NBER WORKING PAPER SERIES

UNIONS, WORKERS, AND WAGES AT THE PEAK OF THE AMERICAN LABOR MOVEMENT

Brantly Callaway
William J. Collins

Working Paper 23516
http://www.nber.org/papers/w23516

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2017

This paper is dedicated to T. Aldrich Finegan, Professor Emeritus at Vanderbilt University. The authors gratefully acknowledge assistance from the staff of the University of Pennsylvania Archives, as well as helpful input from Ran Abramitzky, Al Finegan, Malcolm Getz, Claudia Goldin, Andrew Goodman-Bacon, Barry Hirsch, Ilyana Kuziemko, Robert Margo, Suresh Naidu, Greg Niemesh, John Pencavel, Valerie Ramey, Ariell Zimran, and anonymous referees. NSF funding supported the original data collection (SES 0095943). Francesca Ciliberti, Hannah Moon, Christina Quigley, and Tim Watts provided excellent research assistance with the Palmer Survey. Callaway is an Assistant Professor at Temple University; Collins is the Terence E. Adderley Jr. Professor of Economics at Vanderbilt University and Research Associate of the NBER. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Unions, Workers, and Wages at the Peak of the American Labor Movement
Brantly Callaway and William J. Collins
NBER Working Paper No. 23516
June 2017
JEL No. J5,N12

ABSTRACT

We study a novel dataset compiled from archival records, which includes information on men's wages, union status, educational attainment, work history, and other background variables for several cities circa 1950. Such data are extremely rare for the early post-war period when U.S. unions were at their peak. After describing patterns of selection into unions, we measure the union wage premium using unconditional quantile methods. The wage premium was larger at the bottom of the income distribution than at the middle or higher, larger for African Americans than for whites, and larger for those with low levels of education. Counterfactuals are consistent with the view that unions substantially narrowed urban wage inequality at mid-century.

Brantly Callaway
Temple University
1301 Cecil B. Moore Avenue
Ritter Annex 867
Philadelphia, Pennsylvania 19122
brantly.callaway@temple.edu

William J. Collins
Department of Economics
Vanderbilt University
VU Station B #351819
2301 Vanderbilt Place
Nashville, TN 37235-1819
and NBER
william.collins@vanderbilt.edu

Between 1935 and 1953, union membership in the United States increased from about 13 to 33 percent of nonagricultural employment (Troy 1965). This was the culmination of a tumultuous period in U.S. economic history, as New Deal legislation, the emergence of industrial unions, the exigencies of World War II, and workers' demands for representation led to a sea change in American labor market institutions (Seidman 1953, Lester 1964, Reder 1988, Freeman 1998). The economics and history literatures on unions in this period are large.[1] Yet some fundamental questions about unions and wages at mid-century have proven difficult to answer due to the scarcity of micro-level datasets that record workers' union status. In particular, the federal Census of Population has never inquired about union status, and the Current Population Survey (CPS) first collected data on union membership in 1973, by which time private sector unions were in decline and losing public support (Saad 2015).[2]

Our interest in unions at mid-century is heightened by a concurrent and perhaps related trend in inequality—the high point of American unions coincided with the low point of American wage inequality in the twentieth century (Goldin and Margo 1992). Scholars studying the late twentieth century have found evidence that unions tend to reduce inequality (e.g., Freeman 1980; Freeman and Medoff 1984; DiNardo, Fortin, Lemieux 1996; Card, Lemieux, and Riddell 2004), but far less is known about the relationship between unions and inequality around mid-century. Prominent scholars at the time speculated that unions might raise inequality by placing a wedge between the wages of similar union and non-union workers, by raising wages for workers who might have been already been relatively well paid, and by reducing employment opportunities in unionized firms (Friedman 1962, 123-25; Rees 1962, 98-99), but they did so without the benefit of worker-level data.

In this paper, we compile and study a novel micro-level dataset that includes information on wages and union status circa 1950, as well as an unusually rich set of background characteristics. Such data are extremely rare around mid-century, and to our knowledge, the dataset's combination of information is unique for the period.[3] The original survey was conducted in six cities in early 1951, covering Philadelphia, New Haven, Chicago, St. Paul, San Francisco, and Los Angeles. It was designed to gather information about labor mobility during the 1940s. We transcribed archival

[1] Inter alia, see Millis and Brown (1950), Lewis (1963), Freeman (1978), Lichtenstein (1982), and Kersten (2006).

[2] Data from the Survey of Economic Opportunity became available in the late 1960s. This allowed early studies using microdata (e.g. Ashenfelter 1972, Schmidt and Strauss 1976, and Lee 1978), though the profession shifted to CPS-based analyses once CPS microdata became widely available in the 1970s.

[3] Henry Farber, Dan Herbst, Ilyana Kuziemko, and Suresh Naidu are concurrently working to compile data from early Gallup Polls and American National Election Studies that include information on whether someone in the household belonged to a union.

records for more than 6,900 men in five of the cities.[4] Within cities, the Census Bureau designed the baseline samples to be representative of the local population (Palmer 1954). Although the cities are not representative of the entire United States, they are varied in region, size, specialization, and experience during World War II. We show that the pooled sample of survey microdata closely approximates the economic, demographic, and human capital characteristics of the non-southern urban labor force.

We use the new data to address three basic sets of questions about unions, workers, and wages at mid-century. First, what was the nature of worker selection into union membership at the height of union strength? Were there substantial differences between union and non-union members in age, race, educational attainment, veteran status, or other background or work history variables? Second, what was the conditional-on-observables difference in men's wages depending on their union status, and how did the gap vary across quantiles of the union and non-union wage distributions? Third, how different would the overall distribution of wages have been in these cities circa 1950 if unionized men were paid according to the non-union wage schedule? None of these questions could be addressed effectively in the absence of worker-level data.

We find that male union members circa 1950 were, on average, negatively selected from the labor force in terms of educational attainment and father's occupational status. Union workers earned slightly more than observationally similar non-union workers at the median, though the baseline estimate is only marginally significant (0.029, p-value=0.10). But focusing on the middle of the distribution misses much of the story. The largest union wage gaps appear at the bottom of the wage distribution, and they were larger for African Americans and for less-educated men. It is beyond the scope of this paper to model a full counterfactual economy in which unions did not exist, but we can provide counterfactual estimates of wage inequality based on reweighting techniques developed by DiNardo, Fortin, and Lemiuex (1996), Hirano, Imbens, and Ridder (2003), and Firpo (2007). It is clear that the union distribution was more compressed than the non-union distribution even after accounting for differences in observable characteristics. This is consistent with studies for later periods that find that unions tend to standardize wages within and between establishments and to flatten returns to skill (e.g., Freeman and Medoff 1984). A "no unions" counterfactual wage distribution, which imposes the non-union wage structure on union workers and combines that distribution with the actual distribution for non-union workers, is substantially more unequal than the actual overall wage distribution. We stress that care must be taken in interpreting this result, as the

[4] We discuss the data and the original survey in more detail below. Unfortunately, the original data for men in Chicago are missing from the collection.

setting does not allow strong causal identification, but it is at least consistent with the hypothesis that unions reduced wage inequality in urban areas of the United States at mid-century.

Finally, for the sake of perspective, we compare the results from the Palmer Survey data to a similar analysis using CPS data from 1973 for non-southern cities. This is the first year that union status was recorded in the CPS. In 1973, the union wage premium followed roughly the same pattern as in the early 1950s—there were relatively large union wage advantages at lower percentiles in the distribution and smaller gaps at the middle. Overall, the counterfactual in which union members are paid like non-union members suggests that unions reduced inequality in the 1970s, though perhaps somewhat less than in the 1950s. The difference is partly due to the higher rate of unionization in the 1950s; that is, more workers in 1950 were located in the relatively compressed union wage distribution.

1. Background

Brief history of unions in the U.S. during the twentieth century

Prior to the 1930s, unions in the U.S. were organized primarily along craft lines and had a precarious existence (Rees 1962, Lester 1964). A large, open labor market characterized by high levels of occupational and geographic mobility and ethnic heterogeneity may have forestalled the rise of organized labor in the U.S. In addition, there was no legal requirement for employers to engage in collective bargaining and little restraint on employers' use of strikebreakers, lockouts, retaliatory firing and threats, and other means to oppose unions and prevent their organization. Richard Lester writes, "…vigorous (even ruthless) employer resistance to labor organization continued for almost a century and a half prior to World War II. Such opposition was much more prolonged and bitter here than in other industrial countries….it helped to give our labor relations history its own special strands of physical violence and private warfare" (1964, p. 63). A sharp increase in union membership during World War I, for instance, was reversed after the war in the wake of recession, renewed employer opposition, and shifting government and public support.

The National Labor Relations Act of 1935, also known as the Wagner Act, marked a major change in public policy regarding union organization and collective bargaining. It created the National Labor Relations Board (NLRB) to oversee union representation elections and enforce the Act's provisions, which required firms to engage in collective bargaining with certified unions and refrain from many of the tactics previously used to discourage union formation. The Supreme Court upheld the Act in 1937 (*NLRB v. Jones & Laughlin Steel*), and with substantial revision, it remains the basis for union-management relations in the United States.[5] The new legal framework, in combination with the emergence of the Congress of Industrial Organizations (CIO) and workers' rising demand for representation, led to a rapid increase in union membership in the late 1930s, as shown in Figure 1. It is plausible that popular discontent from the Great Depression was a fundamental factor underlying both the institutional changes of the 1930s and the successful organization of unions that ensued (Ashenfelter and Pencavel 1969, Freeman 1998).

During World War II, temporary policies under the National War Labor Board further strengthened union membership in the interest of promoting maximum production (Seidman 1953, Freeman 1978). Nonetheless, political opposition to organized labor grew during the 1940s (Millis and Brown 1950). In 1947, the Labor Management Relations Act, also known as the Taft-Hartley Act, passed over President Truman's veto. It curbed a variety of union tactics, allowed employers to campaign against union formation, outlawed "closed shops" (in which only union members could be hired), and permitted states to pass "right-to-work" laws. Union density fell slightly in the late 1940s, but rebounded in the early 1950s. Thus, the data we study pertain to workers observed at the high-point of American unionization, a period lasting roughly from the mid-1940s to the late 1950s. Since then, the private-sector union membership rate has declined nearly monotonically.

Closely related research

The mid-century rise of unions in the U.S. motivated a large body of research on whether unions affected wages and overall wage inequality. As mentioned earlier, scholars writing at the time generally did not have recourse to micro-level data on wages. Instead, studies of the union wage gap tended to look for differential changes in average wages within industries as union density changed (e.g., Ross 1948), or across localities with varying degrees of union density (e.g., Sobotka 1953). There are far too many studies to review here, but Lewis (1963) carefully reviews the state-of-the-art circa 1960. In his writing, modern readers will recognize early considerations of classic empirical problems such as selection bias, confounding pre-trends in treatment and control groups, measurement error, and general equilibrium effects.

The empirical evidence from mid-century studies is mixed, with some studies suggesting large union wage effects and others suggesting negligible effects, depending on the industry, method, and time of observation. Lewis notes a tendency for estimates of the average union/nonunion wage gap to be small (less than 5 percent) in studies of the late 1940s (1963, pp. 5, 190), suggesting that

---

[5] The Act did not apply to government employees, agricultural workers, or domestic workers. A separate industry-specific framework covered railroad workers and was an important precursor to the Wagner Act.

post-war inflation may have eroded nominal wages set in collectively bargained contracts. By the mid-to-late 1950s, however, Lewis suggests that the average wage gap had rebounded to approximately 10-15 percent (p. 193). Pencavel and Hartsog (1984, p. 204, 207) follow Lewis's time-series analysis and extend it to cover 1920 to 1980. They confirm that the average union wage gap was positive and that it fluctuated with macroeconomic circumstances, but they emphasize that estimates of period-specific gaps are imprecise. Like Lewis (1963), Rosen (1969) emphasizes that spillovers from union to non-union wages complicate the interpretation of the union wage gap. He examines industry-level wage data within manufacturing from 1958, finding (inter alia) relatively low returns to education associated with unionization and a positive association between the union wage gap and the union coverage rate. Using similar data, Rosen (1970) finds evidence that is consistent with larger union wage gaps for skilled craftsman and unskilled laborers than for semi-skilled operatives.[6]

Scholars also had mixed impressions of whether unions tended to raise or lower inequality (Lewis 1963, p. 292-95). It is plausible that unions exacerbated inequality on some dimensions, as speculated by Friedman (1962, pp. 123-25) and Rees (1962, pp. 98-99). Miller (1958), however, speculates that unions might have lowered inequality in the 1940s because they tended to bargain for across-the-board raises in cents-per-hour, disproportionately raising wages at the bottom. More recently, Goldin and Margo (1992) suggest that unions might have helped sustain the relatively narrow dispersion of wartime wages into the post-war period. Frydman and Molloy (2012) find that executive compensation relative to production workers' pay was negatively correlated with union presence in 1949. And Collins and Niemesh (2016) find that places with industrial concentrations that made them susceptible to unionization after 1939 had relatively large declines in inequality.

Although micro-level data are not a cure-all for the measurement challenges Lewis described, economists shifted strongly to micro-level evidence on the union wage gap and inequality once it became widely available. Early examples using the Survey of Economic Opportunity from 1967 include Ashenfelter's study of racial discrimination and unions (1972) and Lee's study of selection and union wages (1978). Lee's study is of particular note for its effort to model selection into unions and the union wage premium simultaneously.[7] We refer to some of Ashenfelter's and Lee's specific

[6] The measure of unionization in Rosen (1969, 1970) is the share of workers in establishments in which at least 50 percent of all workers were represented in collective bargaining, from Weiss (1966), which may be a noisy proxy for union density. Wages are not observed separately by union status or by occupational category; rather, inferences are drawn based on regressions of industry wages on industry characteristics.
[7] The exclusion restrictions in Lee's analysis are questionable, but the point that it is important to consider union selection and wages together is valid and the effort is insightful. Also see Schmidt and Strauss (1976)

findings in the course of discussing our analysis of the Palmer data.

The subsequent empirical literature in economics focuses on analyses of union wages in CPS microdata. With regard to the wage gap, Freeman and Medoff's landmark study (1984) reports a large average union wage premium during the late 1970s, at approximately 20 to 30 percent for samples of non-agricultural, private-sector, blue-collar workers. They also find heterogeneity across groups of workers: the young, those with low levels of education, and those working in manual labor occupations tended to have the largest union wage premiums.[8] Blanchflower and Bryson (2004) revisit and extend key themes in Freeman and Medoff's inquiry, including evidence on union wage gaps from the 1970s to the early 2000s (see also Hirsch and Macpherson 2016). From the mid-1980s, the wage premium declined along with union membership, which they attribute to increasing competitive pressures in product and labor markets. Nonetheless, Blanchflower and Bryson confirm that many of the patterns across groups—for instance the high union premium among less-educated workers—persist over time. With regard to inequality, Freeman (1980), Freeman and Medoff (1984), DiNardo, Fortin, Lemieux (1996), and Card, Lemieux, and Riddell (2004) present evidence suggesting that unions tended to narrow overall inequality.

