THE DISTRIBUTION OF EMPLOYEES BY ESTABLISHMENT SIZE:
PATTERNS OF CHANGE AND COMOVEMENT IN THE UNITED STATES, 1962-1985

by

Steve J. Davis
Graduate School of Business
University of Chicago
Chicago, IL 60637

John Haltiwanger
Department of Economics
University of Maryland
College Park, MD 20742

November 1989

*We gratefully acknowledge financial support of the National Science Foundation, the Hoover Institution, and the Office of Graduate Studies and Research at the University of Maryland.
I. Introduction

In this paper we document patterns of secular change and cross-industry comovement in the distributions of employees and establishments by establishment size in the United States. Several considerations motivate our study. First, many interesting labor market phenomena are associated with establishment (and/or firm) size or the size distribution of establishments. Important examples include: (a) average compensation is greater at larger establishments and firms, controlling for (other) observed characteristics of workers and employers (see, e.g., Lester (1967), Masters (1969), Mellow (1982), and Brown and Medoff (1989)); (b) unionization rates are higher for larger establishments and firms (see, e.g., Brown, Hamilton and Medoff (1989, chapter 6), and Robinson (1988)); (c) union wage and benefit premiums are a decreasing function of establishment size, controlling for other observed characteristics of workers and employers (see, e.g., Freeman and Medoff (1984), Podgursky (1986), and Bramley, Wunnava and Robinson (1989)); (d) employment is more volatile and death rates are higher for smaller establishments and firms (see, e.g., Evans (1987a,b), Hall (1987), Dunne, Roberts and Samuelson (1989), and Davis and Haltiwanger (1989)); and (e) larger enterprises offer more scope for internal labor market phenomena (see, e.g., Oi (1983)). Second, small business enterprises are frequently claimed to be an increasingly important sector of the U.S. economy. This claim implies leftward shifts over time in the distribution of employees by enterprise size. Third, manufacturing's share of employment has shrunk dramatically over the last forty years. This observation raises the question: to what extent is the decline in manufacturing's share of employment accounted for by the downsizing of manufacturing establishments? Fourth, the last several decades have been associated with large economywide fluctuations in worker turnover, in private sector unionization rates, and productivity. In view of these considerations, it is of obvious interest to investigate the contribution of changes in the size distributions to these phenomena.

We address several questions related to these considerations using a data set that is far more comprehensive in its time-series and industry scope than data sets used in previous work. Our primary data source is the County Business Patterns (CBP) data set, which contains detailed industry-level
counts for establishments and employees by eight to twelve establishment size classes. The CBP data are tabulated from the population of tax-paying private business establishments in the United States. The tabulations are computed for mid-March payroll periods in the years 1962 and 1964 through 1985.

A key focus of our analysis is that the distribution of establishments by establishments size - a (telephone directory) yellow pages concept - generates a strikingly different perspective than does the distribution of employees by establishment size - a "white pages" concept. Much of the literature does not carefully distinguish between these two alternative characterizations of the size distribution. In this paper, the precise relationship between the first and second moments of these two distributions is derived and the importance of the distinction is illustrated in a number of ways. First, we show that there are dramatic differences in the behavior of the two empirical size distributions. Second, we provide an illustration of how theory generates precise guidelines regarding the appropriate measure of size to use for investigating the empirical consequences of changes in the size distribution. This guidance is used in our subsequent empirical analysis.

Our analysis documents several significant shifts in the two size distributions. In 1985, mean establishment size in the U.S. was 17 employees. In the same year, the coworker mean (the first moment of the distribution of employees by establishment size) was bounded between 501 and 915 employees. Mean establishment size rose by 29% from 1962 to 1973, and it rose by a further 10% from 1974 to 1985. In striking contrast, the coworker mean fell by roughly 40% between 1968 and 1985. The dramatic secular decline in the coworker mean reflects a shift away from both the share of employees in large (250+ employees) establishments and the share of employees in small (1-19 employees) establishments. In decomposing these shifts, we find that changes in the industry distribution of employment and movements in the employee size distribution within the average two-digit industry make roughly equal contributions to the secular shift towards mid-sized establishments in the aggregate economy.

Even though the pattern of change in the employee size distribution within the average two-digit industry mimics the pattern of change in the economywide size distribution, fewer than 10% of individual two-digit
industries experienced the same pattern. Simply put, secular shifts in the employee size distribution exhibit remarkable heterogeneity across two-digit industries. Much of this heterogeneity is attributable to differences in the behavior between goods and service producing industries. Since 1967, industries that produce goods have experienced pronounced leftward shifts in the employee size distribution. In contrast, service producing industries have experienced pronounced rightward shifts in the employee size distribution over our entire 1962-1985 sample.

Our analysis documents significant movements in the employee size distribution over the 1962 to 1985 period and considerable cross-industry heterogeneity in the direction and magnitude of these movements. The observed movements and heterogeneity lead naturally to questions regarding their influence for explaining other labor market trends over this period. In this paper, we investigate one such question by examining the ability of movements in the employee size distribution to explain the secular decline in private sector unionization rates. We develop cross-section and time-series regression models that relate industry-level union density to the employee size distribution. We estimate these models by supplementing our CBP data with data on union membership by detailed industry for 1974-80 published by Kokklenberg and Sockell (1985). The time-series and cross-section results provide strong support for the hypothesis that the industry coworker mean is positively related to industry union density. However, we find that only a small amount of the secular decline in aggregate private sector unionization rates over the past twenty years is explained by changes in the employee size distribution.

The paper proceeds as follows. In section II, we develop some key relationships between the two different size distributions and also obtain results useful for calculating bounds on the moments of the distributions. Section III describes the data and measurement methodology.

In section IV, we document the secular movements in the employee and establishment size distributions at an economy-wide level. Of particular interest here is our investigation of the contribution of the within-industry and the between-industry variation in the size distributions to the time-series variation exhibited in the economy-wide size distributions.

Section V investigates the character and source of the heterogeneity of
industry-level movements in the employee size distribution. To explain the observed heterogeneity, we examine three descriptive hypotheses focusing on the significance of (i) the difference in the behavior between goods producing and service producing industries, (ii) the relationship between initial size distribution and the subsequent changes in the distribution, and (iii) the relationship between capital intensity and changes in the distribution.

Having characterized the patterns of change and comovement in the size distributions in the preceding sections, in section VI we examine the role that changes in the employee size distribution played in the secular decline in aggregate private sector unionization rates.
II. The Relationship Between Two Size Distributions

In this section we characterize the relationship between the distribution of employees by establishment size and the distribution of establishments by establishment size. We also explain our procedure for computing the first moment of each distribution.

Let there be $J$ possible sizes of establishments indexed by $j = 1, \ldots, J$, where size is measured by number of employees. Accordingly, let $N_{ijt}$ be the number of establishments of size $j$ in industry $i$ at time $t$. Defining $E_{ijt}$ to be the total number of employees who work at an establishment of size $j$ in industry $i$ at time $t$, we have $E_{ijt} = jN_{ijt}$. An omitted subscript indicates summation over that index unless otherwise noted.

The industry-$i$, time-$t$ empirical probability function of establishment size is given by

$$f_{it}(j) = \frac{N_{ijt}}{N_{it}}, \quad j = 1, \ldots, J. \quad (1)$$

A useful relationship between this function and the empirical probability function of employees by establishment size is given by

$$g_{it}(j) = \frac{E_{ijt}}{E_{it}} = \frac{jN_{ijt}}{N_{it}} = \frac{jf_{it}(j)}{M_{it}}, \quad j = 1, \ldots, J, \quad (2)$$

where $M_{it}$ denotes mean establishment size in industry $i$ at time $t$:

$$M_{it} = \sum_j \left(\frac{N_{ijt}}{N_{it}}\right) j = \frac{E_{it}}{N_{it}}. \quad (3)$$

We also define the variance of establishment size,

$$\sigma^2_{it} = \left[ \sum_j \left(\frac{N_{ijt}}{N_{it}}\right) j^2 \right] - M^2_{it}, \quad (4)$$

and the coworker mean,

$$C_{it} = \sum_j jE_{ijt}/E_{it}. \quad (5)$$

The aggregate counterparts to these moments are

$$M_t = \sum_i M_{it} \frac{N_{it}}{N_t} = E_t/N_t, \quad (6)$$
\[
\sigma_t^2 = \left[ \sum_i \sum_j j^2 N_{ijt} / N_t \right] - M_t^2, \quad \text{and} \quad (7)
\]
\[
C_t = \sum_i (E_{it} / E_t) C_{it}. \quad (8)
\]

Using equation (2), we derive a key expression that relates the coworker mean to the first two moments of the establishment size distribution:
\[
C_{it} = M_{it} + \sigma_{it}^2 / M_{it}, \quad \text{and} \quad (9)
\]
\[
C_t = M_t + \sigma_t^2 / M_t. \quad (10)
\]

Equations (9) and (10) reveal that the coworker mean exceeds mean establishment size unless all establishments are the same size. The accompanying implication is that the magnitude of the difference between the coworker mean and the establishment mean increases in the variance of establishment size. It is useful to note that equations (9) and (10) are equivalently expressed as
\[
C_t = M_t + \sigma_t \kappa_t,
\]
where \( \kappa \) denotes the coefficient of variation in establishment size.

Our data set does not contain employment observations for individual establishments. Instead, it contains counts on the number of establishments and number of employees for eight to twelve establishment size classes. This feature of the data constrains our investigation of the size distributions, and, in particular, it precludes a direct calculation of the coworker mean. The CBP data do admit computation of upper and lower bounds on the coworker mean, as we now show.

Let there be \( S \) size classes indexed by \( s = 1, \ldots, S \). Define \( N_{it}^s \) and \( E_{it}^s \) as the number of establishments and employees in size class \( s \), industry \( i \) at time \( t \). Further, define \( M_{it}^s \) as the within size-class mean for industry \( i \) at time \( t \), and let \( \bar{s} \) and \( \underline{s} \) denote the upper and lower bounds for size-class \( s \). The size classes partition the underlying size distributions, so that \( j = 1, \ldots, J_1, J_1 + 1, \ldots, J_s, \ldots, J_S \), where \( J_s = \bar{s}, J_s + 1 = (s + 1), \) and \( J_0 = 0 \).

