Union Relative Wage Effects: A Survey of Macro Estimates
Author(s): H. Gregg Lewis
Source: *Journal of Labor Economics*, Vol. 1, No. 1 (Jan., 1983), pp. 1-27
Published by: The University of Chicago Press on behalf of the Society of Labor Economists and the NORC at the University of Chicago
Stable URL: http://www.jstor.org/stable/2534928
Accessed: 29-03-2018 09:07 UTC

Union Relative Wage Effects: A Survey of Macro Estimates

H. Gregg Lewis, *Duke University*

This paper surveys union wage effect estimates drawn from empirical cross-section wage equations fitted to macro (aggregated) data in 34 post-1963 studies containing such equations. Each estimate is the partial derivative of the dependent wage variable (in logarithmic units) with respect to the extent-of-unionism (fraction unionized) variable in the equation. This paper shows that these estimates contain a mixture in uncertain ratio of the union/nonunion wage differential or wage gap, on the one hand, and an "extent-of-unionism" effect, on the other. Therefore the Macro estimates should *not* be interpreted as wage gap estimates.

I. Introduction

In the U.S. economy trade or labor unionism currently is present in significant proportions. Roughly one-fourth to one-third of U.S. wage and salary workers are covered at their workplaces by collective bargaining agreements.[1] It has been so for at least 3 decades, and I think

This paper is part of a larger study surveying recent (1963–82) empirical work on union relative wage effects in the United States. In preparing this paper I have benefited from comments of members of the Princeton Labor Workshop and Duke colleagues, especially George Tauchen.

[1] Freeman and Medoff (1979) estimated that, in 1968–72, 30% of private wage and salary workers were employed in establishments in which at least half of the

[*Journal of Labor Economics,* 1983, vol. 1, no. 1]

0734-306X/83/0101-0001$01.50

there is little likelihood that unionism will decline to negligible proportions in the next 30 years.[2] One of the effects of this unionism that many of us would like to measure for recent years is the union/nonunion relative wage differential, as it is often termed, which (after Mincer)[3] I will call the "wage gap." For an individual worker the wage gap is the excess of his real wage if unionized (covered by a collective bargaining agreement) over his real wage if nonunion (not so covered) given his working conditions.

We know far too little about wage determination in all of its fine detail to estimate the wage gap worker by worker. However, the task of estimating the mean wage gap for the whole U.S. labor force and even for some of its large segments is not, I think, so formidable. This is the central task of the larger study of which this paper is a component.

In recent years the wage gap, so understood, usually has been estimated by fitting wage equations to cross-section, individual worker data on wages, union status (union or nonunion), and other variables supposedly controlling for differences among workers in working conditions and worker quality. Most of these equations are encompassed in their form by equation (1):

$$W = a_n + a_{nx}x + a_{ny}y + e_n + U[(a_u - a_n) + (a_{ux} - a_{nx})x + (a_{uy} - a_{ny})y + (e_u - e_n)] , \quad (1)$$

where W is the natural logarithm of a worker's wage, x is a set of variables specifying the working conditions and quality of the worker, y is the fraction of workers who are unionized in the industry (or geographic area or occupation, etc.) in which the worker is employed, U is the union status of the worker (equal to unity if unionized and zero if nonunion), the e's are residuals reflecting left-out variables, and the a's are the estimated coefficients of the equation. Equation (1) may be rewritten and sometimes has been fitted as two separate equations:

$$W_i = a_i + a_{ix}x + a_{iy}y + e_i ; \quad i = u \text{ if } U = 1; \quad i = n \text{ if } U = 0 . \quad (2)$$

I term y an extent-of-unionism variable.

---

workers were covered by collective bargaining agreements and that, in 1973–75, 23% of private wage and salary workers were union members. See their table 6.

[2] My thinking is based on the fact that U.S. Department of Labor estimates of union membership, the labor force, and nonagricultural employment indicate that in 1950 union members comprised 22.0% of the labor force and 31.5% of nonagricultural employment. The corresponding figures for 1978 were 19.7% and 23.6%, respectively. See U.S. Bureau of the Census (1980), p. 429.

[3] See Mincer (1981).

Assume that equations (1) or (2) have been estimated without bias. Then an unbiased estimate of the wage gap M, conditional on or given x and y, is [4]

$$M \equiv E(W_u - W_n|x,y) = a_u - a_n + (a_{ux} - a_{nx})x + (a_{uy} - a_{ny})y\ . \quad (3)$$

Of course if the wage equation contains no cross-product terms Ux or Uy, $M = a_u - a_n$, the coefficient of U in (1). Define $\hat{W}$ as $E(W|x,y,U)$, the expected value of W in (1) conditional on x, y, and U. Then the wage gap M is the partial derivative of $\hat{W}$ with respect to union status U. Thus the presence of the union status variable U on the right-hand side of the wage equation is critical for estimating the wage gap. In this connection, notice that the critical variable is union status U, not extent of unionism y.