The early empirical literature on unions and wages was aware that non-random selection into unions might complicate the interpretation of observed wage gaps between union and non-union workers. This theme of worker mobility and selection links the union wage premium literature with other prominent literatures in labor economics and public economics. A widely applicable insight distilled from the Roy model (1951) is that highly skilled workers may tend to sort into sectors, organizations, or countries that reward skill highly, and may avoid settings where returns to skill are relatively compressed. This idea is central to the modern economics literature on migration patterns, starting with Borjas (1987) and continuing in studies of both international and internal migration in historical and current settings. In the public economics literature, a related and longstanding idea is that heterogeneous households may sort geographically in response to differences in local tax and redistribution policies (Stigler 1957, Epple and Romer 1991, Feldstein and Wrobel 1998). Similar insights apply to organizations with relatively compressed wage structures, such as worker-managed firms (Burdin 2016) and kibbutzim (Abramitzky 2008, 2009). Although the relevance of selection to analyses of unions is clear, the topic has been difficult to address in the early and formative period of the American labor movement due to data limitations. This paper develops a new dataset to

for an early effort to combine consideration of selection and wages. In the absence of determinants of union status that are credibly excludable from wages in the Palmer data, we do not follow this line of modeling.

[8] Lewis (1986) also summarized the micro-level evidence on the wage effects of unions available at that time.

overcome some of those fundamental limitations.

2. The Palmer Survey data

The data examined in this paper were originally collected to study labor force mobility during the 1940s. Gladys L. Palmer, under the auspices of the Social Science Research Council, orchestrated the project, which was funded by the U.S. Air Force. The Bureau of the Census designed the samples, drawing primarily on dwelling units that were enumerated in the 1950 census ("stratum I") and making additional efforts to sample newly constructed dwellings and group quarters. Palmer describes the approach as a "three-stage cluster sample" (1954, p. 148). For example, in stratum I, a sample of census enumeration districts was selected in each city (stage 1), one or more clusters of 18 dwelling units were drawn from each district (stage 2), and finally three dwelling units were selected from each cluster (stage 3). Census enumerators implemented the survey in six cities in January and February 1951. A first interview collected basic information on all household residents age 14 and over; a second interview collected more detailed work history information for those age 25 and over and employed full time for at least one month in 1950. Ultimately, enumerators collected work history information directly from approximately 9,000 male and 4,000 female workers in Philadelphia, New Haven, Chicago, St. Paul, Los Angeles, and San Francisco. More detailed information about the survey design and implementation is provided in Palmer (1954, Chapter 1 and Appendix B).

Fortunately, the University of Pennsylvania Archives holds the handwritten "transcription sheets" from the work history interviews with men in each city except Chicago. For each worker, the original interview's information was written on a "work history schedule." The transcription sheet is like a large index card with boxes corresponding to particular questions or pieces of information retrieved and coded from the work history schedule. We rely entirely on data retrieved from the extant archival transcription sheets. In comparison with counts reported in the Palmer study (1954, p. 152), our sample contains about 99 percent of the original work history surveys for men in the five cities with available data.[9] Portions of the Palmer Survey data were studied in Goldin (1991), which focuses on women's employment during the 1940s, and Collins (2000), which focuses on African American men's occupational upgrading during the 1940s. The dataset used in this paper includes

[9] We do have a small number of duplicate transcription sheets for Chicago, but we have not located the original collection for Chicago. An electronic dataset from UCLA's archives appears to contain some, but not all, men from Chicago (cities are not directly identified). We have not attempted to integrate the data from UCLA here because we have not determined why some men are in that dataset and others are not; unfortunately, the accompanying documentation is incomplete.

thousands of men and several variables that were not previously transcribed or studied, most importantly union status. The original Palmer study (1954) does not examine the union wage premium or any other aspect of union membership, nor does Goldin (1991) or Collins (2000).

Table 1 reports summary statistics for the pooled sample of men whose information we retrieved from the archives. For consistency with subsequent analyses, the sample is restricted to men of ages 25 and over without missing information on educational attainment, union status, or earnings. The survey did not collect wage and salary information for self-employed men (approximately 20 percent of the full sample), and therefore they are omitted from the table. The original Palmer tabulations included some men's work histories more than once as a way to substitute for households that were not found or did not participate in the survey (1954, p. 152). We weight the observations that served as "duplications" accordingly (approximately 5 percent of sample), though doing so has little effect on the summary statistics.

In the pooled sample, without adjustments for city size (discussed below), about half the men belonged to a union at the time of the survey (51 percent). The sample is predominantly white (90 percent), reflecting the demographic composition of the non-southern cities in the sample. Roughly one-third of the men were World War II veterans. They were predominantly engaged in blue-collar occupations in 1950 (66 percent) as were their fathers before them (68 percent), where "blue collar" is defined as an occupation in the craftsmen, operative, laborer, and farmer categories. Manufacturing industries were the largest employers (35 percent), but other major industrial categories are well represented in the sample. Unfortunately, the survey did not collect information on hours of work, though the sample frame was intended to include only full-time workers. We bring industry-level data on hours worked into consideration later in the paper. Also, the survey did not collect information on non-wage benefits such as paid vacation time and health insurance.

The original Palmer Survey collected data for a similar number of households in each city, even though the cities were different sizes. Weighting the observations to reflect differences in the size of each city's male labor force (column 3) gives more influence to Philadelphia and Los Angeles and less to New Haven and St. Paul. Yet the mean values of most variables are only slightly affected (e.g., union status falls from 0.51 to 0.50 and weekly wages rise from 72 to 74).[10]

Relying on the Palmer data entails some shortcomings in terms of geographic coverage, most

[10] We applied weights that are the ratio of each city's male labor force, age 25 and above, to the number of observations for each city in the Palmer sample including duplications (i.e., observations entered more than once into the original Palmer analysis). Data on labor force size are from the published census volumes for 1950 (U.S. Department of Commerce 1952).

notably the omission of southern cities and rural areas. We conclude, however, that the Palmer data are likely to provide a useful characterization of cities outside the South. Appendix Table 1 shows that the five cities available for study in the Palmer Survey are fairly representative of cities outside the South in 1950 according to IPUMS census data (Ruggles et al. 2015).[11] For men, ages 25 and over, who were classified as wage and salary workers, we see that a variety of demographic, human capital, income, inequality (standard deviation of income), occupation, and industry variables were similar in magnitude across census samples for the "Palmer cities" and for "all non-southern cities." Most differences are statistically significant but small. The most notable differences are that the Palmer cities' employment was less concentrated in manufacturing (32 percent versus 37 percent in the "all non-southern cities" column), and the Palmer cities' workers had somewhat higher educational attainment (9.9 compared to 9.6 years). Based on evidence from the American National Election Study (ANES) for 1952 (Campbell and Gurin 2015) and a back-of-envelope calculation based on evidence from Troy (1965) and the 1973 CPS, we believe that 50 percent is a reasonable estimate of union density for male wage and salary workers in cities outside the South.[12] For reference, the table's last column adds southern cities to the sample, which pulls the summary statistics in the expected directions (e.g., lower education and income).

Appendix Table 2 compares the Palmer Survey data with the IPUMS census data for the same five cities. The goal is to see whether measures derived from the Palmer Survey are comparable to figures drawn from the more familiar census microdata. In almost all cases, the Palmer data's values are close to the census-based data, despite some unavoidable incongruities (e.g., the census data do not distinguish St. Paul from Minneapolis; census enumerators did not ask specifically for weekly earnings so they must be calculated from self-reported annual wages and weeks worked; and the census was conducted 9-10 months before the Palmer Survey).

[11] The IPUMS 1-percent sample for 1950 does not include an *urban* variable. We use the *city* variable, which identifies approximately 97 cities. Lorain, Ohio is the smallest city identified (population 43,400).

[12] The ANES sample for 1952, when restricted to male-headed households residing in a non-southern city, where the head was at least 25 years old and not self-employed (or retired, etc.), yields a union membership estimate of 54.5 percent. This is based on a question about whether the respondent or another member of the household belonged to a union. The ratio of union density for men, age 25+, in non-farm employment in metropolitan areas outside the South over the density for all workers in non-farm employment is 1.70 in the 1973 CPS; multiplying 1.70 by 31.2 percent (national non-farm union density in 1950 in Troy (1965)) is 53 percent. Suresh Naidu kindly provided an estimate for non-farm, non-southern households in Gallup Poll data circa 1950 at 35 percent (personal communication); the ANES data suggest that shifting to urban residents and omitting the self-employed, retired workers, etc., would scale that number up substantially.

3. Selection into unions at mid-century

Because micro-level data on union membership at mid-century are rare, so are statistical studies of selection into unions at that time. Selection into unions is interesting in its own right historically, and understanding selection is an important first step to interpreting differences in wages between union and non-union workers. Around mid-century, joining a union often went hand-in-hand with employment in a unionized establishment, reflecting the establishment's collectively bargained rules (e.g., "union shops" and "maintenance of membership" policies), peer pressure, or the worker's support for the union's goals (Seidman, London, and Karsh 1951; Hammond and Nix 1953). It was common for establishments to recruit new employees simply by word-of-mouth (Lester 1954). So, even though union density was high in this period, not all workers were equally well positioned to learn about job openings in unionized establishments. Seniority considerations may have dampened mobility into unionized establishments, as new workers would often start at the bottom of the job assignment and compensation ladder regardless of experience elsewhere (Lester 1954, p. 30). It is likely that, "…accidents, friendships, hearsay, and the timing of job openings" played important roles in determining the distribution of workers over establishments and, therefore, union status.[13] Of course, this does not imply that union membership was randomly assigned. The Palmer data provide new insight regarding how different union and non-union men were circa 1950.

Table 2 provides a simple characterization of selection by splitting the Palmer sample into union and non-union subsamples for comparison. Observations are weighted as described above to reflect differences in the size of cities. Among men with reported earnings, the pattern of selection into unions on the basis of personal characteristics was fairly sharp. There were negligible differences in union membership on the basis of age, race, and years of residence in the area, and only moderate differences in veteran status (4 p.p. lower for union members) and marital status (4 p.p. higher for union members). There were, however, strong differences in terms of educational attainment, father's occupation, and nativity. Union members, on average, had about 1.7 years less education than non-members, and they were far less likely to have completed high school or

[13] Based on interviews with industrial firms in Trenton in the early 1950s, Lester concluded that: "Companies differ from one another in many respects…. So significantly do they differ and so uncertain are the future employment prospects for individual employees that a calculating new job seeker would find it exceedingly difficult, even with the knowledge possessed by those interviewing for this study, to select the one (or even half dozen) of the 80-odd manufacturing firms studied, in which his earnings or job satisfaction would be greatest during his work life. Such difficulties, along with lack of information, help to explain why much of the process of application for, and acceptance of, jobs is governed by accidents, friendships, hearsay, and the timing of job openings" (1954, p. 29).

college.[14] Union members' fathers were much less likely to have been white-collar workers in their "longest job" than the fathers of non-members. Union members were also more likely to be foreign born (4 p.p.). This same pattern holds when city fixed effects are included in regressions comparing union and non-union men (results not shown), implying that the differences are not due to differences across cities in union density and worker characteristics. In terms of their industrial distribution, Table 2 reports that union members were disproportionately represented in manufacturing, construction, and transportation, and were under-represented in trade, finance, business, and government. We do not control for industry in our baseline analyses since industry and union status are likely to be jointly determined, but it is worth noting that the selection patterns described above hold even when controlling for broad industrial categories.

Overall, the comparisons of personal and background characteristics strongly suggest that on average union members were negatively selected from the urban male labor force. A subtler question about selection is whether workers were positively selected into union membership *conditional on observables* such as educational attainment (Card 1996, Hirsch and Schumacher 1998). Selection on unobservables is inherently difficult to identify. The work history information in the Palmer Survey does not track union status over time, and so strategies that rely on worker fixed-effects are infeasible. Nonetheless, there are some background variables in the Palmer Survey that are useful in this context, variables that are rarely available in cross-sectional datasets. For instance, one can see whether *fathers'* occupational status, as gauged by the IPUMS *occscore* variable, was positively associated with workers' union status conditional on workers' educational attainment, age, race, foreign-born status, and city of residence ("basic observables"). It was not. Rather, union members' fathers held slightly lower-status occupations when controlling for the sons' basic observables.[15] Additionally, one can use the Palmer Survey's work history information to see whether men who belonged to unions in 1951 were differentially unemployed in 1940 compared to non-members, or whether their first full-time job was differentially high- or low-status conditional on basic observables. In both cases, the answer is no. There is no statistically significant difference between union members and non-members in these aspects of work history after conditioning on basic observables.

---

[14] The highest grade completed variable in the Palmer Survey is not continuous (0, <5, 5-7, 8, 9-11, 12, 1-3 years of college, 4 years of college, or more). We use midpoints of categories when necessary.

[15] We assigned the fathers occupational income scores based on the IPUMS *occscore* variable, which in turn measures median income in each three-digit occupational category in 1950. The conditional difference between union and non-union members' fathers' status is -0.889 (s.e.=0.393) in a regression that includes fixed effects for age, educational attainment, race, foreign-born status, and city. Mean fathers' status is 27.6.