Using this notation, rewrite the industry coworker mean as
\[
C_{it} = \sum_{s=1}^{S} \frac{E_{it}^s}{E_{it}} \sum_{j=J_s-1+1}^{J_s} j \frac{E_{ijt}}{E_{it}^s} = \sum_{s=1}^{S} C_{it}^s \frac{E_{it}^s}{E_{it}}, \quad (11)
\]
where $C_{it}^*$ is the size-class $s$ coworker mean for industry $i$ at time $t$. Our problem in measuring the industry coworker mean is that we cannot directly calculate $C_{it}^*$. However, we can bound $C_{it}^*$ using available information by employing the type of relationship described in (9) and (10). Define $\sigma_{it}^2(s)$ as the variance of establishment size in size-class $s$,

$$\sigma_{it}^2(s) = \sum_{j=\underline{s}}^{\bar{s}} j^2 \frac{N_{ijt}}{N_{it}^s} - (M_{it}^s)^2.$$ (12)

Analogous to (9) and (10), we have

$$C_{it}^* = M_{it}^s + \frac{\sigma_{it}^2(s)}{M_{it}^s}.$$ (13)

Since we observe $E_{it}^*, N_{it}^*$, and $M_{it}^*$ and know $\underline{s}$ and $\bar{s}$ (for all but the largest size class), we can calculate upper and lower bounds for $C_{it}^*$ using (13). The lower bound for $C_{it}^*$, found by setting $\sigma_{it}^2(s) = 0$, equals $M_{it}^s$. To calculate an upper bound we first choose the $N_{ijt}$ to maximize $\sigma_{it}^2(s)$, subject to $\underline{s} < j < \bar{s}$, $M_{it}^s$, $E_{it}^s$, and the integer constraint on $N_{ijt}$. We then substitute the maximized value of $\sigma_{it}^2(s)$ into (3). To obtain economywide lower and upper bounds on the coworker mean, we aggregate over size classes and industries using (8) and (11).\(^1\)

To summarize, in the empirical work that follows, we use: (a) equations (3) and (6) to compute industry and economywide establishment means; (b) equations (11) and (13) to compute bounds on the industry coworker means, coupled with (8) to compute bounds on the economywide coworker mean; and (c) employee counts to compute the distribution of employees by establishment size class. As equations (6) and (8) indicate, changes in the economywide establishment and coworker means decompose into changes in industry establishment and employment shares, respectively, and changes due to intra-industry variation in the size distributions. An analogous decomposition holds for the distribution of employees by establishment size.

III. The Size Distribution Data

\(^1\)In future work, we plan to obtain the $\sigma_{it}^2(s)$ by drawing on other data sources or using distributional assumptions when necessary. For example, we can compute the $\sigma_{it}^2(s)$ for manufacturing industries from the Longitudinal Research Datafile described in Davis and Haltiwanger (1989). Given the $\sigma_{it}^2(s)$, we can compute the industry-level coworker means directly from (11) and (13).
County Business Patterns (CBP) data contain annual observations on number of employees and number of establishments in eight to twelve establishment size-class categories, cross-classified by detailed industry and geographic criteria. An establishment is defined as “a single physical location where business is conducted or where services or industrial operations are performed.” From 1974 to 1985, CBP provides establishment and employee counts for nine establishment size classes—“1-4”, “5-9”, “10-19”, “20-49”, “50-99”, “100-249”, “250-499”, “500-999”, and “1000+”. From 1962 to 1973, CBP provides establishment and employees counts for eight size classes—“1-3”, “4-7”, “5-9”, “10-19”, “20-49”, “50-99”, “100-249”, “250-499”, and “500+”. In the “1000+” (“500+” prior to 1974) category, establishment counts only are available on a finer size-class breakdown—“1000-1499”, “1500-2499”, “2500-4999”, and “5000+”.

The CBP tabulations are based on the population of tax-paying private business establishments in the United States with one or more paid employee, excluding agricultural production, railroad, and household employment. Establishment size-class designations and employee counts are based on paid employment in the mid-March pay period of each year, with one important exception that we describe presently. Since 1983, the size group “1 to 4” includes establishments that reported no employees during the mid-March pay period but paid wages to at least one employee during the year. Prior to 1983, establishment counts were based on whether an establishment was active in the fourth quarter. Appendix D of County Business Patterns, U.S. Summary Volume, 1982, reports establishment counts by two-digit industry for 1982 under both tabulation methods. We use this information to splice the establishment counts for the “1-4” size class across the 1982/1983 definitional change.

An important definitional change in the CBP data occurred in 1974. Prior to 1974, all non-manufacturing establishments in the same county and four-digit industry owned by the same firm are treated as a single “reporting unit” in CBP tabulations. Due to this definitional change and the 1974 break in the definition of size classes, the levels of most reported measures are not strictly comparable between the pre-1974 and post-1973 periods. For further discussion of CBP methods for counting establishments and employees, see United States Department of Commerce (1987).

The results in this paper rely on CBP data from 1962 and 1964 to 1985 for two-digit industries at the economywide level. While CBP size-class data are available at the level of individual counties and four-digit industries, the employment data for particular size-class/industry/region cells are often suppressed to prevent the disclosure of employment at individual establishments. Finer classification in terms of industry and geographic criteria leads to more severe data suppression problems. In view of this tradeoff, and our desire to
splice the establishment counts for the "1-4" size class across the 1982/1983 definitional change, we chose to focus on a breakdown by two-digit industries.

At the level of two-digit industries with no geographic breakdown, about one-quarter of one percent of all industry/size-class employment observations are suppressed. The suppressions are concentrated in industries that have a very small number of large establishments. We estimate the suppressed employment observations as the product of the number of establishments in the industry/size-class cell (which is never suppressed) and an estimate for the establishment mean in the industry/size-class cell. We estimate the industry size-class means using a modification of linear interpolation between (unsuppressed) nearby years for the same industry and size-class. The modifications insure that the size-class numbers sum to the industry total, and that the estimated employment numbers are consistent with the size-class boundaries. A fortran program that performs the estimation is available upon request.

IV. Secular Movements at the Economywide Level

A. Summary Statistics

Table 1 reports summary statistics on the economywide size distributions and their first moments for the 1974 to 1985 period. Three simple points deserve mention. First, the distribution of establishments by establishment size is sharply skewed toward the small establishments. Fifty-five percent of all establishments have fewer than five employees, and seventy-six percent have fewer than ten employees. Second, employees are distributed relatively evenly across the size classes. Twenty-seven percent of employees work in establishments with 1-19 employees, forty-two percent in establishments with 20-249 employees, and thirty-one percent in establishments with 250 or more employees. Third, the cross-industry variation in the establishment mean, the coworker mean, and the size class shares is an order of magnitude larger than the time-series variation within the average two-digit industry.

B. Secular Movements in the First Moments

Figures 1 and 2 show the time-series movements in the first moments of the economywide size distributions. Mean establishment size rose by 29% from 1962 to 1973, and it rose by a further 10% from 1974 to 1985 to stand at 17 employees per establishment. The years 1979 to 1983 represent the only extended interruption to the secular increase in mean establishment size over the 1962 to 1985 period. In contrast to the rising establishment mean, the lower (upper) bound on the coworker mean fell by 17% (13%) from 1968 to 1973, and it fell by a further 19% (21%) from 1974 to 1985. Over the 1968 to 1985 period,
the lower bound on the coworker mean fell in every year except 1979. In 1985 our data
bound the economywide coworker mean between 501 and 915 employees.

The sharply different time-series behavior of the establishment mean and coworker
mean—two natural measures of average size—highlights an important point. From equation
(10), the patterns in Figures 1 and 2 indicate that the variance of establishment size fell
significantly after 1967. Thus, no single summary statistic is likely to adequately capture
important secular changes in the establishment size distribution, at least over the last
twenty years. As a related point, it is often unclear whether the establishment mean or
the coworker mean (or some other statistic) is an appropriate summary measure of size.
Robinson (1988), for example, uses the establishment mean as a summary measure of size
in a regression framework designed to explain union employment density. The use of the
establishment mean as a summary measure of size is common, presumably because the
data requirements for its computation are minimal.

Despite the incomplete picture provided by a time-series of first moments over
the sample period, average size—whether from the perspective of the employee or the
establishment—is an interesting statistic. For some applications, average size is a sufficient
statistic; we present an example in section VI. Hence, we consider two natural questions
that arise in connection with the pronounced secular movements in the economywide
coworker and establishment means: (a) to what extent do the movements reflect the ex-
perience of the average industry? and (b) to what extent do they reflect changes in the
industrial distribution of aggregate employment or numbers of establishments?

We address these questions by separately tracing out the contributions of within-
industry and between-industry changes to time-series movements in the economywide es-
establishment and coworker means. For example, the time series

\[ MBI_t = \sum_i \bar{M}_i \frac{N_{it}}{N_t}, \quad t = 1962, \ldots, 1985, \]

which fixes the industry establishment means at their time averages, traces out the con-
tribution of between-industry changes to movements in the economywide establishment
mean. The time series

\[ MW_{I_t} = \sum_i M_{it} \left( \frac{N_i}{N} \right), \quad t = 1962, \ldots, 1985, \]

which fixes the industry shares of total establishments, traces out the contribution of
changes in two-digit industry establishment means to movements in the economywide
establishment mean. Likewise,

\[ CBI_t = \sum_i \bar{C}_i \frac{E_{it}}{E_t}, \quad t = 1962, \ldots, 1985, \]

traces out the contribution of shifts in industry employment shares and

\[ CWI_t = \sum_i C_{it} \left( \frac{E_i}{E} \right), \quad t = 1962, \ldots, 1985, \]

traces out the contribution of changes in two-digit industry coworker means to movements in the economywide coworker mean. In tracing out these time series, we accommodate the 1974 definitional changes in the CBP data by allowing the average industry means and shares to vary between the 1962-1973 and 1974-1985 periods.

The contribution of between-industry and within-industry changes to the time-series movements in the economywide establishment mean are shown in Figure 1. Virtually all of the secular increase in the economywide establishment mean reflects increases within two-digit industries. The MWI\(_t\) series rises by 26% from 1962 to 1973, and it rises by a further 10% from 1974 to 1985. The MB\(_I_t\) time series exhibits modest higher frequency variation over the 1974-1985 period, but the contribution of between-industry changes to secular movements in the economywide establishment mean is essentially zero.