Let $\bar{x}$ and $\bar{y}$ denote the means of the right-hand variables x and y among workers included in the fitting of the wage equations (1) or (2). Then, from (3), the estimate of the mean wage gap among included workers is

$$\bar{M} = a_u - a_n + (a_{ux} - a_{nx})\bar{x} + (a_{uy} + a_{ny})\bar{y}\ . \quad (4)$$

Of course, if the fitted wage equation contains no interaction terms Ux or Uy, the estimated mean wage gap is simply $a_u - a_n$.

The union status variable U is a *micro* concept of collective bargaining coverage. For each worker it distinguishes between two states, unionized and nonunion. Its presence on the right-hand side of wage equations such as (1) or (2) above permits the estimation of the wage gap M. I term wage equations that include U as Micro equations, whether they also include or, more often, exclude extent-of-unionism y variables.[5]

Extent of unionism y, whether measured by industry, occupation, locality, etc., is a group or macro concept of union status or collective bargaining coverage. There is a strong presumption, I think, that in the general equilibrium of the economy in the presence of unionism the relative wage of each worker depends, not only on his union status, sex, color, schooling, experience, and like variables, but also on the extent of unionism in the whole work force and the distribution of workers by union status among work-force sectors. This argues for the presence on the right-hand side of Micro wage equations of extent-of-unionism variables characterizing this distribution, though it does not, of course, settle the question of the proper specification of these variables.

[4] To avoid cumbersome notation, I use the symbol M both for the concept of the wage gap and for its estimate. The text will make the distinction between concept and estimate wherever confusion otherwise might arise.

[5] In general the inclusion or exclusion of extent-of-unionism y variables in Micro wage equations has nonnegligible, but small, effects of uncertain sign on estimates of the mean wage gap M.

We now have a substantial stock of Micro wage equations that include extent-of-unionism variables, a sample of which is reported in Section IV. The estimated coefficients a_{uy} and a_{ny} of these variables often are numerically large, of uncertain sign, and have values that are sensitive to the way y is measured, the specification of the wage equation, and the data set used. The large dispersion of the estimated a_y's argues against any simple "threat" or "spillover" interpretation of the wage effects picked up in the data by these coefficients. Indeed, I am not by any means convinced that these estimated wage effects are mostly effects of unionism rather than mostly effects of left-out variables correlated with the included y variables.

For what follows, nothing of great importance is lost if equation (1) is written more simply as

$$W = a + a_x x + a_y y + \bar{M}U + e . \qquad (5)$$

This is a Micro wage equation because it includes the micro union status variable U. Omit U from the equation but retain the extent-of-unionism variable y, a macro union status concept. I term the resulting equation a Macro equation.

Until about 1965 there were no large random samples with broad coverage of the U.S. work force containing information on wages and numerous worker and employment characteristics that also classified workers by their union status U. In the absence of the union status data, either one or the other of two alternative procedures were followed in the earliest of the post-1963 studies and in numerous later studies emulating them:

1. The wage equation (5) or similar equation was fitted to individual worker data on W, x, and y omitting union status U. I write the fitted Macro wage equation as

$$W = c + c_x x + c_y y + e' . \qquad (6)$$

Frequently in such equations y was interacted with one or more of the x variables. This is what I call a "Weiss-type" Macro equation.[6]

2. The wage equation (5) or similar equation, thought of as pertaining to individual workers, was aggregated by industry (or geographic area, etc.) across individual workers and the resulting aggregate or Macro equation was fitted to observations on the industry (or area, etc.) aggregates. Assume that the aggregation is by industry and denote the industry mean of a variable by an asterisk superscript. Also assume that y is exactly equal to U^*. Then the aggregation of (5) by industry is

[6] After Leonard W. Weiss who was the first, I think, to fit such wage equations. See study no. 33 in table 1.

$$W^{*} = a + a_{x}x^{*} + (a_{y} + \bar{M})y + e^{*} . \tag{7}$$

There is no assurance, of course, that the residual e^{*} in (7) is uncorrelated with x^{*} even though the residual e in the individual-worker equation (5) is uncorrelated with x, y, and U. Therefore, the fitted Macro equation

$$W^{*} = d + d_{x}x^{*} + d_{y}y + e'' \tag{8}$$

may have coefficients that differ considerably from those of (7). Some of the fitted Macro equations similar to (8) included one or more yx^{*} interactions. In addition, medians rather than means were often used in the aggregation and usually W^{*} was measured as the logarithm of the arithmetic rather than geometric mean wage in the industry. I call (8) an aggregate Macro equation.

The authors of the studies containing Macro wage equations differ in the way they interpret the unionism content of these equations. However, the most frequently reported statistic is the partial derivative—call it W_{y}—of W in (6) or W^{*} in (8) with respect to extent of unionism y. In this connection recall that the wage gap M is the partial derivative of W in (1) with respect to union status U. Thus W_{y} conceptually has some resemblance to the wage gap M and I read some authors as interpreting W as though it were M.

