Volume Title: The Service Economy

Volume Author/Editor: Victor R. Fuchs, assisted by Irving F. Leveson

Volume Publisher: UMI

Volume ISBN: 0-87014-475-8

Volume URL: http://www.nber.org/books/fuch68-1

Publication Date: 1968

Chapter Title: INTERINDUSTRY AND INTERSECTOR DIFFERENCES IN HOURLY EARNINGS

Chapter Author: Victor R. Fuchs

Chapter URL: http://www.nber.org/chapters/c1160

Chapter pages in book: (p. 128 - 157)

6

INTERINDUSTRY AND INTERSECTOR DIFFERENCES IN HOURLY EARNINGS

In Chapter 3 we saw that earnings in most service industries have not grown as rapidly as in the Industry sector since 1929. The differential rate of growth of earnings was interpreted as implying a differential rate of growth of the quality of labor, although it was indicated that a portion of the gap in earnings could be attributed to the growth of unions in Industry. Some limited evidence on changes in demographic characteristics supported the inferences based on earnings trends.

In Chapter 5, questions concerning earnings, demographic characteristics, and labor quality were again raised. David Schwartzman's hypothesis of a decline in the service provided by retail trade per constant dollar of sales is closely tied to the hypothesis of a decline in the quality of labor in retailing. In the case study of barber and beauty shops, inferences concerning labor quality and idle time are dependent upon the interpretation of earnings data.

This chapter presents a much more systematic look at interindustry and intersector differences in earnings, particularly in relation to demographic characteristics and other variables, such as unionization. Use of the 1/1,000 sample of the *1960 Census of Population* permits systematic analysis of the *level* of earnings, and provides the basis for the estimates concerning unionization, labor quality, and demographic characteristics that were used in Chapters 3 and 5.

The richness of the data makes possible more comprehensive comparisons than those based on economic censuses or sample surveys. All service and nonservice industries can be included, and in considerable detail. Also, it is possible to include salaried employees and self-employed workers as well as production workers. This is particularly important in the Service sector where more than half of all employed persons are either salaried or self-employed.

Average hourly earnings are estimated for 1959 for every detailed industry, major industry group, and the Service and Industry sectors. Intersector differences in earnings are also examined by color, age, sex, and education. We find that hourly earnings in the Industry sector are, on average, 17 per cent higher than in the Service sector; much of the chapter is devoted to exploring and explaining this difference.

Differentials in Earnings

The Estimation of Actual and "Expected" Hourly Earnings

The basic data source for this chapter is the 1/1,000 sample of the *1960 Census of Population and Housing*. A detailed description of the 1/1,000 sample and of the statistical procedures followed in preparing the earnings estimates is given in Appendix E. The principal points to be made here are:

1. The population studied includes all persons who were employed in nonagricultural industries during the Census "reference" week (varying weeks in April) in 1960, and who had some earnings in 1959. The total number of persons covered in the sample was 56,247. Persons employed in agriculture were excluded because average hourly earnings for such persons present special problems of reliability and interpretation.[1]

2. Average hourly earnings for each industry are estimated in the following manner. Average annual hours are obtained by multiplying the number of weeks worked in 1959 by the number of hours worked in the Census reference week in April 1960, for each worker, and then summing for all workers in an industry. This method provides a more accurate estimate of total man-hours than would be obtained by multiplying average weekly hours by the average number of weeks worked because there is a positive correlation between number of weeks and hours per week across individuals. Although the use of hours for a single week in 1960 and inaccuracy in reporting of hours may produce considerable error in the estimate of annual hours (and hourly earnings) for any single worker, much of this error is random in nature and tends to cancel out for large groups.[2] Sampling errors are greatest when there are few observations; all measures that are based on fewer than fifty observations are identified. Aggregate annual earnings for 1959 for each

[1] Hours worked and earnings are both less likely to be reliably reported by agricultural workers. Moreover, some labor income may be earned in kind (unreported), or reported earnings may include substantial returns to land and capital as well as labor income.

[2] The reliability of the hours data is examined in Appendix F.

TABLE 45

Average Hourly Earnings of Nonagricultural Employed Persons, by Demographic Characteristics, United States, 1959

Characteristics	Dollar Earnings Per Hour	Characteristics	Dollar Earnings Per Hour
Color		Sex	
Whites	2.58	Males	2.79
Nonwhites	1.61	Females	1.70
Age		Years of schooling	
14–19	1.38	0–4	1.66
20–24	1.73	5–8	2.09
25–34	2.38	9–11	2.26
35–44	2.72	12	2.40
45–54	2.71	13–15	2.92
55–64	2.62		
65 and over	2.50	16 and over	3.96
		All	2.50

Source: *U.S. Censuses of Population and Housing: 1960. 1/1,000 Sample;* calculations by the author.

industry are divided by the aggregate annual hours for that industry to obtain the average hourly earnings.[3]

3. Past studies of industry differences in earnings have often tried to take account of industry differences in labor quality. One customary approach is to standardize for industry differences in occupation. This is a useful method, but deficient to the extent that there are labor quality differences within the same occupation across industries. Also, there are many occupations that are specific to only a few industries. An alternative, and possibly more direct, approach to the problem is to look at such labor quality proxies as color, age, sex, and education, since it is well-known that there are significant wage differentials at the national level associated with each of these variables. Table 45 summarizes these differentials in gross form; detailed tables are presented in Appendix E. It is readily apparent that industry differences in the composition of the

[3] This is equivalent to calculating average hourly earnings for each worker and then calculating a weighted average for all workers in an industry where the weights are the annual hours of each worker.

labor force with respect to these variables could be an important source of industry differentials in hourly earnings.

All nonagricultural employed persons who had some earnings in 1959 were grouped into 168 cells by color (two classes), sex (two classes), age (seven classes), and years of schooling (six classes). National hourly earnings rates for each cell were calculated by the method described in the preceding section. "Expected" hourly earnings for each industry were then calculated by assuming that each worker in the industry had an hourly earnings rate equal to the national rate for workers with his color, age, sex, and education. To the extent that labor quality is associated with these characteristics, differences in average "expected" earnings across industries measure differences in labor quality; differences between actual and "expected" earnings measure differences in wages holding labor quality constant.[4]

This quality adjustment is necessarily imperfect. Two shortcomings worth noting are: (1) the failure to take account of differences in ability within educational classes, and (2) the open-ended class, "16 years and over," includes workers with very varied degrees of schooling. Despite these and other shortcomings, however, this approach to standardization yields reasonable and useful results. Moreover, it has the advantage of allowing for important interactions between the various quality proxies, such as those between education and color and education and age.

Sector Differences in Earnings

Table 46 presents average hourly earnings in 1959 for each sector and major industry group. Also presented are the "expected" hourly earnings based on the color, age, sex, and education composition of the labor force. The last column in each category shows the ratio of actual to expected earnings.

We see that actual earnings in the Industry sector are 39 cents, or 17 per cent, higher than in the Service sector, and that this difference is not explained by the composition of the labor force since the expected earnings are almost identical in the two sectors.