A sharply different story emerges with respect to the coworker mean lower bound, as Figure 2 shows. (Similar patterns hold for the coworker mean upper bound.) Here, both between-industry and within-industry effects contribute to the pronounced secular decline in the economywide coworker mean after 1968.

The CBI\(_t\) time series declines by 8% from 1968 to 1973 and by a further 10% from 1974 to 1985. In view of the oft-remarked decline in manufacturing’s share of aggregate employment since the early 1970’s, the decline in the CBI\(_t\) time series is unsurprising. But two remarks are worth making. First, in terms of their impact on the coworker mean, sectoral shifts in employment shares are not primarily a phenomenon of the 1970’s. The decline in the CBI\(_t\) time series during the 1980’s is as impressive as the decline in the 1970’s, and the biggest single-year decline occurred between 1969 and 1970. Second, while the sharp decline in manufacturing’s share of aggregate employment since the early 1970’s is highly concentrated in recession years, the decline in the CBI\(_t\) series is smooth and persistent. Shifts in employment shares between industries impart no obvious cyclical pattern to fluctuations in the economywide coworker mean.
The most surprising aspect of Figure 2 is the significant decline in the economywide coworker mean due to coworker mean declines within two-digit industries. Changes within two-digit industries contribute as much to the decline in the economywide coworker mean as shifts in the industry distribution of aggregate employment. The CWI_{t} series declines by 10% from 1968 to 1973 and by a further 10% from 1973 to 1985. In other words, the significant decline in the economywide coworker mean since 1968 also reflects the experience of the average (weighted by industry employment shares) two-digit industry.

Finally, by combining the observed behavior of the CWI_{t} and MWI_{t} series, we infer that the coefficient of variation in establishment size declined after 1968 within the average industry. To draw this inference, recall that mean establishment size rose and the coworker mean fell after 1968 within the average two-digit industry. The inference then follows immediately from equation (10). Thus, in proportion to the establishment mean, the distribution of establishments by establishment size has become more tightly clustered about its first central moment since 1968 in the average industry.

C. Secular Movements in the Distribution of Employees by Establishment Size

Using the nine size classes for which we have both employee and establishment counts, we now document several important shifts in the economywide distribution of employees by establishment size over the 1962 to 1985 period. We focus on the size distribution of employees, rather than establishments, for two reasons. First, given that employment is much more evenly distributed than establishments across the CBP size classes, our data are better suited to an investigation of the employee size distribution. Second, the employee

\footnote{Recall that the CBP data contain both employee and establishment counts for only eight size classes during the 1962 to 1973 period. For this period, we use an imputation algorithm to divide the (known) employee count for the 500+ size class between two finer size classes: 500-999 and 1000+. The imputation algorithm is identical to the one we use to compute lower bounds on the coworker mean; hence, it understates the fraction of employees in the 1000+ size class. The imputation algorithm is unlikely to distort inferences about the the time-series behavior of the fraction of employees falling in the 500-999 and 1000+ size classes within the 1962-1973 subsample, but comparisons between levels of these fractions in the pre- and post-1973 samples are inappropriate. As we noted earlier, comparisons between the two subsamples are also hampered by a change in the definition of an “establishment” in 1974. However, inspection of Figure 3 indicates that this definitional change is a relatively minor impediment to comparisons of levels across the two samples. Hence, we occasionally take the liberty of making cross-sample comparisons of levels when the change in the definition of an establishment is the only consideration.}
size distribution provides a more natural perspective on most labor market phenomena than the establishment size distribution.

Figure 3 depicts the evolution of the employee size distribution over the 1962 to 1985 period. The fraction of employees who work in large establishments dropped dramatically from 1967 to 1985. The fraction of private sector employees who work in establishments with 1000+ employees declined by 13% from 1967 to 1973 and by a further 18% from 1974 to 1985. During the 1974 to 1985 period, a comparable decline occurred in the fraction of employees working in establishments with 500-999 employees, and a milder decline occurred in the fraction in establishments with 250-499 employees. In sum, the fraction of employees in establishments with 500+ employees fell from 27.6% of all employees in 1967 to 20.2% in 1985. By 1985 only 13.1% of employees worked in establishments with 1000+ employees, and only 7.0% worked in establishments with 500-999 employees.

The fraction of employees in the middle size classes—20-49, 50-99, and 100-249 employees—shows a strong, persistent rise throughout the entire sample period, especially after 1967. The fraction of employees in the middle size classes rose from 37.6% of all employees in 1967 to 40.0% in 1973 and 43.8% in 1985.

Secular movements are more mixed in the three smallest size classes. The fraction of all employees declined slightly in the 4-7 and 8-19 size classes during the earlier sample period, while the share of employees in the smallest establishments fell precipitously. The fraction of employees in the smallest size class fell from 7.4% in 1962 to 4.9% in 1973. During the 1974 to 1985 period, the share of employees in the smallest size class declined slightly, but the fraction of employees in the 5-9 and 10-19 size classes grew modestly. In 1985 the fraction of employees who worked in establishments with 1-19 employees stood at 26.9%.

We summarize our findings regarding secular changes in the employee size distribution with five points: (1) Since 1967 the fraction of private sector employees working in the largest establishments—1000+ employees—declined dramatically. (2) Since 1974 the fraction of private sector employees working at establishments that fall into the next two largest size classes—500-999 and 250-499 employees—also declined dramatically. (3) Offsetting these declines are large increases in the fraction of employees in middle-sized establishments. The fraction of employees in the 20-49, 50-99, and 100-249 size classes grew persistently from 1967 to 1985. (4) The fraction of employees in the smallest size class fell precipitously between 1962 (or 1967) and 1973, and it fell slightly between 1974 and 1985. (5) Combining the previous points, the mass in the distribution of employees by establishment size shifted away from the tails and towards the center between 1967 and 1973. This shift toward the center continued between 1974 and 1985, but it was accompanied by shifts in mass from the upper third of the distribution to the second and third smallest size classes.
D. Between-Industry and Within-Industry Effects on the Employee Size Distribution

As before, we can ask to what extent these changes reflect between-industry shifts in employment shares and to what extent they reflect the experience within the average two-digit industry. To address this issue, we proceed as before by separately tracing out the contribution of between-industry and within-industry changes to shifts in the employee size distribution. Figure 4 illustrates the contribution of changes in the industry distribution of total employees to movements in the economywide size distribution, and Figure 5 illustrates the contribution of changes in the employee size distributions within industries.

Clear, and quite distinct, patterns emerge in Figures 4 and 5. Turning first to Figure 4, from 1968 to 1985 between-industry shifts in the distribution of private-sector employees imply strong and persistent leftward movements in the distribution of employees by establishment size. The fraction of employees in each of the four largest size classes steadily shrunk after 1968, while the fraction of employees in each of the five remaining size classes steadily grew after 1968. Using the series plotted in Figure 4, between-industry shifts alone caused the fraction of employees in establishments with 100+ employees to fall from 50.2% in 1968 to 43.3% in 1985. Thus, Figure 4 confirms the view that the relative importance of employment in large establishment sectors has declined since the end of the 1960’s. Interestingly, this trend continues unabated through 1985.

Turning to the within-industry effects, the most striking aspect of Figure 5 is the pronounced, persistent rise in the fraction of employees in the three middle size classes. This trend continues throughout the entire 1962 to 1985 sample period. Using the series plotted in Figure 5, within-industry shifts alone caused the fraction of employees in establishments with 20-249 employees to rise from 36.4% of all employees in 1962 to 43.8% in 1985. In other words, the relative importance of mid-sized establishments consistently grew between 1962 and 1985 within the average (fixed employment weights) two-digit industry.

Figure 5 also shows—within the average two-digit industry—sharp declines in the relative importance of the smallest establishments since 1962, modest declines in the importance of establishments with 4-19 employees between 1962 and 1973, sharp declines in the importance of the largest establishments since 1967, modest declines in the importance of establishments with 500-999 employees, and modest rises in the importance of establishments with 250-499 employees since 1962. Roughly speaking, since 1967 the average two-digit industry has experienced a shift of the mass in the tails of its employee size distribution to the middle of the distribution.
V. Heterogeneity of Industry-Level Movements in the Employee Size Distribution

The preceding analysis documents a secular rise in the relative importance of mid-sized establishments both at the economywide level and within the average two-digit industry. The analysis paints a picture of increasing homogeneity in the size of employees' workplaces, as employee shares shift from the tails of the size distribution toward the center. In light of this picture, the following question arises: does the shift away from large and small establishments toward mid-sized establishments that occurred in the average two-digit industry also reflect the experience of most or many two-digit industries? The answer turns out to be a resounding no. There is remarkable cross-industry heterogeneity in movements of the employee size distribution, and very few industries exhibit the pattern of change experienced by the average industry. After documenting the magnitude of this cross-industry heterogeneity, we characterize its determinants in terms of two simple factors.

A. The Extent and Pattern of Cross-Industry Heterogeneity

An elementary fact about the employee size distribution is that it differs greatly across industries. For example, an employment shift toward establishments with 250-499 employees would represent a very different change in the size distribution for a typical manufacturing industry than for a typical retail industry. As a general point, when characterizing the nature of change in industry-level employee size distributions, appropriate definitions of small, mid-sized and large establishments are industry specific.

We define the industry-specific size cutoffs for small, mid-sized and large establishments as follows. We first choose a reference year for each subsample—1967 for the 1962-1973 period, and 1979 for the 1974-1985 period. In the reference year, we then select size boundaries to distribute the employee mass as evenly as possible across the small, mid-sized and large establishment size classes. Given establishment-level data, we would define size boundaries such that one-third of the employee mass fell into each of the three size classes during the reference year. In practice, we use an algorithm that apportions the employee shares in the eight or nine CBP size classes across small, medium and big size classes to minimize the maximum absolute deviation of the shares from one-third. The result of applying this algorithm is a three-point distribution during the reference year for each two-digit industry. The distribution of mass across the three size classes is as even as we can make it during the reference year, given the size-class breakdown in the CBP data.\(^3\)

\(^3\)In the 1967 reference year, application of the algorithm resulted in six industries with less than 10% of employment in one of the size classes. We deleted these industries for the empirical analysis in this section.
Using the industry-specific size-class boundaries, we computed changes in employee shares for small, mid-sized and large establishments over various time intervals. Figures 6(a) and 6(b) present scatterplots of these changes over the 1962-1973 and 1974-1985 intervals, respectively. 1967 is the reference year for Figure 6(a), and 1979 is the reference year for Figure 6(b).