These findings are relevant to the interpretation of how U.S. labor market institutions worked at mid-century and may at first seem counterintuitive. One might hypothesize that if unions forced employers to pay higher wages, then employers would recruit higher quality workers from the non-union pool, leading to a union wage premium that is illusory and due to positive selection (Lewis 1962, p. 327-28). It is possible, however, that this mechanism takes a long time to play out such that positive selection was not evident by 1951, or that unions were so prevalent at this time that a strategy of systematically replacing current employees with higher quality workers from the non-union pool was infeasible. In any case, as discussed above, there is no evidence of substitution toward more skilled workers in the union sector based on the workers' educational attainment, their family background, or their employment history. Although limited in precision, there is some evidence in the Palmer data that better educated workers tended to leave industries that experienced sharp increases in unionization during the 1940s. At the same time, workers who joined those industries during the 1940s were not better educated than those who had already worked there in 1940 (and still worked there in 1950), conditional on age and city.[16]

From a different perspective, the mid-century pattern of negative selection into union status makes sense because unions are generally believed to compress wage structures by standardizing wages and reducing returns to skill (Freeman and Medoff 1984). In the Palmer sample, the wage premium for high school graduates relative to non-graduates in the union sector was only 4 log points, compared to 16 log points among those not in unions, controlling for age, race, foreign birth, and city fixed effects, and assigning top-coded earnings \$125 per week (the topcode is \$100). Ceteris paribus, following the Roy model intuition mentioned above, this would increase the incentive for highly skilled workers to avoid or select out of union jobs. To our knowledge, this interpretation is not prevalent in the literature on unions circa 1950, but it is plausible and consistent with our findings. It is also consistent with findings reported by Schmidt and Strauss (1976) and Lee (1978) in their analyses of unions and wages in 1967, with more recent studies of selection out of organizations with compressed wage structures (Burdin 2016; Abramitzky 2008, 2009), and with some of the literature on international migration, as reviewed by Abramitzky and Boustan (2016).

[16] The industries with large changes in union density were metals, textiles, leather products, and transportation, communication, and utilities (Troy 1957). For the subset of men who worked in these industries in 1940, a regression of education on indicators for age, city, and "leave" (=1 for those who left that set of industries by 1950) yields a positive coefficient on "leave" (0.40, s.e.=0.22). Omitting from the sample leavers who were union members in 1951 yields a larger coefficient (1.17, s.e.=0.28). Separately, for the subset of wage earners who worked in these industries in 1950 and belonged to a union, the coefficient on "joiner" (=1 if worked in other industries in 1940) is small and statistically insignificant (-0.23, s.e.=0.28).

4. Union and non-union wages

Measuring the union wage premium

Our approach to characterizing the union wage premium emphasizes differences in the distributions of union and non-union wages, after adjusting for differences in men's observable characteristics. As mentioned above we follow re-weighting approaches developed by DiNardo, Fortin, and Lemiuex (1996), Hirano, Imbens, and Ridder (2003), and Firpo (2007). For our purposes, this analytical approach is useful for three reasons. First, the previous section showed that union members differed from non-union members on dimensions that were relevant to earnings, most notably educational attainment. Taking account of these differences is empirically important when comparing workers' wages. Second, we will see below that the union wage premium in the Palmer data is not sufficiently characterized by conditional differences at the middle of the distribution; a better understanding comes from a distribution-wide perspective. Third, approximately 14 percent of the men in the sample have top-coded weekly earnings (at $100). Whereas analysis of mean differences by OLS would require a strong assumption about the earnings of those with top-codes, the analysis of quantiles below the top-code does not.

Most of the results we report come from comparing various quantiles of the union earnings distribution to quantiles of the non-union earnings distribution, where both distributions are adjusted for differences in characteristics relative to the full sample of male workers. Figure 2 provides an illustration of how the technique works, and the appendix describes the technique more formally. Figure 2A shows (i) the observed density of earnings for *union* workers and (ii) a counterfactual density of union earnings created by "weighting-up" union observations that have characteristics similar to those most frequently observed in the overall sample of male workers and "weighting-down" union observations that have characteristics that are uncommon in the overall sample of male workers.[17] Figure 2B shows the same plots but for (i) observed *non-union* earnings and (ii) the counterfactual density of non-union earnings re-weighted to have the same distribution of characteristics as the overall sample of male workers.

Ultimately, we are interested in comparing the re-weighted union and non-union earnings distributions. Figure 2C plots these two densities, which are exactly the same as the counterfactual densities in Figures 2A and 2B. Figure 2C also superimposes the differences between the 20th, 50th, and 80th percentiles of each density. Finally, Figure 2D shows the type of results that we emphasize throughout this part of the paper. Instead of plotting earnings densities, Figure 2D plots quantile

[17] Figure 2 shows kernel density estimates using a Gaussian kernel and Silverman's rule of thumb bandwidth. Note also that the second hump, with log(earnings)=4.6, is actually the top-coded earnings at $100/wk.

treatment effects (QTEs)—the difference between quantiles of the re-weighted distributions.[18] In essence, the horizontal lines in Figure 2C become vertical distances plotted in Figure 2D. From these, one can easily see that after adjusting for observed characteristics, the union earnings premium was largest in the lower part of the distribution. It decreases substantially throughout the distribution and is somewhat negative by the 80th percentile. Above the 80th percentile, results are not available due to the top-coding of wages.[19]

As described above, the comparison of union and non-union wages has a long tradition in labor economics, but this setting does not lend itself to strong causal inference about the effect of union membership. The central challenges to a strong causal interpretation of the union wage premium are (1) unobserved characteristics of workers or jobs that might be correlated with union status, (2) potential endogeneity of some covariates to union status or expectations thereof, and (3) potential spillovers in the form of threat effects (whereby non-union firms raise their wages to reduce their workers' likelihood of organizing or leaving), crowding effects (whereby unions' demands for higher wages displace workers into the non-union sector, depressing wages in that sector), or other general equilibrium effects (e.g., whereby demand shifts toward products from non-union firms).[20] In the presence of such spillovers, the union wage premium would still be an accurate measure of wage differences between observationally similar workers circa 1950, but caution would have to be attached to consideration of counterfactuals in which unions do not exist. Below, we refer to "quantile treatment effects" to be clear that we are following methods developed in the econometrics

---

[18] We focus on quantile comparisons in the same vein as "unconditional quantile treatment effects," as defined in Doksum (1974) and Lehmann (1975). See Firpo (2007) and Frolich and Melly (2013) for recent treatments. Despite similar terminology, QTEs are fundamentally different from the results of *quantile regression*, where the coefficient on a binary treatment variable is the difference between the *conditional* quantiles for treated and untreated individuals. For example, workers with high education in the lower part of the conditional distribution may be in the middle or even upper part of the overall earnings distribution. Instead, our results correspond to workers who are actually in the lower or upper part of the distribution of earnings. Abstracting from issues related to selection into union membership, QTEs correspond to comparing the distribution of earnings for workers randomly assigned to be in a union to workers randomly assigned not to be in a union.

[19] There are advantages to reporting QTEs relative to plotting the densities. First, it is easier to see effects at different points in the distribution from the QTEs than from density plots. Second, plotting densities requires choosing a bandwidth to smooth the data. In practice, results can be sensitive to this choice. QTEs do not require smoothing, so there is no concern about bandwidth selection.

[20] Two additional measurement concerns deserve mention. First, misclassification of union status can bias regression estimates of the union wage premium toward zero (Freeman 1984, Card 1996, Farber and Western 2001, Hirsch 2004). This concern is less pressing in settings where union density approaches 50 percent, as is the case in the Palmer data. Moreover, the Palmer Survey questions were posed directly to the workers and pertained to current status; so, we expect a low level of misreporting. Second, the modern literature on unions relies heavily on CPS data, and the CPS imputes income to some workers. Because the CPS imputation does not match donors on union status, the procedure attenuates differences between union and non-union members (Hirsch and Schumacher 2004). In the Palmer data, there is no imputation of missing wage data.

literature under that name, but we do so with appropriate caution regarding causal inference.[21]

Another measurement concern is that the union wage premium might not fully capture the compensation premium, which would include things like health insurance and pensions. Unfortunately, as mentioned above, the Palmer Survey did not inquire about non-wage compensation. We located no direct comparisons of the value of union versus non-union workers' nonwage compensation circa 1950, but it is clear that unions actively negotiated for improvements in such compensation and that it became more important over the 1940s, albeit from a low base. In the late 1940s and early 1950s, employer supplements to wages and salaries were about 5 percent of employees' total compensation (Bauman 1970, chart 2). To the extent that union members enjoyed better benefits than others circa 1950, the union wage premium may understate the compensation premium.

Baseline results and further exploration

The first set of results illustrates differences between the quantiles of union and non-union workers' earnings, adjusted for differences in covariates across the two groups of workers. The covariates are age (cubic), race, marital status, foreign-born status, city of residence, World War II veteran status, educational attainment (less than high school, some high school (grades 9-12), or at least some college), and whether the person had resided in the area for less than 10 years. We first describe estimates of the union wage premium in the full sample. Then, for comparison, we focus on two groups of particular interest—men with low levels of education and African-Americans. Both groups tend to be close to the bottom of the U.S. income distribution and their outcomes are of particular interest to scholars and policy makers concerned with poverty and inequality. The results may provide a quantitative sense of whether union membership was particularly valuable to such men.

Figure 3 plots baseline estimates of the QTEs—this replicates the information from Figure 2 but rescales it for easier comparison with subsequent figures. Clearly, the largest union wage premiums were concentrated in the lower part of the earnings distribution. The difference between the 10th percentiles of adjusted union earnings and non-union earnings was 20.3 log points. The difference at the medians was smaller, at 2.9 log points (p-value .104), and the difference between

[21] DiNardo and Lee (2004) use a regression discontinuity approach to measure union effects on firms' wages in settings where union representation elections were just barely won or lost after 1984. They find no effect on firm-level average wages. Frandsen (2012) follows a similar approach and finds sizable individual-level wage effects lower in the wage distribution. We are not aware of similar data for mid-century.

the 80th percentiles is negative 5.4 log points.[22] The basic result illustrates how important it is in this case to look at the distributions rather than just the median or mean differences.

We can explore further to see whether the baseline results are due to workers systematically sorting into different kinds of jobs. Workers may select into union status and job types simultaneously. Therefore, controlling for job type does not necessarily leads to a preferable estimate of the union wage premium, let alone a causal estimate of the effect of unionization. Rather, the idea is simply to see whether the baseline patterns are largely driven by differences in the distribution of workers across job types. To start, we add indicators for four broad occupational categories and four broad industrial categories.[23] Although the occupation and industry controls are useful predictors of union membership, including them has only a small effect on estimates of the union premium. For comparison, Table 3 reports the baseline results (column 1A) and results with occupation and industry controls (column 1B). At the 10th percentile, union earnings were 18.2 log points higher than non-union earnings (2.1 log points lower than without occupation and industry controls). At the median, the earnings premium was 5.9 log points (3.0 log points higher than in the baseline). At the 80th percentile, the union premium was zero.

Next, we collected data on detailed industry-specific injury rates in 1949 and 1950 (U.S. Department of Labor 1952), and mapped that information to 1950 industry codes.[24] The goal is to see whether the union wage premium reflected a particular disamenity—the risk of injury—that might be correlated with higher rates of unionization. It does appear that on average unionized men in the Palmer data worked in industries with higher injury rates (17.8 compared to 14.2 per million employee hours). We also calculated median weekly hours of work for each three-digit industry from the 1950 census microdata (Ruggles et al. 2015) to control for potential differences in length of the work week between union and non-union workers. Not every industry is covered by the injury rate data, and so column 2A of Table 3 reports results for a reduced sample of men. The specification in 2A includes controls for broad occupation and industry categories, and so the

---

[22] For reference, ad hoc assignments of $100 or $150 to topcoded earners in OLS regressions result in union wage gaps of 5.3 or 3.7 log points, respectively.

[23] Occupation categories are: professional, managers, clerical, and sales; craftsmen and similar (plus a small number of farmers); operatives and some service jobs; laborers, farm laborers, and some low-skill service jobs. Industry categories are: manufacturing and mining (plus a small number in agriculture); construction and transportation/utilities; wholesale and retail trade, personal service; and finance, business, government.

[24] The injury frequency rate is the average number of disabling work injuries for each million employee-hours worked, where a disabling injury is one that results in death, permanent physical impairment, or makes a worker unable to perform job duties on one or more days after the injury (U.S. Department of Labor 1952, p. 1). Injury rates are the industry-level average of the 1949 and 1950 rates reported in U.S. Department of Labor (1952); 1950 railroad industry rates are taken from U.S. Department of Commerce (1976, p. 388).

differences relative to column 1B are due solely to the change in sample composition. Column 2B adds the controls for injury rates and weekly hours. This has little effect on the estimated union wage premium pattern.

In addition, and again for a reduced sample, we can add controls for background characteristics that are typically unobservable in cross-sectional datasets, such as whether the worker was unemployed in 1940 (if in the labor force), the occupational status of his first full-time job, and his father's occupational status (if reported). Occupational status is measured according to the IPUMS *occscore* variable, which in turn is keyed to median occupation-specific income in the 1950 census (Ruggles et al. 2015). In principle, these background characteristics might capture aspects of selection and productivity that the baseline specification for the full sample does not. Results are given in columns 3A (reduced sample with baseline specification) and 3B (reduced sample with additional control variables). Throughout the distribution, the QTE results are similar with or without the additional background controls. None of the estimates in column 3B is statistically different from the analogous estimate in column 3A. [25]

Results for men with low educational attainment and for African Americans

Figure 4 plots QTEs for the subset of workers who did not attend high school. This group makes up 41 percent of the sample. Its members tended to be older, were more likely to be foreign-born or African American, and earned substantially less on average than men with more education. To be clear, the results in Figure 4 compare the distribution of union earnings to the distribution of non-union earnings for the subsample of workers with less than a high school education; union and non-union men with less than a high school education are both re-weighted to have the same distribution of characteristics as all workers with less than a high school education. Similar to the results for the full sample, for less educated workers, the union earnings premium was largest in the lower part of the distributions. At the 10th percentile, union workers earned 19.1 log points more than comparable non-union workers. The wage gap decreases in the middle and upper part of the distributions, but it does not fall to zero. The difference in median earnings for union and non-union, less educated men was 8.0 log points; the difference at the 90th percentiles was 4.3 log points. Thus, the evidence is consistent with the hypothesis that less-educated men benefited substantially from

[25] Obtaining similar QTE estimates, however, does not imply that remaining unobservables are unimportant. As illustrated by Oster (2016) in an OLS setting, coefficient stability is not by itself indicative of robustness to omitted variable bias. In the QTE setting, the additional background characteristics are jointly significant correlates of union status (in logit) and earnings conditional on union status (in OLS), but as shown in Table 3, the augmented specification yields very similar results to the baseline.

union membership.

