Geographic Spillover of Unionism

Thomas J. Holmes*

University of Minnesota,
Federal Reserve Bank of Minneapolis,
and National Bureau of Economic Research

ABSTRACT

Unionism in the United States is contagious; it spills out of coal mines and steel mills into other
establishments in the neighborhood, like hospitals and supermarkets. The geographic spillover
of unionism is documented here using a newly constructed establishment level data on unionism
that is rich in geographic detail. A strong connection is found between unionism of health care
establishments today and proximity to unionized coal mines and steel mills from the 1950s.

Unionization rates vary substantially across states in the United States. In South Carolina, only 5 percent of workers are covered by union contracts while in West Virginia, 16 percent are covered and in Pennsylvania, 18 percent. One factor underlying these differences is industry composition. The heavily unionized mining and steel industries are nonexistent in South Carolina but are major parts of the economies of West Virginia and Pennsylvania. But differences in industry composition are not the whole story. Even within the same industries, unionization rates are lower in South Carolina. In Pennsylvania and West Virginia it is easy to find unionized supermarkets and hospitals as well as mines and steel mills. These are virtually nonexistent in South Carolina.

The thesis of this study is that hospitals and supermarkets are unionized in Pennsylvania and West Virginia because unionism in coal mines and steel mills spilled over into other establishments in the area. More generally, the thesis is that unionism is contagious at the geographic level.

In recent years, there has been much discussion in the economics literature about social interactions of all kinds, including peer group effects in the classroom, knowledge spillovers across firms, and network effects in technology adoption.[1] Geographic spillover of unionism is a type of social interaction. In the prototypical model of this literature, a decision-making agent is more likely to select a particular option if other agents also choose it because of positive spillover benefits. The identification problem that pervades this literature is distinguishing the existence of spillovers from other explanations for parallel behavior. Individuals who are near each other often share characteristics that lead them to make similar choices. For example, if we see a neighborhood where every household owns a luxury automobile like a Mercedes-Benz or Lexus, this is probably more an indication of the affluence of the neighborhood than a result of spillovers in car purchase decisions.

This same identification issue comes up when analyzing spillovers of unionism. Unionization rates tend to be low in the South. Historically, the South has had little employment in union-prone industries like steel and automobiles. Hence, the absence of spillovers from such industries may well be why hospitals and supermarkets are nonunion in the South. But southern states share three other characteristics that also might matter. First, the South has had a substantially different racial history than the North. It is widely believed that management in the South used race to break unions, playing poor whites off against blacks. Differences in unionization rates between the North and the South today may reflect this different history of race relations.[2] Second, early in the 20th century, the South was quite poor compared to the North. To attract industry, the southern states pursued various anti-

---

[1]See, for example, Manski (1993); Glaeser, Sacerdote, and Scheinkman (1996); Brock and Durlauf (2001); and Gowrisankaran and Stavins (2004).

[2]See Marshall (1967) for a history of unionism in the South and a discussion of the role of race.

union policies, such as right-to-work laws, that contributed to keeping unionization rates low (Cobb 1993, Holmes 1998). Third, the South received fewer immigrants than the North. Such immigrants, like those from Finland, often brought with them what were considered radical ideas about socialism, and these ethnic differences in attitudes towards unions contributed to differences in unionism between the North and the South.

In addition to dealing with these confounding factors, any analysis of how the location of industries affects unionism must confront the fact that effects can go the other way. The extent of unionism at a location may influence which industries locate there. Industries that are prone to unions might avoid areas where unionism is strong. In the 1960s and 1970s, for example, the domestic automakers General Motors and Ford pursued a southern strategy of locating plants in the South with the specific intent of running nonunion operations. The United Auto Workers fought back hard and eventually organized all the southern plants of GM and Ford. Because of the selection of where such plants might locate, in the data we might not see a positive connection between the location of such industries and the extent of unionism in other industries, even if positive spillovers exist.

To address these issues, I undertake the following analysis. First, to circumvent endogeneity issues regarding the location of industry, I determine spillovers coming from coal mines, metal mines, and steel mills in the 1950s. The primary determinant of the location of mines is nature. Moreover, in the time period I consider, virtually all coal and metal mines were unionized regardless of location, so union avoidance is likely not to have been a primary location factor. Though less so than mines, the location of steel mills is heavily determined by natural factors such as access to raw materials. And like mines, steel mills were unionized virtually everywhere in the 1950s, including in the South.

Second, to control for the fact that the South is regarded as different in many ways from the North, I focus the analysis on the South in isolation. Thus my results are coming from variation *within* the South, rather than a comparison of the North and South. The question of the paper can then be posed as follows: Though there weren't many such places, selected areas of the South did have mining and metal. Is there evidence of union spillovers in those places? While the main focus of the paper is the South, I obtain similar results for other regions of the country.

Third, to control for confounding factors that might vary across states, even within the South, my analysis compares locations within the same state. This fixed-effects approach holds constant state-level anti-union policies and any state-level differences in things like race. Undertaking such an analysis requires data on unionism at a fine level of geographic and industry detail. Since such data are not available from standard sources, a substantial undertaking was required to collect the data.

Fourth, to make within industry, within state comparisons as clean as possible, I focus

on nontraded industries. This focus ensures that industries are diffuse enough throughout each state to make comparisons possible (since every location has hospitals and grocery stores). Also, the basic technology for nontraded goods like health care facilities and supermarkets tends to be similar across locations. For traded goods, technologies more likely differ, as locations specialize according to comparative advantage.

I find a significant connection between the probability that a nontraded good establishment is unionized today and proximity to mining and metal establishments from the 1950s. While overall in the South there was little of this activity, there were pockets in the South with substantial amounts of such activity. In particular, Birmingham, Alabama, was a major steel and coal center. And in Birmingham, one finds unionized nursing homes and supermarkets, unlike most other places in the South where these are unheard of.

I. Background

This section provides some background information that is needed before getting started. It begins by discussing unionism in U.S. coal mines and steel mills circa 1950. It then discusses the possible channels through which this unionism might have spilled over into other industries.

The strength of unions varies over time and across industries. The sources of variation over time are many, but of particular importance are variations over time in the legal environment. It is widely recognized that the labor environment in the late 1930s was very positive to unions and contributed to high unionization rates at mid-century. (See Freeman 1998.) In contrast, the legal environment is regarded as less friendly to labor now, in addition to other time trends working against labor (Neumann and Rissman 1984, Farber and Western 2000). The strength of unions varies across industries for various reasons, but certainly one of them is differences in capital intensity and sunk costs. Mining and steel production are capital-intensive industries with high sunk costs and are, hence, highly vulnerable to unionization. These industries also have other characteristics, e.g., working conditions, that make them prone to unionization.

Given the favorable legal environment and favorable industry conditions, at mid-century virtually all of the coal mines and steel mills were unionized throughout the country. In 1958, 95 percent of employment in the primary metals (two-digit SIC) industry in the South was in establishments with a majority of workers under a union agreement (Douty 1960). Regarding mines, the labor economist H. G. Lewis (1963, p. 74) wrote, "In 1934 the United Mine Workers, with the assistance of the National Industrial Recovery Act, succeeded in organizing almost all of the bituminous coal mining industry, and from that time to the

present the extent of unionism in the industry has been close to 100 percent."[3] That was written in 1963. Today, although unions continue to be strong in these industries, unions no longer have them locked up. The 1980s saw entry of nonunion minimills in the steel industry. Given the strength of unions in mines and mills in the earlier time period, I will focus on the industries in that time period as a source of a spillover.

While there is little in the way of academic research that quantifies spillovers of unionism, parties involved in labor relations believe spillovers matter.[4] Union officials believe that if they successfully organize one plant in an area, it will be easier to organize a neighboring plant. To take an example, the United Auto Workers has been trying very hard to organize the Mercedes Plant in Vance, Alabama. According to industry observers, "Vance represents what could be the first domino in a row of non-union transplant facilities for the UAW. . . . If Vance is organized, Honda's forthcoming plant in Alabama could be vulnerable."[5] This view is held by management as well. The Chamber of Commerce of Virginia (a state where unionization is low) offers seminars on "Protecting Our Positive Labor Climate" and publishes annually the details of all the (relatively few) unionization attempts in the state.[6] The Chamber's approach of promoting vigilance and mapping union outbreaks is analogous to the way the Centers for Disease Control might approach an infectious disease.

What is the transmission mechanism through which unionism in one establishment spills over into another establishment? There are several possibilities.