A few remarks serve as useful aids to interpreting the scatterplots. Qualitatively speaking, there are six possible ways to redistribute the mass in a three-point distribution. Each way corresponds to one of the six regions in Figures 6. These regions are defined by three lines, each of which bisects the plane: the horizontal axis, the vertical axis, and a line passing through the origin with slope minus one. Points to the northeast of the line with slope minus one correspond to a decline in the share of employees at mid-sized establishments. An observation falling into region I, for example, corresponds to an industry that experienced a shift in mass from mid-sized to both small and large establishments. Region II corresponds to a shift in mass from mid-sized and big establishments to small establishments, and so on. In addition to documenting the wide disparity across two-digit industries in shifts of the employee size distribution, Figures 6 also provide information on the character of this heterogeneity. In particular, note that each of the two-digit industry observations plotted in in Figures 6 is coded by broad industry group.

We draw several inferences from Figures 6. To begin with the question posed at the outset of this section, few two-digit industries mimic the pattern of change exhibited by the aggregate economy and the average two-digit industry. While the aggregate economy and average two-digit industry observations fall into region IV, only 10% of two-digit industries fall into this region over the 1962-1973 interval and only 5% do so over the 1974-1985 interval.

During both intervals, most industries experienced an increase in the relative importance of mid-sized establishments, but this change is associated with leftward shifts in the employee size distributions in many industries and rightward shifts in many others. There is no pattern of change in the employee size distribution that typifies the experience of most two-digit industries. On the whole, movements in the employee size distributions exhibit considerable heterogeneity across industries.

However, Figures 6 reveal clear patterns of change within broad industry sectors. In the Transportation and Public Utilities sector, all nine industries experienced a decline in the big-establishment share and an increase in the small-establishment share over the 1974-1985 interval. Eight of the industries also experienced an increase in the mid-sized share. Although not illustrated in Figures 6, the leftward shift in the employee size distribution began around 1967 in most of these industries. Over the 1967-1973 interval, seven of nine
industries in the Transportation and Public Utilities sector experienced a decline in the big-establishment share.

Most industries in the Manufacturing sector experienced a similar pattern of change. Figure 6(b) indicates that eighteen of twenty-one manufacturing industries saw a decline in the big-establishment share and an increase in the small-establishment share over the 1974-1985 interval. As in the Transportation and Public Utilities sector, the secular shift away from big establishments began around 1967—fifteen of twenty-two manufacturing industries experienced a shift away from employment in big establishments between 1967 and 1973. Over both the 1962-1973 and 1974-1985 intervals, all but three two-digit manufacturing industries experienced an increase in the share of employees at mid-sized establishments.

The remaining goods-producing industries are grouped into the Other category in Figures 6. This category includes industries in the Agriculture, Mining and Construction sectors. The shifts in the employee size distributions are less uniform for these goods-producing industries, but the secular shift away from employment at big establishments after 1967 holds for most of these industries. Seven (eight) of thirteen industries in the Other category experienced a decline in big establishments’ employee share between 1967 and 1973 (1974 and 1985).

Turning to the service-producing sectors of the economy, we find a sharply different pattern of change in the employee size distribution. Over the 1962-1973 interval, thirty of the thirty-three industries in the FIRE, Trade and Services sectors experienced a decline in the fraction of employees at small establishments. Big establishments’ employee shares rose in two-thirds of the service-producing industries, and mid-sized establishments’ employee shares rose in all but five industries. Over the 1974-1985 interval, more than two-thirds of the service-producing industries experienced further declines (increases) in the share of employees accounted for by small (large) establishments.

B. Explanations for the Pattern of Cross-Industry Heterogeneity

The nature of cross-industry heterogeneity of movements in the employee size distributions, evident in Figures 6, suggests three descriptive hypotheses:

\( H_1 \): Goods-producing industries experienced leftward shifts in the employee size distribution, especially after 1967. Service-producing industries experienced rightward shifts in the employee size distribution over the entire sample period.

\( H_2 \): Industries dominated by employment at large establishments experienced leftward shifts in the employee size distribution, especially after 1967. Industries dominated by employment at small establishments experienced rightward shifts in the employee size distribution over the entire sample period. Here, we define “large” and “small” establishments relative to the economywide size distribution.
$H_3$: Capital-intensive industries experienced leftward shifts in the employee size distribution, especially after 1967. Labor-intensive industries experienced rightward shifts in the employee size distribution over the entire sample period.

The second hypothesis draws on the observation that most goods-producing industries are traditionally dominated by employment at large establishments, while the third hypothesis recognizes that goods-producing industries are relatively capital intensive. Taken together, the descriptive hypotheses point to three factors associated with secular shifts in the employee size distribution: product type, the initial size distribution and capital intensity. We now quantify the effects associated with these three factors and investigate whether any one factor offers marginal explanatory power, controlling for the other two factors.

Table 2 reports regressions of changes in the industry-specific size class shares on indicators of dominant establishment size, product type and capital intensity. The dependent variables in the regressions are changes in the employee shares of small, mid-sized and big establishments, defined in an industry-specific way as described above. The changes in size-class shares are calculated over the 1962-1967, 1967-1973, 1974-1979 and 1979-1985 time intervals. Changes over the 1962-1967 and 1974-1979 intervals are adjusted to a six-year basis to make them comparable to changes calculated over the other two intervals.

Panels A and B regress changes in the size-class shares on dummy variables that indicate the dominant establishment size class at the beginning of the interval. $SDOM$ equals one when the fraction of employees in establishments with 1-19 employees exceeds both the fraction in establishments with 20-249 employees and the fraction in establishments with 250+ employees; otherwise, $SDOM$ equals zero. Likewise, $MDOM$ ($BDOM$) equals one when establishments with 20-249 (250+) employees account for the largest fraction of employees. The column labelled "initial size test" reports the marginal significance level for a test of the null hypothesis that the coefficients on $SDOM$, $MDOM$, and $BDOM$ are equal. Rejection of this hypothesis provides evidence of a statistically significant relationship between changes in the employee size distribution and the initial size distribution.

The hypothesis of no relationship between changes in the employee size distribution and the initial size distribution is overwhelmingly rejected in Panel A. The marginal significance level of the test statistic is zero to three digits for the $CSMALL$ and $CBIG$ regressions. To interpret the magnitude of the changes implied by the regression coefficients, consider the $CSMALL$ regression. According to the coefficient on $SDOM$, the average two-digit industry dominated by establishments with 1-19 employees experienced a decline in small establishments' (industry-specific definition) employee share of 8.0 percentage points in a twelve-year interval about the reference year. In striking contrast, the average industry dominated by establishments with 250+ employees experienced a
2.2 percentage point increase in small establishments’ employee share in a twelve-year interval about the reference year. The \( CBIG \) regression coefficients also imply sharply different shifts in the employee size distribution for industries dominated by establishments with 1-19 employees as compared to industries dominated by establishments with 250+ employees. The \( CMED \) regression indicates significant and roughly equal shifts toward employment in mid-sized establishments, regardless of initial size distribution.

A potential problem in the interpretation of the regressions in Panel A involves the rather coarse breakdown on the employee size distribution available in the County Business Patterns data. Despite limiting ourselves to three categories in the industry-specific size class definitions, we are often unable to apportion the employee mass evenly across the three categories during the reference year. Uneven distribution of employee mass across the size-class categories during the reference year could lead to a spurious relationship between changes in the size distribution and the initial size distribution. To determine whether this coarseness problem underlies the results in Panel A, we restricted the sample to industries with an even distribution of mass across size classes during the reference year. In particular, we included only those industries for which each of the small, mid-sized and big establishment shares fell within the interval \([1/3 - 1/10, 1/3 + 1/10]\). Results based on the restricted sample appear in Panel B, and they largely confirm the results in Panel A. The only notable difference between the two panels is a modest attenuation in the explanatory power of \( SDOM \) in the \( CSMALL \) and \( CBIG \) regressions.

As a further check on whether the coarseness problem drives the results in Panel A, we truncated \( CSMALL \), \( CMED \) and \( CBIG \) to binary variables. The binary variables reflect only the direction of change in the industry-specific employee size-class shares and, hence, should not lead to biased inferences as a result of the coarseness problem. Cross-tabulations of these binary variables and the establishment dominance indicators appear in Table 3. Table 3 also reports the marginal significance level for a Chi-square test (two degrees of freedom) of no association between the shifts in the employee size distribution and the initial employee size distribution. The results in Table 3 strongly confirm the finding of a quantitatively and statistically significant relationship between shifts in the size distribution and the initial size distribution.

Panels C and D include a product-type indicator in the regressions of changes in employee size-class shares. The Panel C sample includes all four time intervals, whereas the Panel D sample drops the 1962-1967 interval in line with hypothesis \( H_2 \). The regression results in these panels point to sharp differences across goods-producing and services-producing industries in movements of the employee size distribution, controlling for the initial size distribution. Conditional on the initial size distribution and using the results in Panel D, the share of employees in small establishments increased by 3.1 percentage points
more in the average goods-producing industry than in the average services-producing industry over a six-year period. The relative increase in the small establishment share occurred largely at the expense of employment at big establishments.

The initial size distribution retains its explanatory power in Table 2 when we control for product type, as indicated by the initial size test results in Panels C and D. We found no evidence of important interaction between initial size distribution and product type in explaining movements in the size distributions. In sum, the Panel C and D results provide strong support for hypotheses \( H_1 \) and \( H_2 \).

Turning to the role of capital intensity, our results are less conclusive. For capital intensity measures, we rely on quality-adjusted labor and capital inputs as estimated by Jorgenson, Fraumeni and Gollop (1987). Unfortunately, the JFG data are disaggregated only to the one-digit level for the services-producing industries. Thus, we lose much of the cross-industry variation in the initial size distribution that is present in the CBP data, and we also lose the ability to distinguish product-type effects from other effects. The loss of information is apparent in Panel E, where we use the same specification and time intervals as in Panel A, but where we also limit the industry breakdown to the one available in the JFG data. Panel E shows only weak evidence that the initial size distribution matters, in contrast to Panel A.

