. 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,[58] 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.

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 twentieth century (1940-60, 1940-70, 1940-80, etc.). In essence, this simply extends our basic empirical approach to 2010; 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 τ is still the coefficient on ΔU_{ir} and where the dependent variable is the change in inequality measured over varying periods of length.

---

[57] Western and Rosenfeld, 'Unions, norms'.
[58] Reder, 'Rise and fall of unions'; Freeman, 'Contraction and expansion'.

To be clear, ΔU_{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 burst of union growth in the United States. Thus, estimates of τ correspond to the long-run association between union exposure 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 speculative, these correlations may reflect persistent variation in union density and influence across space. 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.[59] 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 reinforced wage patterns established in the 1940s and 1950s. Space does not allow for further exploration here. But we suggest that more research in this area might provide useful perspective on the legacy of US unions and the historical roots of cross-place variation in inequality.

VII. 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 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; instead, 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-

[59] The correlation between 1964 and 2000's state-level density is 0.82 according to data from Hirsch, Macpherson, and Vroman, 'Estimates of union density.' 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, 'Impact of forced migration'.

trends in local inequality, but 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 production contracts and local industrial composition 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 τ for the 1939-49 period at face value, moving from the 25^{th} to 75^{th} 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 large change compared to the median local 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 also could 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 finds no evidence that such endogenous sorting underpins the baseline results, but more research is warranted. 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 merit 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 local 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. If the rise of unions and compression of local wages eventually led to lower employment levels (e.g., by raising the cost of less-skilled labor), the medium- to long-run welfare implications would be complex. 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 the economic history of inequality and labor market institutions.

References

Abramitzky, R., 'The effect of redistribution on migration: evidence from the Israeli kibbutz.' *Journal of Public Economics* 93 (2009), pp. 498-511.

Acemoglu, D. and D.H. Autor. 'Skills, tasks and technologies: implications for employment and earnings', in O. Ashenfelter and D. Card, eds., *Handbook of labor economics*, Vol. 4, (Amsterdam: Elsevier-North Holland, 2011), pp. 1043-1171.

Acemoglu, D., D.H. Autor, and D. Lyle, 'Women, war, and wages: the effect of female labor supply on the wage structure at midcentury.' *Journal of Political Economy* 112 (2004), pp. 497-551.

Acemoglu, D. and P. Restrepo, 'Robots and jobs: evidence from US labor markets.' NBER Working Paper 23285 (Cambridge, MA, 2017).

Ashenfelter, O. and J.H. Pencavel, 'American trade union growth: 1900-1960.' *Quarterly Journal of Economics* 83 (1969), pp. 434-448.

Atkinson, A.B. and B. Nolan, 'The changing distribution of earnings in Ireland, 1937 to 1968.' *Economic History Review* 63 (2010), pp. 479-499.

Autor, D.H. and D. Dorn, 'The growth of low-skill service jobs and the polarization of the US labor market,' *American Economic Review* 103 (2013), pp. 1553-97.

Autor, D.H., D. Dorn, and G.H. Hanson, 'The China syndrome: local labor market effects of import competition in the United States.' *American Economic Review* 103 (2013), pp. 2121-2168.

Autor, D.H., L. Katz, and M. Kearney, 'Trends in U.S. wage inequality: revising the revisionists.' *Review of Economics and Statistics* 90 (2008), pp. 300-323.

Bartik, T.J., *Who benefits from state and local economic development policies?* (Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 1991).

Bogue, D.J. *State economic areas: a description of the procedure used in making a functional grouping of counties in the United States*. (Department of Commerce, Bureau of the Census. Washington, DC: Government Printing Office, 1951).

Bureau of Labor Statistics, United States Department of Labor. *Handbook of Labor Statistics, 1950 edition*. Bulletin No. 1016. (Washington, DC: GPO, 1951).

Callaway, B. and W.J. Collins. 'Unions, workers, and wages at the peak of the American labor movement.' NBER Working Paper 23516 (Cambridge, MA, 2017).

Card, D., T. Lemieux, and W.C. Riddell. 'Unions and wage inequality.' *Journal of Labor Research* 25 (2004), pp. 519-559.

Carter, S. et al. *Historical statistics of the United States, millennial edition*. (New York: Cambridge University Press, 2006).

Cohen, L., C. Malloy, and Q. Nguyen. 'The impact of forced migration on modern cities: evidence from the droughts of the 1930s.' (Cambridge, MA: Harvard Business School, working paper, 2016).

DiNardo, J., N.M. Fortin, and T. Lemieux. 'Labor market institutions and the distribution of wages, 1973-1992: A Semiparametric Approach.' *Econometrica* 64 (1996), pp. 1001-1044.

Dunlop, J.T. 'An appraisal of wage stabilization polices.' In *Problems and policies of dispute settlement and wage stabilization during World War II*, 155-86. U.S. Department of Labor Bulletin No. 1009. (Washington, DC: Government Printing Office, 1950).

Ellwood, D.T. and G.A. Fine. 'The impact of right to work laws on union organizing.' *Journal of Political Economy* 95 (1987), pp. 250-273.

Farber, H.S. 'Nonunion wage rates and the threat of unionization.' *Industrial and Labor Relations Review* 58 (2005), pp. 335-352.

Feigenbaum, J.J. 'Intergenerational mobility during the Great Depression.' Unpublished working paper (2015).

Freeman, J., 'Delivering the goods: industrial unionism during World War II.' *Labor History* 19 (1978): 570-593.

Freeman, R.B., 'Unionism and the dispersion of wages.' *Industrial and Labor Relations Review* 34, (1980), pp. 3-23.

Freeman, R.B., 'Contraction and expansion: the divergence of private sector and public sector unionism in the United States.' *Journal of Economic Perspectives* 2 (1988): 63-88.

Freeman, R.B. 'Spurts in union growth: defining moments and social processes.' In M.D. Bordo, C. Goldin, and E.N. White (eds.), *The defining moment: the Great Depression and the American economy in the twentieth century*, pp. 265-295. (Chicago, IL: University of Chicago Press, 1998).

Freeman, R.B. and J.L. Medoff. *What do unions do?* (New York: Basic Books, 1984).

Friedman, M. *Capitalism and freedom*. (Chicago, IL: University of Chicago Press, 1962).

Frydman, C. and R. Molloy, 'Pay cuts for the boss: executive compensation in the 1940s.' *Journal of Economic History* 72 (2012), pp. 225-251.

Gazeley, I., 'The leveling of pay in Britain during the Second World War.' *European Review of Economic History* 10 (2006), pp. 175-204.

Goldin, C. and L.F. Katz. *The race between education and technology*. (Cambridge, MA: Harvard University Press, 2008).

Goldin, C. and R.A. Margo, 'The Great Compression: the wage structure in the United States at mid-century.' *Quarterly Journal of Economics* 107 (1992), pp. 1-34.

Goldsmith, S., G. Jaszi, H. Kaitz, and M. Liebenberg, 'Size distribution of income since the mid-1930s.' *Review of Economics and Statistics* 36 (1954), pp. 1-32.

Haines, M.R., and Inter-university Consortium for Political and Social Research. *Historical, demographic, economic, and social data: The United States, 1790-2002* [computer file]. ICPSR02896-v3. (Ann Arbor, MI: ICPSR 2010).

Hirsch, B.T., D.A. Macpherson, and W.G. Vroman, 'Estimates of union density by state.' *Monthly Labor Review* 124, 7 (2001): pp. 51-55.

Jaworski, T. and G.T. Niemesh. 2018. 'Revisiting the Great Compression: wage inequality in the United States, 1940-1960.' *Historical Methods: A Journal of Quantitative and Interdisciplinary History* (forthcoming).