This paper reports and reacts to estimates of W_{y} that I have retrieved from the Macro equations and underlying data in 34 post-1963 studies containing such equations. The studies are listed in the Appendix to this article, and in table 1 I present the estimated values of W_{y} for each study. Readers who are familiar with the recent empirical literature on union-relative wage effects will observe that there is a noticeable, though not universal, tendency for the estimates of $\bar{W}_{y}$ in table 1 to exceed by a considerable margin the typical wage gap $\bar{M}$ estimates from Micro equations. Thus, if $\bar{W}_{y}$ is interpreted as estimating the union/nonunion wage gap, these 34 studies give a rather different impression of the mean size of the gap than one is likely to get from the $\bar{M}$ estimates drawn from Micro equations.

The critical question, of course, is: What does W_{y} estimate? This question is the subject of Section III. I show there that if Macro and corresponding Micro wage equations are fitted by ordinary least squares (OLS), the estimate of W_{y} in the Macro equation contains a combination of the extent-of-unionism (EU) effect a_{y} and the wage gap effect $\bar{M}$. Indeed, under certain ideal conditions (exact aggregation), W_{y} as estimated from an aggregate Macro equation is exactly the sum of a_{y} and $\bar{M}$. Thus with OLS estimation of the wage equations, W_{y} estimates neither a pure wage gap effect nor a pure EU effect.

This raises the question, Is the relation of estimates of W_{y} to corre-

Table 1
Estimates of W_y

LINE NO. (1)	STUDY NO. (2)	ESTIMATE AND YEAR: W_y (3)	ESTIMATE AND YEAR: Year (4)	CROSS SECTION (5)	COVERAGE (6)
1	1	.38–.43	1960	Industry	Production Workers (PW), manufacturing (MFG)
2	2	.36–.40	1967	Industry	White males, private sector, urban
3	2	.65–.74	1967	Industry	Line 2, blue-collar (BC) only
4	2	.42 *W*	1967	Industry	Line 2
5	3	.37–.41	1966–75	City	Private, professional hospital workers
6	3	.30–.42	1966–75	City	Private, nonprofessional hospital workers
7	3	.04–.08	1966–75	City	Government, professional hospital workers
8	3	−.02–.02	1966–75	City	Government, nonprofessional hospital workers
9	4	.1 *W*	1969	Industry	Clerical and BC, MFG and utilities
10	4	.1 *W*	1969	Industry	Clerical and BC, MFG only
11	5,6	.02–.16	1967	City	Municipal workers, noneducational, 10 departments
12	5,6	.06–.16	1967	City	Municipal workers, noneducational, seven departments
13	7	.10–.22	1966–72	City	Nonprofessional hospital workers
14	7	.18–.29	1966–72	City	Nonprofessional hospital workers, private
15	7	.09–.20	1966–72	City	Nonprofessional hospital workers, public
16	8	.19–.21	1959	Industry	Nonagricultural workers
17	8	.19–.25	1959	Industry	Goods-producing industries
18	9	.00	1969–70	State	Public schoolteachers
19	10	.06–.10	1960–61	City	Selected occupations and cities, MFG only, full-time
20	10	.13–.19	1963–64	City	Line 19
21	10	.03–.06	1966–67	City	Line 19
22	11	.15–.18	1963–71	City	PW, MFG, in selected cities
23	12	.08–.09 *W*	1969	Industry	Male, seven selected occupations, private sector, full-time, MFG and transportation, communication, utilities (TCU)
24	13	.11–.27	1969	Industry	Males, private sector
25	14	.03	1969	Industry	Males, private sector, MFG
26	14	.17	1969	Industry	Males, private sector, non-MFG

Line 1.—See text for 2SLS, 3SLS estimates.

Line 4.—See Sec. IV for corresponding wage gap ($\bar{M}$) estimates from Micro wage equations.

Lines 5–8.—W_y estimates also available by year for 1966, 1969, 1972, 1975. See text for discussion of col. 3 estimates.

Lines 9, 10—Study reported coefficients (W_y) of y variables to only one decimal place, and for several wage equations I could not calculate W_y. W_y estimates by level of industry concentration ratio also available.

Lines 11, 12.—See text for discussion of col. 3 estimates. W_y estimates also available by municipal department, city size, and union affiliation.

Lines 13–15.—W_y estimates also available by year for 1966, 1969, 1972.

Lines 16, 17—Column 3 omits several wage equations for which I could not calculate W_y.

Lines 19–21.—See text for discussion of col. 3 estimates. W_y estimates by sex and occupation also available.

Line 22.—See text for discussion of col. 3 estimates. W_y estimates available by year 1963–71.

Line 23.—W_y estimates available by occupation.

Line 24.—See text for 2SLS, 3SLS estimates. W_y estimates available by type of extent-of-unionism y measure (representation vs. membership).