In Panel F we include the natural log of the (quality-adjusted) capital-labor ratio. The capital intensity variable is highly significant in each regression and the magnitude of the regression coefficients are nontrivial. For example, according to the CBIG regression in Panel F, a two standard deviation increase in the capital intensity variable is associated with a 3.8 percentage point decline over a six-year period in the fraction of employees in big establishments, controlling for the initial size distribution. However, inspection of the regression residuals suggested that the estimated capital intensity effect is highly sensitive to a few observations. Panels G and H show that if we exclude three industries (Metal Mining, Anthracite Mining, and Pipe Lines except Natural Gas) from the sample, or if we exclude the first time interval, the estimated effects of the capital intensity variable differ insignificantly from zero in all three regressions. In view of this mixed pattern of results, we conclude that the data offer no compelling support for hypothesis \( H_3 \).

D. Summary

\(^4\)Initial size/product type interaction terms have no marginal explanatory power when added to the specification in Panels C and D, while the product type fixed effect continues to offer marginal explanatory power, especially in Panel D sample. Using this sample, and allowing for interaction terms, the marginal significance level on product-type indicator is .03 in the CMSALL regression and .11 in the CBIG regression.
Secular shifts in the employee size distribution exhibit remarkable heterogeneity across two-digit industries. While the average industry experienced a shift in employee mass away from the tails and toward the center of the distribution over the 1962-1985 period, fewer than 10% of individual two-digit industries experienced the same pattern. Much of the cross-industry heterogeneity in the secular movement of the employee size distribution can be explained in terms of two simple factors: product type and initial size distribution. Industries that produce goods (services), and industries dominated by establishments with 250+ (1-19) employees, experienced pronounced leftward (rightward) movements in the employee size distribution.

VI. The Employee Size Distribution and Union Density

The analysis in the preceding two sections documents significant movements in the employee size distribution over the 1962 to 1985 period and considerable cross-industry heterogeneity in the direction and magnitude of these movements. These findings point to a potentially important role for changes in the employee size distribution in explaining other secular trends in labor market behavior and institutional structure—recall our opening recitation of stylized facts about the relationship between employer size and important labor market phenomena. Motivated by these considerations, we now investigate the contribution of changes in the employee size distribution to one of these secular trends—the decline in private-sector union density. For a nice review and discussion of other efforts to explain the secular decline in private-sector union density in the United States, we refer the reader to Hirsch and Addison (1986). In an analysis closely-related to ours, Robinson (1988) develops and investigates an empirical model that relates cross-industry and time-series variation in mean establishment size to variation in industry-level union density using Canadian data.

A. Theoretical Considerations and Empirical Specification

Drawing on the May 1983 CPS and Herr’s (1989) tabulations of 1974-1976 BLS survey data, Brown et al. (1989) present evidence that the fraction of employees who are unionized is a sharply increasing function of establishment and firm size. This relationship between employer size and unionization apparently reflects the greater efforts that unions devote to organizing large employers. Brown et al. report that the annual election probability for establishments with 100+ employees is three to four times as high as for establishments with fewer than 50 employees, based on BLS and NLRB data from 1974-1976. Union win percentages, however, decline in establishment size.
These facts are consistent with the idea that organizing costs per potential union member decline with establishment size. Based on a small survey of United Textile organizers, Herr (1989) presents direct evidence that organizing costs per potential union member fall sharply with establishment size over the range of 1-400 employees. Even if organizing costs per potential employee are unrelated to establishment size, the higher death rates of small establishments imply that organizing costs are amortized over a shorter expected life. Both the fixed cost component of organizing establishments and the shorter expected life of small establishments imply a probability of unionization that increases in establishment size, other things equal. Formally,

$$PR(\text{establishment } e \text{ is unionized}) = h(j_e, x_e), \quad h_1(\cdot, \cdot) > 0,$$

(14)

where $j_e$ denotes the employee size of establishment $e$, and $x_e$ denotes other characteristics of establishment $e$ that influence union status.

Some factors conceivably lead to a negative relationship between establishment size and the probability of unionization. For example, several studies provide evidence that union wage and benefit premiums decrease in establishment size. Without taking a stand on the source of this size effect on the union wage differential, it is difficult to draw conclusions about its implications for the relationship between establishment size and union status. However, the evidence suggests that unionization offers greater rent capture for employees of small establishments than for employees of large establishments. This observation, in turn, suggests that unorganized employees in small establishments might more aggressively seek union representation. That the national union itself would more aggressively seek to organize small establishments is less clear, because union revenues in the form of dues are based on wages, not rents captured.

In sum, we carry a presumption that establishment size is positively related to the probability of unionization. Our objective here is to investigate this hypothesis and explore its ability to account for the decline in private-sector union density in the United States. Ideally, we would investigate the relationship between establishment size and union status using a large cross section or panel on individual establishments. Establishment-level data would enable us to control for various factors, in addition to establishment size, that impinge on establishment union status. In practice, we are limited to industry-level data on union density and the employee size distribution. Fortunately, there is considerable cross-industry and time-series variation in both union density and the employee size distribution at the industry level. Unfortunately, our ability to control for factors other than the size distribution that affect union density is limited; we rely entirely on various industry-level and sectoral fixed effects.
To develop an empirical specification relating union density and the employee size distribution at the industry level, we proceed as follows. Assuming that $h(\cdot, \cdot)$ is separable in $j$ and $x$, write union density in industry $i$ as

$$UD_i = \frac{N_i \sum_{j \in i} j f_i(j) h(j)}{E_i} + F_i,$$

where $f_i(j)$ is the empirical probability function over establishment size introduced in section II, $N_i$ and $E_i$ are the numbers of establishments and employees in $i$, and $F_i$ is an industry fixed effect. $F_i$ reflects the distribution of the $x$’s over establishments in $i$ and the effects of $x$’s on the probability of unionization in (14). (Strictly speaking, since $h(\cdot, \cdot)$ represents the probability of unionization, $UD_i$ denotes the expected union density in industry $i$. We ignore this distinction in the exposition.

Using equation (2), rewrite union density in terms of the empirical probability function of employees by establishment size:

$$UD_i = \frac{M_i \sum_{j \in i} g_i(j) h(j)}{E_i / N_i} + F_i = \sum_{j \in i} g_i(j) h(j) + F_i.$$  

Thus, given data on the employee size distribution, $g_i(j)$, industry-level union density is a straightforward transformation of the unionization probability function.

Equation (16) is useful in deriving a sensible specification for a union density regression equation. For example, if the unionization probability function is linear in establishment size–$h_i(j) = F_i + bj$–then equation (16) implies

$$UD_i = F_i + bC_i.$$  

In words, the coworker mean is a sufficient statistic for the employee size distribution if the unionization probability function is linear in establishment size.

We can approximate a nonlinear unionization probability function by writing $h(j)$ as a quadratic in $j$. From (16), we have

$$h_i(j) = F_i + bj + dj^2 \Rightarrow UD_i = F_i + bC_i + d(C_i^2 + V_i),$$

where $V$ denotes the variance of the employee size distribution.

Alternatively, we can approximate a nonlinear unionization probability function by writing $h(j)$ as a step function of establishment size. For example, the simple step function

$$h_i(j) = \begin{cases} 
  b_1, & \text{if } j \leq B, \\
  b_2, & \text{if } j > B,
\end{cases}$$
yields the union density equation

\[ UD_i = b_2 + (b_1 - b_2)G_i(B), \]  

(19)

where \( G_i(B) \) denotes the cumulative empirical probability function of employees by establishment size in industry \( i \).

In our empirical work on union density, we considered regression equations of the form (17)-(19). We also considered regression equations that relate industry union density to mean establishment size in the industry. We report results based on the establishment mean to facilitate comparisons with other work. It is difficult, however, to see what, if any, unionization probability function in (15) leads to a specification with mean establishment size as the appropriate explanatory variable.

In practice, we are unable to control for a full set of industry fixed effects in the cross-section regressions below. To the extent that cross-industry differences in the fixed effects, \( F_i \), are correlated with cross-industry differences in the employee size distribution, our regression analysis yields biased estimates of the slope coefficients. We attempt to control for this source of bias in our cross-section regressions by estimating fixed effects within groups of industries defined by product type.

Alternatively, time-series regressions enable us to control for a full set of industry-level fixed effects, but they suffer other problems that we regard as equally or more serious. Three problems with the time-series regressions deserve mention. First, if the \( h(\cdot, \cdot) \) function in (15) is nonseparable, then changes over time in the distribution of the \( x \)'s across establishments—or changes in the cross-partial derivatives of (15) for a fixed distribution of the \( x \)'s—can bias the estimated slope coefficients in the time-series regressions. Second, measurement error is likely to present a more serious problem when we estimate the regression coefficients from time-series variation rather than cross-industry variation. Although not reported here, we find evidence that temporary spurts (declines) in industry employment growth rates are associated with temporary rightward (leftward) shifts in the employee size distribution. It follows that much of the time-series variation in industry-level employee size distributions is likely to reflect cyclical factors, whereas the theory described above predicts a relationship between secular shifts in the employee size distribution and union density. Finally, the limitations of our data preclude any serious treatment of dynamic aspects of industry-level unionization responses to changes in the employee size distribution. To reduce the force of the measurement error problem and to minimize problems associated with potential misspecification of the dynamics, we use three-year averages of the explanatory variables, and we difference the observations over the longest time interval possible with our data.
B. Union Density Regression Results

To investigate the relationship between union membership and the employee size distribution, we use the industry-level union membership figures in Kokklenberg and Sockell (1985). The KS figures are tabulated from the 1973-1981 May Current Population Surveys. We converted the three-digit Census Industry Code observations in KS to two-digit SIC codes and merged them with CBP data on the employee size distribution. The matched data set contains observations on 65 two-digit industries. We note that the published KS union membership percentages are three-year averages of CPS tabulations; thus, our union density time series spans the period from 1974 to 1980.

Table 4 reports cross-industry regressions of union density on the coworker mean and the establishment mean using the Kokklenberg and Sockell data. The union density and size variables in Table 4 represent averages of the 1974-1980 observations.

The first two lines of Table 4 report regressions of union density on the coworker mean and the establishment mean, respectively. Both the coworker mean and the establishment mean exhibit a positive and statistically significant relationship with union density. The results indicate that a simple summary size measure alone accounts for 20-25% of the cross-industry variation in union density. However, if unobserved industry characteristics that cause high union density are positively correlated with the coworker and establishment means, then the simple bivariate specifications in lines (1) and (2) yield upwardly biased estimates of the size effect on union density. We attempt to control for this bias by considering several variants of the basic specification.