Juhn, C. `Wage inequality and demand for skill: lessons from five decades.' *Industrial and Labor Relations Review* 17, 1 (1999), pp. 424-443.

Katz, L.F. and K.M. Murphy. 'Changes in relative wages, 1963-1987: supply and demand factors.'

Quarterly Journal of Economics, 107, 1 (1992), pp. 35-78.

Kuznets, S. and E. Jenks. *Shares of upper income groups in income and savings*. (New York: NBER, 1953).

Lester, R.A., *Economics of labor*. (New York: The MacMillan Company, 1964).

Levy, F. and P. Temin. 'Inequality and institutions in 20th century America.' NBER Working Paper 13106 (Cambridge, MA 2007).

Lewis, H.G., *Unionism and relative wages in the United States: an empirical inquiry*. (Chicago, IL: University of Chicago Press, 1963.)

Miller, H.P., 'Changes in the industrial distribution of wages in the United States, 1939-1949.' In *An appraisal of the 1950 census income data*. National Bureau of Economic Research, Studies in Income and Wealth, Volume 23. (Princeton, NJ: Princeton University Press, 1958).

Millis, H.A. and E.C. Brown. *From the Wagner Act to Taft-Hartley: a study of national labor policy and labor relations*. (Chicago: University of Chicago Press, 1950).

Piketty, T. and E. Saez. 'Income inequality in the United States, 1913-1998.' *Quarterly Journal of Economics* CXVIII (2003), pp. 1-39.

Reder, M., 'The rise and fall of unions: the public sector and the private.' *Journal of Economic Perspectives* 2 (1988), pp. 89-110.

Rees, A. *The economics of trade unions*. (Chicago: University of Chicago Press, 1962).

Roback, J., 'Wages, rents, and the quality of life.' *Journal of Political Economy* 90 (1980), pp. 1257-1278.

Rosen, S., 'Wages-based indexes of urban quality of life.' In P. Mieszkowski and M. Straszheim (eds.) *Current issues in urban economics*. (Baltimore: Johns Hopkins University Press, 1979).

Roy, A.D. 'Some thoughts on the distribution of earnings.' *Oxford Economic Papers* 3 (1951), pp. 135-46.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek. *Integrated public use microdata series: version 6.0* [Machine-readable database]. (Minneapolis: University of Minnesota, 2015).

Seidman, J., *American labor from defense to reconversion*. (Chicago, IL: University of Chicago Press, 1953).

Troy, L., *The distribution of union membership among the states, 1939 and 1953*. (New York: National Bureau of Economic Research, 1957).

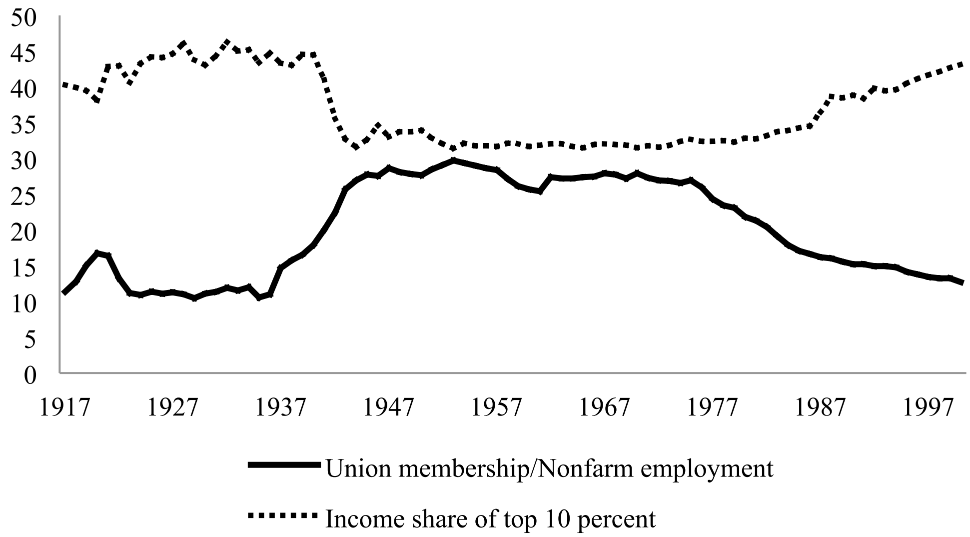
Troy, L. 'Trade union membership, 1897-1962.' NBER Occasional Paper 92 (1965).

United States Department of Labor, Bureau of Labor Statistics, 'Cost of living in the United States, 1917-1919' ICPSR08299-v5, (1986), Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

Wachter, M.L. 'The striking success of the National Labor Relations Act.' University of Pennsylvania, *Faculty Scholarship*, paper 493 (2012). http://scholarship.law.upenn.edu/faculty_scholarship/493

Western, B. and J. Rosenfeld, 'Unions, norms, and the rise in U.S. wage inequality.' *American Sociological Review* 76 (2011), pp. 513-537.

Figure 1: Unions and income inequality trends in the 20th-century United States

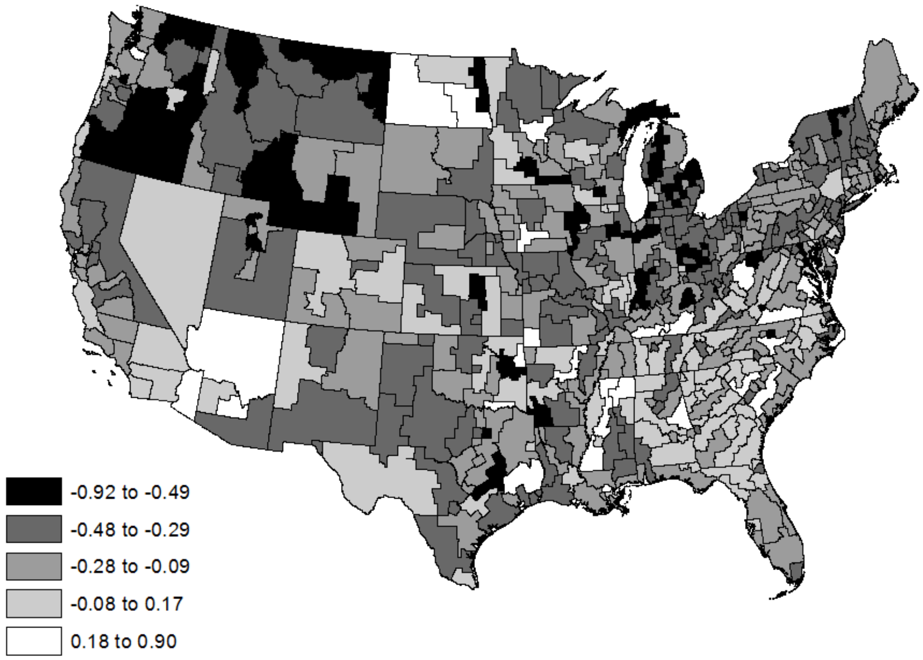


Sources: The series for income share of the top 10 percent of tax units is from Piketty and Saez, 'Income inequality' (updated to 2014, Table A1, income shares excluding capital gains, accessed August 3, 2015); 1917 is the first year available in the income inequality series. The union membership data up to 1983 are the Troy-Sheflin series, minus Canadian membership of U.S. unions, as reported in Carter et al., *Historical statistics* (series Ba4785 and Ba4786). The civilian nonfarm employment data up to 1983 are also from Carter et al., *Historical statistics* (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