With one exception, every major industry group in the Industry sector has a ratio of actual to expected of over 1.0, whereas in the Service sec-

[4] Systematic differences in national hourly earnings rates by color, age, sex, and education suggest that these variables do, to some extent at least, measure labor quality. The white-nonwhite differences are probably due in part to market discrimination, but color is relevant to quality because of the likelihood that, at given levels of education, nonwhites have received poorer quality schooling and less on-the-job training than have whites.

TABLE 46

Actual and Expected [a] Hourly Earnings, by Sector and Major Industry Group, 1959

	Total			Males			Females		
	Actual	Expected	Ratio of Actual to Expected	Actual	Expected	Ratio of Actual to Expected	Actual	Expected	Ratio of Actual to Expected
All Industries [b]	$2.50	$2.50	1.00	$2.79	$2.79	1.00	$1.70	$1.70	1.00
Industry sector	2.70	2.50	1.08	2.88	2.69	1.07	1.84	1.63	1.12
Service sector	2.31	2.51	.92	2.69	2.94	.91	1.63	1.74	.94
Mining	2.89	2.62	1.10	2.94	2.66	1.10	1.91 [c]	1.83 [c]	1.05 [c]
Construction	2.87	2.58	1.11	2.90	2.61	1.11	2.02	1.78	1.13
Manufacturing	2.67	2.47	1.08	2.91	2.70	1.08	1.80	1.61	1.12
Durable	2.79	2.54	1.10	2.94	2.69	1.09	1.96	1.65	1.19
Nondurable	2.51	2.38	1.05	2.86	2.72	1.05	1.69	1.58	1.07
Transportation	2.70	2.57	1.05	2.75	2.63	1.04	2.01	1.73	1.16
Communications and public utilities	2.63	2.45	1.07	2.90	2.75	1.06	1.96	1.71	1.15
Communications	2.69	2.28	1.18	3.40	2.87	1.18	2.00	1.70	1.18
Public utilities	2.58	2.59	1.00	2.68	2.70	.99	1.84	1.77	1.04
Postal service	2.58	2.78	.93	2.61	2.87	.91	2.24	1.77	1.27

Wholesale trade	2.88	2.71	1.06	3.04	2.89	1.05	1.95	1.73	1.13
Retail trade	1.96	2.37	.83	2.22	2.68	.83	1.34	1.63	.82
Finance, insurance, and real estate	2.92	2.63	1.11	3.58	3.17	1.13	1.82	1.72	1.06
Business and repair services	2.43	2.56	.95	2.50	2.70	.92	2.04	1.79	1.14
Personal services	1.36	1.82	.74	1.93	2.57	.75	1.00	1.35	.74
Entertainment and recreation	2.28	2.40	.95	2.39	2.62	.91	1.94	1.71	1.13
Professional and related services	2.63	2.77	.95	3.31	3.58	.92	1.97	1.99	.99
Public admin.	2.49	2.71	.92	2.64	3.02	.88	2.06	1.86	1.11
Industry not reported	2.35	2.29	1.03	2.61	2.58	1.02	1.72	1.61	1.07

Notes to Table 46

Note: The ratios of Actual to Expected Earnings in this and succeeding tables were computed from more detailed data than are shown in the tables.

Source: 1/1,000 Sample.

[a] Expected earnings are obtained by assuming that each worker in each industry had hourly earnings equal to the national average for his color, age, sex, and education.

[b] Excludes agriculture, forestry, and fisheries, as do all the tables in this chapter.

[c] Based on fewer than fifty observations.

tor every group except two has a ratio of below 1.0. Service sector earnings are particularly depressed by personal services and retail trade; communications, construction, mining, and durable manufacturing show the highest ratios in the Industry sector.

The sector differences in the ratios of actual to expected are similar for males and females, 17 per cent and 19 per cent, respectively. Within the Service sector, however, we find that females have relatively high earnings in several industry groups in which male earnings are low.

Table 47 explores the intersector differences in earnings for specific demographic groups. The difference for each group is summarized in the last column where the ratio of actual to expected in the Industry sector is shown relative to the actual/expected ratio in the Service sector. We find that the intersector difference is appreciably greater for nonwhites than for whites, and is greater the lower the level of education. Moreover, these results are confirmed for educational groups within each color and for white-nonwhite comparisons at each level of education.

It should be noted that the same data can be used to analyze absolute differentials, as shown in Table 48. In these terms, the Industry-Service differential is larger for males than for females, and is the same for whites as for nonwhites. The absolute differential first decreases with additional schooling, then rises, and increases steadily with age. In the regression analyses presented in the next section of this chapter, both relative and absolute differentials are studied.

Examination of the measures for the detailed industries reveals that industries differ much more with respect to actual than to expected earnings.[5] This is true within each sector as well as for the total. Charts 9 and 10 show that there is a very marked central tendency for expected earnings, around $2.50 per hour, whereas actual hourly earnings tend to be highly dispersed. Chart 9 is based on simple frequency distributions of industries, whereas Chart 10 is based on distributions in which each industry is weighted by its number of employees. The midpoints for the open-ended classes are based on the median and weighted median industries in those classes, respectively.

Comparison of the distributions for the two sectors shows that the service industries tend to be much more heterogeneous with respect to both actual and expected earnings. At the lower extreme are the personal services, while at the upper end are such high-wage, high-skill services as medical and legal.

[5] Appendix Table I-1 presents actual and expected hourly earnings for each of 139 industries. This represents the entire *Census of Population* industry list with the exception of a few "not specified" industries.

When we look at the ratios of actual to expected earnings, we find that most industries (73 per cent) in the Industry sector are greater than 1.0 and most service industries (66 per cent) are below 1.0. The cumulative frequency distributions, weighted and unweighted, are shown in Chart 11. A chi square test shows that the two sectors differ significantly from each other at the .001 level of confidence.

Regression Analysis

This section reports the results of regression analyses of interindustry differences in earnings. The primary purpose is to test certain hypotheses concerning industry characteristics that are believed to affect earnings. The use of expected earnings as one of the independent variables permits a rigorous test of these hypotheses after allowing for industry differences in demographic characteristics, i.e., it tests the influence of industry characteristics on standardized wages. The identification of wage-related industry characteristics helps to explain the sector differentials in earnings described in the previous section.

Demographic Characteristics (X_1)

Table 49 shows the results when actual earnings are regressed on expected earnings. We see that well over half the interindustry differences in hourly earnings can be explained by differences in demographic characteristics alone.[6] In the weighted logarithmic run, as much as 70 per cent of interindustry variation is explained by this one variable. In both the linear and logarithmic runs, weighting the industries by their total man-hours tends to improve the correlation because those industries with the greatest residuals are typically small industries. In the multiple regressions that follow, weighting by industry man-hours is applied throughout. This tends to minimize the disturbances introduced by the inclusion of industries whose estimated earnings are based on few observations.

The regression coefficient in the linear equation shows the number of cents change in actual earnings per hour that is associated with a one cent change in expected earnings. The logarithmic form, which gives a slightly better fit, shows the percentage change in actual earnings asso-

[6] By comparison, it may be noted that when actual earnings were regressed on expected earnings across twenty-eight region-city size cells, the $\bar{R}^2$ (adjusted coefficient of determination) was only .36. Demographic characteristics explain twice as much of interindustry variation in earnings as they do of geographical variation in earnings.