Figure 5 plots estimates of QTEs for black workers. Again, the union wage premium was relatively large, and it remained so throughout the black income distribution. At the 10th percentile, the difference between union earnings and non-union earnings was 26.5 log points for workers with similar observed characteristics. At the median, the difference was 14.8 log points, and at the 90th percentile the difference was 12.3 log points. Because the sample contains a relatively small number of black workers, the confidence intervals are wider than in previous figures, but nearly all the estimates are statistically significant. Figure 5's clear implication is that black men who gained entry to union jobs earned more than observationally similar black men who did not.[26]

This finding is important because the history of unions is replete with examples of racial discrimination and exclusion (Northrup 1944, Hill 1967, Zieger 2007). The emergence of industrial unions after 1935, which sought to unionize production workers along industry rather than craft lines, likely opened more union job opportunities for African Americans, as did the labor demand shock of World War II in combination with federal anti-discrimination policies (Collins 2001). Because unions tended to standardize wages without regard to race, it is plausible that black men in unions earned pay that was far higher than their non-union peers. This is consistent with the longstanding efforts of Civil Rights organizations to pry open access to union jobs (Zieger 2007). It also possible, however, that selection on unobservables into union jobs was more prevalent for black workers, or that black workers in unions were paid compensating differentials for especially unpleasant aspects of work (c.f., Foote, Whatley, and Wright 2003). These are interesting questions for future research.

Results within detailed sub-groups

Space and sample size do not allow detailed QTE descriptions for every subgroup of the labor force. Figure 6, however, conveys a sense of how much heterogeneity there was in the union wage gap at mid-century. It reports QTEs and confidence intervals at the 20th percentile (left panel) and median (right panel) for 12 different groups defined by interactions of education and age. It also reports separate estimates for white, black, blue-collar, and white-collar workers. All the estimates are conditional on age, city, race, education, years residing in the area, marital status, foreign-born status, and veteran status (omitting covariates when they are used to define the group).

---

[26] See Ashenfelter (1972) for a careful discussion of race and unionism, which concludes that circa 1967 industrial unions tended to narrow black-white wage differentials whereas craft unions tended to do the opposite. He too finds that the union/nonunion wage gap was relatively large for black men (p. 450-51). Lee (1978, Table 3) also finds a relatively large union premium for nonwhite men.

For those with less than a high-school education, blue-collar workers, and black workers, the union wage premium tended to be large at the median and even larger at the 20th percentile. For those with some high school education (or more) and white-collar workers, the union wage gaps were smaller and point estimates are sometimes negative. For the subsample of all white men, the union wage premium at the median was zero, though it was significant and positive at the 20th percentile. For reference, the full set of median results with additional information on sub-sample sizes, union membership rates, and wage levels are reported in Appendix Table 3.

When the sample is split and QTEs are estimated separately for each city, the basic pattern of a large union premium at low quantiles, a small premium at the median, and an even smaller or negative premium near the top quantiles generally holds (Appendix Table 4). There are notable differences in point estimates across cities, but the estimates are often imprecise, which makes it difficult to make inferences from the cross city comparisons.

5. Unions and inequality at mid-century

As noted above, the peak of union density in the U.S. coincided with the low point of twentieth-century inequality. This might have been entirely coincidental. For instance, it is likely that the relative supply and demand for skilled workers was a key determinant of the "Great Compression" (Goldin and Margo 1992). But it is also plausible that the rise of unions and their tendency to compress the wage distribution mattered. This would be consistent with studies of the fall of unions and the rise of inequality in recent decades, as well as with the observations of Miller (1958), Frydman and Molloy (2012), and Collins and Niemesh (2016) regarding mid-century wages.

A first clue regarding the effect of unions on overall inequality comes from skill-group differences in median earnings and their correlation with the union wage premium. The discussion above noted that groups that earned relatively low wages (e.g., those with less than high school education, black men, those in blue-collar occupations) also had large union wage premiums. This is consistent with unions moving the earnings of some workers in low-wage groups closer to the middle of the overall distribution, tending to reduce inequality.

To develop more direct evidence, we modify the reweighting approach outlined in the previous section to address the following question: How different would overall inequality have been in 1950 (in the Palmer sample) if union workers had been paid according to the non-union wage schedule? This entails keeping the non-union workers' wages fixed, but replacing the union workers' wages with a counterfactual distribution. The counterfactual for union workers maintains their characteristics (such as age, education, and race), but the wage structure corresponds to that of non-union workers with similar characteristics. The combination of the actual non-union and counterfactual union workers' wage distributions is then a "no unions" counterfactual distribution for the full sample. This approach cannot take into account general equilibrium effects, and it is subject to the caveats expressed above regarding the absence of randomized union status and the potential endogeneity of some covariates. But it does provide a quantitative sense of the importance of differences in union versus nonunion wages in comprising the overall wage distribution and its level of inequality.

We take the difference between the 80th percentile and the 10th percentile as a simple measure of overall inequality, the difference between the 50th percentile and the 10th percentile as a measure of lower tail inequality, and the difference between the 80th percentile and 50th percentile as a measure of upper tail inequality. Recall that top-coding makes it difficult to analyze the earnings distribution above the 80th percentile.

Table 4 reports the results. In Panel A, wages for union members are based on the wages for non-union members adjusted for differences in age, race, education, years residing in the area, marital status, veteran status, and city of residence—key traits that workers bring to the labor market. Overall (80-10) inequality is 21 percent (=0.150/0.715) higher in the counterfactual "no union" scenario than in reality. Most of the increase in inequality occurs in the lower part of the distribution. The 50-10 differential is 24 percent (=0.103/0.427) larger in the "no unions" counterfactual than in the actual distribution, reflecting the relatively large union wage premium at the low end of the wage distribution compared to the median. In Panel B, we add adjustments for broad occupation and industry categories, so that the counterfactual distribution of workers' wages maintains the same broad occupation and industry mix as the actual distribution. Although informative, it is not clear that Panel B is preferable to Panel A's thought experiment because workers' choices of occupation and industry are endogenous to the wage structure. Again, the results are consistent with unions reducing the overall level of wage inequality, yielding a combined 80-10 inequality change of 13 percent. In sum, it seems reasonable to conclude that unions circa 1950 tended to narrow the 80-10 wage dispersion by 13 to 21 percent in the set of cities for which we have data, subject to the caveats mentioned above.

6. Comparing the 1950s and 1970s

We noted earlier that the CPS first collected data on union status in May 1973, and this has been the starting point for much of the modern literature on unions and wages. The Palmer data provide an opportunity to bridge 1950 to the 1970s, spanning most of the period when the American labor movement was at its height but starting to decline. The comparisons of Palmer and CPS data are imperfect due to differences in sample design, variables' definitions and availability, and geographic coverage. But we believe the datasets can provide useful first-order comparisons. We use the May 1973 CPS file that is available from the National Bureau of Economic Research.[27]

Despite many changes in the US economy between 1950 and 1973, the basic patterns that we have described above for the Palmer data are also apparent in the 1973 CPS data when restricted for comparability to men, ages 25 and over, who worked full time for wages or salary, and resided in metropolitan areas outside the South. Thus, after the sharp rise in union membership during the late 1930s and 1940s, it appears that unions' basic imprint on the labor market's wage structure circa 1950 endured until at least the mid-1970s. We cannot, of course, rule out fluctuations in the meantime, and we are mindful of Lewis's (1963) suggestion that the union premium might have increased over the 1950s (also see Pencavel and Hartsog 1984). For the sake of brevity, we will not describe results for the 1973 at length, but we will point out particularly interesting features in light of the earlier discussion.

Table 5 reports characteristics of union and non-union workers in the 1973 CPS, where the sample is restricted as described above. In the wake of the high school movement (Goldin and Katz 2008), the average level of education for U.S. workers increased substantially between 1950 and 1973, but the education gap between union and non-union workers in 1973 was comparable to that in 1950—a little less than two years. It is notable that the typical union member in 1973 was a high school graduate, whereas in 1950 only about 30 percent of union members had four years of high school. By 1973, black workers were somewhat over-represented among union members, another notable change from 1950, reflecting in part decades of civil rights groups' efforts to open union jobs to black workers but also the movement of better educated whites out of union-intensive sectors.

Figure 7 shows QTEs of union membership on usual weekly wages in the 1973 data. The downward slope is reminiscent of Figure 3 for 1950, where the union wage gap was largest at the lower quantiles and drifted toward zero at higher quantiles. Thus, the general pattern of union wage premia in 1950 was roughly similar to that observed in 1973. The largest and only statistically significant difference between the estimates for 1950 and 1973 is at the 10th percentile where the union premium is estimated to be 15.4 log points in 1973, which is 4.9 log points smaller than the corresponding estimate for 1950.

The estimated union wage premium at the median is low compared to estimates of the

[27] The data are posted at: http://www.nber.org/data/cps_may.html (accessed September 9, 2016).

average premium found elsewhere in the literature on the 1970s (Freeman and Medoff 1984, Card 2001). Hirsch and Macpherson (2016), for example report a union wage premium of about 16 log points in 1973. Investigation reveals that the sample restrictions we impose to improve comparability with the Palmer data—omitting those under age 25, part-time workers, residents of non-metropolitan areas, and residents of the South—substantially reduce the estimated average wage premium in the CPS data analyzed with OLS (to approximately 3 log points).

Despite the visual similarity of the QTE patterns in 1950 and 1973, the difference between counterfactual and actual inequality levels appears to have been somewhat smaller in 1973 than in 1950. Following a procedure similar to that described in the previous section, our baseline estimates indicate that in a counterfactual without unions, 80-10 inequality would have been 6.4 log points (or 7.5 percent) higher than what was actually observed in 1973. This is smaller than the most comparable results for 1950 (Table 4), where unions were associated with a decline in inequality by 15.0 log points (21 percent in the baseline specification). The difference reflects the smaller fraction of workers who were union members in 1973, the smaller union wage premium at the low end of the distribution, and the higher level of inequality (implying a larger denominator when expressed in percentage terms). We remind readers that the 1950 and 1973 data are not perfectly comparable, and so the differences between 1950 and 1973 should not be overemphasized. Rather, the main patterns appear to be fairly consistent across the years, even if unions' influence on inequality appears somewhat weaker by 1973.

7. Conclusions

This paper examines a novel dataset, originally collected in early 1951, that we retrieved from archival sources. The extant "transcription sheets" from the survey cover only five U.S. cities, but we find that workers in those cities were observationally similar to those in a sample of residents from all non-southern cities drawn from census records. To our knowledge, the survey's combination of data on union status, weekly wages, city of residence, and extensive background information is unique for the time. It provides a new view of workers, their wages, and union membership at the height of the American labor movement.

The peak of unionization coincided with the low-point of American wage inequality at mid-century. The results in this paper suggest that this was not merely a coincidence. Selection into unions was negative in the sense that union workers had lower levels of education and had fathers with lower levels of occupational status than non-union workers. Union workers appear to have earned a wage premium in comparison with observationally similar men below the median of the wage distribution. Based on this evidence, counterfactuals suggest that the overall wage distribution was considerably narrower in 1950 than it would have been if union members had been paid like non-union members with similar characteristics. We caution that selection on unobservables could confound some of these interpretations. But it is worth reiterating that conditional on what we can observe, there is no evidence that men who belonged to unions in 1951 came from better off families, had better employment outcomes in 1940, or had better jobs when they first entered the labor force than others.

Our historical interpretation is that in the wake of the Great Depression, workers sought and policymakers delivered institutional reforms to labor markets that promoted unions, reduced inequality, and helped lock in a relatively narrow distribution of wages that lasted for a generation. There is some evidence, albeit imperfect, that by 1950 better educated workers were sorting out of industries that had experienced large changes in union density during the 1940s, which is consistent with seeking higher rates of return to investments in education. But a better understanding of the dynamics and the implications of worker sorting in this period awaits future research.

By 1973, when the CPS began collecting information on union status, the prevalence of unions had started to decline in the United States. Nonetheless, the basic patterns we observed for 1950 were still discernable in the 1973 data. Male union members still had substantially lower levels of education than non-members even though the entire distribution of educational attainment had shifted in a positive direction after 1950. The union wage premium was still relatively large at lower quantiles in the earnings distribution. Finally, the evidence was still consistent with unions compressing the overall wage distribution relative to a counterfactual in which union workers are paid like similar non-union workers.

In the modern literature's broad discussion of rising inequality, the supply and demand for skills has taken center stage, but institutional factors appear to have been important as well (DiNardo, Fortin, and Lemieux 1996; Card, Lemieux, and Riddell 2004). With a longer-run view in mind, and at a time when concerns about inequality are salient, the underpinnings of the post-1940 period of wage compression and broadly shared economic growth merit closer examination. Unions were an important and frequently studied feature of the U.S. economy at that time, yet there is still much to learn about their role in shaping workers' outcomes, firms' decisions, and labor-related policy.

References

Abramitzky, Ran. 2008. "The Limits of Equality: Insights from the Israeli Kibbutz." *Quarterly Journal of Economics* 123, 3: 1111-1159.

Abramitzky, Ran. 2009. "The Effect of Redistribution on Migration: Evidence from the Israeli Kibbutz." *Journal of Public Economics* 93: 498-511.

Abramitzky, Ran and Leah Platt Boustan. 2016. "Immigration in American Economic History." NBER Working Paper 21882.

Ashenfelter, Orley. 1972. "Racial Discrimination and Trade Unionism." *Journal of Political Economy* 80, 3: 435-464.

Ashenfelter, Orley and John H. Pencavel. 1969. "American Trade Union Growth: 1900-1960." *Quarterly Journal of Economics* 83, 3: 434-448.

Bauman, Alvin. 1970. "Measuring Employee Compensation in the United States." *Monthly Labor Review* 93, 10: 17-24.

Blanchflower, David and Alex Bryson. 2004. "What Effect Do Unions Have on Wages Now and Would Freeman and Medoff Be Surprised?" *Journal of Labor Research* 25, 3: 383-414.

Borjas, George J. 1987. "Self-Selection and the Earnings of Immigrants." American Economic Review 77, 4: 531-53.

Burdin, Gabriel. 2016. "Equality under Threat by the Talented: Evidence from Worker-Managed Firms." *Economic Journal* 126, 594: 1372-1403.

Campbell, Angus and Gerald Gurin. *American National Election Study 1952 Time Series Study*. ICPSR07213-v4. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2015-11-10. https://doi.org/10.3886/ICPSR07213.v4.

Card, David. 1996. "The Effects of Unions on the Structure of Wages: A Longitudinal Analysis." *Econometrica* 64, 4: 957-979.

Card, David. 2001. "The Effect of Unions on Wage Inequality in the U.S. Labor Market." *Industrial & Labor Relations Review* 54, 2: 296-315.

Card, David, Thomas Lemieux, and W. Craig Riddell. 2004. "Unions and Wage Inequality." *Journal of Labor Research* 25, 4: 519-559.

Collins, William J. 2000. "African American Economic Mobility in the 1940s: A Portrait from the Palmer Survey." *Journal of Economic History* 60, 3: 756-781.

Collins, William J. 2001. "Race, Roosevelt, and Wartime Production: Fair Employment in World War II Labor Markets." *American Economic Review* 91, 1: 272-286.