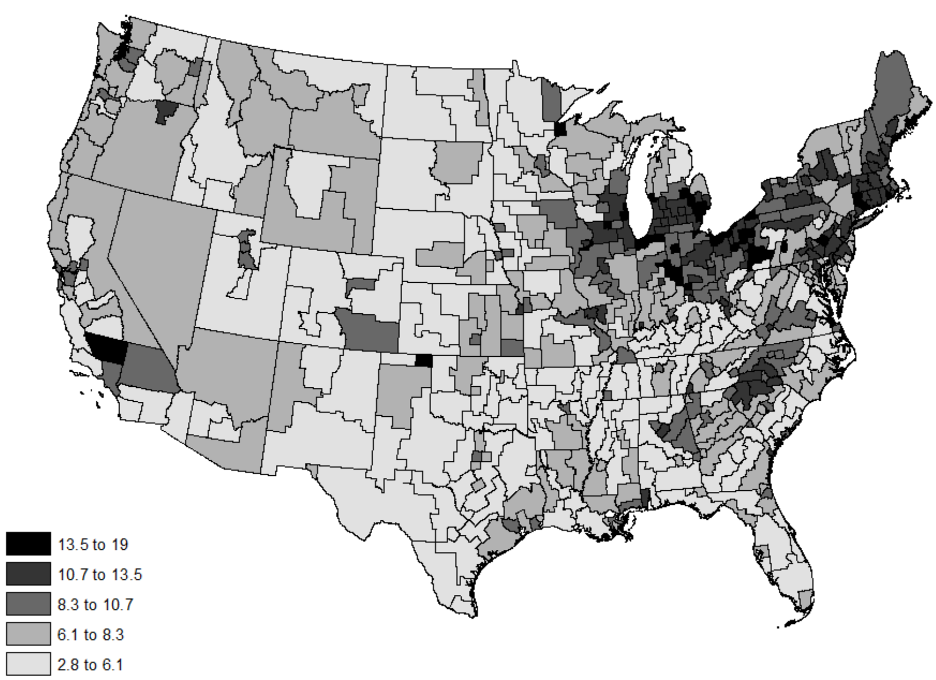
Notes: The earnings sample includes men reporting wage and salary employment in the reference week, aged 18 to 64 years who reported positive wage and salary income in the year prior to the census, worked more than 39 weeks, and earned more than half the minimum wage at a full-time basis (weekly wages of \$6 in 1940 and \$8 in 1950, \$26 in 1970, and \$27 in 1990).
Sources: Earnings data from decennial census microdata provided by Ruggles et al., *Integrated public use microdata series*, 1940 complete count (100 percent), 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: Ruggles et al., *Integrated public use microdata series*, 1940complete count and 1950 1 percent 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 ΔU_i variable for state economics areas (SEAs).

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

Panel A: Inequality and union exposure	Mean	St. Dev	25th pctile	Median	75th pctile
1940-1950					
Δ 90-10 wage differential	-0.224	0.229	-0.372	-0.248	-0.107
Δ 90-50 wage differential	-0.188	0.129	-0.266	-0.182	-0.105
Δ 50-10 wage differential	-0.036	0.234	-0.194	-0.073	0.089
Δ Variance of log wages	-0.048	0.075	-0.094	-0.055	-0.010
Δ Gini	-0.057	0.036	-0.080	-0.057	-0.033
1940-1960					
Δ 90-10 wage differential	-0.249	0.153	-0.356	-0.244	-0.158
Δ 90-50 wage differential	-0.199	0.104	-0.273	-0.199	-0.129
Δ 50-10 wage differential	-0.050	0.160	-0.160	-0.069	0.058
Δ Variance of log wages	-0.069	0.047	-0.103	-0.068	-0.038
Δ Gini	-0.052	0.026	-0.071	-0.053	-0.035
1939-1953					
ΔU_{ir}	8.2	3.1	5.7	7.5	10.1

Panel B: Control variables	Mean	St. Dev	25th pctile	Median	75th pctile
Median log weekly wage 1939	3.006	0.248	2.795	3.033	3.219
Initial log 90-10 gap in 1939	1.418	0.179	1.287	1.435	1.566
Per capita WWII contracts (1940$)	1.172	1.805	0.192	0.467	1.347
Male employment rate in 1940	0.674	0.054	0.642	0.679	0.707
Percent urban 1940	0.417	0.245	0.214	0.364	0.622
Percent males HS grads	0.248	0.054	0.208	0.249	0.285
Local demand shock index	1.896	0.262	1.725	1.865	2.051
Local skill-specific demand shock index	-0.002	0.019	-0.013	-0.004	0.007
State WWII mobilization rate	0.469	0.032	0.449	0.468	0.490
Percent of males earning below minimum wage in 1939	0.313	0.178	0.161	0.279	0.453
State passed right to work law by 1950	0.259	0.439	0	0	1

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 as described in the text and provided by Ruggles et al., *Integrated public use microdata series*. Exposure to unions (ΔU_{ir}) is calculated by interacting industrial employment data from the 1940 full count census and industry-level changes in union density from Troy, *Distribution of union membership.* The employment rate, share of population in urban areas, and per capita war expenditure variables are calculated with data from Haines and ICPSR, *Historical, demographic.* War mobilization is from Acemoglu, Autor, and Lyle, 'Women, war, and wages.' Right to work laws are from Ellwood and Fine, 'Impact of right to work laws.' More detailed information is provided in the supplemental appendix.

Table 2: Regressions results, changes in wage inequality and unionization

	(1) Δ90-10	(2) Δ90-50	(3) Δ50-10	(4) ΔGini	(5) ΔVar
Panel A: 1940-50					
(1) Base specification	-0.0232	-0.00303	-0.0202	-0.0041	-0.0066
	(0.00618)	(0.00326)	(0.00434)	(0.0011)	(0.0019)
(2) Base with controls for mobilization and policy	-0.0191	0.00017	-0.0193	-0.0031	-0.0049
	(0.00627)	(0.00310)	(0.00439)	(0.0010)	(0.0016)
(3) Base with state f.e.	-0.0201	-0.00411	-0.0160	-0.0035	-0.0057
	(0.00650)	(0.00329)	(0.00506)	(0.0010)	(0.0019)
(4) Base w/ pre-trend control	-0.0232	-0.0018	-0.0195	-0.0037	-0.0062
	(0.0062)	(0.0032)	(0.0046)	(0.0011)	(0.0019)
(5) Base, restrict to urban SEAs (>=50%)	-0.0208	0.00027	-0.0210	-0.0030	-0.0034
	(0.0103)	(0.0056)	(0.0070)	(0.0018)	(0.0028)
(6) Base, unweighted	-0.0199	-0.00319	-0.0167	-0.0040	-0.0074
	(0.00520)	(0.00254)	(0.00458)	(0.0009)	(0.0019)
Panel B: 1940-60					
(1) Base specification	-0.0215	-0.00477	-0.0166	-0.0032	-0.0057
	(0.00474)	(0.00248)	(0.00461)	(0.00081)	(0.00148)
(2) Base with controls for mobilization and policy	-0.0212	-0.00214	-0.0190	-0.00270	-0.00563
	(0.00516)	(0.00261)	(0.00486)	(0.00083)	(0.00160)
(3) Base with state f.e.	-0.0260	-0.00681	-0.0191	-0.00370	-0.00695
	(0.00489)	(0.00266)	(0.00487)	(0.00081)	(0.00160)
(4) Base w/ pre-trend control	-0.0215	-0.0041	-0.0154	-0.0030	-0.0057
	(0.0046)	(0.0025)	(0.0046)	(0.00085)	(0.0015)
(5) Base, restrict to urban SEAs (>=50%)	-0.0166	-0.0009	-0.0157	-0.0029	-0.0035
	(0.0087)	(0.0035)	(0.0080)	(0.0013)	(0.0024)
(6) Base, unweighted	-0.0181	-0.00543	-0.0126	-0.0025	-0.0053
	(0.00527)	(0.00230)	(0.00539)	(0.0007)	(0.0015)