Lines 25, 26.—See text for 2SLS, 3SLS estimates.

Table 1—*Continued*

LINE NO. (1)	STUDY NO. (2)	ESTIMATE AND YEAR: W_y (3)	ESTIMATE AND YEAR: Year (4)	CROSS SECTION (5)	COVERAGE (6)
27	15	.18	1970–73	City	Selected occupations in selected cities, private sector
28	15	.22	1978–79	City	Line 17
29	16	.45	1960	Industry	PW, MFG
30	17	.24	1960	Industry	PW, MFG
31	18	−.06–.10	1966–68	State	Public schoolteachers
32	19	.26–.30	1977	Industry	Nonagricultural full-time workers
33	20	.38–.39	1966–75	City	Private, professional hospital workers
34	20	.33–.38	1966–75	City	Private, nonprofessional hospital workers
35	20	.05–.09	1966–75	City	Government, professional hospital workers
36	20	−.01–.02	1966–75	City	Government, nonprofessional hospital workers
37	21	.28	1976	Industry	Selected mining (MIN), MFG, transportation (TRANS) industries
38	22	.13	1972	Industry	PW, MFG
39	23	−.01–.20	1969	Occupation	Male, full-year, 16 years of age and older
40	24	.32	1959	Industry	MFG
41	25	.15–.42	1969	State	Full-year workers
42	26	.27	1960	Industry	PW, MFG
43	26	.20	1960	Industry	PW in MIN, MFG, TCU, and construction (CONST)
44	27	.19–.66	1959	Industry	Male laborers (n.e.c.), MFG
45	28	.17–.29	1959	Industry	Males in MIN, MFG, CONST, TCU
46	29	.13–.28	1958	Industry	PW, MFG
47	30	.26–.34	1958	Industry	PW, MFG
48	30	.26–.37	1963	Industry	PW, MFG
49	31	.22	1950	Industry	PW in MIN, CONST, MFG, and trade
50	31	.26	1960	Industry	Line 49
51	32	.08–.08	1975	City	Municipal policemen
52	32	.00–.07	1975	City	Municipal firemen
53	32	.07–.14	1975	City	Municipal refuse collection workers
54	33	.25–.28 *W*	1959	Industry	Males in BC in MIN, MFG, CONST, TCU
55	34	.82	1972	State	PW, MFG

Lines 27, 28—See text for discussion of col. 3 estimates.
Line 29.—See text for 2SLS estimate.
Line 30.—See text for 3SLS estimate.
Line 31.—Omits equations that include square of y variable and equations that author describes as 2SLS estimates.
Line 32.—W_y estimates available by type of dependent (mean vs. median wage) variable.
Lines 33–36.—See text for discussion of col. 3 estimates. W_y estimates avialable by year 1966, 1969, 1972, 1975.
Line 39.—See text for discussion of col. 3 estimates.
Line 40.—See text for 2SLS estimate.
Line 45.—See text for discussion of col. 3 estimates.
Lines 51–53. See text for discussion of col. 3 estimates.
Line 54.—W_y estimates available by occupation and by regulated vs. unregulated industries.
Line 55.—See text for 2SLS, 3SLS estimates.

sponding estimates of $\bar{M}$ sufficiently stable that one can estimate $\bar{M}$ reliably from knowledge of the estimate of W_y? In Section IV I report experiments designed to answer that question. The answer is negative. Thus the conclusion to be drawn from Sections III and IV is that the union wage effect measured by W_y is not the union/nonunion wage gap $\bar{M}$ and $\bar{M}$ cannot be estimated reliably from W_y.

In the concluding section I consider and reject an alternative interpretation of W_y, namely, that for a group of workers whose extent of unionism is y, the average absolute wage *gain* attributable to unionism is yW_y.

II. Survey of Estimates of W_y from Macro Equations

Because W in (6) and W^* in (8) are measured in logarithmic units, the partial derivative W_y also is in these units. Readers who wish to convert values of W_y to percentage units should calculate $100(e^{W_y} - 1)$. When the left-hand wage variable is W or W^* and there are no yx^* interactions on the right-hand side, the value of W_y is the estimated coefficient, c_y in (6) and d_y in (8), of the extent-of-unionism variable y. However, when such y interactions are present, W_y depends on the values of the interacted x or x^* variables. For all such equations I have evaluated W_y at the mean values (over the observations to which each equation was fitted) of the interacted x or x^* variables.

In some of the Macro equations discussed in this section the dependent wage variable was in its natural arithmetic units (i.e., the dependent variable was e^W or e^{W^*} rather than W or W^*). The calculation of W_y then involves two steps. First calculate $\partial e^W/\partial y$ or $\partial e^{W^*}/\partial y$ and, if necessary, evaluate the partial derivative at means of interacted x or x^* variables. Then divide this derivative by the mean (over the covered observations) of the dependent variable.