Collins, William J. and Gregory T. Niemesh. 2016. "Unions and the Geography of the Great Compression of American Inequality, 1940-1960." Unpublished working paper.

DiNardo, John and David S. Lee. 2004. "Economic Impacts of New Unionization on Private Sector Employers: 1984-2001." *Quarterly Journal of Economics* 119, 4: 1383-1441.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64, 5: 1001-1044.

Doksum, Kjell. 1974. "Empirical Probability Plots and Statistical Inference for Nonlinear Models in the Two-Sample Case." *The Annals of Statistics* 2, 2: 267-277.

Epple, Dennis and Thomas Romer. 1991. "Mobility and Redistribution." *Journal of Political Economy*

99, 4: 828-858.

Farber, Henry S. and Bruce Western. 2001. "Accounting for the Decline of Unions in the Private Sector, 1973-1998." *Journal of Labor Research* 22, 3: 459-485.

Feldstein, Martin and Marian Vaillant Wrobel. 1998. "Can State Taxes Redistribute Income?" *Journal of Public Economics* 68: 369-396.

Firpo, Sergio. 2007. "Efficient Semiparametric Estimation of Quantile Treatment Effects." *Econometrica* 75, 1: 259-276.

Foote, Christopher L., Warren C. Whatley, Gavin Wright. 2003. "Arbitraging a Discriminatory Labor Market: Black Workers at the Ford Motor Company, 1918-1947." *Journal of Labor Economics* 21, 3: 493-532.

Frandsen, Brigham R. 2012. "Why Unions Still Matter: The Effects of Unionization on the Distribution of Employee Earnings." Unpublished working paper. MIT.

Freeman, Joshua. 1978. "Delivering the Goods: Industrial Unionism during World War II." *Labor History* 19: 570-593.

Freeman, Richard B. 1980. "Unionism and the Dispersion of Wages." *Industrial & Labor Relations Review* 34, 1: 3-23.

Freeman, Richard. 1998. "Spurts in Union Growth: Defining Moments and Social Processes." In Michael D. Bordo, Claudia Goldin, and Eugene N. White (eds.), *The Defining Moment: The Great Depression and the America Economy in the Twentieth Century*, pp. 265-295. Chicago, IL: University of Chicago Press.

Freeman, Richard and James L. Medoff. 1984. *What Do Unions Do?* New York, NY: Basic Books.

Friedman, Milton. 1962. *Capitalism and Freedom*. Chicago, IL: University of Chicago Press.

Frolich, Markus and Blaise Melly. 2013. "Unconditional Quantile Treatment Effects under Endogeneity." *Journal of Business and Economic Statistics* 31, 3: 346-357.

Frydman, Carola and Raven Molloy. 2012. "Pay Cuts for the Boss: Executive Compensation in the 1940s." *Journal of Economic History* 72, 1: 225-251.

Goldin, Claudia. 1991. "The Role of World War II in the Rise of Women's Employment." *American Economic Review* 81, 4: 741-756.

Goldin, Claudia and Lawrence F. Katz. 2008. *The Race between Education and Technology*. Cambridge, MA: Harvard University Press.

Goldin, Claudia and Robert A. Margo. 1992. "The Great Compression: The Wage Structure in the United States at Mid-Century." *Quarterly Journal of Economics* 107, 1: 1-34.

Hammond, Cordy and James C. Nix. 1953. "Union Status Provisions in Collective Agreements, 1952." *Monthly Labor Review* 76, 4: 383-387.

Hill, Herbert. 1967. "The Racial Practices of Organized Labor—The Age of Gompers and After." In Arthur M. Ross and Herbert Hill (eds.), *Employment, Race, and Poverty*, pp.365-402. New York: Harcourt, Brace, & World.

Hirano, Keisuke, Guido W. Imbens, Geert Ridder. 2003. "Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score." *Econometrica* 71, 4: 1161-1189.

Hirsch, Barry T. 2004. "Reconsidering Union Wage Effects: Surveying New Evidence on an Old Topic." *Journal of Labor Research* 25, 2: 233-266.

Hirsch, Barry T. and David A. Macpherson. 2016. *Union Membership and Earnings Data Book.*

Arlington, VA: Bureau of National Affairs, Inc.

Hirsch, Barry T. and Edward J. Schumacher. 1998. “Unions, Wages, and Skills.” Journal of Human Resources 33, 1: 201-219.

Hirsch, Barry T. and Edward J. Schumacher. 2004. “Match Bias in Wage Gap Estimates Due to Earnings Imputation.” Journal of Labor Economics 22, 3: 689-722.

Kersten, Andrew E. 2006. *Labor’s Home Front: The American Federation of Labor during World War II*. New York: New York University Press.

Lee, Lung-Fei. 1978. “Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables.” *International Economic Review* 19, 2: 415-433.

Lehmann, Erich L. 1975. *Nonparametrics: Statistical Methods Based on Ranks*. Holden-Day.

Lester, Richard A. 1954. *Hiring Practices and Labor Competition*. Princeton, NJ: Industrial Relations Section, Princeton University.

Lester, Richard A. 1964. *Economics of Labor*. New York: The MacMillan Company.

Lewis, H. Gregg. 1962. “The Effects of Unions on Industrial Wage Differentials.” In National Bureau of Economic Research, *Aspects of Labor Economics*, pp. 319-344. Princeton, NJ: Princeton University Press.

Lewis, H. Gregg. 1963. *Unionism and Relative Wages in the United States: An Empirical Inquiry*. Chicago, IL: University of Chicago Press.

Lewis, H. Gregg. 1986. *Union Relative Wage Effects: A Survey*. Chicago, IL: University of Chicago Press.

Lichtenstein, Nelson. 1982. *Labor’s War at Home: The CIO in World War II*. Cambridge, UK: Cambridge University Press.

Miller, Herman P. 1958. “Change in the Industrial Distribution of Wage 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: Princeton University Press.

Millis, Harry A. and Emily Clark Brown. 1950. *From the Wagner Act to Taft-Hartley: A Study of National Labor Policy and Labor Relations*. Chicago, IL: University of Chicago Press.

Northrup, Herbert. 1944. *Organized Labor and the Negro*. New York, NY: Harper.

Oster, Emily. forthcoming. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business Economics and Statistics*.

Palmer, Gladys L. 1954. *Labor Mobility in Six Cities: A Report on the Survey of Patterns and Factors in Labor Mobility, 1940-1950*. New York, NY: Social Science Research Council.

Pencavel, John and Catherine E. Hartsog. 1984. “A Reconsideration of the Effects of Unionism on Relative Wages and Employment in the United States, 1920-1980.” *Journal of Labor Economics* 2, 2: 193-232.

Reder, Melvin W. 1988. “The Rise and Fall of Unions: The Public Sector and the Private.” *Journal of Economic Perspectives* 2, 2: 89-110.

Rees, Albert. 1962. *The Economics of Trade Unions*. Chicago, IL: University of Chicago Press.

Rosen, Sherwin. 1969. “Trade Union Power, Threat Effects and the Extent of Organization.” *Review of Economic Studies* 36, 2: 185-196.

Rosen, Sherwin. 1970. “Unionism and the Occupational Wage Structure in the United States.”

International Economic Review 11, 2: 269-286.

Ross, Arthur M. 1948. "The Influence of Unionism upon Earnings." *Quarterly Journal of Economics* 62, 2: 263-286.

Roy, A.D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3, 2: 135-146.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. *Integrated Public Use Microdata Series: Version 6.0* [Machine-readable database]. Minneapolis: University of Minnesota, 2015.

Saad, Lydia. 2015. "Americans' Support for Labor Unions Continues to Recover." http://www.gallup.com/poll/184622/americans-support-labor-unions-continues-recover.aspx. Accessed May 27, 2016.

Schmidt, Peter and Robert P. Strauss. 1976. "The Effect of Unions on Earnings and Earnings on Unions: A Mixed Logit Approach." *International Economic Review* 17, 1: 204-212.

Seidman, Joel, Jack London, and Bernard Karsh. 1951. "Why Workers Join Unions." *Annals of the American Academy of Political and Social Science* 274 (March): 75-84.

Seidman, Joel. 1953. *American Labor from Defense to Reconstruction*. Chicago: University of Chicago Press.

Sobotka, Stephen P. 1953. "Union Influence on Wages: The Construction Industry." *Journal of Political Economy* 61, 2; 127-43.

Stigler, George J. 1957. "The Tenable Range of Functions of Local Government." In *Federal Expenditure Policy for Economic Growth and Stability: Papers Submitted by Panelists Appearing before the Subcommittee on Fiscal Policy*, November 5, 1957. Joint Economic Committee, Washington, DC.

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

Troy, Leo. 1965. "Trade Union Membership, 1897-1962." NBER Occasional Paper 92.

U.S. Department of Commerce, Bureau of the Census. 1952. *Census of Population: 1950*, Volume II. Washington, DC: GPO.

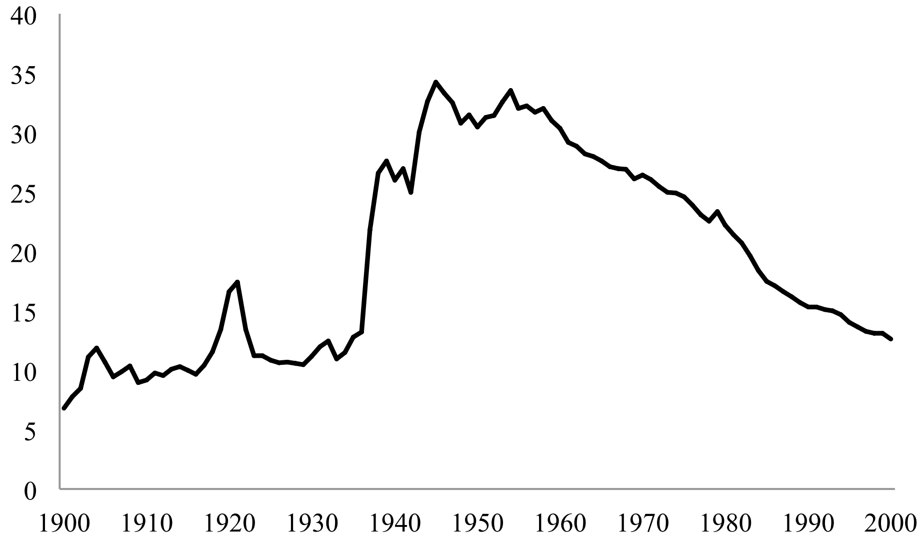
U.S. Department of Labor. 1952. "Work Injuries in the United States During 1950: A Collection of Basic Work-Injury Data for Each of the Major Industries in the United States." Bulletin Number 1098. Washington, DC: GPO.

U.S. Department of Commerce, Bureau of the Census. 1976. *Statistical Abstract of the United States: 1976*. Washington, DC: GPO.

Weiss, Leonard W. 1966. "Concentration and Labor Earnings." American Economic Review 56, 1/2: 96-117.

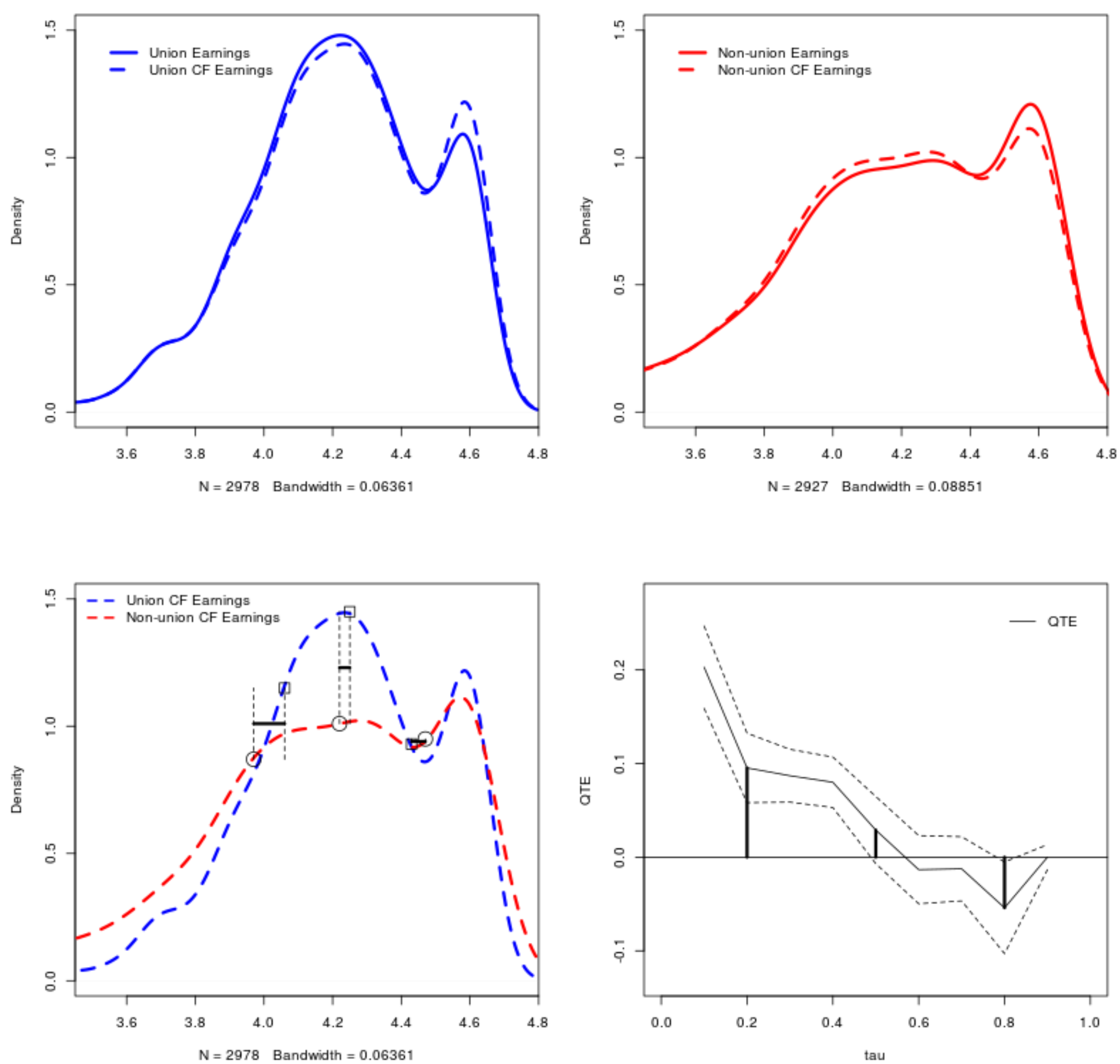
Zieger, Robert H. 2007. *For Jobs and Freedom: Race and Labor in America since 1865*. Lexington, KY: University Press of Kentucky.