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Baseline regressions: Adding Δ share HS+ and Δ ln(pop) as controls				
Dep. variable	Δ share HS+	Δ ln(pop)	Δ90-10	Δ90-50	Δ50-10	ΔGini	ΔVar
1940-50	-0.0029	0.0072	-0.0228	-0.0024	-0.0204	-0.0040	-0.0061
	(0.0017)	(0.0040)	(0.0063)	(0.0032)	(0.0044)	(0.0011)	(0.0019)
1940-60	0.0020	0.0267	-0.0229	-0.0042	-0.0187	-0.0032	-0.0061
	(0.0017)	(0.0084)	(0.0048)	(0.0026)	(0.0045)	(0.0008)	(0.0015)

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

	(1) 1940-50	(2) 1940-60	(3) 1940-70	(4) 1940-80	(5) 1940-90	(6) 1940-2000	(7) 1940-2010
Long differences from 1940							
90-10	-0.0232	-0.0215	-0.0255	-0.0167	-0.0121	-0.0186	-0.0137
	(0.0062)	(0.0047)	(0.0037)	(0.0030)	(0.0051)	(0.0062)	(0.0064)
90-50	-0.0030	-0.0048	-0.0046	-0.0040	-0.0060	-0.0040	-0.0066
	(0.0033)	(0.0025)	(0.0027)	(0.0026)	(0.0024)	(0.0027)	(0.0033)
50-10	-0.0202	-0.0166	-0.0209	-0.0127	-0.0061	-0.0146	-0.0071
	(0.0043)	(0.0046)	(0.0037)	(0.0035)	(0.0047)	(0.0058)	(0.0064)
Gini	-0.0041	-0.0032	-0.0030	-0.0026	-0.0018	-0.0033	-0.0027
	(0.0011)	(0.0008)	(0.0008)	(0.0008)	(0.0011)	(0.0012)	(0.0015)
Var	-0.0066	-0.0057	-0.0063	-0.0046	-0.0036	-0.0077	-0.0070
	(0.0019)	(0.0015)	(0.0014)	(0.0012)	(0.0021)	(0.0028)	(0.0031)

Sources: Income inequality is calculated from the census microdata, Ruggles et al., *Integrated public use microdata series.* The ΔU_i variable is described in the text; it combines information from Troy, *Distribution of union membership,* and the 1940 complete count U.S. census microdata for 1940.

Unions and the Great Compression of Wage Inequality in the United States at Mid-Century

Collins and Niemesh

Supplemental Appendix

A**1. Inequality Measures**

We make extensive use of decennial census microdata provided by IPUMS: the 1940 full count, the 5 percent extract for 1960, and the 1 percent extracts for 1950, and 1970-2000.[60] Sample restrictions follow those in Goldin and Margo's study of the Great Compression: men aged 18-64, reporting at least 40 weeks of work as primarily a wage and salary worker, and an annual income at least half the minimum wage, on average, earned on a full-time basis for the reported number of weeks worked.[61] In practice, this consists of limiting the sample to observations with calculated weekly earnings above the following yearly cutoffs:

	1939	1949	1959	1969	1979	1989	1999
Min wage per hour ($)	0.30	0.40	1.00	1.30	3.10	3.35	5.15
Per week at (1/2) full-time basis (40 hrs.) ($)	6	8	20	26	62	67	103

Weeks worked are reported in intervals in the 1960 and 1970 censuses. We impute weeks as the middle point of each interval in these years. Top-coded annual incomes are multiplied by 1.4 for 1940-1980. Top-coded values for 1990 and 2000 are set as the median or mean value of top-coded values in the respondent's state. 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.

	1940	1950	1960	1970	1980	1990	2000
Top-code cutoff for annual income ($1,000)	5	10	25	50	75	140	175
Share of sample topcoded (%)	1.1	1.2	0.4	0.2	0.5	0.7	1.3

Observations are weighted by person weights (sample line weight for 1950). The sample includes the continental United States. In each year and location, we calculate the variance, gini, and 90-10, 90-50, and 50-10 percentile difference in log weekly wages.

[60] Ruggles et al., *Integrated public use microdata series*.
[61] Goldin and Margo, 'Great compression'

A2. Geographic Identifiers

Recent literature on the post-1970s increase in inequality uses commuting zones or metropolitan statistical areas as the relevant geographic unit.[62] We use a definition of local labor market that captures the geographic distribution of economic activity at mid-century, the time-period under consideration. While MSAs were defined for the 1940 Census, only a small number existed relative to today (140 in 1940 and 382 in 2010). Moreover, the changes in earnings inequality outside of metropolitan areas play an important role in the Great Compression that we would like to capture.

For these reasons, we use the State Economic Area (SEA) as our definition of a local labor market. The Census Bureau created SEAs to capture and tabulate data for single counties or contiguous groups of counties that were economically similar, and thus provide a similar notion to a local labor market, although different from Commuting Zones. In addition to economic characteristics, the Census Bureau used social, industrial, commercial, demographic, climatic, physiographic, and cultural factors into account when delaminating SEAs. SEAs are "*relatively homogenous subdivisions of States. They consist of similar economic and social characteristics. The boundaries of these areas have been drawn in such a way that each State is subdivided into a few parts, with each part having certain significant characteristics which distinguish it from the other areas which it adjoins*".[63]

The 1940 full count dataset contains county identifiers, which are easily aggregated into SEAs. The 1950 1 percent sample provides SEA identifiers as the smallest geographical unit. After 1950, neither SEA nor county identifiers are provided in the microdata samples. The smallest geographic unit in the 1960 sample is the mini public-use microdata area (mini-PUMA), created by combining census tracts and untracted counties to form areas of at least 50,000 persons or more.[64] The smallest geographic identifier in 1970 and 1980 is the county group (CNTYGP97 and CNTYGP98), each designed to contain at least 250,000 and 100,000 persons, respectively. The 1990 and 2000 samples use public-use microdata areas (PUMAs), again designed to contain at least 100,000 persons.

We construct SEA level indicators for 1960 and later samples following a procedure developed in Autor and Dorn by probabilistically matching county-groups and PUMAS in the census

[62] Autor and Dorn, 'Growth of low-skill service'; Autor, Dorn, and Hanson, 'The China syndrome'; Moretti, 'Local labor markets'; Baum-Snow and Pavan, 'Inequality and city size'.
[63] Bogue, *State economic areas*
[64] See https://usa.ipums.org/usa/volii/1960geotools.shtml for a detailed discussion of how mini-PUMAs were constructed by IPUMS.

public use files to SEAs.[65] The following description is heavily paraphrased from Dorn.[66] Each MINIPUMA, PUMA, or CG j = 1,...,J is related to every SEA s = 1,...,467 by computing the probability that a resident of j lives in s using the following equation,

$$\alpha_{js} = \sum_{c=1}^{C} \frac{r_{jc}}{r_j} \frac{r_{cs}}{r_c}$$

where r_j is the number of residents in MINIPUMA j, r_c is the number of residents in county c, r_{jc} is the number of residents in the overlap between MINIPUMA j and county c, and r_{cs} is the number of residents in the overlap between county c and SEA s. The second term is either zero or one as each county is fully contained in a single SEA.

The overlap of MINIPUMA j (or PUMA or CG) and county c (r_{jc}) is unobservable in the microdata. For 1960, IPUMS provides a crosswalk of MINIPUMA population in each county, which can be used in conjunction with total population counts for counties from Haines to compute r_{jc}.[67] The alpha factor weight would be equal to 1 for MINIPUMAs contained within a single SEA, even if spread over multiple counties. The alpha factor weight on a single census observation will be strictly between 0 and 1 when a MINIPUMA spans multiple SEAs. For example, suppose 30 percent of MINIPUMA j's population resides in SEA 1, and 70 percent resides in SEA 2. Then each observation in the MINIPUMA j will have two copies in the dataset, one in SEA 1 a weighting factor of $\alpha_{j1} = 0.30$, and the other with $\alpha_{j2} = 0.70$. A similar procedure is completed for 1970 through 2000 where the MINIPUMA is replaced by the CG or PUMA.