Unionism can be transmitted from a coal mine to a grocery store when an individual with experience working as a unionized coal miner either changes jobs and works in retail or has children who work in retail. As discussed by Freeman and Medoff (1984), union members tend to have more positive views towards unions than nonmembers do. It is plausible that workers retain these positive attitudes towards unions when they switch jobs. There is a large literature in psychology and organizational behavior that emphasizes the importance of workers' prior attitudes in determining union organization success. These are formed in a process of *socialization* and affect the *union commitment* of a worker. (See, for example, Fullagar et al. 1995 and Hester and Fuller, 2001.) There is evidence that positive attitudes towards unions can be transmitted through the family. Deshpande and Fiorito (1989) report that individuals who have family members in a union are more likely to have positive views

---

[3]Additional evidence on industry shares is presented in Freeman and Medoff (1979). Table 1 of that study reports estimates of production worker covered by union contracts by industry. For anthracite and bituminous coal mining, the percentages were 100 and 89 percent. For iron ores and copper ores, the percentages were 100 percent. For blast furnace and basic steel products, the percentage was 98 percent.

[4]There is some related academic research. Kremer and Olken (2001) present a model of evolutionary competition of unions. Geographers are interested in the geography of unionism, e.g., Herod (1998), but don't frame the question the way I do here.

[5]See "UAW Looks for Membership AdVANCEment," *Ward's Auto World,* July 1999.

[6]See http://www.vachamber.com/pubs/labor_info.htm.

4

towards unions.

This discussion of attitudes can be given an economic interpretation. It is plausible that workers who have never worked in a unionized facility and who don't know many union members are uncertain about the benefits of unions. Unions may be an *experience good* that one can learn about either through direct contact or that of family and friends. With risk aversion, following the logic of Erdem and Keane (1996), we can expect the market share of a product—in this case, union representation—to be higher, the greater the experience with the product.

The above discussion emphasizes the personal connection of a worker with unionism though either his or her own experience or the experience of a family member. But even if none of the workers in a particular establishment has a personal connection to unionism, the workers will benefit more from a union if other people outside the establishment in the surrounding area have such a connection. The ability to strike is a crucial component of union success. For a strike to be successful, it is important that replacement workers and customers not cross picket lines. These lines are less likely to be breached in an area where a large fraction of population either works in coal mines or steel mills or has close family members who do. For example, Marshall (1967, p. 177) discusses how a strike in 1937 to obtain a union contract at a garment factory in Tennessee was successful because of the support of the large number of members of the United Mine Workers in the area. Marshall also noted that, "The workers in the garment plants were mostly women from the miners' families."

Another channel for spillovers besides attitudes has to do with union infrastructure and staff. Locations with many unionized mills and coal mines tend to be the homes of numerous professional union staff people with training in organizing and other specialized skills related to unionism. In addition to tending to their responsibilities for existing unionized plants, these staff members are on the lookout for new opportunities, and these opportunities can cross industry boundaries. As will be shown later, at locations where the steel industry is concentrated, it is not uncommon for nursing homes and hospitals to be organized by the United Steelworkers of America (USWA) union. Of course, this spillover of union organizing staff need not occur through the same union, since individuals with specialized skills can switch affiliations. This is illustrated by the case of one Jon Hitchings from Pennsylvania. According to his union, "He had formerly worked for the United Mine Workers working underground for 22 years. He had been the President of the local mine committee and has a good background for protecting workers' rights."[7] This individual is now on the staff of the Services Workers union, representing social workers.

---

[7] See the minutes from a meeting of Chapter 5 of Local 668 of the Service Employees International Union (SEIU) from October 10, 2002, posted at http://seiu668.localsonline.org/ourlocal/chapter5minutes10102002.cfm.

II. The Data

The first part of this section uses the standard data source on unions to describe regional and state-level differences in unionization rates. The standard source is not geographically precise enough for my purposes, however. Therefore, the second part of the section describes the two alternative data sets that I have constructed in order to provide the needed geographic detail.

A. From the Standard Source: Regional and State Differences in Unionism

The standard source about unionization in the United States is the Current Population Survey (CPS) pro'duced jointly by the U.S. Bureaus of Labor Statistics and the Census. This is a monthly survey of individuals that asks them whether or not they are union members or covered by union contracts. The top section of Table 1 reports the percentage of workers who have responded that they are covered by a union contract, by region, for selected industries in a recent period. I take an average over three years of data, 1999 through 2001, rather than use only one year, since otherwise the sample size is too low for some of the industry/location cells in the table.

The main focus of this paper is the *South*, here defined as the 11 states that joined the Confederacy.[8] For the sake of making some initial comparisons, I also include two other regions in Table 1. The *North* is the region of 21 states bounded on the west by the industrial states of Minnesota, Wisconsin, Illinois, and Missouri.[9] This region is the traditional industrial heartland of the United States and was once commonly called the "manufacturing belt." The *Rest of U.S.* is the remaining states, less Alaska and Hawaii (a total of 17).

The table also reports unionization rates for several selected states. North and South Carolina from the South are selected because they are the least unionized states in the country. As will be shown later, they also happen to have negligible mining and metal sectors. Alabama, West Virginia, and Pennsylvania are selected because of their connections to the coal mining and steel industries. These three states vary in their connection to the south. Alabama is, of course, Deep South. Pennsylvania is unmistakably North. West Virginia is classified as in the North because it wasn't part of the Confederacy. But in other respects it is like the South, and no one would confuse it with New York. West Virginia will be left out of the baseline estimates for this paper which defines the South as just those states in the Confederacy. But I emphasize that the case of West Virginia, with its relatively high unionization rates in health care and groceries, provides strong prima facie evidence of my

---

[8] The South is Arkansas, Alabama, Florida, Georgia, Texas, Louisiana, Mississippi, South Carolina, Tennessee, Texas, and Virginia.

[9] The North is Connecticut, Delaware, the District of Columbia, Illinois, Indiana, Kentucky, Missouri, Minnesota, Maine, Maryland, Massachusetts, Michigan, New Hampshire, New Jersey, New York, Ohio, Pennsylvania, Rhode Island, West Virginia, Vermont, and Wisconsin.

thesis that unionism spills out of coal mines. And I get similar results if I were to include West Virginia in the South.

Looking first at all industries combined (the first row), we see that the union coverage rate in the South is 7.7 percent, and this is substantially less than the 19.5 and 15.3 percent rates in the North and the rest of the United States. The rates in the Carolinas are only about 5 percent. Even though Alabama is in the Deep South, it has a unionization rate of 11.3 percent. West Virginia's rate is higher still. Alabama and West Virginia have more mining and metal activity than the Carolinas, and this industry composition contributes to their higher unionization rates. But even within the same industries, Alabama and West Virginia tend to have higher unionization rates than North and South Carolina.

To make comparisons within the same industry, I focus attention on *nontraded* industries with significant unionization activity. By focusing on nontraded industries, I ensure broad geographic coverage and, as discussed in the introduction, make it more likely that the industries are similar across locations. To produce a list of industries, I took the 2000 CPS and selected industries in the construction, wholesale trade, retail trade, and service sectors, because industries in these sectors are primarily not traded. I eliminated educational services and government services in order to focus on the private sector. I selected only those industries with a union coverage rate above 10 percent in 2000 and with at least 500,000 total employees in order to focus on industries with significant union activity. I eliminated miscellaneous industries or "those not elsewhere classified." This selection process resulted in six industries: construction, wholesale groceries, retail groceries, hotels and motels, nursing homes, and hospitals.[10]

As is clear in Table 1, unionization rates within these industries are substantially higher in the North than in the South. Within the South, unionization rates in Alabama are higher than in North or South Carolina in the same industry. West Virginia's rates are higher still. The thesis of this study is that the relatively high unionization rates in Alabama and West Virginia in nontraded industries are in part a consequence of union spillovers from the mining and metal sectors in these states.

B. From Alternative Sources: Richer Geographic Detail . . .

The empirical analysis of this project requires data on unionization at a fine level of geographic detail. The CPS data are inadequate for this purpose. One problem is that the CPS is only a tiny sample of the U.S. population, so sample sizes can be too small when we go to a high level of geographic and industry detail. For example, in Table 1, I have observations on only 24 workers in the hotel industry in Alabama. Another problem is that for a large portion of

---

[10]The CPS treats all of construction as one industry; unionization rates at a finer level of detail are not available. Since unionization is heavy throughout all the industries in the construction sector, I take the sector as a whole and treat it as an industry.

the CPS sample, geographic location within the state is not disclosed. Location in the large metropolitan statistical areas (MSAs) is disclosed, but location in the smallest MSAs and rural counties is generally not disclosed. This is a particular problem for me because many of the mining areas that I will be interested in are in rural counties.

My need for geographic data at a fine level of detail leads me to construct two alternative data sets that use administrative data from federal agencies. The first data set covers nursing homes and hospitals. It is a very clean data set that has been put together in a labor-intensive process. The second data set is cruder, but is broader in coverage, including all six industries listed above. The results for hospitals and nursing homes are very similar across the two data sets.