None of the values of the partial derivative W_y presented below should be attributed to the authors of the papers from which I have derived them as *their* estimates of some type of union-induced relative wage effect. The numbers are *my* reading and calculations from the wage equations and associated data some of which were obtained by correspondence. They are a convenient way of summarizing the unionism content of these equations. What they mean remains to be seen.

The Appendix lists the 34 studies from which I have drawn the estimates of W_y summarized in column 3 of table 1.[7] Notice that a third of the studies appeared more than a decade ago.

The "Study No." in column 2 of table 1 corresponds to that given in the Appendix. Thus, study number 1 reported on line 1 of table 1 is the

[7] In table 1 I have excluded Macro equations fitted to time-series data, some others for which I could not calculate W_y, and undoubtedly others that I have missed in my search of the literature. I have also excluded several Macro equations that I have fitted simply for the purpose of trying to discover how to interpret W_y. These are presented in Sec. IV.

Ashenfelter-Johnson study, the first entry in the Appendix. The estimates of W_y appear in column 3. More than half of the studies provided more than one estimate of W_y. In all such cases, column 3 shows the *range* of the estimates. For example, from study number 1 (on line 1) I drew three estimates ranging from 0.38 to 0.43. The letter *W* on some lines of column 3 indicates estimates that come from Weiss-type equations. The year to which an estimate pertains is given in column 4. Column 5 shows whether the extent-of-unionism (y) variable used in the Macro wage equation was by industry, city, state, or occupation. Column 6 briefly describes the worker coverage of the observations on the dependent wage variable. The notes to table 1 identify the Macro studies that provide simultaneous-equation (two-stage and three-stage least squares, hereafter 2SLS and 3SLS) estimates discussed later in this section, or contain interesting detail in the estimates of W_y that is not shown in table 1, or use wage equation models requiring the special discussion that immediately follows.

In five of the studies (nos. 3, 5 and 6, 20, 32), all involving "city" cross-sections, some of the W_y estimates were derived from wage equations that included more than one right-hand extent-of-unionism y variable, one for the group of workers covered by the equation, the "own-y" variable, and the others for other groups of workers in the same city. For all such equations the estimates of the partial derivative W_y in column 3 are with respect to the own-y variable.

For the Cain et al. study (no. 3), each of the W_y estimate ranges on lines 5–8 covers three separate estimates, each from an equation fitted to the pooled 1966, 1969, 1972, 1975 data. The equations differ in their y variable as follows: equation (1), own-y only; equation (2), own-y and y for other occupation (professional vs. nonprofessional); and equation (3), own-y and y for same occupation but other hospital class (private vs. government). The detailed W_y estimates are:

	Eq. (1)	Eq. (2)	Eq. (3)
Private, professional	.39	.37	.41
Private, nonprofessional	.42	.30	.42
Government, professional	.04	.08	.04
Government, nonprofessional	.02	−.02	.01

The column 3 ranges for the Ehrenberg and Ehrenberg and Goldstein studies (nos. 5 and 6) on lines 11 and 12 cover separate estimates by municipal department, 10 departments on line 11 and seven on line 12. The underlying wage equations for these estimates included only one extent-of-unionism variable, that for own y. However, the authors also fitted wage equations comparable to those for line 12 that included y variables for several other departments as well as that for own department. The range of the seven departmental estimates was 0.01–0.07, about half of that on line 12.

Each of the estimate ranges on lines 33–36 for the McLaughlin study (no. 20) covers a pair of estimates from a pair of equations fitted to pooled 1966, 1969, 1972, and 1975 data. Both equations included an own-y variable. One of the two equations also included a y variable for the other (professional vs. nonprofessional) occupational group of hospital workers.

In the Victor study (no. 32), the ranges for policemen and firemen on lines 51 and 52 cover two equations, one of which included a y variable for the other occupation (police or fire); the other did not. The range on line 53 for refuse collectors covers four equations with W_y estimates and y variable specifications as follows: equation (1) ($W_y = 0.14$), own y only; equation (2) ($W_y = 0.07$), own y and police y; equation (3) ($W_y = 0.12$), own y and fire y; equation (4) ($W_y = 0.08$), own y, police y, and fire y. All of the wage equations in this study were estimated by 2SLS in an equation system in which the second equation was for employment.