A3. Union Density Rates

Troy provides industry and state-level union density measures for 1939 and 1953.[68] We construct local measures of industry structure using the full-count 1940 census microdata provided by IPUMS. The local change-in-union exposure variable for SEA s (ΔU_s), is a weighted average of national industry-level changes in unionization as a percent of employment for each industry j (ΔU_j), where the weights (ω_{js}) correspond to the local mix of employment in 1940 measured as the employment in

[65] Autor and Dorn, 'Growth of low-skill service'
[66] Dorn, 'Essays on Inequality'
[67] Haines, *Historical, Demographic*
[68] Troy, *Distribution of union membership*

industry j in SEA s divided by total employment in SEA s $\omega_{js} = \frac{E_{js}}{E_s}$, where employment is based on the same sample restrictions as used to make the inequality measures.

$$\Delta U_s = \sum_{j=1}^{N} \Delta U_j \times \omega_{js}$$

Table A1 reports how we applied union density rates by industry reported in Troy to the 1950 industry codes used in the 1940 full count census.[69]

A4. Construction of labor demand control variables

We construct two labor demand indexes from IPUMS census microdata. The first is a Bartik style index using initial local-level industry mix in 1940 and national-level industry growth rates.[70] The second demand index allows for skill-specific changes in demand at the local level following the index used in Goldin and Margo.[71]

Let j denote industry, s denote SEA, and t denote year. The first demand index (non-skill specific) is constructed according to the following equation:

$$D_s^{Bartik} = \sum_j \left(\frac{E_{js1940}}{E_{s1940}}\right)\left(\frac{E_{j1950} - E_{j1940}}{E_{j1940}}\right)$$

For a given SEA, the first term captures the percent of total employment made up by industry j, while the second term captures the national growth rate in employment for industry j. SEAs in which fast-growing industries (nationally) make up a larger share of employment will have a higher value for this index.

The second index captures the fact that the skill-mix varies by industry and within an industry j the skill-mix varies across SEAs. Thus, the relative demand for skill deriving from employment changes in that industry would vary across SEAs for a given industry growth rate, causing differential impacts on the local wage distribution. Following Goldin and Margo, we let i denote skill category, and can find the local demand for skill group i in year t by,[72]

$$D_{ist} = \sum_j \left(\frac{E_{jt}}{E_t}\right)\left(\frac{E_{ijs40}}{E_{js40}}\right)\left(\frac{E_{js40}}{E_{s40}}\right)$$

The first term captures the national employment share of each industry. Varying over time as industries wax and wane, it provides the sole source for changes in the index over time within an

---

[69] Troy, *Distribution of union membership*
[70] Bartik, *Who benefits*.
[71] Goldin and Margo, 'Great compression'
[72] Goldin and Margo, 'Great compression'

SEA. The second term is fixed at 1940 base levels and captures the SEA-specific skill mix within each industry. The third term captures the industry mix at the SEA level, and is fixed at 1940 base levels. The index can be simplified by canceling out the denominator in the second term and the numerator in the third term to get.

$$D_{ist} = \sum_j \left(\frac{E_{jt}}{E_t}\right)\left(\frac{E_{ijs40}}{E_{s40}}\right)$$

The index above is skill specific for each SEA. To get a relative skill demand index for year *t*, we use the following ratio of skilled to unskilled,

$$D_{s,t}^{GM} = \frac{D_{(skilled)st}}{D_{(unskilled)st}} = \frac{\sum_j \left(\frac{E_{jt}}{E_t}\right)\left(\frac{E_{(skilled)js40}}{E_{s40}}\right)}{\sum_j \left(\frac{E_{jt}}{E_t}\right)\left(\frac{E_{(unskilled)js40}}{E_{s40}}\right)}$$

Finally, the decadal change in local relative skill demand is found by taking $D_{s,1950}^{GM} - D_{s,1940}^{GM}$.

A5. Construction or sources of other control variables

Using IPUMS samples:[73]

- Share of males that are high school graduates in 1940, 1950, 1960: percent of males fitting sample restrictions for inequality measures that reported completing four years of high school.
- Share of wage and salary workers earning below the minimum wage in 1939: percent of males aged 18-64, earning positive wage and salary income, reporting wage and salary employment as their primary work, that earn less than the minimum wage on a weekly full-time basis (or less than \$12 weekly in 1939 - \$0.30 per hour times 40 hours).
- Variables measuring changes in inequality in occupational standing from 1920 to 1940: We use two measures of occupational standing. The first, provided by IPUMS, is the standard *occscore* variable, which "assigns each occupation in all years a value representing the median total income (in hundreds of 1950 dollars) of all persons with that particular occupation in 1950." Using the 1950 income distribution bakes in the large cross-occupation compression in wages of the "Great Compression." We create a second and more flexible occupational standing variable from the 1940 income distribution of male workers that satisfy the sample restrictions used in the main inequality measures. For observations in industry-by-occupation cells with more than 50 observations, we create an *occscore40* variable that

[73] Ruggles et al., *Integrated public use microdata series*.

assigns the median weekly earnings of that cell. For cells with 50 observations or less, the national median weekly earnings for the occupation is assigned. Each of the occscore measures are then assigned to observations in 1920, 1930, and 1940 based on the occupational code. We calculate inequality measures from 1920-40 and 1930-40: change in log 90-10, change in log 90-50, change in log 50-10, change in the Gini of the log occscore, and the change in the variance of log occscores.

Calculated from Haines[74]

- The male employment rate in 1940 (the percent of males aged 14+ that are employed)
- Percent of total population that resides in an urban area in 1940
- Per capita WWII contracts in dollars

BLS Cost of Living Report 1917-1919[75]

- City-level inequality in 1919 for a reduced sample of 83 cities: In combination with the 1940 full count IPUMS data, we estimate the change in local wage inequality from 1919 to 1939. The BLS survey provides micro-level income data for a sample of more than 12,000 families in 99 cities. Unfortunately for the purposes of this study, the sample frame was narrowed to focus on urban middle-class families. The BLS limited the sample to husband-wife families with at least one child were surveyed, families not living in "slums", and to families earning less than $2,000 a year (though they were not always). Families were to have resided in the place for at least one year, were not to have more than three boarders or lodgers, and were not to be non-English speakers who had resided in the US for less than five years. Feigenbaum estimates that the BLS sample is concentrated between the 30th to 70th percentiles of the national income distribution.[76] Note that the data were made available (in part) by the Inter-university Consortium for Political and Social Research (ICPSR). The data for *Cost of Living in the United States, 1917-1919* were originally collected by the Bureau of Labor Statistics. Neither the collector of the original data nor the consortium bear any responsibility for the analyses or interpretations presented here.

Other sources:

---

[74] Haines, *Historical, Demographic*
[75] U.S. Department of Labor, 'Cost of living' (ICPSR Study 8299),
[76] Feigenbaum, 'Intergenerational mobility'

- State mobilization rate in WWII.[77]
- Right to work state in 1950: Indicator variable for whether the state had passed “right to work” legislation by 1950.[78]

[77] Acemoglu, Autor, and Lyle, ‘Women, war, and wages.
[78] Ellwood and Fine, ‘Impact of right to work laws’, Table 1.