B.1. The FMCS/POS Data on Health Care Facilities

The starting point of the first data set is a directory of the universe of nursing homes and hospitals in existence in the United States in 1996. This data set is the "Provider of Services" file (POS) for 1996 released by the Centers for Medicare and Medicaid Services. There are 15,014 nursing homes and 6,906 hospitals.[11] The POS data include address information and facility characteristics, like organizational form (nonprofit, for profit, or government) and number of beds.

In order to determine which of these facilities are unionized, I matched data on expiring union contracts to establishments in the POS file. The Federal Mediation and Conciliation Service (FMCS) is the source of the contract expiration data. The National Labor Relations Act requires that notice of contract expiration be filed with this agency. This notice gives the agency a heads up so that it can be ready if it has to mediate any strikes.

I matched 26,650 contracts in the FMCS data base from the period 1986–2003 to facilities in the POS file. My matching algorithm used the company name, address information, and phone information found in both files. The process required an extensive amount of manual matches, and here Google proved invaluable. For example, by Googling a phone number in the FMCS file, I could find additional information about the entity from the FMCS file that would enable me to complete the match in the POS file. The overwhelming majority of contracts in the health care industry cover a bargaining unit at a single establishment. There do exist multi-unit contracts, particularly in the New York area, where industry consortia such as the League of Voluntary Hospitals negotiate industry-wide contracts covering sometimes dozens of facilities. In such cases, I obtained additional information from the U.S. Department of Labor, including the actual contracts themselves, to identify the establishments covered by the contracts. I also used Web sites of union locals to verify matches. See

---

[11] I exclude Alaska and Hawaii here and throughout the rest of the study. The 15,014 figure includes only nursing homes that are not part of hospitals.

the data appendix for details about the data and summary statistics. All of the data for this paper are posted on the Web, and the appendix provides the web address.

DiNardo and Lee (2004) have also recently used the FMCS contract data to study a different question. They use the data in a wholesale fashion, using the data from all industries, trying to match up all of the contracts. My approach is less ambitious than theirs in terms of industry coverage. But by focusing on only one industry, I am able to do things in terms of matching by hand and bringing in other sources of information which would not be feasible if one were to look at all industries together.

I define a health care facility as unionized if there was at least one bargaining unit under contract in the facility over the period of the contract data, 1986–2003. There are 3,555 unionized facilities in the POS file, according to this definition. The median unionized facility has 6 expiring contracts. The maximum number of contracts for any one facility is 50, and this is for Washington Hospital in San Francisco. A contract with a given bargaining unit will appear multiple times in the data set as it expires and is renewed on a two- or three-year cycle. In addition, many facilities have multiple bargaining units with a contract for the registered nurses, a contract for food service workers, etc. Washington Hospital even has a contract for the radiologists.

Establishment unionization rates, weighted by beds, are reported in the middle section of Table 1. The table shows that 20.2 percent of all nursing home beds are in unionized establishments. This is about twice as large a portion as the 10.8 percent of workers in the industry that are covered by unions from the CPS. A difference of this order of magnitude is to be expected since in any given unionized establishment, a sizable fraction of workers would not be covered by a union contract.[12]

The patterns with the FMCS/POS data are qualitatively very similar to the patterns found with the CPS. Unionization rates in the South overall are much lower than in the North and the rest of the country. The rates in North and South Carolina are particularly low.

B.2. The NLRB/CBP Data on Union Elections

To construct the FMCS/POS data, I made use of the publicly available data on the universe of health care establishments. For the other four industries I am looking at, such as retail groceries and construction, there are no analogous public data on the universe of establishments. Even if such data were available, say, for grocery stores, additional complications would make it difficult to link them to the FMCS. Nursing home and hospital union contracts

---

[12]For North Carolina, the establishment rate of 1.6 is actually smaller than the worker rate from the CPS of 1.9. There are several potential explanations for this. First, the CPS estimate is based on a small sample, so there can be sampling error. Second, the underlying population of establishments is somewhat broader in the CPS sample than in the POS file that I use. See the data appendix for further discussion about the set of establishments used in the POS file.

usually cover a single establishment, and the specific address and name of the establishment are provided in the FMCS file. In contrast, contracts in the retail grocery industry usually cover multiple grocery stores in an area, the location information in the FMCS file is often vague, like "Chicago area," and the number of establishments covered is often unspecified.

In order to have something to say about these other industries, I turn to data on certification of representative (RC) elections supervised by the National Labor Relations Board (NLRB). By law, these elections take place at the establishment level. I have data on all elections over the period 1980–1999, and for each election, the three-digit SIC code of the establishment is provided as well as the county where it is located.

The NLRB election data measure a flow while the FMCS contract data measure a stock. If an industry were 100 percent union at a particular location and if there were no entry and exit, then no elections would be observed there. The FMCS data on contracts would pick up that this location is 100 percent union, but the NLRB data would find no indication of unionism. While this limitation of the NLRB data should be kept in mind, I think the NLRB data are a useful proxy of the extent of unionism in a particular industry and location. Even heavily unionized areas are rarely 100 percent union, so there is room for a union to expand. State overall unionization rates from the CPS are highly correlated with NLRB election rates ($\rho = 0.83$).[13]

My goal here is to evaluate the rate at which NLRB elections in an industry and location occur relative to the number of establishments in this industry and location. The U.S. Census Bureau's data series County Business Patterns (CBP) provides establishment counts by industry, county, and employment size category. I linked the 1980–1999 NLRB election data to the 1992 CBP data, using the industry, county, and size information. In the first step of the matching process, I linked NLRB records of the same establishment, using address and company name information. This is important since 37.5 percent of elections are for establishments with multiple elections. I then matched the NLRB establishment-level data to the CBP data, by three-digit SIC, county, and establishment size.[14] One issue is that there are establishments with elections before 1992 in the NLRB data that are shut down by 1992, so my process will match them to some other establishment that does exist in 1992. This is not a serious problem because I am not actually using any information in the CBP record that the NLRB record doesn't already have. The purpose of the exercise is to create a denominator for the probability of the event of an organizing election, to set up a pool of potential election targets from which the given one was drawn. Since the CBP population

---

[13]The correlation is between state total unionization rates from the CPS for 1990 (Hirsch and MacPherson 2003, 2005) and NLRB RC elections 1980–1999 per 1990 CPS state employment, weighted by employment.

[14]The NLRB data set does not provide information about establishment size. But it does list the size of the bargaining unit. I estimated establishment size in the NLRB data by setting it to twice bargaining unit size.

moves slowly over time, the underlying pool for 1980 is roughly similar to the pool for 1992. Table 4 below reports establishment counts by industry for establishments in the CBP in the South group of states. Further details about the data are provided in the appendix.

An establishment in the NLRB/CPB data is defined as *unionized* if it has at least one election over the period 1980–1999. Unionization rates, per 1,000 establishments, are reported in the bottom section of Table 1. The nursing home and hospital industries have relatively high rates overall, at 121 and 211 per 1,000 establishments. These rates are close in magnitude to the rates in the FMCS/POS data (if we take into account that one is per 100, the other per 1,000). The rates in the remaining four industries are relatively small. Unlike the health care industry, most of the unionization of the grocery and construction industries took place prior to 1980. Nonetheless, the qualitative patterns in the NLRB/CBP data set are consistent with the patterns in the CPS data. For each industry, the unionization rates in the South in the NLRB/CBP data are about half or smaller than the levels in the North, just as occurs in the CPS. At the state data, compared with the Carolinas, unionization is high in Alabama in nursing homes, groceries, and construction, in both the NLRB/CBP data and the CPS data. The rates in West Virginia and Pennsylvania are higher still.

C. . . . And Other Variables

As explained in the introduction, I focus on the mining and primary metals sectors from the mid-20th century, the heyday of unionism, as a source of contagion, or spillover. I use data from the 1958 Census (U.S. Bureau of the Census 1961a, b) to determine the distribution of employment by county. The data published by the Census include cell counts of establishments by employment size classifications, e.g., an employment range of 250–499 employees. I use these establishment-level size ranges to estimate employment. For mining, I include metal mining (SIC 1000), anthracite coal mining (SIC 1100), and bituminous coal mining (SIC 1200). Metal manufacturing is defined here as primary metals manufacturing (SIC 3300). See the data appendix for more details.

Table 2 reports employment shares in the mining and metal industries for 1958 (calculated as a percentage of 1960 total employment) for the nation as a whole and for selected regions and states. Observe that in 1960, the employment share in these sectors was three times higher in the North than in the South. These industries were essentially nonexistent in North and South Carolina. Note, however, that the industries were well-represented in Alabama, at nearly twice the national average. Pennsylvania was the leading steel-producing state, and West Virginia was a leading coal-producing state.