Figure 1: U.S. Union Membership as a Share of Non-agricultural Employment



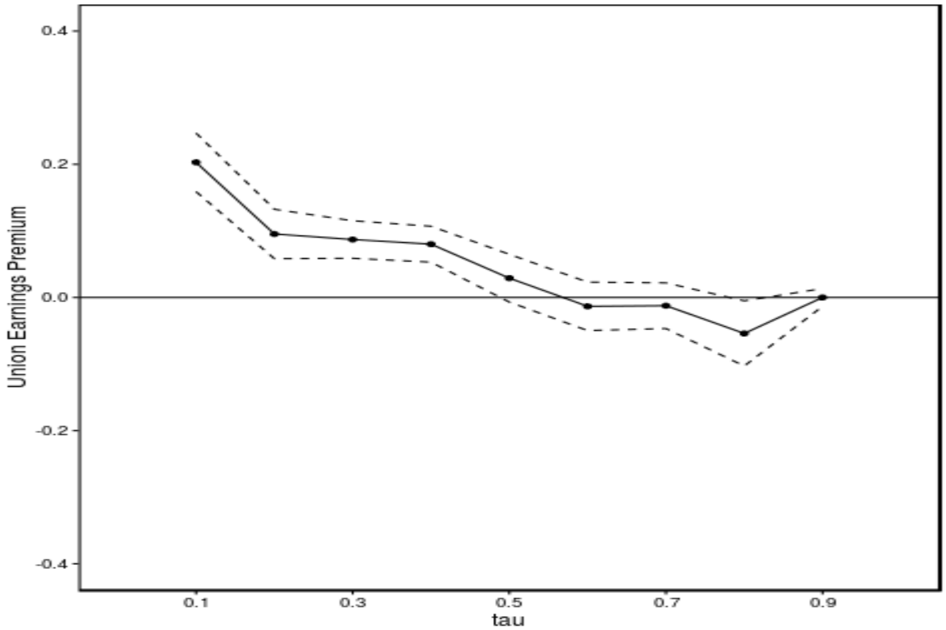
Notes and sources: The series for 1900 to 1995 is from Freeman (1998). For 1995-2000, we used BLS data series LUU0204899600, which pertains to wage and salary workers in all industries (www.bls.gov/data, accessed Sept. 7, 2016). We spliced the BLS series to the Freeman series in 1995.

Figure 2A-D: Constructing QTE Estimates



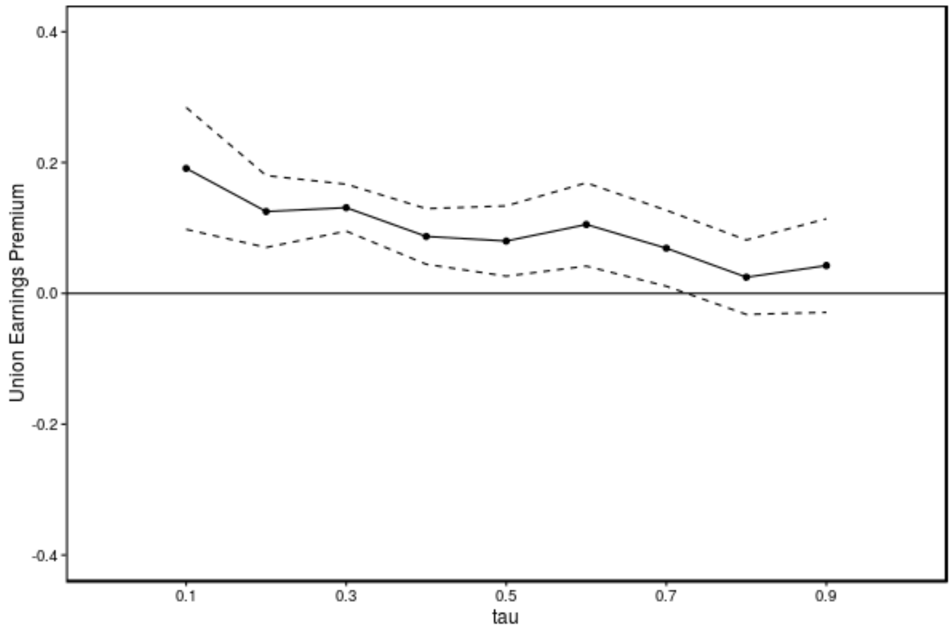
Notes: Panel A (upper left) shows the observed density of earnings for union members and their counterfactual distribution of earnings after re-weighting so that they have the same distribution of observable characteristics as in the population. Panel B (upper right) shows the observed density of earnings for non-union workers and their counterfactual distribution of earnings after re-weighting to have the same distribution of characteristics as the population. See the appendix for more information on reweighting techniques. Panel C (lower left) compares union and non-union counterfactual densities. Circles and squares are placed at the 20th, 50th, and 80th percentiles of each distribution. The horizontal distance between each pair of circles and squares gives the Quantile Treatment Effect (QTE) for $\tau = (0.2, 0.5, 0.8)$, respectively. The plot of the QTE (which is the standard result reported in the rest of the paper) is given in Panel D (lower right). At $\tau = (0.2, 0.5, 0.8)$, the vertical distance in the plot is the same as the horizontal distance in Panel C. The distributions in panels A-C are smoothed for presentation, but the QTE estimates do not require or use smoothing.
Sources: Palmer Survey, as described in the text.

Figure 3: QTE Estimates of the Union Wage Premium for Men



Notes: This figure plots QTEs after re-weighting the distributions of union and non-union earnings so that they have the same distribution of observable characteristics based on a first step estimation of the propensity score. The propensity score is estimated by logit; the dependent variable is union status and the explanatory variables are a cubic in age, a dummy variable for completing at least some high school but no college, a dummy variable for completing at least some college, race dummy variables, a dummy variable indicating whether an individual has lived in the area for less than 10 years, marital status dummy variables, a dummy variable indicating whether an individual was born outside of the U.S., and a dummy variable for whether an individual was a World War II veteran. The analysis uses sampling weights as described in the text. Due to the censoring of earnings, the results are available up to the 80th percentile. The 90th percentile is exactly equal to zero due to top-coding of both union and non-union earnings. Pointwise confidence intervals are obtained using the bootstrap with 1,000 iterations.
Sources: Palmer Survey, as described in the text.

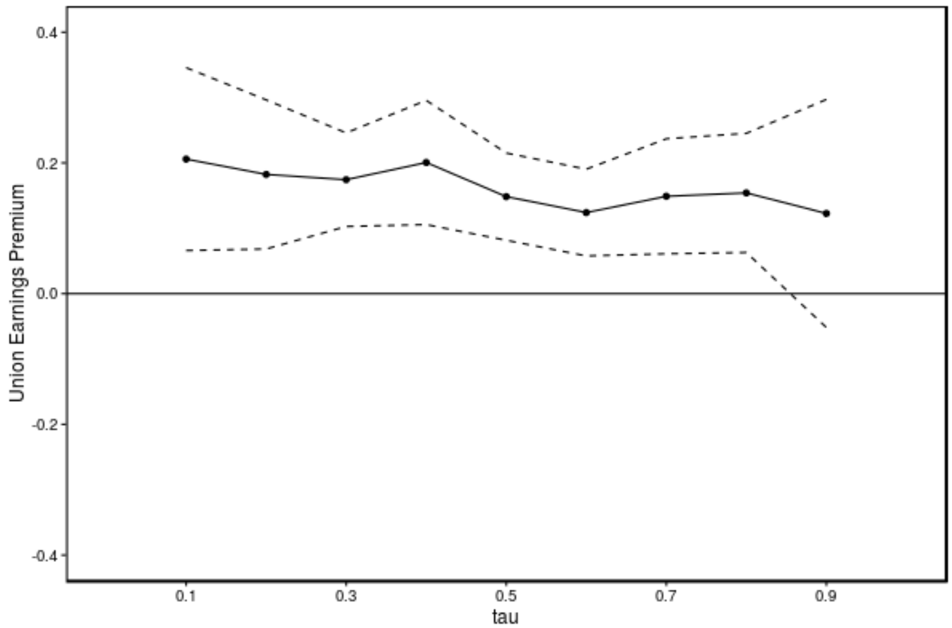
Figure 4: QTE Estimates of Union Wage Premium for Men with Less Than High School Education



Notes: This figure plots QTEs after re-weighting the distributions of union earnings and non-union earnings for less-educated men to be the same as the distribution of observed covariates for all men who never attended high school. The estimates depend on a first step estimation of the propensity score, as described in the text and Figure 3 except that education control variables are not included here. The analysis uses sampling weights as described in the text. Pointwise confidence intervals are obtained using the bootstrap with 1,000 iterations.

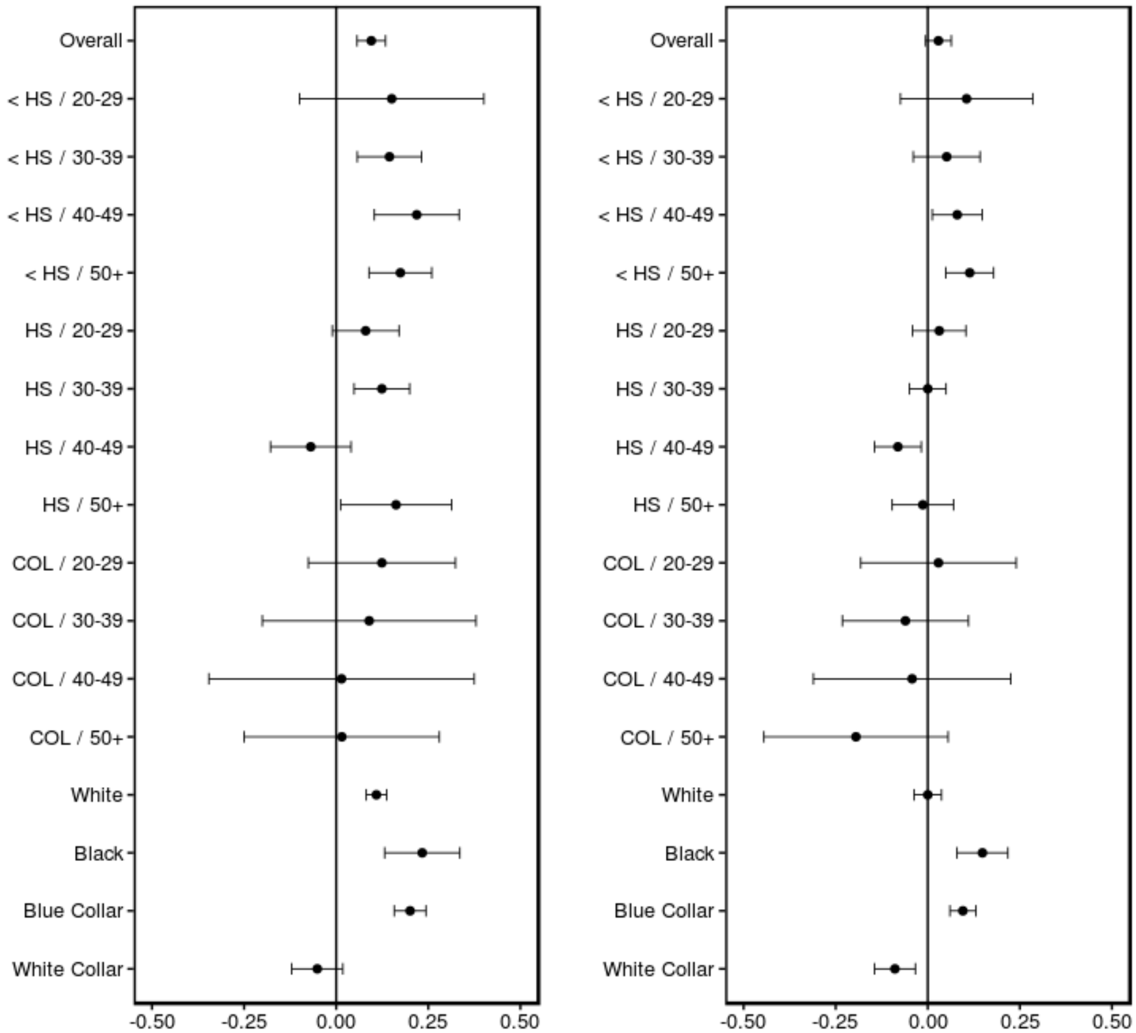
Sources: Palmer Survey, as described in the text.

Figure 5: QTE Estimates of Union Wage Premium for Black Men



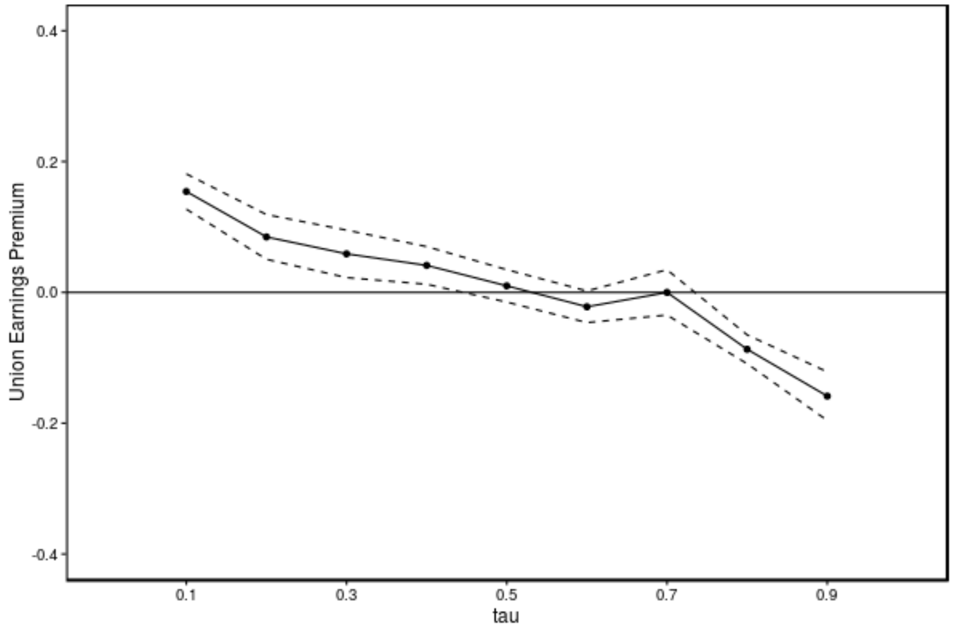
Notes: This figure plots QTEs after re-weighting the distributions of union earnings and non-union earnings for African Americans to be the same as the distribution of observed covariates for all African American men. The estimates depend on a first step estimation of the propensity score, as described in the text and Figure 3 except that race control variables are not included here. The analysis uses sampling weights as described in the text. Pointwise confidence intervals are obtained using the bootstrap with 1,000 iterations.
Sources: Palmer Survey, as described in the text.