A6: Tables

Table A1: Crosswalk of industry categories reported in Troy to 1950 industry codes in IPUMS census microdata[79]

Iindustry categories	1950 industry codes
Manufacturing	
Metals	336 - 399
Clothing	448 - 449
Food, liquor, and tobacco	406 - 429
Paper, printing and publishing	456 - 459
Leather and leather products	487 - 489
Chemicals, rubber, clay, glass, and stone	316 - 328, 466 - 478
Textiles	436 - 446
Lumber and lumber products	306 - 309
Transportation (exclusive of railways), communications, and public utilities	507 - 598
Railway transportation	506
Building and construction	246
Mining, quarrying, and oil	206 - 239
Public service	906 - 936
Services	606 - 899
Apply the manufacturing average to "non specified manufacturing"	499

[79] Troy, *Distribution of union membership*

Table A2: Full estimation results for base specification (Row 1 of Table 2)

Panel A: 1940-50	(1)	(2)	(3)	(4)	(5)
	Δ90-10	Δ90-50	Δ50-10	ΔGini	ΔVar
Change in union exposure	-0.0232	-0.0030	-0.0202	-0.0041	-0.0066
	(0.0062)	(0.0033)	(0.0043)	(0.0011)	(0.0019)
Median log weekly wage 1939	-0.2631	0.2670	-0.5300	0.0148	-0.0547
	(0.0514)	(0.0304)	(0.0543)	(0.0120)	(0.0222)
Initial log 90-10 gap in 1939	-0.5544	-0.2709	-0.2837	-0.1046	-0.2038
	(0.0654)	(0.0365)	(0.0583)	(0.0113)	(0.0211)
Per capita WWII contracts (1940$)	0.0024	-0.0021	0.0045	-0.0002	-0.0009
	(0.0052)	(0.0026)	(0.0052)	(0.0009)	(0.0017)
Male employment rate in 1940	0.6176	0.0002	0.6172	0.1627	0.2755
	(0.1987)	(0.1001)	(0.1573)	(0.0433)	(0.0798)
Percent urban 1940	0.0338	-0.0869	0.1205	-0.0016	-0.0250
	(0.0478)	(0.0328)	(0.0458)	(0.0095)	(0.0183)
Percent males HS grads	-0.5045	-0.0370	-0.4680	-0.1055	-0.1455
	(0.1271)	(0.0970)	(0.1384)	(0.0450)	(0.0671)
Local demand shock index	0.0955	0.0162	0.0791	0.0219	0.0321
	(0.0615)	(0.0273)	(0.0462)	(0.0087)	(0.0171)
Skill-specific demand shock index	1.4179	0.7751	0.6419	0.1870	0.3142
	(0.4679)	(0.2359)	(0.3767)	(0.0737)	(0.1480)
Midwest census region	-0.0079	0.0070	-0.0148	0.0017	0.0057
	(0.0142)	(0.0120)	(0.0164)	(0.0050)	(0.0064)
South census region	0.0393	0.0337	0.0057	0.0069	0.0017
	(0.0220)	(0.0165)	(0.0170)	(0.0046)	(0.0064)
West census region	0.0465	-0.0132	0.0598	0.0064	0.0169
	(0.0233)	(0.0135)	(0.0237)	(0.0062)	(0.0092)
Constant	1.0344	-0.5773	1.6121	-0.0456	0.2571
	(0.2382)	(0.1375)	(0.2232)	(0.0520)	(0.0949)
Observations	467	467	467	467	467
R-squared	0.3223	0.4483	0.5013	0.2858	0.2973

Continued Table A2: Full estimation results for base specification (Row 1 of Table 2)

Panel B: 1940-60	(1)	(2)	(3)	(4)	(5)
	Δ90-10	Δ90-50	Δ50-10	ΔGini	ΔVar
Change in union exposure	-0.0215	-0.0048	-0.0166	-0.0032	-0.0057
	(0.0047)	(0.0025)	(0.0046)	(0.0008)	(0.0015)
Median log weekly wage 1939	-0.1719	0.3570	-0.5291	0.0346	-0.0266
	(0.0495)	(0.0411)	(0.0431)	(0.0107)	(0.0170)
Initial log 90-10 gap in 1939	-0.6034	-0.3040	-0.2994	-0.1251	-0.2343
	(0.0546)	(0.0420)	(0.0603)	(0.0109)	(0.0177)
Per capita WWII contracts (1940$)	0.0135	0.0050	0.0085	-0.0001	0.0016
	(0.0059)	(0.0026)	(0.0042)	(0.0011)	(0.0016)
Male employment rate in 1940	0.3373	0.1853	0.1515	0.0924	0.0767
	(0.1037)	(0.0600)	(0.1134)	(0.0241)	(0.0345)
Percent urban 1940	0.1784	-0.0827	0.2607	0.0100	0.0342
	(0.0324)	(0.0223)	(0.0295)	(0.0084)	(0.0127)
Percent males HS grads	-0.3679	-0.1088	-0.2594	-0.0878	-0.1294
	(0.1522)	(0.1106)	(0.1554)	(0.0291)	(0.0521)
Local demand shock index	0.1151	0.0094	0.1054	0.0191	0.0466
	(0.0412)	(0.0127)	(0.0439)	(0.0052)	(0.0106)
Skill-specific demand shock index	0.7307	0.5302	0.2024	0.1893	0.2847
	(0.4915)	(0.2165)	(0.4614)	(0.0655)	(0.1166)
Midwest census region	-0.0580	-0.0225	-0.0353	-0.0001	-0.0017
	(0.0136)	(0.0091)	(0.0113)	(0.0051)	(0.0062)
South census region	0.0689	0.0175	0.0516	0.0213	0.0312
	(0.0199)	(0.0123)	(0.0206)	(0.0044)	(0.0064)
West census region	0.0817	-0.0149	0.0969	0.0123	0.0387
	(0.0290)	(0.0109)	(0.0259)	(0.0057)	(0.0108)
Constant	0.8413	-0.8836	1.7261	-0.0422	0.2508
	(0.2304)	(0.1336)	(0.1954)	(0.0385)	(0.0672)
Observations	467	467	467	467	467
R-squared	0.5679	0.7927	0.6095	0.6429	0.6586

Notes: Each column represents the results from a separate regression and corresponds to the full estimation output of the base specification in row (1) of Table (2) in the main text. Region effects are relative to the Northeast census region.
Sources: Inequality measures, wage measures, demand indices, high school educational attainment, and are derived from the 1940, 1950, or 1960 IPUMS samples from Ruggles et al., '*Integrated public use microdata series'* as described in the text. Exposure to unions (ΔU_{ir}) is calculated by interacting industrial employment data from the 1940 full count census and industry-level changes in union density from Troy, '*Distribution of union membership'*. The employment rate and share of population in urban areas are calculated with data from Haines, *Historical, Demographic*.

Table A3: Changes in wage inequality and unionization, robustness to sample definitions