TABLE 47

Actual and Expected Hourly Earnings in Industry and Service Sectors, by Sex, Color, Education, and Age, 1959

	Industry Sector			Service Sector			
	Actual Hourly Earnings (A)	Expected Hourly Earnings (E)	A/E	Actual Hourly Earnings (A)	Expected Hourly Earnings (E)	A/E	Industry A/E / Service A/E
All nonagricultural employed persons	$2.70	$2.50	1.078	$2.31	$2.51	.921	1.17
Males	2.88	2.69	1.072	2.69	2.94	.914	1.17
Females	1.84	1.63	1.124	1.63	1.74	.941	1.19
Whites	2.75	2.56	1.075	2.41	2.62	.923	1.16
Nonwhites	1.97	1.73	1.137	1.36	1.53	.888	1.28
Years of schooling							
0–4	1.95	1.76	1.109	1.26	1.54	.817	1.36
5–8	2.38	2.19	1.090	1.69	1.97	.857	1.27
9–11	2.56	2.38	1.078	1.91	2.14	.892	1.21
12	2.67	2.50	1.070	2.15	2.31	.928	1.15
13–15	3.30	3.08	1.072	2.68	2.82	.951	1.13
16 and over	4.42	4.14	1.068	3.75	3.90	.961	1.11

Whites							
0–4	2.05	1.88	1.093	1.47	1.80	.820	1.33
5–8	2.43	2.24	1.086	1.80	2.10	.858	1.27
9–11	2.60	2.41	1.077	1.99	2.22	.893	1.21
12	2.69	2.52	1.068	2.19	2.36	.930	1.12
13–15	3.33	3.12	1.071	2.74	2.88	.951	1.12
16 and over	4.45	4.16	1.070	3.79	3.95	.960	1.11
Nonwhites							
0–4	1.68	1.44	1.166	.95	1.17	.811	1.44
5–8	1.91	1.66	1.148	1.12	1.32	.847	1.36
9–11	2.02	1.80	1.121	1.29	1.47	.876	1.28
12	2.19	1.91	1.145	1.49	1.65	.899	1.27
13–15	2.27	1.96	1.159	1.74	1.82	.953	1.22
16 and over	2.47 [a]	2.85 [a]	.864 [a]	2.71	2.71	.999	.86 [a]
Age							
14–19	1.45	1.39	1.045	1.29	1.38	.934	1.12
20–24	1.91	1.75	1.092	1.59	1.72	.923	1.18
25–34	2.56	2.38	1.076	2.19	2.40	.911	1.18
35–44	2.89	2.71	1.065	2.54	2.73	.928	1.15
45–54	2.89	2.69	1.076	2.52	2.73	.925	1.16
55–64	2.88	2.63	1.096	2.41	2.62	.918	1.19
65 and over	2.94	2.50	1.177	2.28	2.52	.906	1.30

Note: See note to Table 46.

[a] Based on fewer than fifty observations.

TABLE 48

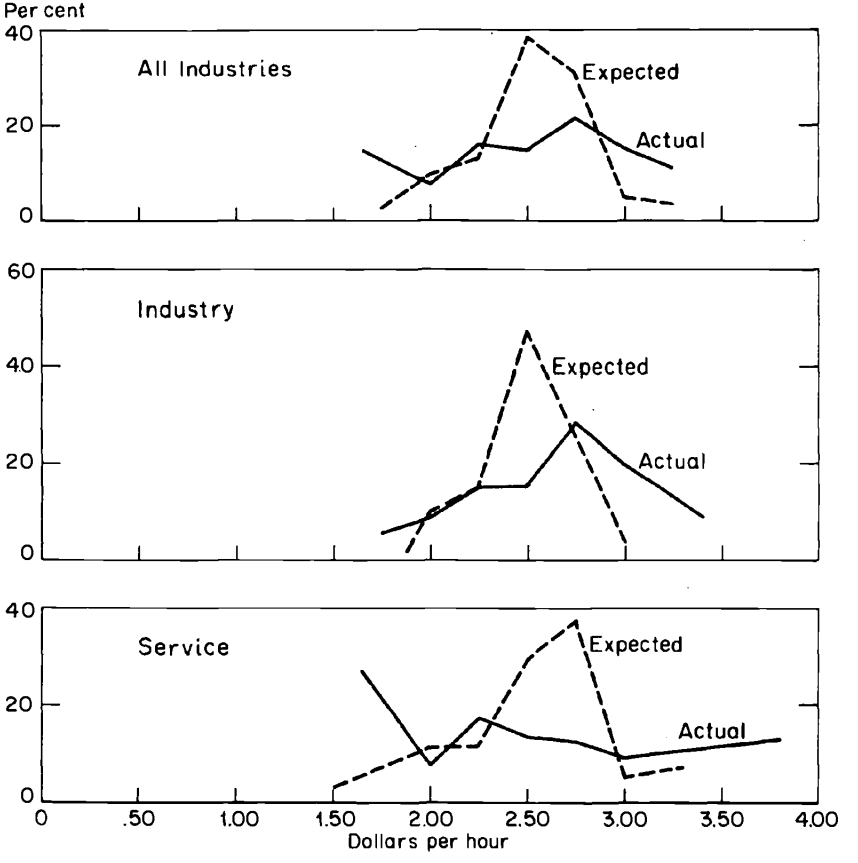
Absolute Differentials in Hourly Earnings in Industry and Service Sectors, by Sex, Color, Education, and Age, 1959

	Actual Minus Expected Earnings		Column 1 Minus Column 2
	Industry (1)	Service (2)	
All	$.20	$−.20	$.40
Males	.19	−.25	.44
Females	.21	−.11	.32
Whites	.19	−.21	.40
Nonwhites	.24	−.17	.41
Years of schooling			
0–4	.19	−.28	.47
5–8	.19	−.28	.47
9–11	.18	−.23	.41
12	.17	−.16	.33
13–15	.22	−.14	.36
16 and over	.28	−.15	.43
Whites			
0–4	.17	−.33	.50
5–8	.19	−.30	.49
9–11	.19	−.23	.42
12	.17	−.17	.34
13–15	.21	−.14	.35
16 and over	.29	−.16	.45
Nonwhites			
0–4	.24	−.22	.46
5–8	.25	−.20	.45
9–11	.22	−.18	.40
12	.28	−.16	.44
13–15	.31	−.08	.39
16 and over	−.38 [a]	.00	−.38 [a]
Age			
14–19	.06	−.09	.15
20–24	.16	−.13	.29
25–34	.18	−.21	.39
35–44	.18	−.19	.37
45–54	.20	−.21	.41
55–64	.25	−.21	.46
65 and over	.44	−.24	.68

[a] Based on fewer than fifty observations.