As discussed above, immigrants may have brought positive attitudes towards unions with them. This is potentially an interesting kind of spillover to look at, but not one that I will focus on here. In my analysis, I will control for differences in immigration across locations. For my immigration measure, I go all the way back to the 1910 Census. Immigration was

closed in the 1920s. In Table 2, the "foreign" statistic reported is the percentage of the white population with at least one parent born outside of the United States. Note the dramatic difference between the North and South in this measure: 49.1 percent of whites in the North and only 8.1 percent of whites in the South.

The last row of the table highlights the well-known historical differences between the North and South in matters of race. In 1910, 35 percent of the population in the South was black compared to only 3 percent in the North.

III. The Statistical Model

Now I examine the cross-section relationship between an establishment's unionization status in the late 20th century and the degree to which it is located near where mining and metal establishments existed in the 1950s.

A. Notation

In terms of notation, for an establishment e at location ℓ, let s_ℓ^m denote *unionism spillover from mining and metal* industries at the establishment's location. Let u_e be an indicator variable, so that $u_e = 1$ if the establishment is unionized and $u_e = 0$ if not. I examine the relationship between u_e and s_ℓ^m, holding fixed various other characteristics of the establishment and the location.

To define the spillover, let q_ℓ^m be the share of total employment that is in the unionized mining and metal establishments in the 1950s at location ℓ. The spillover s_ℓ^m *collected* at ℓ is a weighted average of the spillovers being emitted nearby. Let $y_{k\ell}$ denote the distance in miles between location k and location ℓ. Let $N_\ell^{\overline{y}} = \{k, y_{k\ell} \leq \overline{y}\}$ be the neighboring locations within $\overline{y}$ miles of ℓ. Define the spillover by

$$s_\ell^m(\delta, \overline{y}) = \frac{\sum_{k \in N_\ell^{\overline{y}}} \exp(-\delta y_{k\ell}) n_k q_k^m}{\sum_{k \in N_\ell^{\overline{y}}} \exp(-\delta y_{k\ell}) n_k}. \tag{1}$$

This specification weights nearby spillovers being emitted both by population n_k as well as an exponential decay function that declines with distance. It is similar in spirit to a gravity specification in trade models. The parameter δ governs how things decay with distance. The parameter $\overline{y}$ governs how far one goes out to collect spillovers. In effect, beyond the distance $\overline{y}$ there is infinite decay of the spillover.

Let x_e denote a vector of establishment characteristics, e.g., a measure of size or a form of organization (like profit vs. nonprofit). Let z_ℓ denote a vector of location ℓ characteristics. This will include a vector of state fixed effects, to capture differences across states in labor market polices and race policies and mean differences across states in things like immigration.

The vector z_ℓ also includes location characteristics that potentially vary within a state, like demographic variables.

I model the decision to unionize in terms of a discrete choice framework. Suppose that the value of unionization status to establishment e is

$$v_e = x'_e\beta + z'_\ell\gamma + \alpha s_\ell^m + \epsilon_e. \tag{2}$$

The establishment is unionized if $v_e \geq 0$ and is not if $v_e < 0$. The key parameter here is α, the coefficient on the unionism spillover variable.

B. Variables

In the FMCS/POS data set, an establishment is unionized, $u_e = 1$, if it has at least one union contract over the period 1986–2003. In the NLRB/CBP data, an establishment is unionized if it has at least one NLRB election over the period 1980–1999. I later consider alternative definitions.

In the FMCS/POS data set, the establishment characteristics x_e are size (number of beds and number of beds squared), dummy variables for establishment organization form (nonprofit, profit, or government), and a dummy variable for facility type (nursing home or hospital). There are a few establishments with a missing value for organization form; these get a separate dummy variable. In the NLRB/CBP data set, there are two characteristics: employment size (with three size categories) and industry. Summary statistics are provided in a table in the appendix.

Location ℓ is defined at the level of the county. The location characteristics z_ℓ at the county level are population density (population in the year 2000 divided by land area), share of whites in 1910 with at least one foreign-born parent, and the black population share in 1910.

Recall that q_ℓ^m is defined to be the share of total employment in the 1950s in unionized mining and metal establishments at location ℓ. As discussed earlier, at that time virtually the entire mining and metal industries were unionized. So I assume here that they are both all union, and I use the share of employment in 1958 in county ℓ to approximate q_ℓ^m.

I consider two alternative specifications for the spillover function (??). In the *30-mile-spillover specification*, I assume that the decay $\delta = 0$ and the distance $\overline{y} = 30$. The assumption of $\delta = 0$ makes things easy to estimate. In the *continuous decay specification*, I free up δ and impose that $\overline{y} = 75$.

C. Estimates

Table 3 reports the estimates of the statistical model with the FMCS/POS data on health care facilities. The first three columns of figures are based on the data from the South, the

main focus of the paper. For comparison purposes, the last three columns report estimates for the remaining contiguous United States. Table 6 below considers additional geographic groupings.

Using the data from the South, the estimate of α in the 30-mile case is 13.5 and the standard error is 2.5. To estimate standard errors, I cluster observations by county, using the robust estimator of variance. (This makes a difference—without clustering, the standard error is 2.0.) I discuss the magnitude of α further below. Here I just note that if the probability of unionization is very small, as is the case here, then the probability is approximately proportionate to $e^{\alpha s^m}$. So at $\alpha = 13.6$, an increase from spillover $s^m = 0$ to $s^m = 0.1$ is associated with a multiplicative increase in probability on the order of $e^{1.36} = 3.9$ times.

For the continuous decay specification and the South data, the estimate of α is 31.3 and spillovers decay at a rate of 1.5 percent per mile. It is not surprising that the spillover coefficient α is larger in the continuous decay specification. If there is decay, then treating all spillovers within a 30-mile radius the same is a kind of measurement error.

The estimates for α using the data from the rest of the contiguous United States are about half of what they are for the South. This finding is likely related to the fact that this model is nonlinear and that unionization rates are much higher outside the South. So a given proportionate increase outside the South is a much bigger increase in levels. To explore this issue further, I report estimates of the linear probability model for both regions. With this specification, the estimated coefficient on s^m is actually larger outside the South than in the South. So a given change in s^m outside the South is associated with a bigger change in level, but a smaller change in percentage terms, outside the South as compared to the South.

The main conclusion to draw from Table 3 is that over several specifications and two geographic populations of establishments, there is a strong positive correlation between unionization of health care establishments and proximity to mining and metal industries in the 1950s.

It is worth noting that the probability of unionization depends upon other variables as well. The model is estimated with state-level fixed effects (not reported) that are important quantitatively. Examination of the coefficients in Table 3 yields the following conclusions: In the South, unionization is positively correlated with locations being historically black. Outside the South, it is correlated with locations with large immigrant populations in the early 20th century. In the South, hospitals are less likely to be unionized compared to nursing homes. Outside the South, this is reversed.

Table 4 reports the estimates for the NLRB/CBP data in other industries in the South. For the sake of comparison, I also include the 30-mile spillover estimates for the FMCS/POS data. I report only the coefficients on the spillover variables, but all of these estimates

14

include establishment characteristics and location characteristics in addition to state fixed effects. For nursing homes and hospitals, the 30-mile spillover coefficient α estimate from the NLRB/CBP data is 10.5, which is fairly close to the estimate of 13.5 from the FMCS/POS data. When I group the remaining four industries in the NLRB/CBP data, the estimated spillover coefficient is 5.1. Examining each industry individually, I find that both retail groceries and construction have coefficients of 5.1, and both are statistically significant. For wholesale groceries and hotels, there are relatively few observations, and the standard errors are high.

Table 5 provides information about magnitudes. The first column of figures takes the fitted probability of an establishment being unionized from the model with a 30-mile spillover. For the FMCS/POS data (the first estimate in Table 4), I report the predicted probability by state in the South. The second column of figures re-evaluates the fitted probabilities, setting the spillover $s^m = 0.065$, which is the average mining and metal share for all of Pennsylvania in Table 2. The last column sets $s^m = 0.149$, the average share in West Virginia. With the Pennsylvania level of spillover, the fitted probabilities on average double, from 3.2 to 6.2, and with the West Virginia level, they more than triple. In the estimates with the NLRB data, for nursing homes and hospitals, the results are quantitatively the same, a doubling with Pennsylvania's level and a tripling with West Virginia's. For the other industries in the NLRB data, the relationship is more like a 50 percent increase with Pennsylvania and a doubling with West Virginia.