	(1) Δ90-10	(2) Δ90-50	(3) Δ50-10	(4) ΔGini	(5) ΔVar
Panel A: 1940-50					
(1) Base	-0.0232	-0.0030	-0.0202	-0.0041	-0.0066
	(0.0062)	(0.0033)	(0.0043)	(0.0011)	(0.0019)
(2) Drop Ag workers	-0.0155	-0.0020	-0.0134	-0.0028	-0.0033
	(0.0051)	(0.0032)	(0.0040)	(0.0010)	(0.0016)
(3) Census division f.e.	-0.0219	-0.0035	-0.0183	-0.0034	-0.0058
	(0.0063)	(0.0035)	(0.0043)	(0.0010)	(0.0018)
(4) Intervaled weeks (midpoint)	-0.0205	-0.0058	-0.0147	-0.0041	-0.0068
	(0.0066)	(0.0034)	(0.0045)	(0.0011)	(0.0019)
(5) Intervaled weeks (1950 average)	-0.0218	-0.0059	-0.0159	-0.0041	-0.0067
	(0.0063)	(0.0032)	(0.0043)	(0.0011)	(0.0019)
(6) Bottom code obs earning half min wage	-0.0268	0.0005	-0.0266	-0.0035	-0.0059
	(0.0084)	(0.0038)	(0.0081)	(0.0011)	(0.0024)
(7)) Δ1919-39 in inequality (83 cities)	0.0020		-0.0076	0.0096	-0.0034
	(0.0117)		(0.0098)	(0.0082)	(0.0040)
Panel B: 1940-60					
(1) Base	-0.0215	-0.0048	-0.0166	-0.0032	-0.0057
	(0.0047)	(0.0025)	(0.0046)	(0.0008)	(0.0015)
(2) Drop Ag workers	-0.0131	-0.0036	-0.0094	-0.0019	-0.0027
	(0.0049)	(0.0025)	(0.0049)	(0.0008)	(0.0015)
(3) Census division f.e.	-0.0203	-0.0050	-0.0153	-0.0030	-0.0053
	(0.0047)	(0.0026)	(0.0047)	(0.0008)	(0.0015)
(4) Intervaled weeks (midpoint)	-0.0195	-0.0050	-0.0144	-0.0032	-0.0057
	(0.0049)	(0.0025)	(0.0048)	(0.0008)	(0.0015)
(5) Intervaled weeks (1950 average)	-0.0204	-0.0053	-0.0151	-0.0032	-0.0067
	(0.0050)	(0.0024)	(0.0050)	(0.0008)	(0.0019)
(6) Bottom code obs earning half min wage	-0.0366	-0.0032	-0.0329	-0.0036	-0.0059
	(0.0075)	(0.0031)	(0.0078)	(0.0008)	(0.0024)

Notes: Each figure in the table is the estimated coefficient (τ) on the change-in-unions variable from a separate regression of the base specification on different samples. The base specification's control variables are described in the text; they 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. Row (1) reports results from the base specification from Table 2. Row (2) drops all wage and salary farmers, farm managers, and farm laborers. Row (3) includes census division fixed effects instead of census regions. Row (4) imposes the week intervals (and midpoints) of the 1960 census on the 1940 and 1950 census to calculate weekly earnings. Row (5) uses the mean weeks worked in each interval of the 1950 weeks distribution. Instead of dropping observations that earn less than half the minimum wage at a full-time weekly basis, row (6) bottom codes at half the minimum wage. The dependent variable in row (7) is the change in inequality from 1919-1939 for a set of 83 cities in the BLS 1919 Cost of Living Survey.

Sources: See notes to table A2. We derive 1919 inequality from the Bureau of Labor Statistics 1919 Cost of Living Survey.

Table A4: Sensitivity of results to varying sample restriction on weeks worked

	(1) 90-10	(2) 90-50	(3) 50-10	(4) Gini	(3) Var
Panel A: 1940-50					
(1) Base (restrict weeks >= 40)	-0.0232	-0.0030	-0.0202	-0.0041	-0.0066
	(0.0062)	(0.0033)	(0.0043)	(0.0011)	(0.0019)
(2) Restrict weeks >= 26	-0.0176	-0.0102	-0.0075	-0.0045	-0.0067
	(0.0050)	(0.0031)	(0.0034)	(0.0010)	(0.0017)
(3) Positive weeks worked	-0.0176	-0.0126	-0.0050	-0.0049	-0.0069
	(0.0052)	(0.0028)	(0.0038)	(0.0020)	(0.0023)
Panel B: 1940-60					
(1) Base (restrict weeks >= 40)	-0.0215	-0.0048	-0.0166	-0.0032	-0.0057
	(0.0047)	(0.0025)	(0.0046)	(0.0008)	(0.0015)
(2) Restrict weeks >= 26	-0.0206	-0.0098	-0.0108	-0.0043	-0.0069
	(0.0046)	(0.0027)	(0.0041)	(0.0007)	(0.0013)
(3) Positive weeks worked	-0.0156	-0.0113	-0.0043	-0.0041	-0.0063
	(0.0041)	(0.0028)	(0.0040)	(0.0008)	(0.0013)

Notes: Each figure in the table is the estimated coefficient (τ) on the change-in-unions variable from a separate regression of the base specification on different samples. The base specification's control variables are described in the text; they 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. Row (1) reports results from the base specification from Table 2, which limits the sample to observations having worked at least 40 weeks in the previous year. Row (2) relaxes that restriction to at least 26 weeks worked, and row (3) removes the restriction on weeks worked entirely (positive weeks worked is implicit due to division by zero in the weekly earnings calculation). Row (4) increases the restriction on weeks to 50 or greater.
Sources: See notes to table A2.

Table A5: Sensitivity of base results to controls for share foreign born and percent black

	(1)	(2)	(3)	(4)	(3)
	90-10	90-50	50-10	Gini	Var
Panel A: 1940-50					
(1) Base	-0.0232	-0.0030	-0.0202	-0.0041	-0.0066
	(0.0062)	(0.0033)	(0.0043)	(0.0011)	(0.0019)
(2) Add control for change in percent black (1950-40)	-0.0227	-0.0030	-0.0197	-0.0039	-0.0062
	(0.0061)	(0.0033)	(0.0043)	(0.0010)	(0.0018)
(3) Add control for 1940 share foreign born	-0.0234	-0.0034	-0.0200	-0.0043	-0.0069
	(0.0061)	(0.0032)	(0.0046)	(0.0011)	(0.0019)
Panel B: 1940-60					
(1) Base	-0.0215	-0.0048	-0.0166	-0.0032	-0.0057
	(0.0047)	(0.0025)	(0.0046)	(0.0008)	(0.0015)
(2) Add control for change in percent black (1960-40)	-0.0219	-0.0044	-0.0175	-0.0031	-0.0057
	(0.0048)	(0.0026)	(0.0045)	(0.0008)	(0.0015)
(3) Add control for 1940 share foreign born	-0.0187	-0.0036	-0.0150	-0.0029	-0.0051
	(0.0051)	(0.0030)	(0.0048)	(0.0008)	(0.0015)