CHART 9

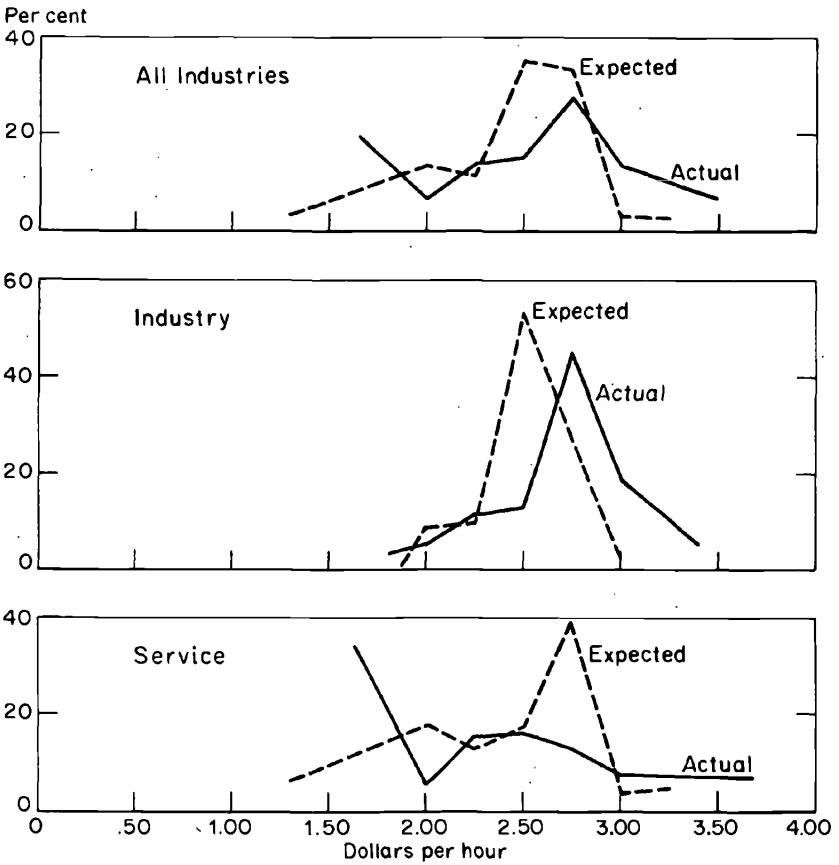
Distribution of Industries, by Actual and Expected Hourly Earnings



ciated with a 1 per cent change in expected earnings. A priori, this coefficient might be expected to be equal to 1.0, but we observe that it is significantly above unity in all four runs. In regressions described later in this chapter, inclusion of additional variables brings the demographic characteristics regression coefficient closer to 1.0 and, for runs limited to the Industry sector, it falls to slightly below 1.0. In the Service sector, however, there are several industries at either end of the skill scale that account for the coefficient exceeding unity. At the upper end, security and commodity brokers, medical except hospital, and legal all have high expected earnings, but actual earnings far exceed expected. At the lower end, several of the personal services and retail trades have low

CHART 10

Weighted Distribution of Industries, by Actual and Expected Hourly Earnings



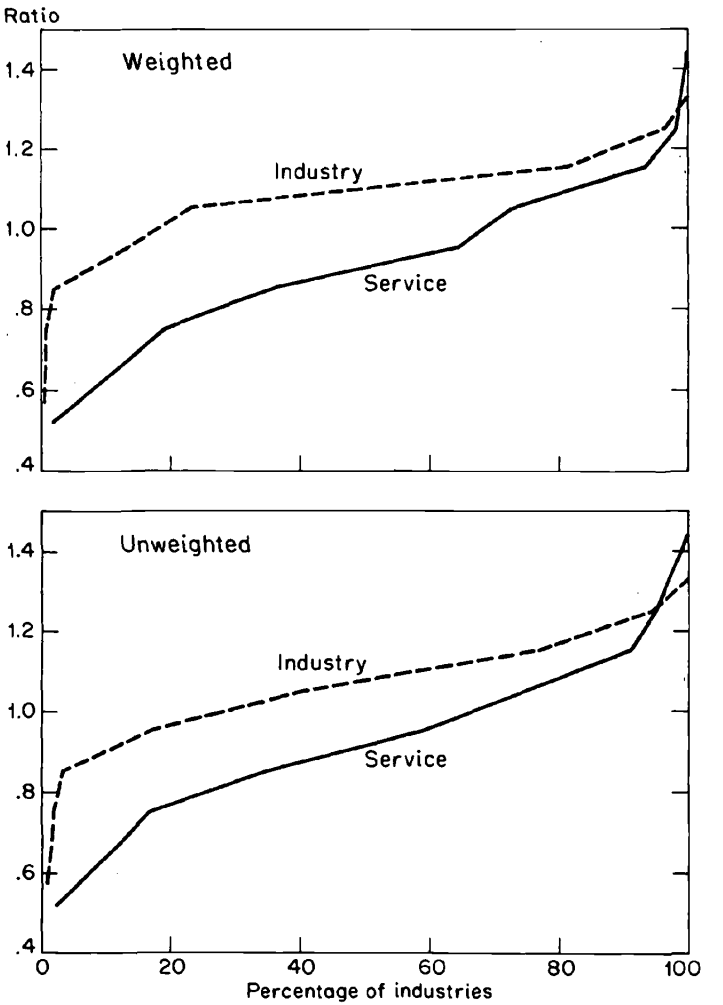
expected earnings but even lower actual earnings. These deviations suggest the possibility that the six education classes do not provide an adequate measure of skill at the extreme ends; it is also possible that there are errors in the reporting of earnings.[7]

The demographic characteristics variable goes a long way toward explaining interindustry differences in earnings generally but it is of no help in explaining the differential between the Industry and Service sectors. When the residuals are examined for the log-weighted run, we find

[7] For instance, reported earnings for food stores, eating and drinking places, and private households might omit income in kind.

CHART 11

Cumulative Frequency Distribution of Industries, by Ratio of Actual to Expected Earnings



that 74 per cent of the industries in the Industry sector have observed earnings that exceed predicted values, while only 30 per cent of the Service sector industries are in that category. Because mean expected earnings are almost identical in the two sectors, it is not surprising that demographic characteristics cannot explain intersector differences in actual earnings. To do so it is necessary to identify variables that have

TABLE 49

Results of Regression of Hourly Earnings on Demographic Characteristics Across All Industries
($N = 138$)

	$\bar{R}^2$	Regression Coefficient	t Value [a]
Linear: unweighted	.572	1.456	13.58
weighted	.641	1.448	15.67
Logs: [b] unweighted	.576	1.462	13.69
weighted	.698	1.547	17.83

Note: Demographic characteristics are measured by expected hourly earnings. Industry 879, "welfare and religious services" was excluded from all regression analyses because of special problems of measuring and interpreting the compensation of priests, nuns, and other religious workers.

Source: Appendix Table I-1.

[a] These are equal to the regression coefficient divided by its standard error. These values indicate extremely high statistical significance.

[b] All logarithms of variables are in natural logs.

different values in the two sectors and are thought to be significantly related to earnings. The most promising is the extent of unionization.[8]

Unionization (X_2)

It is well-known that the workers in much of the Industry sector are highly organized, whereas unions are weak or nonexistent in most service industries. For the Industry sector as a whole, nearly 50 per cent of all employed persons were union members in 1960. For the Service sector as a whole, the comparable figure was less than 10 per cent.[9] It is widely (though far from universally) believed that workers in unionized industries receive higher wages than do those in other industries.[10]

[8] Another variable that is often mentioned along with unionization is industry concentration; however, the evidence presented by L. W. Weiss in "Concentration and Labor Earnings," *American Economic Review*, March 1966, convincingly rejects the hypothesis of a systematic relation between concentration and earnings, *ceteris paribus*. Adrian Throop reports a similar finding in his Stanford Ph.D. dissertation, 1967, "Sources of Inflationary Bias in the American Economy," Chapter 5.