Figure 6: QTE Estimates at the 20th Percentile and Median, by Subgroup



Notes: The panel on the left plots QTE estimates at the 20th percentiles of earnings for union and non-union workers by subgroup after adjusting for differences in covariates. The panel on the right does the same at the medians. The results depend on a first step estimation of the propensity score, as described in previous figures with the exception that when a particular covariate is used to form the group, that covariate is omitted from the propensity score specification. The analysis uses sampling weights as described in the text. Standard errors are computed using the bootstrap with 1,000 iterations.
Sources: Palmer Survey, as described in the text.

Figure 7. Estimate of Quantile Treatment Effects in 1973



Notes: The figure plots QTEs for the union earnings premium in 1973. The distributions are calculated based on re-weighting the union earnings distribution and non-union earnings distribution so that each has the same distribution of covariates as in the population of workers in 1973. The sample includes residents of metropolitan areas that are outside of the southern census region. The results depend on a first step estimation of the propensity score that includes a cubic in age, race dummy variables, whether an individual completed some high school, whether an individual completed some college, whether an individual is married, whether an individual is a veteran, and dummy variables for each metropolitan area. The estimates also apply CPS sampling weights.
Sources: CPS, May 1973 from NBER (http://www.nber.org/data/cps_may.html).

Table 1: Summary Statistics for the Palmer Survey Data

Variable	Mean, no weights	Std. dev., no weights	Mean, weighted	Std. Dev., weighted
Union	0.51	0.50	0.50	0.50
Age	43.34	12.22	43.02	11.77
White	0.90	0.30	0.86	0.34
Black	0.08	0.28	0.12	0.32
Other	0.01	0.12	0.02	0.13
Foreign born	0.17	0.38	0.16	0.37
Education	9.79	3.59	9.83	3.63
WWII veteran	0.32	0.47	0.32	0.47
Father, white collar	0.26	0.44	0.28	0.45
Father, blue collar	0.68	0.47	0.66	0.47
Father, missing	0.06	0.23	0.06	0.24
Years in area	26.22	16.59	24.09	16.51
Married	0.82	0.38	0.82	0.39
Income, weekly	71.73	26.09	73.93	27.20
Ln(income)	4.21	0.38	4.23	0.38
White collar	0.34	0.47	0.34	0.48
Blue collar	0.66	0.47	0.66	0.48
Manufacturing	0.35	0.48	0.34	0.48
Construction	0.08	0.27	0.09	0.28
Transportation, utilities	0.14	0.35	0.12	0.33
Trade	0.19	0.39	0.19	0.39
Finance, business, government	0.21	0.41	0.22	0.42
Personal service	0.02	0.14	0.02	0.15
Ag., mining	0.01	0.07	0.01	0.08
PH	0.22	0.41	0.39	0.49
NH	0.22	0.41	0.03	0.17
SP	0.21	0.41	0.06	0.23
SF	0.18	0.39	0.16	0.36
LA	0.17	0.38	0.36	0.48

Notes: The sample is restricted to men age 25 and older with information on earnings, education, and union status. The number of observations is 5,569 (of this, three are missing information on birth place and one is missing information on occupation). Columns 1 and 2 do not weight observations by duplicate status or by city size. In Columns 3 and 4, the weights are the ratio of each city's male labor force (age 25 and above) to the sample size in the survey including duplicates as described in the text. This gives more weight to larger cities like Philadelphia and Los Angeles. Union membership pertains to status at the time of interview (January or February 1951). Earnings, occupation, and industry pertain to longest job in 1950. Topcoded earnings are assigned $125 here but not in subsequent quantile analyses.
Sources: Archived transcription sheets from the Palmer Survey (Palmer 1954), as described in text. Labor force estimates are from the 1950 Census of Population, Volume II, Table 66 for each city.

Table 2: Selection into Union Membership, Palmer Survey Data

Variable	Union	Non-Union	Difference	p-value on mean difference
Personal characteristics				
Age	42.97	43.07	-0.098	0.81
White	0.87	0.86	0.004	0.77
Black	0.12	0.12	-0.002	0.87
Other	0.02	0.02	-0.002	0.72
Foreign born	0.18	0.14	0.045	0.00
Education	8.95	10.69	-1.740	0.00
High school, 4 yrs	0.29	0.49	-0.204	0.00
College, 4 yrs	0.02	0.15	-0.127	0.00
WWII vet	0.30	0.34	-0.034	0.04
Years in area	24.29	23.89	0.396	0.49
Married	0.84	0.80	0.035	0.01
Background characteristics				
Father white collar	0.21	0.35	-0.132	0.00
Father blue collar	0.72	0.59	0.131	0.00
Father missing occ	0.06	0.06	0.002	0.85
Father occscore	26.28	28.85	-2.560	0.00
Unemployed 1940	0.04	0.03	0.008	0.19
First-job occscore	21.73	23.83	-2.091	0.00
Industrial distribution				
Manu., mining, ag.	0.39	0.31	0.085	0.00
Construct., transport, utilities	0.30	0.12	0.173	0.00
Trade, personal service	0.22	0.27	-0.054	0.00
Finance, business, govt	0.09	0.30	-0.205	0.00

Notes: The sample is restricted to men ages 25 and over with earnings, education, and union status reported. Observations are weighted as described in the text and notes to Table 1. The differences and p-values are from separate regressions for each variable; standard errors are robust to heteroscedasticity. Father's occupation is not available for all men in the base sample (see "Father Missing Occ" variable). The variable *occscore* represents the median income in each three-digit occupation, as described in the IPUMS documentation (Ruggles et al. 2015). Employment status in 1940 is missing for 13 percent of men in the base sample, primarily because they were young and not in the labor force (of those missing, the median age was 28 in 1951). First job's occupation is missing for less than 2 percent of men. Results are similar when regressions include fixed effects for each city.
Sources: Archived transcription sheets of the Palmer Survey, as described in the text.

Table 3: QTE Estimates, Baseline and Specifications with Additional Control Variables

Quantile	(1A)	(1B)	(2A)	(2B)	(3A)	(3B)
0.1	0.203	0.182	0.091	0.113	0.223	0.223
	(0.023)	(0.031)	(0.032)	(0.032)	(0.026)	(0.027)
0.2	0.095	0.095	0.095	0.095	0.113	0.113
	(0.02)	(0.022)	(0.027)	(0.028)	(0.017)	(0.018)
0.3	0.087	0.087	0.087	0.087	0.087	0.087
	(0.015)	(0.021)	(0.023)	(0.022)	(0.011)	(0.011)
0.4	0.080	0.080	0.080	0.080	0.080	0.080
	(0.014)	(0.019)	(0.020)	(0.021)	(0.010)	(0.010)
0.5	0.029	0.059	0.059	0.059	0.014	0.014
	(0.019)	(0.020)	(0.022)	(0.022)	(0.023)	(0.022)
0.6	-0.013	0.027	0.027	0.027	0.000	0.000
	(0.019)	(0.017)	(0.019)	(0.019)	(0.018)	(0.018)
0.7	-0.012	0.000	0.000	0.000	-0.024	-0.012
	(0.018)	(0.025)	(0.026)	(0.026)	(0.018)	(0.016)
0.8	-0.054	0.000	0.043	0.043	-0.063	-0.053
	(0.025)	(0.029)	(0.032)	(0.031)	(0.028)	(0.028)
N	5,553	5,553	4,919	4,919	4,423	4,423

Notes: The table reports QTE estimates for union versus non-union members using subsets of the Palmer data. Column 1A reports the baseline results. Column 1B reports results with additional controls for occupation and industry as described in the text. The next two columns use a subset of the sample for which detailed industry-level injury rates are observed. Column 2A uses the baseline specification plus broad occupation and industry controls. Column 2B adds the industry-level injury rate and median hours worked variables described in the text. The last two columns use a subset of the sample for which employment status in 1940, occupational score in 1940, and father's occupational score are available. Column 3A uses the baseline specification. Column 3B adds 1940 employment status, 1940 occupation score, and father's occupational score as controls. We re-weight the distributions of union and non-union earnings so that they have the same distribution of observable characteristics based on a first step estimation of the propensity score. Because weekly earnings are expressed as whole numbers and there is some bunching, the QTEs are often stable in response to slight changes in weighting across columns. The appendix provides more details on the reweighting procedure. The propensity score is estimated by logit; the dependent variable is union status and the explanatory variables are a cubic in age, a dummy variable for completing at least some high school but no college, a dummy variable for completing at least some college, race dummy variables, a dummy variable for living less than 10 years in the area, marital status dummy variables, a dummy variable indicating whether an individual is born outside of the U.S., and a dummy variable for whether an individual is a World War II veteran. Due to topcoding of income, not much can be learned about differences in quantiles above the 80th. Pointwise confidence intervals are obtained using the bootstrap with 1000 iterations.
Sources: Archived transcription sheets of the Palmer Survey, as described in the text.

Table 4: Unions and Inequality in 1950 Weekly Earnings

	80-10	50-10	80-50
Actual inequality	0.715	0.427	0.288
Panel A			
Counterfactual 1	0.865	0.531	0.334
Difference	0.150	0.103	0.047
Percentage change	20.9%	24.1%	16.2%
Panel B			
Counterfactual 2	0.811	0.501	0.310
Difference	0.096	0.073	0.022
Percentage change	13.4%	17.1%	7.8%

Notes: The first column provides the difference between the 80th and 10th percentiles of log earnings. The second and third columns provide results for other percentiles. Actual inequality is the observed difference between the percentiles. The counterfactuals are estimated using re-weighting techniques as discussed in the appendix. Panel A results are based on re-weighting using the baseline specification. Panel B additionally includes broad occupation (four categories) and industry (four categories) controls. Rows labeled "Difference" are the difference between the counterfactual and actual level of inequality. Rows labeled "Percentage change" are the Difference divided by "Actual Inequality."
Sources: Archived transcription sheets of the Palmer Survey, as described in the text.

Table 5: Selection into Union Membership in 1973 CPS Data

	Union	Non-Union	Difference	p-value on mean difference
White	0.88	0.93	-0.044	0.00
Black	0.10	0.05	0.050	0.00
Age	42.27	41.09	1.174	0.00
Education	11.33	13.23	-1.903	0.00
High School, 4 years	0.64	0.81	-0.175	0.00
College, 4 years	0.06	0.31	-0.242	0.00
Married	0.88	0.86	0.015	0.04
Veteran	0.58	0.57	0.006	0.55
Blue Collar Occupation	0.72	0.30	0.425	0.00
White Collar Occupation	0.19	0.62	-0.424	0.00
N	3,868	5,324		

Notes: The sample is restricted to men, age 25 and over, who work full time as wage and salary workers and reside in a metropolitan area but not in the southern census region. Men without observed union status or earnings are dropped. Education variables pertain to highest grade attended. CPS sampling weights are applied.
Sources: CPS, May 1973, from NBER (http://www.nber.org/data/cps_may.html).

Appendix
A1. Reweighting distributions

The goal is to re-weight observed distributions of earnings for union and non-union workers to what they would be if union members and non-union members had the same distribution of observed characteristics (age, education, race, etc.) as in the population of male workers. To construct these counterfactual distributions, we use a re-weighting approach similar to Dinardo, Fortin, and Lemiuex (1996), Hirano, Imbens, and Ridder (2003), and Firpo (2007). The idea is to build the counterfactual distribution under the restriction that if union workers were not in a union, their non-union wage would come from the distribution of wages observed for non-union workers with the same characteristics. Under this framework, because the distribution of observable characteristics is not the same for union and non-union workers, we re-weight the distribution of non-union outcomes to account for these differences. Let $F_{<J,K>}$ be the distribution of earnings when workers are paid according to the wage schedule of group $J \in \{U, N\}$ and have the distribution of observable characteristics as group $K \in \{U, N, P\}$ where U denotes union, N denotes non-union, and P denotes the population of all workers. Thus, $F_{<U,U>}$ is the observed distribution of earnings for union members, $F_{<N,N>}$ is the observed distribution of earnings for non-union individuals, and $F_{<N,U>}$ is the counterfactual distribution of earnings for union members if they faced the non-union wage schedule. We focus primarily on comparing $F_{<U,P>}$ and $F_{<N,P>}$, the counterfactual distributions of earnings for the population of workers if they faced the union wage schedule or if they faced the non-union wage schedule.

Next, we provide the details for estimating counterfactual distributions. The details of the procedure are very similar for each counterfactual distribution; thus, we focus here on identifying and estimating $F_{<U,P>}$. The definition of $F_{<U,P>}$ is

$$F_{<U,P>}(y) = \int_{X_P} F_{Y_U|X_U}(y|x)\, dF_{X_P}(x)$$

where $F_{Y_U|X_U}(y|x)$ is the union wage structure. This is a counterfactual distribution because we are integrating the union wage structure over the distribution of population (rather than union) characteristics. Using results from Dinardo, Fortin, and Lemieux (1996) and Firpo (2007), one can show that this distribution is equal to

$$F_{<U,P>}(y) = \int_{X_P} F_{Y_U|X_U}(y|x)\Psi(x)\, dF_{X_U}(x)$$

where $\Psi(x) = p/p(x)$ is a weighting function that puts additional weight on the earnings of union workers with characteristics that are relatively more common in the non-union sector (low $p(x)$). Moreover, this can be rewritten as an expectation

$$F_{<U,P>}(y) = E[\,\Psi(X) \cdot 1\{\,Y \leq y\}|\,U = 1]$$

where $p(x) = P(U = 1|X = x)$ is the propensity score, $p = P(U = 1)$ is the unconditional probability that an individual participates in a union, and $1\{\cdot\}$ is the indicator function.