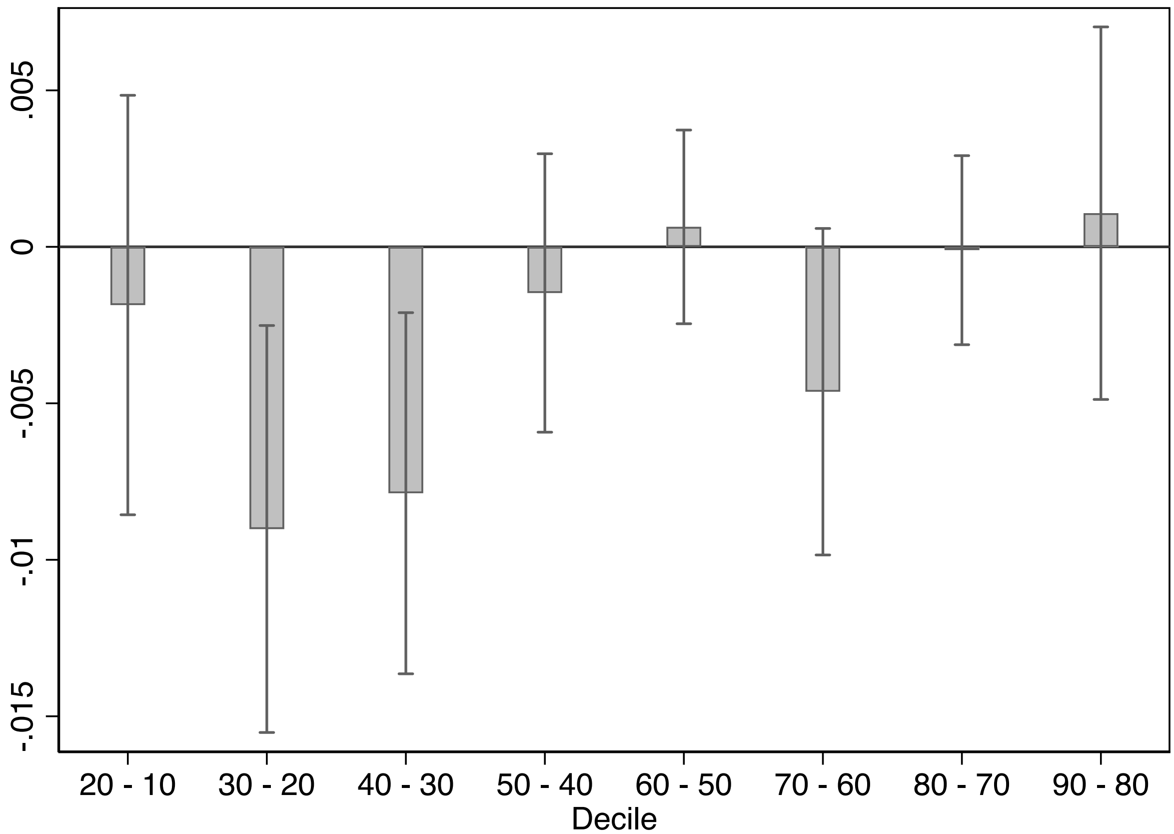
Notes: Each figure in the table is the estimated coefficient (τ) on the change-in-unions variable from a separate regression of the base specification with the inclusion of different control variables. The base specification's control variables are described in the text; they include regional fixed effects, 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. Row (1) reports results from the base specification from Table 2. Row (2) adds the change in the percent black from either 1950-40 or 1960-40 in an SEA. Row (3) includes a control for the share of the population that is foreign born in 1940.
Sources: See notes to table A2. The share foreign born and share black in an SEA is from Haines, *Historical, Demographic*.

Table A6: Effect of union exposure on (residual) within-group inequality

	(1) $d\sigma_{5040}$	(2) $d\sigma_{5040}$	(3) $d\sigma_{6040}$	(4) $d\sigma_{5040}$
Change in union exposure	-0.0015	-0.0022	-0.0020	-0.0029
	(0.0010)	(0.0009)	(0.0008)	(0.0008)
$\sigma_{s,40}$		-0.5542		-0.7368
		(0.1315)		(0.0726)
Median log weekly wage 1939	-0.0350	-0.0190	-0.0225	-0.0013
	(0.0151)	(0.0155)	(0.0108)	(0.0095)
Initial log 90-10 gap in 1939	-0.1069	-0.0109	-0.1232	0.0044
	(0.0152)	(0.0234)	(0.0103)	(0.0162)
Per capita WWII contracts (1940$)	-0.0016	-0.0013	-0.0001	0.0002
	(0.0011)	(0.0011)	(0.0010)	(0.0009)
Male employment rate in 1940	0.2024	0.1729	0.0804	0.0412
	(0.0528)	(0.0483)	(0.0248)	(0.0179)
Percent urban 1940	0.0061	0.0104	0.0288	0.0346
	(0.0113)	(0.0100)	(0.0077)	(0.0072)
Percent males HS grads	-0.0790	-0.0653	-0.0974	-0.0792
	(0.0578)	(0.0594)	(0.0309)	(0.0280)
Local demand shock index	0.0053	-0.0066	0.0180	0.0021
	(0.0126)	(0.0129)	(0.0070)	(0.0058)
Skill-specific demand shock index	0.2438	0.1361	0.2068	0.0636
	(0.0946)	(0.0899)	(0.0687)	(0.0535)
Midwest census region	0.0043	0.0022	0.0026	-0.0002
	(0.0054)	(0.0049)	(0.0037)	(0.0022)
South census region	0.0092	0.0059	0.0225	0.0182
	(0.0050)	(0.0053)	(0.0046)	(0.0033)
West census region	0.0139	0.0044	0.0244	0.0118
	(0.0093)	(0.0091)	(0.0067)	(0.0043)
Constant	0.0962	0.0773	0.1185	0.0934
	(0.0694)	(0.0686)	(0.0384)	(0.0412)
Observations	467	467	467	467
R-squared	0.2064	0.2388	0.6077	0.7246
St. dev of dependent variable	0.0499		0.0271	
St. dev of change in union exp.	3.14			

Notes: Each column reports the full estimation results from a separate regression at the SEA-level. The dependent variable is the change in the variance of the residuals from a wage regression individually estimated for each SEA-by-year. The variance of predicted residuals for SEA s in year t ($\sigma_{s,t}$) is derived from the predicted residuals from a regression of log weekly earnings on experience, experience squared, years of schooling, years of schooling squared, the interaction of experience and years of schooling and an indicator for marital status (experience = age – years of schooling – 6). Columns (2) and (4) add the initial 1940 level of variance of the residuals. Region effects are relative to the Northeast census region.
Sources: See notes to table A2.

Figure A1: Union effect on compression for deciles over 1940-50



Notes: The figure plots the coefficients and 95 percent confidence intervals from regressions that estimate τ for each decile separately for 1940-50 using the controls from the base specification. The 90-10 effect from the base specification is -0.0233. Roughly 75 percent of the union exposure effect on the 90-10 differential comes from compression between the 20th to 40th percentiles; another 20 percent appears between the 60th and 70th percentiles.

References

Acemoglu, D., D.H. Autor, and D. Lyle, 'Women, war, and wages: the effect of female labor supply on the wage structure at midcentury.' *Journal of Political Economy* 112 (2004), pp. 497-551.

Autor, D.H. and D. Dorn, 'The growth of low-skill service jobs and the polarization of the US labor market,' *American Economic Review* 103 (2013), pp. 1553-97.

Autor, D.H., D. Dorn, and G.H. Hanson, 'The China syndrome: local labor market effects of import competition in the United States.' *American Economic Review* 103 (2013), pp. 2121-2168.

Bartik, T.J., *Who benefits from state and local economic development policies?* (Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 1991).

Baum-Snow, N. and R. Pavan. 'Inequality and City Size,' *Review of Economics and Statistics* 95 (2013), pp. 1535-1548.

Bogue, D.J., *State economic areas: a description of the procedure used in making a functional grouping of counties in the United States*. (Department of Commerce, Bureau of the Census. Washington, DC: Government Printing Office, 1951).

Dorn, D., 'Essays on Inequality, Spatial Interaction, and the Demand for Skills' (unpub. Ph.d. thesis, University of St. Gallen no.3613, September. Data Appendix, pp. 121-138, 2009)

Ellwood, D.T. and G.A. Fine. 'The impact of right to work laws on union organizing.' *Journal of Political Economy* 95 (1987), pp. 250-273.

Feigenbaum, J.J. 'Intergenerational mobility during the Great Depression.' Unpublished working paper (2015).

Goldin, C. and R.A. Margo, 'The Great Compression: the wage structure in the United States at mid-century.' *Quarterly Journal of Economics* 107 (1992), pp. 1-34.

Haines, M.R., and Inter-university Consortium for Political and Social Research. *Historical, Demographic, Economic, and Social Data: The United States, 1790-2002* [computer file]. ICPSR02896-v3. (Ann Arbor, MI: ICPSR 2010).

Moretti, E., 'Local Labor Markets', in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, vol 4, (Elsevier, 2011), pp. 1237-1314.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek. *Integrated Public Use Microdata Series: Version 6.0* [Machine-readable database]. (Minneapolis: University of Minnesota, 2015).

Troy, L., *The Distribution of Union Membership among the States, 1939 and 1953*. (New York: National Bureau of Economic Research, 1957).

United States Department of Labor, Bureau of Labor Statistics, 'Cost of Living in the United States, 1917-1919' ICPSR08299-v5, (1986), Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].