[9] These estimates were derived from the data assembled from a number of sources by H. Gregg Lewis and presented in *Unionism and Relative Wages in the United States*, Chicago, 1963.

[10] For comprehensive reviews of this subject, see H. Gregg Lewis, *ibid.*, and George H. Hildebrand, "The Economic Effects of Unionism," in *A Decade of*

TABLE 50

Results of Regression of Hourly Earnings on Demographic Characteristics and Unionization Across All Industries ($N = 138$)

Form of X_2	$\bar{R}^2$	Demographic Characteristics (X_1)		Unionization (X_2)	
		Regression Coefficient	t Value	Regression Coefficient [a]	t Value
		Linear			
$U =$ 0 to 100	.748	1.449	18.71	.0075	7.64
$U =$ 0 to 60	.746	1.451	18.69	.0081	7.58
$U =$ 20 to 60	.773	1.426	19.39	.0125	8.93
$U =$ 20 to 100	.767	1.423	19.12	.0107	8.66
Reciprocal of $U =$ 20 to 100	.774	1.432	19.51	.0108 [b]	
		Logarithmic [c]			
$U =$ 0 to 100	.797	1.524	21.38	.33	8.18
$U =$ 0 to 60	.797	1.524	21.43	.36	8.22
$U =$ 20 to 60	.819	1.506	22.36	.55	9.58
$U =$ 20 to 100	.812	1.504	21.93	.46	9.15
Reciprocal of $U =$ 20 to 100	.820	1.510	22.51	.48 [b]	

[a] The linear runs show the change in dollars per hour and the logarithmic runs in per cent associated with a change of one percentage point in unionization.

[b] At the unionization mean of 35.52 per cent, this is the union effect implied by the regression coefficient.

[c] The dependent variable, hourly earnings (X_0) and X_1 are in natural logarithms; unionization (X_2) is not.

The unionization variable used in this study measures, in principle, the fraction of total employment covered by collective bargaining agreements. In practice it is obtained through a variety of methods and sources (see Appendix I), and the figures for some industries are undoubtedly subject to considerable measurement error.

Despite these measurement problems, the regression results reveal a strong and consistent relation between unionization and earnings, after allowing for demographic characteristics (see Table 50). It is doubtful that the elimination of measurement error would make these results

Industrial Relations Research, Neil W. Chamberlain, Frank C. Pierson, and Theresa Wolfson (eds.), New York, 1958.

weaker or less consistent.[11] Unionization and demographic characteristics together explain up to 82 per cent of all interindustry variation in hourly earnings.

Another important problem concerns the form of the relation between unionization and earnings. The adjusted coefficients of multiple determination ($\bar{R}^2$) indicate that the best fits are obtained when the unionization variable is limited to the range 20 to 60 per cent (all industries below or above that range are set equal to 20 or 60, respectively) or when the reciprocal of unionization is used and limited to the range 20 to 100 per cent.

It should be noted that, if the true relation between earnings and unionization were linear throughout, the grouping of observations at 20 and 60 per cent would tend to lower the $\bar{R}^2$. In fact, the $\bar{R}^2$ is increased, and by an amount that is statistically significant.[12]

The results, therefore, suggest the possibility that the relation between standardized earnings and unionization is not linear throughout. At low levels, changes in unionization seem to have little effect on earnings. This also appears to be true once most of the workers in an industry are covered by collective bargaining agreements. It is possible that measurement error obscures any union effect below 20 per cent, since the measurement of unionization in these industries is particularly poor. However, it should be noted that there is still a statistically significant difference between the fit of unionization 20 to 100 per cent and unionization 20 to 60 per cent in the logarithmic runs.

In the linear runs, the effects of unionization on hourly earnings can be read directly from the regression coefficients. For instance, in the third linear run the coefficient .0125 implies that, over the range 20 to

[11] Random errors of measurement in an independent variable bias the regression coefficient *downward* if the hypothesized relation is positive. See J. Johnston, *Econometric Methods,* New York, 1963, pp. 149–150.

[12] Yoel Haitovsky has suggested to me the following test for statistical significance:

$$F_{(1,n-p-1)} = \frac{R^{*2} - R^2}{2 - R^{*2} - R^2}(n - p - 1)$$

If we apply this test to the linear runs, where unionization can take any value from zero to 100 and where it is restricted to 20 to 60, we obtain the following F value:

$$F = \frac{.776 - .751}{2 - (.776 + .751)}(135) = 7.1,$$

which is highly significant.

I am also grateful to F. Thomas Juster and Jack Johnston for advice concerning the statistical significance of these results.

60 per cent unionization, an increase of 1.25 cents per hour is associated with an increase of one percentage point in unionization. The first linear run indicates a change of .75 cents per hour per percentage point change in unionization over the full range from 0 to 100 per cent. The run in which the reciprocal of unionization is used results in a different unionization effect at different levels of unionization. The effect shown is the one implied at the mean level of unionization.[13]

In the runs where actual and expected earnings are in natural logarithms, the union effect is given by the regression coefficients, with the decimal point shifted two places to the right. The values shown in Table 50 are the percentage change in earnings associated with a one percentage point change in unionization. The value for the last run is again based on the implied change at the mean level of unionization.

It should be noted that the results concerning unionization are open to more than one interpretation. It is clear that, other things being equal, earnings are higher the higher the level of unionization, but this does not necessarily mean that they are higher for workers of equal skill. The adjustment for color, age, sex, and education cannot capture all differences in labor quality, and it is possible that some of the unmeasured quality differences are correlated with the extent of unionization. This would be true if unions exercised some quality control, or if the higher wages that unionized employers pay permits them to be more selective in their hiring within given demographic groups. Other factors, such as the nonpecuniary advantages and disadvantages of different types of employment also enter into the determination of hourly earnings, but it is doubtful that these have a systematic negative correlation with unionization.[14]

The combined explanatory power of the two independent variables is very high; moreover, the unionization variable does explain most of the Industry-Service differential. The residuals now show almost as high

[13] If the regression equation is

$$X_0 = a + b_1 X_1 + \frac{b_2}{X_2},$$

the effect of a 1-unit change in X_2 would be

$$\frac{\partial X_0}{\partial X_2} = -\frac{b_2}{X_2^2}.$$

The mean value of X_2 and the regression coefficient b_2 are substituted to find the effect of unionization on earnings.

[14] It should be emphasized that this analysis is concerned only with the relation between the structure of earnings and unionization, and does not attempt to appraise the over-all impact of unions on the economy.