Line (3) in Table 5 controls for unobserved factors that lead to a higher union density in goods-producing industries than in services-producing industries. GOODS is a dummy variable that assumes the value of one for SIC < 50, and zero otherwise. The estimated coefficients on GOODS in lines (3)-(6) indicate that union density in goods-producing industries is 21 or 22 percentage points higher than in services-producing industries, controlling for differences in the employee size distribution across product type. Note from line (3) and (4) that this simple control for product type leads to a substantial percentage decline in the estimated size effect on union density. Line (7) extends the controls for product type to the six-way breakdown by broad industry groups we used in Figures 6: Manufacturing, Transportation/Public Utilities, Trade, FIRE, Services, and Other. The six-way breakdown leads to results similar to the two-way breakdown. This pattern of results suggests that a simple product-type fixed effect adequately controls for unobserved industry characteristics that both influence union density and are correlated with the size distribution.
We have experimented with several other variants of the basic cross-section regression, many of which are reported in Table 4. Line (5) of Table 4 includes a quadratic term in the coworker mean. (Since we have not yet computed variances for the employee size distributions, we could not estimate the specification (18).) The quadratic term does not improve the explanatory power of the regression nor does it have much effect on the estimated coefficient on the linear term. Line (6) includes both the coworker and establishment means as regressors. This specification has virtually no marginal explanatory power relative to specification (3), and only the coworker mean shows any evidence of a statistically significant effect on union density. We also estimated several variants of equation (19) and hybrids of (17) and (19). While specifications of the form (19) often led to statistically significant effects of employee size-class shares on union density, these effects disappeared in the hybrid specifications whereas the coworker mean term often retained a statistically significant effect. In short, the linear unionization probability function underlying the linear coworker mean regression in (17) seems to offer a satisfactory approximation in the cross-section regressions.

Specifications (8)-(13) in Table 4 consider regressions on selected subsets of the full industry sample. By and large, the pattern of results in (8)-(13) confirms the results in specifications (3) and (7). Restricting the sample to goods-producing industries or manufacturing industries yields coworker mean effects on union density that are statistically significant and roughly the same size as in (3) and (7). Restricting the sample to services-producing industries also yields similar size coworker mean effects on union density, although the estimated standard errors rise substantially. The large standard errors reflect the relatively small amount of variation across services-producing industries in the coworker mean.

In sum, we interpret the cross-section regression results as (a) supporting the hypothesis that establishment size influences establishment union status, (b) indicating that cross-industry variation in the employee size distribution explains part of observed cross-industry variation in union density, and (c) suggesting that the simple regression specification (17) represents a useful way to estimate the relationship between the employee size distribution and union density. Before we discuss the magnitude of the size effects implied by the coefficient estimates we describe the time-series regression results.

Table 5 reports the results of regressing the 1974 to 1980 changes in union density on changes in the employee size distribution. As before, we focus on the coworker mean and the establishment mean as summary measures of the size distribution. Recall that the KS union density data represent three-year averages. The change in the coworker mean and the establishment mean that we use as explanatory variables are calculated as a three-year average centered on 1979 minus a three-year average centered on 1973. As explained in
the preceding section, we take averages and long differences to isolate the low-frequency variation in the size measures.

Since we estimate the time-series regressions as changes, the constant term in specification (1), for example, captures time effects on union density that are equal across industries. We also consider product type and broad industry group effects in Table 5 to allow for time effects that vary across industries.

The results in Table 5 reveal a positive and significant relationship between union density and both the establishment mean and the coworker mean. The coefficient on the coworker mean is typically about one-half as large in the time-series regressions as in the cross-section regressions, except for samples restricted to services-producing industries. Restricting the sample to services-producing industries yields time-series estimates of the coworker mean effect that are of the same magnitude as the cross-section estimates. The smaller estimated size effects for goods-producing industries in the time-series regressions are consistent with the view that goods-producing industries exhibit greater cyclicity of the employee size distribution and that this cyclicity leads to a downward bias in the size effect estimated from the time-series regressions.

On the whole, we find the time-series results supportive of the interpretations that we drew from the cross-section results. In light of the sharp differences in the behavior of the employee size distribution across goods-producing and services-producing industries over the 1974 to 1980 period, we are encouraged that regressions for both types of industries show positive and statistically significant size effects on union density.

C. Movements in the Employee Size Distributions and Union Density

We now turn to the magnitude of the union density response implied by the regression results. We apply the regression results to trace out a predicted path for union density over the 1962 to 1985 period. We also follow in the footsteps of several other researchers by considering the contribution of shifts in the industrial composition of employment to changes in union density. Finally, we consider the combined predicted effect of changes in the employee size distribution and changes in the industry distribution of employment on union density. To generate our predicted union density time series, we use 1973 and 1974 as benchmark years.\(^5\)

Various predicted time series for union density appear in Figures 7-9. Figure 7 depicts the predicted path for the aggregate private nonagricultural union density. Figures 8 and 9

\(^5\)Note that since we do not have 2-digit industry union percentages for 1973, for the 1973 reference year union density we use the density implied by the 1974 industry union densities along with the actual 1973 industry employment distribution.
correspond to the goods-producing and services-producing sectors of the private economy. The effects depicted in these figures are based on the estimated coefficient on the coworker mean from line 8 of Table 4 for the goods-producing industries and line 10 of Table 4 for the services-producing industries. To provide a basis of comparison with the implied changes, we also plot the actual aggregate union density from the Kokklenberg and Sockell data for 1974-80.

Figures 7 and 8 reveal that changes in the coworker mean imply a modest decline in union density over the period. For example, for 1974-85, changes in the coworker mean imply a decrease in the aggregate union density from 25% to 24.5%, a decrease in the union density for goods-producing industries from 40.2% to 39.2%, and a decrease in the union density for services-producing industries from 11% to 10.9%. In contrast, changes in industry composition account for much larger changes in the aggregate union density but quite modest changes in the union density within the two broad sectors. For example, for the same 1974-85 period, changes in industry composition of employment imply a decrease in the aggregate union density from 25% to 21.9%, a decrease in the union density for goods from 40.2% to 39.4%, and a decrease in the union density for services from 11% to 10.9%. Thus, we observe that changes in the coworker mean account for as much of a decline as does the change in industry composition within the goods and services sectors.\(^6\)

The combined influence of changes in the coworker mean and the industry composition of employment is more impressive. For the 1974-85 period, the combined effects imply a decline in aggregate union density from 25% to 21.6%, a decline in the union density for goods industries from 40.2% to 38.6%, and a decline in the union density for services industries from 11% to 10.9%. Even though the combined influence generates significant implied decreases in the union density, comparing the implied changes with the actual changes reveals that much of the decline is still unaccounted for. For example, based on the Kokklenberg and Sockell estimates, the combined influence can account for only about 40% of the decline in aggregate union density from 1974-80. Thus our result indicate that other factors play an important role in the secular decline of union density.

### D. Summary

\(^6\)We also investigated the implied changes in union density using changes in the establishment mean (results not reported in the figures). Using the coefficient from line 9 of Table 4 for goods industries and line 11 of Table 4 for services, changes in the establishment mean imply a change in the aggregate union density from 25% in 1974 to 25.1% in 1980 to 24.9% in 1985.

28
We summarize our findings on the relationship between union density and the employee size distribution as follows. First, we show how to derive an industry-level regression specification from an establishment-level unionization probability function. In particular, for a linear unionization probability function, the industry coworker mean is a sufficient statistic for the employee size distribution in an industry-level regression. In contrast, the establishment mean does not emerge naturally as a relevant measure of size for this purpose. Second, cross-section regression results indicate that the coworker mean has a robust positive and significant relationship with union density. Third, time-series regression results indicate a robust positive and significant relationship between the change in union density and both the change in the coworker mean and the change in the establishment mean. Fourth, using the results from the cross-section regressions, we find that taking into account both changes in the coworker mean and changes in the industry composition of employment explain a substantial amount of the observed decline in union density. However, much of the observed decline remains unexplained.
References


Employment Contracts,” mimeograph.


<table>
<thead>
<tr>
<th></th>
<th>Overall Average</th>
<th>Simple Mean of Industry Time Averages</th>
<th>Stnd. Dev. of Ind/Year Obs. about Overall Average</th>
<th>Stnd. Deviations of Ind. Time Averages</th>
<th>Mean Time-series Std. Deviations (weighted by ind. employment shares)</th>
<th>Mean Time-series Std. Deviations (weighted by industry establishment shares)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coworker Mean Lower Bound</td>
<td>558.0</td>
<td>553.3</td>
<td>713.0</td>
<td>693.8</td>
<td>65.8</td>
<td>42.8</td>
</tr>
<tr>
<td>Coworker Mean Upper Bound</td>
<td>1024.4</td>
<td>887.3</td>
<td>1321.4</td>
<td>1294.4</td>
<td>136.4</td>
<td>78.5</td>
</tr>
<tr>
<td>Employment Size-Class Shares:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-4 employees</td>
<td>.07</td>
<td>.07</td>
<td>.08</td>
<td>.08</td>
<td>.006</td>
<td>.012</td>
</tr>
<tr>
<td>5-9</td>
<td>.09</td>
<td>.08</td>
<td>.08</td>
<td>.08</td>
<td>.005</td>
<td>.007</td>
</tr>
<tr>
<td>10-19</td>
<td>.11</td>
<td>.10</td>
<td>.07</td>
<td>.07</td>
<td>.005</td>
<td>.006</td>
</tr>
<tr>
<td>20-49</td>
<td>.16</td>
<td>.15</td>
<td>.07</td>
<td>.07</td>
<td>.008</td>
<td>.009</td>
</tr>
<tr>
<td>50-99</td>
<td>.12</td>
<td>.11</td>
<td>.04</td>
<td>.04</td>
<td>.008</td>
<td>.008</td>
</tr>
<tr>
<td>100-249</td>
<td>.14</td>
<td>.15</td>
<td>.07</td>
<td>.07</td>
<td>.011</td>
<td>.010</td>
</tr>
<tr>
<td>250-499</td>
<td>.09</td>
<td>.11</td>
<td>.07</td>
<td>.07</td>
<td>.006</td>
<td>.005</td>
</tr>
<tr>
<td>500-999</td>
<td>.08</td>
<td>.08</td>
<td>.06</td>
<td>.06</td>
<td>.007</td>
<td>.006</td>
</tr>
<tr>
<td>1000+</td>
<td>.14</td>
<td>.14</td>
<td>.16</td>
<td>.16</td>
<td>.015</td>
<td>.010</td>
</tr>
<tr>
<td>Establishment Mean</td>
<td>16.4</td>
<td>42.7</td>
<td>54.9</td>
<td>48.0</td>
<td>2.89</td>
<td>1.28</td>
</tr>
<tr>
<td>Establishment Size-Class Shares:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-4 employees</td>
<td>.55</td>
<td>.44</td>
<td>.21</td>
<td>.18</td>
<td>.024</td>
<td>.026</td>
</tr>
<tr>
<td>5-9</td>
<td>.21</td>
<td>.18</td>
<td>.05</td>
<td>.04</td>
<td>.015</td>
<td>.018</td>
</tr>
<tr>
<td>10-19</td>
<td>.13</td>
<td>.14</td>
<td>.05</td>
<td>.04</td>
<td>.010</td>
<td>.011</td>
</tr>
<tr>
<td>20-49</td>
<td>.08</td>
<td>.12</td>
<td>.07</td>
<td>.06</td>
<td>.008</td>
<td>.008</td>
</tr>
<tr>
<td>50-99</td>
<td>.03</td>
<td>.06</td>
<td>.05</td>
<td>.04</td>
<td>.005</td>
<td>.003</td>
</tr>
<tr>
<td>100-249</td>
<td>.015</td>
<td>.045</td>
<td>.054</td>
<td>.05</td>
<td>.004</td>
<td>.002</td>
</tr>
<tr>
<td>250-499</td>
<td>.004</td>
<td>.018</td>
<td>.025</td>
<td>.021</td>
<td>.0013</td>
<td>.0004</td>
</tr>
<tr>
<td>500-999</td>
<td>.002</td>
<td>.007</td>
<td>.011</td>
<td>.009</td>
<td>.0007</td>
<td>.0002</td>
</tr>
<tr>
<td>1000+</td>
<td>.001</td>
<td>.005</td>
<td>.010</td>
<td>.009</td>
<td>.0005</td>
<td>.0002</td>
</tr>
</tbody>
</table>
Table 2
Regressions of Industry-Specific Employee Size-Class Shares on Indicators of:
(A) Initial Size Distribution, (B) Product Type, and (C) Capital Intensity