In the Hamermesh study, number 10 (lines 19, 20, 21), the dependent wage variable was the natural logarithm of the ratio of the average wage of white-collar (WC) workers to that of blue-collar (BC) workers in manufacturing by sex, occupation, and city. There were two right-hand extent-of-unionism (y) variables for each city: y_{WC} for WC workers and y_{BC} for BC workers, both for manufacturing. The estimate of W_y for WC workers was the coefficient of y_{WC} and that for BC was the negative of the coefficient of y_{BC}. The numbers on lines 19, 20, and 21 are employment-weighted means of the WC and BC estimates. The separate estimates for WC and BC are:

	WC	BC
1960–61	−.02–.04	.12–.14
1963–64	.12–.17	.14–.22
1966–67	−.11–−.06	.12–.15

The procedure in the second Hamermesh study (no. 11) was fairly similar. The dependent wage variable was the natural logarithm of the ratio of the wage of *unionized* bus drivers to the average wage of production workers in manufacturing by city and year. There was of course only one right-hand y variable, that for manufacturing production workers, and I took the negative of its coefficient as the estimate of W_y for these workers. The numbers on line 22 are means of the nine yearly figures as follows:

1963	.25–.26	1968	.12–.14
1964	.25–.27	1969	.12–.15
1965	.23–.26	1970	.04–.08
1966	.18–.21	1971	.06–.10
1967	.15–.18	Pooled data	.16

The Hirsch-Rufolo study (no. 15, lines 27, 28) is similar to the first Hamermesh study (no. 10). The numerator wage was for municipal workers, the denominator wage for private sector workers. There were two corresponding right-hand y variables, one for the government workers and the other for private workers in manufacturing. The estimates of W_y in the table are for the private sector workers. The corresponding estimates for the government workers are 0.23 for 1970–73 and 0.25 for 1978–79. In the next section I argue that the procedure followed in these three studies (nos. 10, 11, and 15) is likely to make the estimates of W_y that I have drawn from them somewhat incomparable to other estimates in table 1.

The estimate range on line 39 for the Pashigian study (no. 23) covers estimates from five wage equations. Three of these equations, fitted by OLS and differing slightly in right-hand variables, all yielded W_y estimates of 0.20. Two other wage equations, also differing in right-hand variables, were fitted by 2SLS as part of a simultaneous equations system that also included geographical mobility equations. They yielded W_y estimates of -0.01 and 0.11.

The first Rosen study (no. 28, line 45) fitted by 2SLS a simultaneous equations model in which there were two equations, an hours-of-work-demand equation by employers and an hours-of-work-supply equation by workers, in which the y variable (in the demand equation) was treated as exogenous. The estimates of W_y on line 45 came from the reduced-form wage equations.

The Potthoff study (no. 25, line 41) is one of the few Macro studies in which the dependent wage was expressed in both real (cost-of-living deflated) and nominal terms. The low figure, 0.15, on line 41 corresponds to the equation in which the real wage was dependent and the high figure, 0.42, to the nominal wage equation. He also noticed that the observations for Alaska and Hawaii were outliers. When these two states were omitted, the estimate range went from 0.15–0.42 to 0.09–0.24.

In most of the equations from which table 1 was derived (excluding W-type equations on indicated lines and the Hamermesh and Hirsch-Rufolo estimates on lines 19–22, 27, and 28), the macro wage concept that was used was that of an arithmetic mean or, less often, a median, rather than a gometric mean. Ashenfelter and Taussig (no. 2, lines 2 and 3) experimented with both the arithmetic and geometric means. On line 2 the low figure, 0.36, is for the geometric mean and the high figure, 0.40, for the arithmetic, while on line 3 the reverse is true. Hirsch (no. 13, line 24) tried all three: arithmetic mean, geometric mean, and median. The estimates of W_y for the geometric mean and median versions were nearly equal and roughly 0.1 higher in log units than for the arithmetic mean. Killingsworth (no. 19, line 32) used the arithmetic mean and median, with W_y estimates for the median 0.08–0.09 higher than for the arithmetic mean.

Hirsch (no. 13, line 24) experimented with alternative measures of the extent-of-unionism y variable—extent of collective bargaining *coverage* or representation versus union *membership*—and obtained higher estimates for membership than for coverage. The wide range, 0.11–0.27, of W_y estimates on line 24 (Hirsch, no. 13) is entirely accounted for by differences in the wage concept and extent-of-unionism concept used in the six equations from which the estimates were drawn. The low figure, 0.11, is for an equation with an arithmetic mean wage and a coverage extent-of-unionism measure. The high figure, 0.27, is for a median wage measure coupled with a membership union status concept.

The even wider range, 0.19–0.66, of the estimates drawn from the Rapping study (no. 27, line 44) is the result of differences in right-hand variables in the eight equations covered in the range.

Those who are familiar with the often cited study by Weiss (no. 33, line 54) undoubtedly will not recognize the numbers reported on line 54. The smaller figure, 0.25, is the employment-weighted mean of three W_y estimates by major occupation among *unregulated* industries. The larger figure, 0.28, is the corresponding mean among both unregulated and regulated industries. The detailed estimates of W_y by major occupation and industry coverage are:

Major Occupation	Unregulated	All Industries
Craftsmen	.28	.25
Operatives	.20	.28
Laborers	.34	.33

Weiss also fitted an equation for operatives in manufacturing for which the estimate of W_y is 0.31.