Note that the exercise in Table 5 is intended merely to give a sense of the magnitude of the empirical relationship. It is not an estimate of a *treatment effect,* or prediction, of what would have happened to say, South Carolina, if it had had the same geology and coal fields as West Virginia. If this had been true, other things would have been different that are being held fixed in the model. For example, the exercise in Table 5 holds fixed the state-level fixed effect. If South Carolina had coal like West Virginia, it is likely it would have had a base of powerful mine and steel worker unions that could have potentially blocked anti-union laws, such as the right-to-work laws that were passed in South Carolina. With such a treatment, it might very well be that the resulting unionization rate in South Carolina would be higher than the fitted value of 10.6 percent in Table 5. While the parameter estimates are not structural, they do provide strong evidence that geographic spillover takes place. If spillovers didn't occur, I would expect no empirical connection between unionization rates and proximity to mining and metal, given all of the controls that I have included.

Table 6 looks at the robustness of these results. The top row contains the 30-mile spillover model's estimates for the three data groups: nursing homes and hospitals in the FMCS/POS and NLRB/CBP data and the other industries in the NLRB/CBP data (from Table 4). The next row adds three location-level variables that have been suggested by sem-

inar participants: per capita income (2000 per capita income of the establishment's county), a political variable (the share voting Democratic in the 1988 presidential election), and demographic information more recent than 1910. (I use the black share for 2000, which in the South is highly collinear with the 1910 black share.) Adding these variables makes virtually no difference in the estimate of the spillover coefficient. The second change is to eliminate all right-side variables besides the spillover and the state fixed effects, and the third is to get rid of state fixed effects as well. With these changes, the estimates decrease from the original somewhat, but remain large. The fourth change keeps the right side the same as the original but uses a stricter definition for classifying an establishment as *unionized.* For the FMCS/POS data, I now require that an establishment have at least four contracts while before I required only one. For the NLRB data, I now require an establishment to have at least one election in which the union won while before I required only that it have an election. These further restrictions reduce the number of establishments classified as union on the order of one-half. But for the FMCS/POS data, there is virtually no change in the spillover estimates. For the NLRB/CBP data, the coefficient increases slightly.

The next set of robustness exercises maintains the baseline specification but changes the geographic areas. Some scholars, such as Marshall (1967), use a broader definition of the South that includes West Virginia, Kentucky, and Oklahoma. Adding these states lowers the coefficients, but the magnitude of the relationship remains large. The next geographic area is the North, as defined earlier. The coefficients are approximately half as they are for the baseline estimates for the South. So in percent terms, the effect is approximately one-half. But the base probabilities in the North are on the order of twice as high, so the effects in the North are comparable in level terms with the South. The next region is the *Mountain and Plains States*, the 14-state region between the *Pacific States* (California, Oregon, and Washington) and the North and South. This region, including states like Montana, Colorado, New Mexico, and Arizona, has significant mining activity in some places and so is a natural area to focus on. For this region and for the FMCS/POS data, the estimate of α is 8.3 and the standard error is 1.5, which are remarkably similar to the estimates for the other regions. There is a difference with the estimates using the NLRB data as the coefficients are close to zero. This area is an anomaly in that elsewhere, for the health care sector, the coefficient obtained with the FMCS/POS contract data is similar to the coefficient obtained with the NLRB/CBP data. I don't have an explanation for the difference. But I put more weight on the FMCS/POS data, both because the data are cleaner and more reliable and because conceptually I prefer a unionization measure in levels (i.e., existing contracts) rather than in election rates. The final region considered is the Pacific States. There is little in the way of mining or metal industries to speak of in this area and little variation in spillover for identifying the relationship. (See summary statistics in Table A2.) Unsurprisingly, the

16

standard errors are quite high, in the range of 5 to 7, for this region. The exercise of this paper is not interesting to consider for these states. The last row in Table 6 reports the coefficients when the data for all contiguous United States are pooled together. The bottom line is that, particularly for the FMCS/POS contract data for health care facilities, the relationship between unionization and spillover is strong throughout United States, not just in the South.

The final data exercise broaches the issue of the channels through which spillovers take place. As discussed earlier, one channel is union infrastructure and staff. Unions employ staff people who are on the lookout for new organizing activities. Naturally, a union staff person will be mainly interested in expanding his or her own union. That is, a staff person for the USWA will be working to organize new USWA establishments. A staff person for the United Mine Workers (UMW) will be organizing for the UMW. To the extent that spillovers work through an infrastructure channel, we should see as a necessary condition that health care facilities and grocery stores near mining and metal establishments should be organized by the UMW or the USWA. If we see health care facilities near coal mines represented by the service workers union, the spillover can't all be infrastructure.

To examine this issue, I use the information about union affiliation available in the data. I consider what happens when I exclude representation by the USWA and the UMW. Specifically, all establishments in the FMCS/POS data that have a labor contract with the UMW or USWA or establishments in the NLRB/CBP data that have an election involving one of those unions are classified as *nonunion.* So only those establishments unconnected to the UMW or the USWA and having some other union affiliation are now classified as *union.* In this case, as reported in Table 7, the spillover coefficients for nursing homes and hospitals fall by 30–45 percent, for the two data sets. The coefficient falls for the other industries as well, though by a much smaller amount. These declines are consistent with union infrastructure being a factor. But the results also imply that union infrastructure is not the whole story. More than half of the coefficient remains when we reclassify the UMW and USWA observations. This result is being driven by nursing homes and grocery stores with affiliations like the Service Workers (SEIU) and Food and Commercial Workers (UFCW) unions near mines and metal plants that are different from the affiliations that have organized the mines and factories.

The final row in Table 7 goes the other way and classifies as union only those establishments with a UMW or USWA contract. This increases all the coefficients compared to the original model estimates, especially so for the other industries in the NLRB/CBP, where the coefficient more than triples. It is clear that the USWA and UMW are getting out of steel mills and coal mines into neighboring establishments.

IV. A Visual Presentation of the Data

The previous section uses statistics to summarize the data. This section provides a visual presentation of the raw data. The results parallel the statistical findings.

The majority of mining and metal activity in 1958 was concentrated in the 15-state region illustrated in Figure 1. The region accounted for 70 percent of all coal and metal mining employment and 57 percent of primary metal manufacturing. And, again, virtually all of these establishments were unionized. Figure 1(a) is a map of the location in 1958 of establishments in these industries with 100 or more employees; each dot represents one plant.[15] Consider the point labeled A in Pennsylvania in Figure 1(a). (State names are provided in the bottom right map.) Just below point A is the area around Pittsburgh, which has the heaviest concentration of the mining and metal industries. Next consider point B in Alabama. This is the Birmingham area. This is also a major center of mining and metal manufacturing. Looking between these two points, we see significant mining and metal activity, particularly in West Virginia and eastern Kentucky. This area follows the Appalachian Mountains and is rich in coal seams.

Figure 1(b) plots the locations of the health care facilities in the 1996 POS data that I have classified as unionized based on the FMCS/POS data. The extent to which current unionism in the health care sector follows the distribution of mining and metal activity from the 1950s is remarkable. Note, in particular, the heavy concentration of unionized facilities in West Virginia and eastern Kentucky and in Alabama around Birmingham, all areas where mining and metal activities are concentrated.

The remaining two maps restrict attention to establishments affiliated with the USWA. Figure 2(a) illustrates the location of health care facilities in the FMCS/POS data in the region with at least one USWA contract. There are 54 such facilities in the region. The areas where the USWA gets into health care facilities are all areas where there is a large base of mines and steel mills. Figure 2(b) is based on information from the USWA Web site on the location of grocery stores and beverage distributors that are represented by the USWA. (See data appendix.) The same pattern holds here.

V. Concluding Remarks

This study shows that while unionism in health care facilities and grocery stores is weak in the South overall, those facilities situated near coal mines and steel mills have unionization rates closer to their counterparts in the North. How does unionism spill out of mines and into grocery stores? This study suggests two answers. One is union infrastructure. This is

---

[15]The location information in the map is accurate to the level of the county.

consistent with the fact that some of the health care facilities near mines and steel mills are organized by the USWA. The second answer is that the workers themselves are bringing in the seeds of unionism they have picked up from their own experience in mines or mills or from the experiences of their families and friends. This is consistent with the fact that most health care facilities near mines and mills are organized by unions unrelated to those organizing mines and mills.

It is well-appreciated that the massive decline of heavy industry employment in the United States has had a direct effect in decimating the number of unionized workers in this country. The results presented here suggest that there is a secondary effect that goes beyond those industries. The reduction in geographic spillovers from heavy industry can be expected to also negatively impact unionism in other sectors, like health care and retail trade.

An example can make this more concrete. In 1958, Jefferson County, Alabama, where Birmingham is located, had 30,000 workers in steel mills and coal mines accounting for 14 percent of all employment. As of 2002, this number had fallen to only 9,000 workers, accounting for just 2.5 percent of all employment. Spillovers have been vastly diluted. Lately, several new auto plants have been built in the Birmingham area, and the United Auto Workers has so far been unsuccessful in organizing them. If new auto plants in the Birmingham area can resist unionization, there is little chance that the new nonunion entrants to the grocery business in the area, like Wal-Mart, will ever be unionized.