<table>
<thead>
<tr>
<th>Panel</th>
<th>Dependent Variable</th>
<th>Time Intervals and Sample</th>
<th>Regressors</th>
<th>Initial Size Test</th>
<th># of obs.</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>SDOM</td>
<td>MDOM</td>
<td>BDOM</td>
<td>GOODS</td>
</tr>
<tr>
<td>A</td>
<td>CSMALL</td>
<td>CBP; all four intervals</td>
<td>-.040</td>
<td>-.006</td>
<td>.011</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.005)</td>
<td>(.004)</td>
<td>(.004)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CMED</td>
<td>all four intervals</td>
<td>.009</td>
<td>.007</td>
<td>.011</td>
<td>.758</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.004)</td>
<td>(.004)</td>
<td>(.004)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBIG</td>
<td></td>
<td>.031</td>
<td>-.001</td>
<td>-.022</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.007)</td>
<td>(.006)</td>
<td>(.006)</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>CSMALL</td>
<td>Restricted CBP; all four intervals</td>
<td>-.031</td>
<td>-.009</td>
<td>.012</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.006)</td>
<td>(.005)</td>
<td>(.006)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CMED</td>
<td>all four intervals</td>
<td>.004</td>
<td>.007</td>
<td>.010</td>
<td>.693</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.005)</td>
<td>(.004)</td>
<td>(.005)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBIG</td>
<td></td>
<td>.027</td>
<td>.003</td>
<td>-.022</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.009)</td>
<td>(.007)</td>
<td>(.008)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>CSMALL</td>
<td>CBP; all four intervals</td>
<td>-.044</td>
<td>-.017</td>
<td>-.004</td>
<td>.020</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.006)</td>
<td>(.006)</td>
</tr>
<tr>
<td></td>
<td>CMED</td>
<td>all four intervals</td>
<td>.010</td>
<td>.009</td>
<td>.014</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.006)</td>
<td>(.005)</td>
</tr>
<tr>
<td></td>
<td>CBIG</td>
<td></td>
<td>.034</td>
<td>.008</td>
<td>-.009</td>
<td>-.017</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.007)</td>
<td>(.008)</td>
<td>(.008)</td>
<td>(.008)</td>
</tr>
<tr>
<td>D</td>
<td>CSMALL</td>
<td>CBP; 67-73,</td>
<td>-.039</td>
<td>-.017</td>
<td>-.009</td>
<td>.031</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.006)</td>
<td>(.006)</td>
<td>(.007)</td>
<td>(.007)</td>
</tr>
<tr>
<td></td>
<td>CMED</td>
<td>74-79 and</td>
<td>.005</td>
<td>.007</td>
<td>.011</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.006)</td>
<td>(.005)</td>
</tr>
<tr>
<td></td>
<td>CBIG</td>
<td>79-85</td>
<td>.035</td>
<td>.010</td>
<td>-.002</td>
<td>-.033</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.008)</td>
<td>(.008)</td>
<td>(.010)</td>
<td>(.009)</td>
</tr>
<tr>
<td>Panel</td>
<td>Dependent Variable</td>
<td>Time Intervals and Sample</td>
<td>Regressors</td>
<td>Initial Size Test</td>
<td># of obs.</td>
<td>R²</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------</td>
<td>---------------------------</td>
<td>------------</td>
<td>------------------</td>
<td>-----------</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SDOM</td>
<td>MDOM</td>
<td>BDOM</td>
<td>log(K/L)</td>
</tr>
<tr>
<td>CSMALL</td>
<td>JFG; all four intervals</td>
<td>-.018</td>
<td>.009</td>
<td>.017</td>
<td></td>
<td>.102</td>
</tr>
<tr>
<td></td>
<td>(.017)</td>
<td>(.006)</td>
<td>(.005)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>CMED</td>
<td>.019</td>
<td>.012</td>
<td>.008</td>
<td></td>
<td>.620</td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(.004)</td>
<td>(.004)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBIG</td>
<td>-.000</td>
<td>-.020</td>
<td>-.024</td>
<td></td>
<td>.574</td>
</tr>
<tr>
<td></td>
<td>(.023)</td>
<td>(.008)</td>
<td>(.007)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSMALL</td>
<td>JFG; all four intervals</td>
<td>-.029</td>
<td>.003</td>
<td>.009</td>
<td>.0075</td>
<td>.075</td>
</tr>
<tr>
<td></td>
<td>(.017)</td>
<td>(.006)</td>
<td>(.006)</td>
<td>(.0031)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>CMED</td>
<td>.011</td>
<td>.007</td>
<td>.002</td>
<td>.0054</td>
<td>.598</td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(.005)</td>
<td>(.004)</td>
<td>(.0024)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBIG</td>
<td>.018</td>
<td>-.010</td>
<td>-.012</td>
<td>-.0126</td>
<td>.440</td>
</tr>
<tr>
<td></td>
<td>(.023)</td>
<td>(.008)</td>
<td>(.008)</td>
<td>(.0042)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSMALL</td>
<td>JFG, excluding SIC's 10, 11 and 46; all four intervals</td>
<td>-.020</td>
<td>-.002</td>
<td>.013</td>
<td>.0011</td>
<td>.009</td>
</tr>
<tr>
<td></td>
<td>(.014)</td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.0030)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>CMED</td>
<td>.017</td>
<td>.009</td>
<td>.005</td>
<td>.0011</td>
<td>.437</td>
</tr>
<tr>
<td></td>
<td>(.010)</td>
<td>(.004)</td>
<td>(.004)</td>
<td>(.0022)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBIG</td>
<td>.003</td>
<td>-.007</td>
<td>-.018</td>
<td>-.0021</td>
<td>.236</td>
</tr>
<tr>
<td></td>
<td>(.018)</td>
<td>(.007)</td>
<td>(.006)</td>
<td>(.0039)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSMALL</td>
<td>JFG; 67-73, 74-79 and 79-85</td>
<td>-.011</td>
<td>.014</td>
<td>.021</td>
<td>.0039</td>
<td>.227</td>
</tr>
<tr>
<td></td>
<td>(.019)</td>
<td>(.007)</td>
<td>(.007)</td>
<td>(.0038)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>CMED</td>
<td>.016</td>
<td>.007</td>
<td>.007</td>
<td>.0031</td>
<td>.808</td>
</tr>
<tr>
<td></td>
<td>(.015)</td>
<td>(.006)</td>
<td>(.006)</td>
<td>(.0029)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CBIG</td>
<td>-.005</td>
<td>-.021</td>
<td>-.028</td>
<td>-.0066</td>
<td>.610</td>
</tr>
</tbody>
</table>
Notes:
(1) Changes in industry-specific employee size-class shares, based on the industry size distribution in the reference year. 1967 is the reference year for changes calculated over the 62-67 and 67-73 intervals; 1979 is the reference year for the 74-79 and 79-85 intervals. Changes over the 62-67 and 74-79 intervals have been adjusted to place them on a six-year basis.
(2) The CBP sample includes all two-digit industries in the County Business Patterns data except as follows. When it was impossible to apportion at least 10% of an industry's employees to each of three separate size classes, we deleted the industry from the sample. This criterion resulted in the deletion of SIC's 19, 21, 33, 36 and 37 during the "62-67" and "67-73" time intervals. No industries were deleted in the "74-79" and "79-85" time intervals.
(3) The restricted CBP sample excludes all industries for which Max (small share-1/3, mid-sized share-1/3, big share-1/3) > .1 during the reference year.
(4) The Jorgenson, Fraumeni and Gollop (1987) data are disaggregated only to the one-digit level in the FIRE, Retail, Services, Wholesale, and Construction sectors of the economy.
(5) Define SH1-19 as the fraction of industry employees at establishments with 1-19 employees. Define SH20-249 and SH250+ analogously. SDOM is a dummy variable that takes the value 1 when SH1-19 > Max (SH20-249, SH250+), and zero otherwise. MDOM and BDOM are defined analogously.
(6) GOODS takes the values of one for goods-producing industries, zero otherwise.
(7) log(K/L) equals the log of the quality-adjusted capital-labor ratio at the beginning of the time interval, based on data in Jorgenson, Fraumeni and Gollop (1987). The sample mean and standard deviation of log(K/L) are 1.07 and 1.30, respectively.
(8) The marginal significance level in a test of the null hypothesis that the coefficients on SDOM, MDOM, and BDOM equal. The test has two degrees of freedom.
Table 3