TABLE 51

Results of Regression of Hourly Earnings on Demographic Characteristics and Unionization, Across Industry Sector Only
($N = 81$)

Form of X_2	$\bar{R}^2$	Demographic Characteristics (X_1)		Unionization (X_2)	
		Regression Coefficient	t Value	Regression Coefficient	t Value
		Linear			
$U =$ 0 to 100	.697	1.248	11.65	.0081	6.74
$U =$ 0 to 60	.708	1.248	11.86	.0104	7.06
$U =$ 20 to 60	.710	1.203	11.44	.0122	7.13
$U =$ 20 to 100	.693	1.197	11.04	.0093	6.60
Reciprocal of $U =$ 20 to 100	.726	1.202	11.77	.0061 [a]	7.65
		Logarithmic [b]			
$U =$ 0 to 100	.685	1.166	11.24	.33	6.59
$U =$ 0 to 60	.703	1.166	11.57	.42	7.11
$U =$ 20 to 60	.711	1.123	11.27	.51	7.38
$U =$ 20 to 100	.687	1.118	10.75	.38	6.63
Reciprocal of $U =$ 20 to 100	.727	1.122	11.59	.25 [a]	7.88

[a] At the unionization mean of 51.17 per cent this is the union effect implied by the regression coefficient.

[b] The dependent variable, hourly earnings (X_0) and X_1 are in natural logarithms; unionization (X_2) is not.

a proportion of Service as Industry sector industries with observed earnings greater than predicted.

Because unionization is so much more prevalent in the Industry sector than in the Service sector, there is a possibility that the observed union effect really reflects the impact of some other variable that distinguishes the two sectors. One test of this possibility is to run the same regressions for the Industry sector alone. (See Table 51.) We find that the unionization coefficient is still highly significant and has about the same values as in the all-industries run. These values are all higher than those reported by Lewis.[15] They suggest the need to introduce a number of other inde-

[15] H. Gregg Lewis, *Unionism and Relative Wages in the United States.* He estimates the difference between zero and 100 per cent unionization at 10 to 20 per cent.

TABLE 52

Average Hourly Earnings of All Nonagricultural Employed Persons, by City Size and Region, Standardized for Demographic Characteristics, 1959 [a]
(dollars per hour)

Region	Rural	Urban Places		Standard Metropolitan Statistical Areas			
		Under 10,000	10,000–99,999	Under 250,000	250,000–499,999	500,000–999,999	1,000,000 and Over
Northeast	2.27	2.28	2.36	2.40	2.40	2.48	2.75
North Central	2.12	2.12	2.32	2.57	2.57	2.77	2.89
South	1.95	1.93	2.09	2.24	2.35	2.41	2.65
West	2.27	2.21	2.36	2.52	2.50	2.64	2.87

Source: Victor R. Fuchs, *Differentials in Hourly Earnings by Region and City Size.*
[a] Standardized earnings equals actual minus "expected" plus $2.50. For definition of "expected" earnings, see text.

pendent variables that may be related to earnings and unionization, and therefore may have biased the union effect upwards. These variables are discussed below.

Other Independent Variables

LOCATION—REGION AND CITY SIZE (X_3). A location variable is introduced to take account of interindustry differences in distribution by region and city size. Numerous writers have observed wages to be higher in the non-South than in the South, and higher in large cities than in small towns or rural areas.[16] Some precise measures of the differentials in hourly earnings for all nonagricultural employed persons in 1959 were developed for seven city sizes in each of the four regions.[17] In addition, measures of "expected" earnings were calculated for each region-city size group following the methods described above for calculating "expected" earnings for industries. The difference between "expected" earnings and the national average of $2.50 per hour was subtracted from actual earnings to obtain average standardized earnings. The results are shown in Table 52.

[16] See, for example, Lowell E. Galloway, "The North-South Wage Differential," *Review of Economics and Statistics,* August 1963, pp. 265–272, and Edwin Mansfield, "City Size and Income, 1949," in *Regional Income,* Princeton for NBER, 1957.

[17] The regions are: Northeast, North Central, South, and West. See V. R. Fuchs, *Differentials in Hourly Earnings by Region and City Size, 1959,* New York, NBER, OP 101, 1967.

TABLE 53

Average Hourly Earnings, by Location Within Standard Metropolitan Statistical Areas, 1959
(dollars per hour)

Location of Worker	Actual Hourly Earnings (1)	Expected Hourly Earnings [a] (2)	Differential (col. 1 minus col. 2) (3)	Differential Minus Differential for All SMSA's (4)
Works in central city and lives in same SMSA, same county and same city as place of work	2.47	2.41	.06	−.10
Works in ring of SMSA	2.69	2.56	.13	−.03
Works in central city but lives outside the SMSA, or city or county of place of work	3.11	2.72	.39	.23
All workers in SMSA's	2.68	2.52	.16	—

[a] Based on color, age, sex, and education.

A location adjustment factor for each industry was then calculated by weighting the values of Table 52 by the distribution of industry employment across twenty-eight city size-region groups. Those industries that have a disproportionate share of their employment in large cities, and in the non-South, have a location factor greater than $2.50. Those industries located primarily in small towns and in the South have factors below $2.50.

LOCATION WITHIN STANDARD METROPOLITAN STATISTICAL AREAS (X_4). Location and earnings are related in another way; it has been observed that earnings tend to be lower in residential areas, possibly because workers find working near their homes to be more convenient, more congenial, and less expensive.[18] The higher earnings that have been observed for commuters to the central city may also reflect some ability differentials that are not captured by the adjustment for color, age, sex, and education.

Table 53 shows the earnings and earnings differentials for three groups of workers, covering all those who work in Standard Metropolitan Statis-

[18] I am grateful to Albert Rees and to H. Gregg Lewis for calling this point to my attention.

tical Areas. We see that those who live and work in the central city have the lowest earnings, and those who travel to the central city to work have the highest earnings. Workers in the rings of SMSA's are in an intermediate position.

Using the differentials of the last column of Table 53, an adjustment factor was calculated for each industry based on the distribution of its employment among the three groups within SMSA's. This adjustment for location within SMSA's was then multiplied by the ratio of the industry's employment in SMSA's to the industry's total employment. The result was added to $2.50 (the mean earnings for all industries).

ESTABLISHMENT SIZE (X_5). A positive relation between earnings and size of establishment has often been hypothesized. The variable used here measures the fraction of the industry's employment in establishments with more than 250 employees.

EMPLOYMENT GROWTH (X_6). This is measured by the index of employment in 1960 (1950 = 100). Some writers have argued that rapidly growing industries will pay higher than average wages.

UNEMPLOYMENT RATE (X_7). Industries with high unemployment may have to pay higher hourly wages in order to hold labor. On the other hand, a high unemployment rate may indicate a weak labor market and hence lower earnings. This variable is measured by the average of the unemployment rate in 1950 and 1960.

AVERAGE ANNUAL HOURS PER EMPLOYED MALE (X_8). Industries with short hours may have to offer high hourly wages; on the other hand, long hours may indicate a strong demand for labor.

SELF-EMPLOYMENT INCOME AS PERCENTAGE OF TOTAL EARNINGS (X_9). This variable is included because those industries with a large amount of self-employment may have their earnings inflated by the inclusion of the property income of the self-employed. On the other hand, self-employment may have a negative effect on earnings if it involves opportunities to build up equity through capital gains.

The results of regressing hourly earnings on all independent variables are presented in Table 54.[19]

Inclusion of the additional variables raises the explanatory power in every case. The logarithmic run with unionization limited to 20 to 60 per cent has an adjusted coefficient of multiple determination of .88. The significance of such a result in a cross-section regression with 138 observations is very high. It should be noted that the fit of the 20 to 60 run is again significantly better than the 0 to 100 run.