Data Appendix

Here I describe in more detail the data I collected from various sources to get more geographical detail on unionism in the United States. Data and sample programs for the paper are posted at www.econ.umn.edu/~holmes/data/geo_spill. Summary statistics are presented in Tables A1 and A2.

A. FMCS/POS

I obtained the POS files from the Centers for Medicare and Medicaid Services (CMS). I selected all nursing homes and hospitals in the POS file from 1996. Nursing homes included establishments in the following categories: skilled nursing facilities (SNF), nursing facilities (NF), SNF/NF dually, and SNF/NF distinct. This resulted in 21,920 establishments, excluding any in Alaska and Hawaii.

I obtained the data on expired contracts from the Federal Mediation and Conciliation Service (FMCS). I took the subset of FMCS contracts designated as from the "health care" industry and matched them to the POS universe. For the purpose of matching, I included the 1991 and 2001 POS files as well as the 1996 file. I used the address, phone numbers, and company name information in the FMCS and POS file for matching. The computer did most of the work here, but there was also an extensive amount of manual matching. As discussed in the text, I made frequent use of Google to complete a match. In some cases, I would be able to determine that the FMCS entity was actually a retirement home or a mental health facility and therefore not part of my underlying universe. In other cases, I would be able to obtain additional information about the facility that would enable me to complete the match to the POS file. I was able to match 96.4 percent of the 27,629 health care facility contracts in the FMCS data base.

In New York and Minnesota, there are industry-wide contracts that are negotiated by a consortia of health care providers. In a majority of such cases, I was able to obtain the actual contract from the U.S. Department of Labor and other sources in order to determine the facilities covered from the contract language. There were several important multi-facility contracts that could not be matched this way, particularly the complex set of contracts for the Kaiser health system, a system of hospitals on the West Coast that is essentially all union. For such cases, I went to the Web sites of the local unions negotiating the contracts, where I usually found lists of the covered facilities. Of the 560 facilities listed on union Web sites as represented, in 88 percent of the cases I already had one or more matching contracts for the facility in the FMCS data base. I take this as evidence that the matched data are very reliable.

In determining whether a facility was represented, I excluded in the FMCS data notices of bargaining for first contracts. I used only expiration notices of contracts that have actually

been obtained.[16]

B. NLRB/CBP

I put together NLRB election data for 1980–1999, which I obtained from various sources: Bruce Fallick (Fallick and Hassett 1999), Henry Farber (1984, 2001), the NLRB, and the National Archives. I selected elections from the following six SIC industry codes: 15, contract construction; 5140, wholesale groceries and related products; 5410, grocery stores; 7010, hotels and motels; 8050, nursing and personal care facilities; and 8060, hospitals. Excluding Alaska and Hawaii, this resulted in 13,859 elections.

I then matched elections by address and employer name and identified 11,098 unique establishments. I then set establishment size equal to twice the eligible number of votes in an election. Using the industry, county, and size information in the NLRB data, I matched the 11,098 establishments to the CBP data. I aggregated size classes in the CBP data to three employment size groups: 1–49, 50–249, and 250 and more. In a few cases, there was not an establishment in a given size category in a given industry and county, and in such cases, I assigned the NLRB establishment to a larger CBP establishment or, if none, to a smaller CBP establishment. There were only 87 NLRB establishments for which I could not find a corresponding establishment in the 1992 CBP in the same industry and county, and I eliminated them.

C. Other Variables

As discussed in the text, data on mining and metal employment were obtained from the U.S. Bureau of the Census (1961a, b). For the primary metal industries (SIC 33), the establishment employment size data have seven categories: 1–19, 20–49, 50–99, 100–249, 250–499, 500–999, and 1,000 and more. I set the employment of each establishment equal to the mean of its size group. For the mining industries (SIC 10, 11, 12), the employment size groupings in the tables are more aggregated: 0–19, 20–99, 100–249, and 250 and more. However, there is additional information in the Census publication in the form of county maps of employment that permitted me to break up the 250 and more category to 250–999 and 1,000 and above.

County demographic data from 1910 were obtained from a 1970s study of the Inter-University Consortium for Political and Social Research (ICPSR, 197?, 1988). Use of these data required various cleanings. I went back to the original Census tabulations to fix some mistakes and to fill in left out observations like those of the District of Columbia.

---

[16]When workers vote for a union in an NLRB election, the firm is required to recognize the union and bargain for a year. But in a large percentage of cases, the union never gets a first contract. (See Cooke 1985.) I exclude notices of first bargaining by eliminating records with blank values for contract expiration and reopen dates. Beginning in 2003, the FMCS added an explicit categorical variable that distinguishes notices of first bargaining. Using this new field, I have determined that my approach of identifying notices of initial bargaining is very accurate.

Data on demographic information for 2000 and on per capita income were obtained from the 2000 Census. Data for the 1988 presidential election were obtained from a 1988 ICPSR study. The figure reported is the Democratic share of the combined vote for the Democratic and Republican candidates.

The data for Figure 2(b), the location of retail food and wholesale beverage distribution establishments represented by the USWA, were obtained directly from the USWA Web site on 5/24/2001. This information is no longer posted there, so I have posted this noncopyrighted information at my Web site, along with the rest of the data at my Web site.

References

Brock, William A., and Steven N. Durlauf, "Discrete Choice with Social Interactions," Review of Economic Studies, LXVIII (2001), 235–260.

Cobb, James C., The Selling of the South: The Southern Crusade for Industrial Development, 1936–1990, 2nd ed. (Urbana: University of Illinois Press, 1993).

Cooke, William N., Union Organizing and Public Policy: Failure to Secure First Contracts (Kalamazoo, MI: W. E. Upjohn Institute for Employment Research, 1985).

Deshpande, Satish P., and Jack Fiorito, "Specific and General Beliefs in Union Voting Models," Academy of Management Journal, XXXII (1989), 883–897.

DiNardo, John, and David S. Lee, "Economic Impacts of New Unionization on Private Sector Employers: 1984–2001," Quarterly Journal of Economics, CXIX (2004), 1383–1441.

Douty, H. M., "Collective Bargaining Coverage in Factory Employment, 1958," Monthly Labor Review, LXXXIII (1960), 345–349.

Erdem, Tulin, and Michael P. Keane, "Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," Marketing Science, XV (1996), 1–20.

Fallick, Bruce C., and Kevin A. Hassett, "Investment and Union Certification," Journal of Labor Economics, XVII (1999), 570–582.

Farber, Henry S., "The Extent of Unionization in the United States," in Challenges and Choices Facing American Labor, Thomas A. Kochan, ed. (Cambridge, MA: MIT Press, 1984), pp. 15–43.

————, "Union Success in Representation Elections: Why Does Unit Size Matter?" Industrial and Labor Relations Review, LIV (2001), 329–348.

Farber, Henry S., and Bruce Western, "Round Up the Usual Suspects: The Decline of Unions in the Private Sector, 1973–1998," Princeton University IRS Working Paper 437, 2000.

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

Freeman, Richard B., and James L. Medoff, "New Estimates of Private Sector Unionism in the United States," Industrial and Labor Relations Review, XXXII (1979), 143–174.

————, What Do Unions Do? (New York: Basic Books, 1984).

Fullagar, Clive J. A., Daniel G. Gallagher, Michael E. Gordon, and Paul F. Clark, "Impact of Early Socialization on Union Commitment and Participation: A Longitudinal Study," Journal of Applied Psychology, LXXX (1995), 147–157.

23

Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman, "Crime and Social Interactions," Quarterly Journal of Economics, CXI (1996), 507–548.

Gowrisankaran, Gautam, and Joanna Stavins, "Network Externalities and Technology Adoption: Lessons from Electronic Payments," Rand Journal of Economics, XXXV (2004), 250–276.

Herod, Andrew, ed., Organizing the Landscape: Geographical Perspectives on Labor Unionism (Minneapolis: University of Minnesota Press, 1998).

Hester, Leslie Kim, and Jerry Bryan Fuller, "Building Union Commitment: The Impact of Parental Attitudes and Participation," Labor Studies Journal, XXVI (2001), 17–30.

Hirsch, Barry T., and David A. MacPherson, "Union Membership and Coverage Database from the Current Population Survey: Note," Industrial and Labor Relations Review, LVI (2003), 349–354.

_____, Union Membership and Coverage Database from the CPS, www.Unionstats.com, 2005.

Holmes, Thomas J., "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders," Journal of Political Economy, CVI (1998), 667–705.

Inter-University Consortium for Political and Social Research, "Historical, Demographic, Economic, and Social Data: The United States, 1790–1970" [Computer file], ICPSR 3, 197?