Direction of Change in Employee Size-Class Shares
Cross-Tabulated with Dominant Size Class

<table>
<thead>
<tr>
<th>Dominant Establishment Size Class at Beginning of Interval</th>
<th>Direction of Change in (Industry-Specific) Employee Size Class During the Interval</th>
<th>Chi-Square Test (two d.f.) of No Association Between Dominant Size Class and Direction of Change (Marginal Significance Level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency (Percentage)</td>
<td>CSMALL</td>
<td></td>
</tr>
<tr>
<td>Increase</td>
<td>Decrease</td>
<td>Total</td>
</tr>
<tr>
<td>Big</td>
<td>71 (61%)</td>
<td>46 (39%)</td>
</tr>
<tr>
<td>Mid-Sized</td>
<td>45 (39%)</td>
<td>71 (61%)</td>
</tr>
<tr>
<td>Small</td>
<td>12 (15%)</td>
<td>69 (85%)</td>
</tr>
<tr>
<td>Total</td>
<td>128</td>
<td>186</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CMEC</th>
<th>Increase</th>
<th>Decrease</th>
<th>Total</th>
<th>1.08 (.582)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big</td>
<td>65 (56%)</td>
<td>52 (44%)</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>Mid-Sized</td>
<td>68 (59%)</td>
<td>48 (41%)</td>
<td>116</td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>51 (63%)</td>
<td>30 (37%)</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>184</td>
<td>130</td>
<td>314</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CBIG</th>
<th>Increase</th>
<th>Decrease</th>
<th>Total</th>
<th>29.99 (.000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big</td>
<td>42 (36%)</td>
<td>75 (64%)</td>
<td>117</td>
<td></td>
</tr>
<tr>
<td>Mid-Sized</td>
<td>63 (54%)</td>
<td>53 (46%)</td>
<td>116</td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>61 (75%)</td>
<td>20 (25%)</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>166</td>
<td>148</td>
<td>314</td>
<td></td>
</tr>
</tbody>
</table>
Notes:


(2) GOODS equals one for goods-producing industries, zero otherwise.

(3) Estimated standard errors appear in parentheses.

(4) In line (7), the sectoral dummies are Manufacturing, Transportation/Public Utilities, FIRE, Trade, Services, and Other. In line (12), the sectoral dummies are FIRE, Trade, and Services.
<table>
<thead>
<tr>
<th>Line</th>
<th>Industries</th>
<th>Intercept</th>
<th>$C_1(X10^2)$</th>
<th>$M_1$</th>
<th>$C_1^2(X10^5)$</th>
<th>GOODS</th>
<th>Sectoral Dummies</th>
<th># of Observations</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>All</td>
<td>19.68</td>
<td>0.7567</td>
<td></td>
<td></td>
<td>No</td>
<td></td>
<td>65</td>
<td>0.201</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.44)</td>
<td>(0.1902)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>All</td>
<td>17.98</td>
<td>0.1756</td>
<td></td>
<td></td>
<td>No</td>
<td></td>
<td>65</td>
<td>0.245</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.53)</td>
<td>(0.0388)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td>All</td>
<td>9.524</td>
<td>0.4220</td>
<td></td>
<td>21.71</td>
<td>No</td>
<td></td>
<td>65</td>
<td>0.535</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.418)</td>
<td>(0.1546)</td>
<td></td>
<td>(3.255)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td>All</td>
<td>9.876</td>
<td>0.0776</td>
<td></td>
<td>20.95</td>
<td>No</td>
<td></td>
<td>65</td>
<td>0.516</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.458)</td>
<td>(0.0354)</td>
<td></td>
<td>(3.533)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5)</td>
<td>All</td>
<td>9.545</td>
<td>0.4140</td>
<td></td>
<td>0.0017</td>
<td>No</td>
<td></td>
<td>65</td>
<td>0.535</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.618)</td>
<td>(0.3834)</td>
<td></td>
<td>(3.357)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6)</td>
<td>All</td>
<td>9.486</td>
<td>0.3775</td>
<td></td>
<td>0.0132</td>
<td>No</td>
<td></td>
<td>65</td>
<td>0.535</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.441)</td>
<td>(0.2393)</td>
<td></td>
<td>(3.523)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7)</td>
<td>All</td>
<td></td>
<td></td>
<td></td>
<td>0.4491</td>
<td>Yes</td>
<td></td>
<td>65</td>
<td>0.862</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.1617)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8)</td>
<td>Goods</td>
<td>31.13</td>
<td>0.4321</td>
<td></td>
<td></td>
<td>No</td>
<td></td>
<td>38</td>
<td>0.122</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.094)</td>
<td>(0.1935)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9)</td>
<td>Goods</td>
<td>30.91</td>
<td>0.0764</td>
<td></td>
<td></td>
<td>No</td>
<td></td>
<td>38</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.57)</td>
<td>(0.0427)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10)</td>
<td>Services</td>
<td>9.832</td>
<td>0.3359</td>
<td></td>
<td></td>
<td>No</td>
<td></td>
<td>27</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.975)</td>
<td>(0.3238)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11)</td>
<td>Services</td>
<td>9.224</td>
<td>0.1214</td>
<td></td>
<td></td>
<td>No</td>
<td></td>
<td>27</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.656)</td>
<td>(0.1417)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(12)</td>
<td>Services</td>
<td></td>
<td></td>
<td></td>
<td>0.4920</td>
<td>Yes</td>
<td></td>
<td>27</td>
<td>0.740</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.3225)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(13)</td>
<td>Manufacturing</td>
<td>30.03</td>
<td>0.3910</td>
<td></td>
<td></td>
<td>No</td>
<td></td>
<td>20</td>
<td>0.189</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.757)</td>
<td>(0.1910)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 5
Union Density and the Employee Size Distribution
Time Series Regressions
Dependent Variable: 1974-1980 Change in Two-Digit Industry Union Density

<table>
<thead>
<tr>
<th>Line</th>
<th>Industries</th>
<th>Regression Coefficient On (2,3)</th>
<th>Sectoral Dummies</th>
<th># of Observations</th>
<th>(R^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Intercept (C_i (X \times 10^2))</td>
<td>(M_i) GOODS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td>All</td>
<td>-1.433 (.551)</td>
<td>.1796 (.0653)</td>
<td>No</td>
<td>63</td>
</tr>
<tr>
<td>(2)</td>
<td>All</td>
<td>-1.581 (.553)</td>
<td>.0374 (.0147)</td>
<td>No</td>
<td>63</td>
</tr>
<tr>
<td>(3)</td>
<td>All</td>
<td>.1590 (.8025)</td>
<td>.1826 (.0641)</td>
<td>-2.782 (1.059)</td>
<td>63</td>
</tr>
<tr>
<td>(4)</td>
<td>All</td>
<td>.0122 (.8078)</td>
<td>.0363 (.0141)</td>
<td>-2.788 (1.069)</td>
<td>63</td>
</tr>
<tr>
<td>(5)</td>
<td>All</td>
<td>.1153 (.8084)</td>
<td>.1266 (.1017)</td>
<td>.0159 (.0228)</td>
<td>63</td>
</tr>
<tr>
<td>(6)</td>
<td>All</td>
<td>.1897 (.0634)</td>
<td></td>
<td>Yes</td>
<td>63</td>
</tr>
<tr>
<td>(7)</td>
<td>Goods</td>
<td>-2.646 (.7497)</td>
<td>.1410 (.0735)</td>
<td>No</td>
<td>36</td>
</tr>
<tr>
<td>(8)</td>
<td>Goods</td>
<td>-2.769 (.7384)</td>
<td>.0332 (.0153)</td>
<td>No</td>
<td>36</td>
</tr>
<tr>
<td>(9)</td>
<td>Services</td>
<td>.3494 (.6899)</td>
<td>.4009 (.1325)</td>
<td>No</td>
<td>27</td>
</tr>
<tr>
<td>(10)</td>
<td>Services</td>
<td>.0452 (.7151)</td>
<td>.1399 (.0575)</td>
<td>No</td>
<td>27</td>
</tr>
<tr>
<td>(11)</td>
<td>Services</td>
<td>.3808 (.1378)</td>
<td></td>
<td>Yes</td>
<td>27</td>
</tr>
<tr>
<td>(12)</td>
<td>Manufacturing</td>
<td>-3.312 (.7932)</td>
<td>.1594 (.0675)</td>
<td>No</td>
<td>20</td>
</tr>
</tbody>
</table>
Notes:


(2) GOOD equals one for goods-producing industries, zero otherwise.

(3) Estimated standard errors appear in parentheses.

(4) In line (11), the sectoral dummies are Manufacturing, Transportation, Public Utilities, FIRE, Trade, Services, and Other. In line (11), the sectoral dummies are FIRE, Trade and Services.
Figure 1.
Mean Size of Establishments or Reporting Units, 1962 and 1964-1985

- Mean Size
- Between Industry Effects (2-Digit)
- Within Industry Effects (2-Digit)
Figure 2.
Coworker Mean for Establishments or Reporting Units, 1962 and 1964-1985

- Lower Bound on the Coworker Mean
- Between-Industry Effects on Lower Bound
- Within-Industry Effects on Lower Bound

Size

700

660

620

580

540

500

Year

6 6 6 6 6 6 6 6 6 7 7 7 7 7 7 7 7 7 7 7 7 8 8 8 8 8 8 8 8 8 8 8 8 8 8
2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
Figure 5.
Distribution of Employees by Establishment Size Class: Within Industry Effects (2-digit)

- + 1–3
- × 4–7
- * 8–19
- □ 20–49
- ◊ 50–99
- △ 100–249
- # 250–499
- ♠ 500–999
- ♥ 1000+
- ♦ 1–4
- ○ 5–9
- < 10–19

Percent

Year
Two hidden observations: (.09, -.22, Other) and (.02, -.20, FIRE).
Figure 6(b).
Changes in Industry-Level Employee Size Distributions between 1974 and 1985

Change in Small Establishment Share

□ Tran/PU
◊ Mfg
△ FIRE
☆ Serv
+
Other
γ Trade

Three hidden obs: (.28, -.20, Other), (.25, -.38, Tran/PU), (.07, -.33, Other)