[19] The zero order correlation matrixes are presented in Appendix Table I-3.

TABLE 54

Results of Regression of Hourly Earnings on All Independent Variables, Across All Industries and Industry Sector Only

Form of X_2	$\bar{R}^2$	Demographic Characteristics (X_1)	Unionization (X_2)	Region and City Size (X_3)	Location Within SMSA (X_4)	Establishment Size (X_5)	Employment Growth (X_6)	Unemployment Rate (X_7)	Annual Hours (X_8)	Self-Employment (X_9)
				All Industries ($N = 138$)						
Linear:										
$U = 0$ to 100	.819	1.247**	.0047**	1.098**	8.466**	.001	.061	−.013	−.058**	.005†
$U = 0$ to 60	.819	1.249**	.0052**	1.105**	8.352**	.001	.051	−.014	−.060**	.005†
$U = 20$ to 60	.830	1.232**	.0086**	1.016**	7.668**	.001	.046	−.016	−.053**	.005*
$U = 20$ to 100	.827	1.232**	.0070**	1.049**	7.854**	.001	.062	−.011	−.052**	.005†
Reciprocal of $U = 20$ to 100	.829	1.249**	.0072 [a]**	1.023**	7.582**	.001	.041	−.013	−.054**	.005*
Logarithmic: [c]										
$U = 0$ to 100	.868	1.354**	.21**	.670*	6.088*	.026**	.044	−.053*	−.379**	.020*
$U = 0$ to 60	.868	1.356**	.25**	.703*	6.026*	.026**	.034	−.055*	−.386**	.020*
$U = 20$ to 60	.879	1.341**	.37**	.608†	5.177*	.026**	.031	−.056*	−.331**	.022*
$U = 20$ to 100	.875	1.342**	.30**	.599†	5.292*	.027**	.048	−.049*	−.333**	.022*
Reciprocal of $U = 20$ to 100	.876	1.358**	.31 [a]**	.618†	5.216*	.026**	.028	−.052*	−.344**	.022*

Industry Sector Only (N = 81)

Linear:										
U = 0 to 100	.762	.976**	.0052**	.542†	4.547†	.002	.071	−.010	−.042*	.010
U = 0 to 60	.764	.990**	.0066**	.515†	4.010†	.002	.052	−.010	−.041*	.010
U = 20 to 60	.769	.952**	.0082**	.466†	4.494†	.002	.059	−.013	−.041*	.010
U = 20 to 100	.764	.933**	.0062**	.539†	4.611†	.002	.084	−.011	−.042*	.011†
Reciprocal of U = 20 to 100	.772	.979**	.0040 [b]**	4.11	4.170†	.002	.050	−.009	−.041*	.010
Logarithmic: [c]										
U = 0 to 100	.749	.950**	.25**	.444†	4.963†	.007	.071†	.000	−.315*	.001
U = 0 to 60	.750	.969**	.30**	.462†	4.494†	.011	.051	.000	−.308*	.002
U = 20 to 60	.762	.933**	.37**	.396	4.847†	.013	.055	.004	−.300*	.004
U = 20 to 100	.757	.908**	.28**	.424	4.875†	.010	.079†	.001	−.310*	.004
Reciprocal of U = 20 to 100	.762	.958**	.18 [b]**	.339	4.615†	.010	.050	.001	−.301*	.002

Note: † indicates the regression coefficient is significantly different from zero at .05 level of confidence, * at .01, and ** at .001, on two-tailed test.

[a] At the unionization mean of 35.52 per cent, this is the union effect implied by the regression coefficient.

[b] At the unionization mean of 51.17 per cent, this is the union effect implied by the regression coefficient.

[c] All variables except unionization in natural logarithms.

Inclusion of the additional variables tends to lower the effect of unions on earnings. The best estimates of that effect that emerge from this analysis are that, over the range 20 to 60 per cent unionization, a change of one percentage point in unionization leads to an increase of .8 to .9 cents per hour, or about .37 per cent in earnings. At lower or higher levels of unionization, no effect on earnings is observed.[20]

The runs for the Industry sector alone show almost exactly the same relation between earnings and unionization as do the all-industries regressions. This strengthens our confidence in the relationship and in the conclusion concerning the role of unionization in explaining the intersector difference in earnings.

None of the other variables are as highly and consistently significant as are unionization and demographic characteristics. The two location variables are usually significant in the expected direction, but more so in the linear than in the logarithmic runs, and much more so for the all-industries runs than for the Industry sector alone. This last point suggests that geographical wage differentials may be much smaller for workers in the Industry sector than for those in the Service sector.[21]

Establishment size is significant only in the all-industries logarithmic run; and employment growth is not significant at all. Self-employment income is significant in the all-industries run; the regression coefficients suggest that an industry with 100 per cent of its earnings in self-employment income would have 2 per cent higher hourly earnings than would an industry with no self-employment, all other things being equal.

Both the unemployment rate and annual hours show negative coefficients. These results must be interpreted guardedly. First, there is a negative bias associated with the annual hours variable because a similar variable is implicit in the denominator of the dependent variable. Second, the unemployment rate and the annual hours variable show considerable multicollinearity ($r = .50$). It appears that earnings tend to be lower in industries with long hours, but also tend to be lower in industries with high unemployment rates. Much more investigation of these relationships is needed before any firm conclusions are warranted.

The key role of unionization in explaining the Industry-Service differential in earnings is brought out clearly in Table 55 where the gross sector earnings differential is decomposed according to the regression re-

[20] The less reliable runs with unionization values not restricted indicate a union effect of .47 cents per hour, or .21 per cent per percentage point change over the 0 to 100 range.

[21] Direct analysis of geographical wage differentials by industry, now under way at the NBER, indicates that this is indeed true.

TABLE 55

Decomposition of Industry–Service Earnings Differential According to Regression Results

Variable	Weighted Means		Percentage Difference [a]	Regression Coefficient	Contribution to Sector Differential in Hourly Earnings (percentage points)
	Industry	Service			
X_1 Demographic characteristics	2.5083	2.4946	0.5	1.34	0.7
X_2 Unionization 20–60	48.3919	20.2414	28.2 [b]	0.37	10.4
X_3 Region and city size	2.5006	2.4869	0.5	0.61	0.3
X_4 Location within SMSA	2.5076	2.4948	0.5	5.18	2.6
X_5 Establishment size	52.8330	16.8296	103.4	0.03	3.1
X_6 Employment growth	1.2671	1.3210	–4.2	0.03	–0.1
X_7 Unemployment rate	4.8430	3.3969	35.1	–0.06	–2.1
X_8 Annual hours	19.7745	20.5138	–3.67	–0.33	1.2
X_9 Self-employment	7.6344	21.7752	–96.2	0.02	–1.9
Total contribution of nine independent variables					14.2
X_0 Hourly earnings					15.2
Unexplained residual					1.0

[a] The difference between the Industry and Service means is divided by the average of the two means.