_____, "General Election Data for the United States, 1970–1990," DS14: National, 1988 [Computer file], ICPSR 13, 1988.

Kremer, Michael, and Benjamin A. Olken, "A Biological Model of Unions," NBER Working Paper 8257, 2001.

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

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," Review of Economic Studies, LX (1993), 531–542.

Marshall, F. Ray, Labor in the South (Cambridge, MA: Harvard University Press, 1967).

Neumann, George R., and Ellen R. Rissman, "Where Have All the Union Members Gone?" Journal of Labor Economics, II (1984), 175–192.

U.S. Bureau of the Census, U.S. Census of Manufactures: 1958, Location of Manufacturing Plants by Industry, County, and Employment Size: 1958, Part 7. Primary Metal Industries; Fabricated Metal Products, Special Report MC58(S)-2.7 (Washington, DC: U.S. Government Printing Office, 1961a).

_____, U.S. Census of Mineral Industries: 1958, Vol. II, Area Statistics (Washington, DC: U.S. Government Printing Office, 1961b).

Table 1
Unionization Rates by Industry and Geographic Area

| Data Set | All | Regions | | | Selected States | | | | |
| | | | | | Least Unionized | | Highly Unionized | | |
		South	North	Rest of U.S.	N.C.	S.C.	Ala.	W.Va.	Penn.
CPS (1999–2001)									
Rates per 100 workers									
All Industries (including public sector)	14.9	7.7	19.5	15.3	4.2	5.0	11.3	15.9	18.3
Nontraded Industries with Significant Unionization									
Nursing and Personal Care Facilities	10.8	1.9	16.0	6.5	1.9	.0	3.4	8.3	17.7
Hospitals	15.3	4.7	19.2	20.0	2.7	3.4	5.0	8.0	15.2
Wholesale Trade: Groceries	12.9	4.6	18.5	13.3	1.1	.0	11.7	23.1	23.6
Retail Trade: Groceries	21.3	5.8	26.7	31.8	1.2	.0	9.1	11.8	17.3
Hotels and Motels	10.6	2.5	13.1	14.7	.0	.0	.0	.0	6.9
Construction	20.4	6.7	32.1	19.8	2.5	4.3	10.3	23.9	28.9
FMCS/POS									
Rates per 100 establishments									
Nursing Homes	20.2	4.8	31.8	10.9	1.6	1.9	10.5	35.6	32.3
Hospitals	26.0	3.4	38.5	30.4	4.6	2.4	1.9	12.2	41.2
NLRB/CBP									
Rates per 1,000 establishments									
Nursing and Personal Care Facilities	120.8	54.5	180.4	68.5	20.5	50.5	81.8	350.6	195.7
Hospitals	210.6	27.5	324.7	230.2	17.6	37.0	36.4	246.6	384.2
Wholesale Trade: Groceries	27.8	16.7	32.1	30.8	18.0	10.7	25.7	42.3	28.9
Retail Trade: Groceries	7.2	2.3	7.3	14.8	.6	1.2	2.4	10.7	8.6
Hotels and Motels	18.9	6.7	28.2	19.3	3.1	3.4	2.1	21.9	42.6
Construction	7.0	4.3	8.1	7.9	.7	.7	4.8	9.4	10.1

CPS = Results of Current Population Survey averaged over 1999–2001.

FMCS/POS = Combination of data on health care facilities from the Federal Mediation and Conciliation Service and the "Provider of Services" file of the Centers for Medicare and Medicaid Services.

NLRB/CBP = Combination of data on establishments from the National Labor Relations Board (data on union representation elections) and the U.S. Census Bureau's data series "County Business Patterns" (number of establishments by industry, county, and employment size).

Table 2
Sources of Spillovers

Size and Location of Highly Unionized Industries
and Related Population Characteristics
in Selected Geographic Areas of the United States
in the 20th Century

Category	All	Regions			Selected States				
		South	North	Rest of U.S.	N.C.	S.C.	Ala.	W.Va.	Penn.
Industry Employment Shares in 1958									
Mining	.5	.2	.6	.4	.0	.0	.9	11.1	1.5
Metal Manufacturing	1.7	.7	2.4	.8	.2	.1	4.0	3.8	5.0
Mining + Metal Manufacturing	2.2	1.0	3.0	1.2	.2	.1	4.9	14.9	6.5
Population Shares in 1910									
White Population with Foreign-Born Parent	37.5	8.1	49.1	40.0	1.1	3.5	4.3	10.8	43.4
Black Population	10.7	35.2	3.0	1.8	32.5	54.7	42.7	4.9	2.6

Source: U.S. Bureau of the Census.

Tables 3–4
**Estimating the Correlation Between
Unionization and Proximity to the
U.S. Mining and Metal Manufacturing Industries
in the 1950s**

Table 3
**In Health Care Facilities in Two Geographic Areas
Using the FMCS/POS Data and Three Model Specifications**
(Numbers in parentheses are standard errors.)

Statistic	South			Rest of Contiguous United States		
	30-Mile Spillover	Contin Decay	Linear Probability	30-Mile Spillover	Contin Decay	Linear Probability
Spillover						
Spillover Coefficient α	13.6	31.3	.595	6.6	9.8	.837
	(2.4)	(4.6)	(.215)	(1.0)	(2.0)	(.147)
Decay Rate $\delta*100$ (% per mile)	—	1.4	—	—	4.5	—
		(.6)			(1.9)	
County Variables						
Population Share in 1910						
Black	3.4	4.0	.084	1.5	1.8	.062
	(.7)	(.6)	(.022)	(1.7)	(1.6)	(.122)
White with Foreign-Born	.4	.4	.005	2.8	2.7	.389
Parent	(.9)	(.9)	(.017)	(.4)	(.4)	(.064)
Population Density (2000)	.1	.2	.001	.06	.1	.009
	(.1)	(.1)	(.004)	(.03)	(.0)	(.003)
Establishment Characteristics						
Log Beds	1.1	1.1	−.028	1.7	1.7	.069
	(1.2)	(1.1)	(.020)	(.5)	(.5)	(.035)
$(\text{Log Beds})^2$	−.0	−.0	.005	−.1	−.1	.001
	(.1)	(.1)	(.002)	(.0)	(.0)	(.004)
For-Profit Dummy	.2	.2	.006	.1	.1	.018
	(.2)	(.2)	(.004)	(.1)	(.1)	(.014)
Government Dummy	−1.4	−1.4	−.010	−.2	−.2	−.031
	(.5)	(.5)	(.007)	(.1)	(.1)	(.013)
Type-Missing Dummy	−1.2	−1.2	−.013	−1.9	−1.9	−.210
	(.9)	(.9)	(.009)	(.2)	(.2)	(.020)
Hospital Dummy	−1.0	−1.1	−.022	.6	.6	.089
	(.3)	(.3)	(.009)	(.1)	(.1)	(.011)
Number of Establishments	6,194	6,194	6,194	15,561	15,561	15,561

*All specifications include state-level fixed effects that aren't reported.

Table 4
In Other Industries in the South
Using the NLRB/CBP Data and Two Model Specifications
(Numbers in parentheses are standard errors.)

| | Value for Each Specification and Statistic | | | | |
| | 30-Mile Spillover | Continuous Decay | | Number of | |
Data Set and Industry	α	α	$\delta*100$	Establishments	Unionized
FMCS/POS					
Nursing Homes and Hospitals	13.6 (2.4)	31.3 (4.6)	1.4 (.6)	6,194	199
NLRB/CBP					
Nursing Homes and Hospitals	10.5 (2.3)	20.9 (3.9)	2.1 (.8)	7,557	357
Other Nontraded Industries	5.1 (2.0)	11.1 (3.3)	1.1 (1.7)	228,026	1,048
Wholesale Trade: Groceries	1.0 (5.0)	6.9 (7.0)	—*	10,583	177
Retail Trade: Groceries	5.1 (2.7)	8.7 (8.8)	2.4 (10.4)	47,536	110
Hotels and Motels	4.3 (4.4)	13.9 (5.5)	3.0 (1.6)	12,198	84
Construction	5.1 (1.7)	11.2 (3.9)	1.6 (2.5)	157,344	677

*For this case, the estimate of δ was at the lower bound of zero.