Estimation proceeds in two steps. In the first step, we estimate the propensity score $p(x)$ using a flexible logit specification:

$$p(x) = \Lambda(R(x)'\pi)$$

where Λ is the logit function, $R(x)$ are functions of the observed covariates which can include polynomials and interaction terms (a special case is $R(x) = x$), and π is a vector of parameters to be estimated.[28] Estimating this model results in estimates of $p(x)$. In the second step, we plug in these estimates of the propensity score to obtain estimates of the counterfactual distribution:

$$\hat{F}_{<U,P>}(y) = \frac{1}{n_U}\sum_{i\in U}^{n_U} \hat{\Psi}(X_i) \cdot 1\{Y_i \leq y\}$$

Most of the results presented in the paper are sample quantiles such as the median, the 80th percentile, and the 20th percentile rather than the distribution itself. For some quantile $\tau \in (0,1)$, that τth-quantile can be obtained by inverting the above distribution; that is,

$$\hat{Q}_{<U,P>}(\tau) = \inf\{y : \hat{F}_{<U,P>}(y) \leq \tau\}$$

Results throughout the paper given by subgroup can be obtained by following exactly the same

28 Hirano, Imbens, and Ridder (2003) and Firpo (2007) show that similar re-weighting estimators are efficient when the propensity score is estimated nonparametrically using a series logit estimator. Our flexible logit model follows the same idea but is simpler in terms of choosing which higher order terms to include in the model.

procedure while conditioning on membership in the subgroup. For example, in calculating the adjusted median earnings for whites, we limit the sample to whites, calculate the first-step propensity score for whites only (the race variable is dropped from X), and then calculate the distribution function and median from the reweighted distribution on the white subset.

Counterfactual inequality

Constructing counterfactual inequality measures is very similar to the re-weighting procedure outlined above though the weights are slightly different. The goal is to compare the observed distribution of earnings in 1950 to a counterfactual distribution of earnings where there are no unions, union workers retain the same observable characteristics such as age and education, and union workers are paid according to the non-union wage schedule for workers with the same observable characteristics. Let $F_W(w)$ denote the observed distribution of earnings in 1950. This distribution can be directly estimated from the Palmer data using all observations – both union and non-union. Let $F^*_{W_N}(w)$ denote the counterfactual distribution where no workers are unionized. Our main counterfactual results come from adjusting the distribution of earnings for non-union workers to account for differences in observed characteristics. To see this, notice that

$$\begin{aligned} F^*_{W_N}(w) &= P(W^*_N \leq w) \\ &= P(W \leq w \mid U = 0)(1-p) + P(W^*_N \leq w \mid U = 1)p \end{aligned}$$

where p is the fraction of unionized workers. The first term is equal to $E[(1-U)1\{W \leq w\}]$. Using the same types of arguments as in the previous section, the second term is equal to $E\left[\frac{p(X)}{1-p(X)}(1-U)1\{W \leq w\}\right]$. Combining these implies that the counterfactual distribution is given by

$$F^*_{W_N}(w) = E\left[\frac{1}{1-p(X)}(1-U)1\{W \leq w\}\right]$$

Quantiles of this distribution (and differences between quantiles of this distribution) along with those of the observed distribution of earnings are what are used to construct the counterfactual inequality measures reported in the paper.

A2. Descriptive statistics and detailed results

Appendix Table 1: Similarity of Palmer Cities and All Cities in 1950 IPUMS

	Palmer cities	All non-southern cities	All US cities
Age	43.08	43.18	42.81
White*	0.88	0.90	0.87
Black	0.10	0.10	0.13
Foreign born*	0.17	0.20	0.17
Education*	9.91	9.64	9.58
WWII veteran*	0.34	0.33	0.34
Married*	0.77	0.79	0.79
Occscore	27.53	27.41	27.36
Average income, weekly	67.53	66.99	65.69
Median income, weekly	62.50	62.50	61.00
Std. dev. income, weekly	27.39	26.40	27.03
White collar job*	0.33	0.31	0.32
Blue collar job*	0.67	0.68	0.68
Manufacturing*	0.32	0.37	0.35
Construction*	0.09	0.07	0.08
Transportation, utilities	0.14	0.14	0.14
Trade*	0.19	0.17	0.18
Finance, business, government*	0.22	0.20	0.21
Personal services	0.03	0.03	0.03
N	3,720	22,074	27,236

Notes: Figures are means unless otherwise labeled. Asterisks indicate variables for which there are statistically significant differences between men in the "Palmer cities" and men in other "non-southern cities," based on separate bivariate OLS regressions of each variable on a dummy for "Palmer city" in the full non-southern sample. Specifically, * denotes p-values under 0.05 with standard errors robust for heteroskedasticity. Such regressions are not applied to the sample summary statistics for the standard deviation or median of weekly income. All data are from the 1950 IPUMS 1-percent sample (Ruggles et al. 2015). In this table, the "Palmer cities" are Philadelphia, New Haven, San Francisco/Oakland, Los Angeles, and St. Paul/Minneapolis. Although the Palmer Survey pertains specifically to St. Paul and San Francisco, the 1950 IPUMS does not provide separate codes for Minneapolis and St. Paul nor for San Francisco and Oakland. Topcoded annual wage and salary income was assigned a value of 1.25 times the topcode (10,000*1.25). The sample includes all men classified as wage or salary workers, ages 25 and over, who lived in cities identified in the IPUMS (97 total cities), who worked at least four weeks in 1949 (to approximate the Palmer Survey's restriction to those who worked at least one month in 1950) and earned more than $10 per week and less than $240 per week (approximately omitting the bottom 1 percent and top 1 percent). Sample-line weights are applied. IPUMS provides the "occscore" variable, which represents the median total income of persons employed in each detailed occupation in 1949 according to the 1950 Census of Population. The industry variables reported in the table are not exhaustive (i.e., do not sum to 1), and a small number of men are included in the samples who do not have occupation codes (their occscore is set to missing).
Sources: Authors' calculations from IPUMS 1950 microdata sample (Ruggles et al. 2015).

Appendix Table 2: Similarity of Palmer Survey Data and 1950 Census Data for Palmer Cities

	Palmer Survey data, weighted	IPUMS
Age	43.02	43.08
White	0.86	0.88
Black	0.12	0.10
Foreign born	0.16	0.17
Education	9.83	9.91
WWII veteran*	0.32	0.34
Married*	0.82	0.77
Occscore*	28.27	27.53
Income (weekly)*	73.93	67.53
Std. dev. of income	27.20	27.39
White collar job	0.34	0.33
Blue collar job	0.66	0.67
Manufacturing*	0.34	0.32
Construction	0.09	0.09
Transportation, utilities	0.12	0.14
Trade	0.19	0.19
Finance, business, government	0.22	0.22
Personal service	0.02	0.03
PH*	0.39	0.35
NH*	0.03	0.02
SP*	0.06	0.13
SF*	0.16	0.20
LA*	0.36	0.30
Union	0.50	---
N	5,569	3,720

Notes: Figures are means unless otherwise labeled. Asterisks indicate variables for which there are statistically significant differences between men in the Palmer sample and men in the 1950 IPUMS sample, based on separate bivariate OLS regressions of each variable on a dummy for "Palmer" in a pooled sample. Specifically, * denotes p-values under 0.05 with standard errors robust for heteroskedasticity. Regressions are not applied to the summary statistics for the standard deviation of weekly income. The Palmer sample is restricted to men age 25 and over with information on earnings, education, and union status; weights are the ratio of each city's male labor force (age 25 and above) to the sample size in the survey including duplicates as described in the text. For comparison, the IPUMS sample is restricted to men, ages 25 and over who were wage or salary workers and worked at least four weeks in 1949 (the Palmer Survey includes only men who worked at least one month in 1950); sample line weights are applied. The IPUMS cities correspond as closely as possible to those in the Palmer Survey for men, but some incongruities are unavoidable. The relatively large weight for "St. Paul" in the IPUMS data compared to the Palmer data reflects that the same city code applies to both St. Paul and Minneapolis in the IPUMS (they cannot be separated); a similar caveat applies to San Francisco and Oakland. Also, the weekly earnings variables are not closely comparable for two reasons. (1) The Palmer Survey question specifically sought weekly earnings at the end of 1950's longest job, whereas the IPUMS-based variable is derived by dividing annual wage and salary in 1949 by the number of weeks worked. (2) The Palmer variable is topcoded at 100, whereas annual income in the IPUMS is topcoded at 10,000. In the table above, both topcodes (weekly for Palmer, annual for IPUMS) are multiplied by 1.25. The median value of weekly earnings in the Palmer Survey is 69; the median value in the IPUMS is 62.5. IPUMS weekly income is trimmed at $10 and $240, as in the previous table.

Sources: The archived transcription sheets of the Palmer Survey, as described in the text. The IPUMS 1950 1-percent sample (Ruggles et al. 2015).

Appendix Table 3: Union Earnings Premium at Medians, by Subgroups

	Raw Difference	(1)	(2)	Median non-union ln(w)	% U	N
Overall	0.0000	0.0290	0.0588	4.1744	0.5043	5,553
	(0.0202)	(0.0178)	(0.021)			
Educ.-by-age						
< HS / 20-29	0.1054	0.1054	0.1054	3.9608	0.6169	201
	(0.0714)	(0.0916)	(0.118)			
< HS / 30-39	0.1515	0.0514	0.0690	3.989	0.6104	444
	(0.0406)	(0.0462)	(0.0515)			
< HS / 40-49	0.0953	0.0800	0.0800	4.0943	0.6329	621
	(0.0257)	(0.0345)	(0.034)			
< HS / 50+	0.1139	0.1139	0.0800	4.0073	0.6296	1,015
	(0.0298)	(0.0331)	(0.0355)			
HS / 20-29	0.0473	0.0313	0.0163	4.0943	0.5568	555
	(0.0309)	(0.0369)	(0.0387)			
HS / 30-39	0.0000	0.0000	0.0426	4.2341	0.5912	866
	(0.0305)	(0.0253)	(0.0338)			
HS / 40-49	-0.054	-0.0811	-0.0690	4.3241	0.5595	563
	(0.0326)	(0.0324)	(0.0409)			
HS / 50+	-0.0134	-0.0134	0.0136	4.3175	0.4662	444
	(0.0411)	(0.0425)	(0.0672)			
COL / 20-29	0.0290	0.0290	0.0439	4.1744	0.2287	188
	(0.0798)	(0.1077)	(0.1583)			
COL / 30-39	0.0791	-0.0607	0.1424	4.4427	0.2669	296
	(0.0707)	(0.0873)	(0.1314)			
COL / 40-49	-0.1278	-0.0425	-0.0425	4.5539	0.2674	187
	(0.099)	(0.1367)	(0.1621)			
COL / 50+	-0.2877	-0.1949	0.0833	4.5643	0.2672	116
	(0.0978)	(0.1276)	(0.1272)			
Race						
White	-0.0420	0.0000	0.0142	4.2341	0.5069	5,012
	(0.0173)	(0.0191)	(0.0202)			
Black	0.131	0.1485	0.1484	3.912	0.4765	459
	(0.0361)	(0.0352)	(0.0375)			
Occupation						
Blue Collar	0.1420	0.0953	NA	4.0431	0.6308	3,657
	(0.021)	(0.0179)				
White Collar	-0.1013	-0.0892	NA	4.382	0.1391	1,896
	(0.0269)	(0.0283)				

Notes: The Column "Raw Difference" is the difference in medians of union and non-union earnings by subgroup. Column (1) adjusts for differences age, education, race, years in the area, marital status, foreign born status, and World War II veteran status (omitting these covariates when they are used to form the group). Column (2) additionally adjusts four broad occupation and industry categories. Occupation categories are (i) professional, managers, clerical, sales, (ii) craftsmen and similar, plus farmers, (iii) operatives and some service, laborers and some service, and farm labor. Industry categories: (i) agriculture, mining, manufacturing, (ii) construction and transportation/utilities, (iii) wholesale and retail trade, personal service, and (iv) finance, business, government. Column "Median non-union ln(w)" gives the median log earnings for non-union members of each subgroup. Column "N" provides the sample size for each subgroup.
Sources: Palmer Survey, as described in the text.

Appendix Table 4: The Union Wage Premium by City

	LA	NH	PH	SF	SP
Panel A					
0.1	0.223	0.147	0.169	0.063	0.198
	(0.045)	(0.056)	(0.051)	(0.055)	(0.052)
0.2	0.160	0.105	0.105	0.036	0.095
	(0.034)	(0.020)	(0.031)	(0.042)	(0.025)
0.3	0.095	0.095	0.113	0.065	0.069
	(0.026)	(0.020)	(0.024)	(0.030)	(0.024)
0.4	0.044	0.105	0.087	0.015	0.049
	(0.025)	(0.024)	(0.013)	(0.031)	(0.025)
0.5	0.013	0.050	0.064	0.000	-0.029
	(0.029)	(0.031)	(0.023)	(0.020)	(0.035)
0.6	0.012	0.095	0.000	-0.025	-0.069
	(0.023)	(0.036)	(0.035)	(0.023)	(0.020)
0.7	0.011	0.057	-0.041	-0.035	-0.089
	(0.043)	(0.030)	(0.022)	(0.034)	(0.037)
0.8	0.000	0.065	-0.086	-0.032	-0.134
	(0.010)	(0.030)	(0.038)	(0.032)	(0.033)
Panel B					
0.1	0.219	0.100	0.318	0.083	0.178
	(0.050)	(0.088)	(0.149)	(0.063)	(0.196)
0.2	0.160	-0.069	0.223	0.087	0.201
	(0.039)	(0.059)	(0.116)	(0.048)	(0.160)
0.3	0.095	0.077	0.182	0.031	0.070
	(0.029)	(0.055)	(0.098)	(0.035)	(0.130)
0.4	0.059	0.087	0.051	0.072	0.033
	(0.027)	(0.052)	(0.082)	(0.035)	(0.118)
0.5	0.041	0.065	-0.080	0.041	-0.059
	(0.032)	(0.051)	(0.088)	(0.023)	(0.097)
0.6	0.025	0.029	-0.145	0.026	-0.133
	(0.025)	(0.060)	(0.101)	(0.027)	(0.089)
0.7	0.034	-0.041	-0.134	0.000	-0.182
	(0.047)	(0.062)	(0.096)	(0.038)	(0.099)
0.8	0.000	-0.118	-0.204	-0.021	-0.069
	(0.013)	(0.068)	(0.108)	(0.035)	(0.078)

Notes: Panel A estimates the union premium within each of the five Palmer cities. The distributions of union and non-union earnings are re-weighted within cities to have the same distribution of covariates as all male workers in that particular city. Panel B re-weights union and non-union earnings distributions to have the same distribution of covariates as are found across all cities, not just that particular city. This setup distinguishes the within city union premium from differences in the union premium across cities that may be due to differing distributions of covariates across cities. In practice, this involves weighting up observations that have characteristics that occur infrequently in a particular city but are commonly found across all cities and weighting down observations that have characteristics that occur frequently in a particular city but are uncommon in other cities.
Sources: Archived transcription sheets of the Palmer Survey, as described in the text.