In seven of the table 1 studies the wage equation, modeled as a part of a simultaneous equations system containing an extent-of-unionism equation and sometimes other equations in addition to the wage equation, was estimated by simultaneous equations (SE) methods (2SLS or 3SLS or both). The SE estimates of W_y, which are not shown in table 1, are reported below in table 2. Column 1 identifies the study, column 2 the

Table 2
Simultaneous Equations (SE) Estimates of W_y

STUDY NO. (1)	TABLE 1 LINE NO. (2)	ESTIMATES OF W_y: Table 1 (OLS) (3)	ESTIMATES OF W_y: SE (4)
1	1	.38–.43	−.09–.18
13	24	.11–.27	−.21–.12
14	25	.03	.10–.17
14	26	.17	.02–.07
16	29	.45	.78
17	30	.24	.55
24	40	.32	.24
34	55	.82	.60–.78

table 1 line number, column 3 repeats the W_y estimates (by OLS) given in table 1, and column 4 gives the corresponding SE estimates. In five of the eight lines of table 1, the SE estimates are lower than the corresponding OLS estimates, but the SE estimates are more dispersed than the OLS estimates.

Twelve of the studies in table 1, most of which were published before 1973, give estimates of W_y for 1958, 1959, or 1960—see lines 1, 16, 17, 29, 30, 40, 42–47, 50, and 54. The estimates of W_y from these 12 studies for the 3 years average 0.26 (or 30%) and range from 0.13 to 0.45 even when only the low sides of estimate ranges are included. (The average is 0.33, or 40%, when high sides of estimate ranges are substituted for low sides.) In 1963 I estimated that the mean wage gap $\bar{M}$ in the U.S. labor force in 1957–58 was 10%–15%, or 0.10–0.14 in log units (see Lewis 1963, p. 193). If estimates of W_y are interpreted as estimates of the mean wage gap $\bar{M}$ for the workers covered, as at first I did, then these 12 studies suggest that I underestimated the economy-wide average wage gap in 1957–58 by a factor of one-half or more. Moreover, table 1 indicates that there was nothing special about the years 1958–60. The estimates of W_y for later years are roughly at their 1958–60 level.

Of course, the table 1 figures are disproportionately for workers in manufacturing industries. Is the mean wage gap for these workers well above the economy-wide average? Neither my book nor table 1 contains much evidence on this question, but what there is does not suggest that the gap in manufacturing is unusually high. Indeed, recent estimates of the mean gap by industry derived from Micro wage equations indicate that the manufacturing wage gap is *lower* than the all-industry average wage gap.

After 1965, as several large micro-data sets became available that contained information on the union status U of individual workers, estimates of the union wage effects came in rapidly increasing proportions from Micro, rather than Macro, wage equations. As these Micro equations appeared, I noticed that the wage gap estimates that I was obtaining from them were considerably smaller than the W_y estimates I had drawn from the earlier Macro equations.

The earlier estimation of Macro wage equations (other than Weiss type) usually involved laborious assembly and processing of data from a variety of sources. As a consequence, the Macro equations (except Weiss type) underlying table 1 commonly include relatively few right-hand variables. In contrast, the new micro-data sets made it easy to fit Micro wage equations with numerous right-hand variables using data from a single source. Can the differences between the Macro W_y estimates and the Micro $\bar{M}$ estimates be accounted for by differences in right-hand variables?

The answer to this question is negative. In Section IV I compare Micro and Macro equations fitted to data from the same source with the same right-hand variables, except of course that in the Micro equation the union status variable U is for individuals and in the Macro equation it is replaced by an extent-of-unionism variable y, a macro version of U. Large differences emerge between estimates of $\bar{M}$ and corresponding estimates of W_y.

What is it, then, that W_y estimates? That is the subject of the next three sections.

III. What Do These Estimates of W_y Estimate?

Weiss-Type Equations

I begin with discussion of estimates of W_y from Weiss-type equations. Return to equations (5) and (6):

$$W = a + a_x x + a_y y + \bar{M}U + e \tag{5}$$

and

$$W = c + c_x x + c_y y + e' , \quad \text{where} \quad W_y = c_y . \tag{6}$$

Both are fitted to data for individual workers. The first, (5), is a Micro wage equation; it includes the union status dummy variable U (equal to unity for union workers and zero for nonunion workers). The second, (6), is a Weiss-type equation because it excludes U. It follows from the omitted variable theorem that if both equations are fitted by OLS, then

$$W_y = a_y + \bar{M}b_{Uy\cdot x} = J\bar{M} , \quad \text{where} \quad J \equiv \frac{a_y}{\bar{M}} + b_{Uy\cdot x} , \tag{9}$$

where $b_{Uy\cdot x}$ is the partial regression coefficient of union status U on extent of unionism y in a regression that also includes the x variables on the right-hand side.