[b] The percentage change in earnings implied by the regression coefficient is applied to the absolute difference because this variable was not run in logarithmic form.

sults.[22] The contribution of each independent variable to the differential is obtained by finding the weighted mean of the variable for each sector and multiplying the percentage difference between the means by the regression coefficient for that variable. For instance, the percentage difference in the sector means for X_1 (demographic characteristics) is .5 per cent. The regression results show that earnings rise 1.34 per cent for every one percentage point difference in this variable; therefore, a sector differential in earnings of .7 per cent ($.5 \times 1.34$) would be expected on account of this variable alone.

The contribution of the unionization variable, 10.4 percentage points, is obtained by using the absolute difference between the sector means because this variable was not converted to logarithms. We see that of the total sector earnings differential of 15.2 per cent, more than two-thirds is explained by the sector differential in unionization. The second most important variable in terms of explaining the sector difference in earnings is establishment size. All the variables taken together explain 14.2 percentage points, leaving an unexplained residual of only one percentage point.

Differences in Earnings for Three Groups of Industries

One way of extending and testing the preceding analysis is to run regressions and make comparisons for separate groups of industries. The 138 industries were divided into three groups according to the percentage of the industry's employment that is male. This variable is a good proxy for the "heaviness" or difficulty (in a physical sense) of work. Earnings may be related to physical difficulty; by running regressions across industries with similar sex mixes, we are, in effect, allowing for this possible relationship.

The three groups of industries are: those with less than 60 per cent male ($N = 28$); those with 60 to 79.9 per cent male ($N = 47$); and those with 80 per cent or more male ($N = 63$).

The regression results are shown in Table 56. The equations are the same as those used for the all-industries run, except that employment growth was dropped as an independent variable because it had no effect on any of the results. Also, only the results for the two unionization forms with the highest explanatory power are reported.

We see that the equations do an excellent job of explaining inter-industry variation in earnings within each group of industries. The lowest

[22] The logarithmic run with unionization restricted to 20 to 60 per cent, which gives the highest $\bar{R}^2$, is used.

TABLE 56

Results of Regression Analysis for Three Groups of Industries Classified by Percentage Male

Variables	Per Cent Male (linear runs)			Per Cent Male (logarithmic runs)		
	Under 60 ($N = 28$)	60–79.9 ($N = 47$)	80 and Over ($N = 63$)	Under 60 ($N = 28$)	60–79.9. ($N = 47$)	80 and Over ($N = 63$)
	Unionization = 20 to 60 Per Cent					
$\bar{R}^2$	.852	.861	.842	.916	.858	.842
X_1 Demographic characteristics	1.1833**	1.8715**	1.7796**	1.4040**	1.6130**	1.8311**
X_2 Unionization	.0108†	.0119**	.0049†	.0067*	.0037**	.0017
X_3 City size and region	1.4356†	1.8430*	.0571	.8573	2.2261**	−.0460
X_4 Location within SMSA	7.6885	5.9657	9.6469**	.3380	6.5215†	10.2162**
X_5 Establishment size	.0032	−.0037	−.0000	.0311†	.0026	.0245*
X_7 Unemployment rate	−.0347	−.0141	.0292†	−.0934	.0364	.0585†
X_8 Annual hours	−.0481	−.0679	−.0601**	−.2809	−.2091	−.4293**
X_9 Self-employment	.0074	−.0040	−.0088*	.0194	.0082	−.0069
	Reciprocal of Unionization = 20 to 100 Per Cent [a]					
$\bar{R}^2$	.846	.871	.844	.914	.869	.842
X_1 Demographic characteristics	1.1982**	1.9973**	1.8029**	1.4327**	1.7289**	1.8693**
X_2 Unionization	.0163	.0145**	.0028†	1.10*	.47**	.09
X_3 City size and region	1.5146†	1.5367†	.0555	.9250	1.9609**	−.0563
X_4 Location within SMSA	7.8707	4.9233	9.5246**	.1853	5.7684†	10.2825**
X_5 Establishment size	.0034	−.0039†	−.0001	.0332†	.0026	.0226†
X_7 Unemployment rate	−.0313	−.0144	.0306†	−.0933	.0410	.0633†
X_8 Annual hours	−.0494	−.0760†	−.0593**	−.3309	−.2390	−.4323**
X_9 Self-employment	.0072	−.0044	−.0084*	.0208	.0095	−.0078

Note: † indicates the regression coefficient is significantly different from zero at .05 level of confidence, * at .01, and ** at .001, on two-tailed test.

[a] Effect on earnings at mean levels of unionization: under 60 per cent male = 25.30 per cent; 60–79.9 per cent male = 31.53 per cent; 80 per cent male and over = 45.33 per cent.

TABLE 57

Effect of Unionization on Hourly Earnings, Various Regressions

Form of X_2	All Industries Regressions	Weighted Average of Regressions Across Industries, Grouped by Percentage Male
Linear runs		
$U = 20$ to 60 per cent	.0086	.0080
Reciprocal of $U = 20$ to 100 per cent	.0072	.0064
Logarithmic runs [a]		
$U = 20$ to 60 per cent	.37	.30
Reciprocal of $U = 20$ to 100 per cent	.31	.24

Source: Tables 54 and 56.
[a] All variables in natural logarithms except unionization.

adjusted coefficient of multiple determination is .84, and some are as high as .91. The effect of unionization on earnings seems to be stronger in the two groups with smaller percentage males. In fact, for the industries with 80 per cent or more males, the unionization coefficient is not always statistically significant.

These runs also afford an opportunity to examine the union effect on earnings after allowing for the fact that unionization is correlated with the percentage male. By comparing a weighted (by man-hours) average of the unionization coefficients in each group with the coefficients from the all-industries run, we can see whether this correlation significantly affects the relation between earnings and unionization.

Table 57 indicates that the results are similar when the regressions are run across industries grouped by percentage male. The effect of unions on earnings is reduced about one-fifth compared with results that were obtained when each regression was run across all industries, but this difference is not statistically significant.

Summary

Average hourly earnings in the Industry sector were 39 cents, or 17 per cent higher than in the Service sector in 1959. Multiple regression analysis across industries was used to identify the sources of interindustry

and intersector differences in earnings. A demographic characteristics variable ("expected" earnings) based on color, age, sex, and education explains more than half of the interindustry variation in earnings, but explains none of the sector differential because "expected" earnings were equal in the two sectors.

The principal explanatory variable of the sector differential is unionization. Approximately half of the workers in Industry are union members; fewer than one-fifth of the Service workers are organized. Multiple regression analysis across industries reveals a significant positive relation between hourly earnings and extent of unionization after taking account of demographic characteristics, location of industry, and many other variables.

The effect of unions on wages seems to be most pronounced in the range of 20 to 60 per cent. Below and above that range no systematic relation between changes in unionization and hourly earnings was observed. Within that range, hourly earnings rise by about .3 or .4 per cent with each increase of one percentage point in unionization.

Differences in size of establishments is the second most important variable in explaining the Industry-Service earnings differential. Service sector employment is mostly in small establishments, and the multiple regression analysis reveals a significant relation between industry earnings and the fraction of employment in establishments with more than 250 employees.

Although service earnings are typically low, a few industries show high earnings, and two groups—wholesale trade and finance, insurance and real estate—show high earnings relative to those that would be "expected" based on demographic characteristics. In general, it was found that the industries in the Service sector are much more heterogeneous than those in the Industry sector with respect to both actual and "expected" earnings.