Table 5
The Magnitude of the Fitted Empirical Relationship

Mean Fitted Probability of an Establishment Being Unionized
in Various Southern States and Industries,
at Alternative Spillover Levels, Using Estimates from 30-Mile Spillover Model

| | Fitted Probability with Spillover of | | |
| | | | West |
Geographic Area and Industry	Actual	Pennsylvania	Virginia
By State			
FMCS/POS Health Care Facilities			
(per 100 establishments)			
Alabama	6.1	6.6	16.5
Arkansas	4.2	8.5	20.3
Florida	7.6	15.6	34.8
Georgia	1.7	3.9	10.6
Louisiana	2.4	5.3	14.0
Mississippi	4.2	8.7	20.3
North Carolina	1.4	3.2	9.0
South Carolina	1.7	3.9	10.6
Tennessee	3.9	7.5	18.3
Texas	1.1	2.0	5.8
Virginia	3.3	5.3	14.0
All of the South			
FMCS/POS Health Care Facilities			
(per 100 establishments)	3.2	6.2	15.1
NLRB/CBP Industries			
(per 1,000 establishments)			
Nursing Homes and Hospitals	47.2	79.0	159.6
Other Nontraded Industries	4.6	6.0	9.1
Wholesale Trade: Groceries	16.7	17.6	19.1
Retail Trade: Groceries	2.3	3.0	4.6
Hotels and Motels	6.9	8.8	12.3
Construction	4.3	5.7	8.5

Table 6
Robustness of the Model's Spillover Coefficient Estimates

Model's Estimates of Spillover Coefficients with Proposed Model Changes
and with Alternative Data Sets and Industries
(Numbers in parentheses are standard errors.)

| | Value with Specified Data Set and Industries | | |
| | | NLRB/CBP | |
Specification	FMCS/POS Nurs/Hosp	Nurs/Hosp	Other Industries
Baseline 30-Mile Spillover Model	13.6	10.5	5.1
	(2.4)	(2.3)	(2.0)
Model with Changes			
Additional Location-Level Variables	12.4	10.4	4.8
	(2.7)	(2.2)	(2.0)
Only State-Level Fixed Effects	9.7	7.3	4.2
	(2.1)	(2.0)	(1.5)
No Controls	8.6	6.4	4.6
	(2.9)	(2.5)	(1.3)
Stricter Definition of Union Status	12.9	12.8	7.0
	(2.8)	(2.4)	(1.6)
Alternative Geographic Areas			
Expand South to include W. Va., Ky., and Okla.	9.6	7.6	3.0
	(1.4)	(1.2)	(1.0)
North	7.0	6.4	2.6
	(1.2)	(1.1)	(.7)
Mountain and Plains States	8.3	1.2	1.0
	(1.5)	(1.5)	(.9)
Pacific States (Calif., Ore., and Wash.)	−4.9	−14.5	−.7
	(6.9)	(7.1)	(4.9)
All of Contiguous United States	7.5	5.5	1.6
	(1.0)	(.9)	(.6)

Table 7
Examining a Potential Channel of Spillover:
Affiliation with Mining and Metal Manufacturing Unions

Model's Estimates of Spillover Coefficients with the Classification of Establishments
as Unionized Linked to Their Affiliation with Either of Two Large Unions:
United Mine Workers or United Steelworkers of America
(Numbers in parentheses are standard errors.)

	Value with Specified Data Set and Industries		
		NLRB/CBP	
Specification	FMCS/POS Nurs/Hosp	Nurs/Hosp	Other Industries
30-Mile Spillover Model	13.6	10.5	5.1
	(2.4)	(2.3)	(2.0)
Model with Change in Definition of Union Status			
Excludes UMW and USWA	9.2	6.0	4.3
	(2.5)	(2.5)	(1.6)
Includes only UMW and USWA	15.4	11.7	15.6
	(3.4)	(3.3)	(3.2)

Appendix Table A1
Detailed Summary Statistics for Establishments in South

| | | FMCS/POS Health Care Facilities | | Value with Alternative Data Sets and Industries NLRB/CBP | | | |
| | | | | Nursing Homes and Hospitals | | Other Nontraded Industries | |
Variable		Mean	Std. Error	Mean	Std. Error	Mean	Std. Error
Union		.03	.18	.05	.21	.005	.07
Establishment Size							
Beds		136.03	154.20				
Employment	1–49	—		.34	.47	.948	.22
	50–249	—		.51	.50	.049	.22
	250+	—		.14	.35	.004	.06
Establishment Organization Form							
Nonprofit		.21	.40				
For-Profit		.64	.48				
Government		.12	.33				
Type Missing		.04	.19				
Industry							
Nursing Homes		.64	.48	.73	.44		
Hospitals		.36	.48	.27	.44		
Wholesale Trade: Groceries		—	—			.05	.21
Retail Trade: Groceries		—	—			.21	.41
Hotels and Motels		—	—			.06	.23
Construction		—	—			.69	.46
Population Share 1910							
White Foreign		.12	.15	.13	.15	.13	.15
Black Share		.30	.19	.31	.18	.32	.17
Population Density (2000)		.55	.85	.66	.89	.72	.89
Neighboring Mining/Metal Share		.01	.03	.01	.03	.01	.02
Number of Establishments		6,194		7,557		228,026	

Appendix Table A2
Selected Summary Statistics by Region

	Value with Alternative Data Sets and Industries					
	FMCS/POS Health Care Facilities		NLRB/CBP			
			Nursing Homes and Hospitals		Other Nontraded Industries	
Region and Variable	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error
South						
Union	.032	.176	.047	.212	.005	.068
Neig. Mine/Metal Share	.009	.027	.008	.026	.008	.024
Number of Establishments	6,194		7,557		228,026	
North						
Union	.261	.439	.212	.409	.010	.100
Neig. Mine/Metal Share	.030	.042	.030	.041	.027	.037
Number of Establishments	10,003		13,522		362,470	
Mountain and Plains States						
Union	.041	.198	.066	.249	.012	.110
Neig. Mine/Metal Share	.012	.042	.015	.042	.021	.052
Number of Establishments	3,231		3,603		89,067	
Pacific States (Calif., Ore., Wash.)						
Union	.263	.449	.162	.368	.010	.098
Neig. Mine/Metal Share	.009	.009	.009	.009	.009	.009
Number of Establishments	2,327		3,263		123,729	
Contiguous United States						
Union	.163	.370	.143	.350	.009	.093
Neig. Mine/Metal Share	.019	.037	.019	.037	.018	.034
Number of Establishments	21,755		27,946		803,292	

Figures 1-2
A Visual Display of the Data:
Unionism Spills Over from Mining and Manufacturing
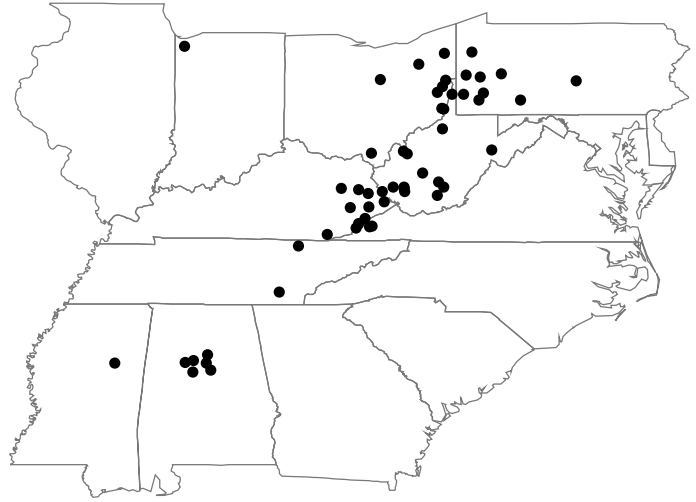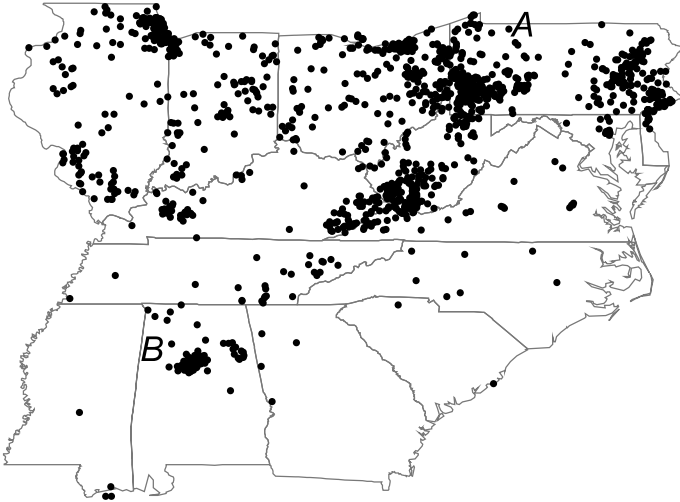## into Other Industries

Locations of Establishments in the 15 States with the Most Mining
and Manufacturing
(Each dot represents one establishment.)

Figure 1
Unionized Establishments
in the Region

Figure 2
Unionized Establishments
Represented by the USWA

1a. Mining and Metal Plants in 1958
With 100 or More Employees

2a. Health Care Facilities

1b. Health Care Facilities

2b. Grocery Stores and Beverage Distributors

Sources: FMCS/POS, USWA, Census 1958