Therefore, unless J is close to unity, W_y and $\bar{M}$ will not be nearly equal. However, even if J were not near unity, it would still be possible to infer $\bar{M}$ from W_y if J varied little across data sets and y variables so that its magnitude could be estimated from a few experiments of the kind reported in the next section. Unfortunately, these experiments suggest that J is not unity and is not by any means invariant to the choice of y variable.

Aggregate Macro Equations

All of the Macro equations except the Weiss type were fitted to aggregates, say by industry, of individual-worker data. Go back to equation (5), interpret it as a Micro equation, assume that it has been fitted by OLS, but modify its format as follows:

$$W = a + a_x x + a_x^* x^* + a_y U^* + \bar{M} U + e \,. \tag{10}$$

The modifications consist only of adding to the list of right-hand variables the industry means x^* of all of the x variables and replacing the extent-of-unionism y variable by the industry mean U^* of the union status dummy variable U. Aggregate (10) by industry to obtain

$$W^* = a + d_x x^* + W_y U^* + e^* \,; \quad d_x \equiv a_x + a_x^* \,, \\ \text{and} \quad W_y \equiv a_y + \bar{M} \,. \tag{11}$$

Since (10) has been fitted by OLS, the residual e is uncorrelated across the individual worker observations with the industry means x^* and U^*. Therefore, the industry mean residual e^* in (11) also is uncorrelated with the industry means x^* and U^* in (11), provided only that each industry observation is weighted by the number of covered workers employed in the industry. Hence, if (11) were fitted by employment-weighted least squares, the coefficient W_y of the extent-of-unionism variable $y = U^*$ would be exactly equal to the sum $a_y + \bar{M}$ of the coefficients a_y for $y = U^*$ and the wage gap $\bar{M}$ for the union status dummy variable U in the Micro equation.

The argument of the preceding paragraph assumes, of course, that the extent-of-unionism variable y used in the Micro and Macro equations is the exact aggregation U^* of the union status dummy variable U in the Micro equation. However, the only Macro equations in table 1 for which this was true are those on lines 2 and 3 (Ashenfelter-Taussig study) and line 32 (Killingsworth study). In all of the other studies the extent-of-unionism variable was not exactly the same as U^*. Of course, if U^* in (11) is measured inaccurately by the extent-of-unionism variable y that is used in the Macro equation, the coefficient W_y of y in the fitted equation will not be exactly equal to $a_y + \bar{M}$. Indeed, if the measurement errors are uncorrelated with U^*, W_y will be less than $a_y + \bar{M}$.

There are other reasons why in the table 1 Macro equations the interpretation of W_y as the sum $a_y + \bar{M}$ of coefficients in an underlying, but not observed, Micro equation is not exact. Most of the Macro equations reported in the table were fitted without employment weighting of the observations. Frequently some of the right-hand x^* variables were inexact aggregates of their x counterparts in the underlying Micro equation. The same was often true of the left-hand wage variable.

The key question, of course, is whether the coefficient a_y of the extent-of-unionism variable U^* in the Micro equation (10) typically is positive and so large as to make W_y in the Macro equation (11) overstate substantially the wage gap $\bar{M}$ for the workers covered in the Macro equation, even in the presence of errors in measuring U^* and other variables. The next section reports several experiments addressed to this question.

I turn now to the slightly special problem encountered in the interpretation of the Macro equations of the Hirsch-Rufolo study (no. 15) and the two Hamermesh studies (nos. 10 and 11). In these studies the left-hand variable (in logarithmic units) was the wage difference between two different groups of workers and there were two right-hand extent-of-unionism variables in the Macro equations, one for each of the two groups of workers. Denote the two groups by subscripts 1 and 2 and write the Micro wage equation for each of the two groups as follows:

$$W_k = a_{k1}U_1^* + a_{k2}U_2^* + \bar{M}_k U_k , \quad k = 1, 2. \tag{12}$$

Each equation also includes an intercept and an error term along with x and x^* variables, but I have left them out because their presence is not essential to what follows. Aggregate each of the equations (12) and then subtract one equation from the other:

$$W_1^* - W_2^* = W_{y1}U_1^* - W_{y2}U_2^* ; \quad W_{yk} = a_{kk} - a_{jk} + \bar{M}_k ; \quad j \neq k = 1, 2. \tag{13}$$

Thus, if the a's are all positive, W_{yk} will overstate $\bar{M}_k$ not by a_{kk} but by the smaller amount $a_{kk} - a_{jk}$. This may account in part for the lowness of the estimates of W_y from these studies relative to the others in table 1. I suspect, however, that measurement error in the extent-of-unionism variables used in these studies and in the Hendricks study, line 23, was the chief reason for their low W_y figures.

Let us return to equations (10) and (11). Notice that in the Micro equation (10) there are two quite distinct "unionism" variables, extent of unionism ($U^* = y$) and the dummy variable U classifying the observations by their union status: unionized or nonunion. It is the presence of U rather than y that is critical for the estimation of the wage gap $\bar{M}$. (It is not at all essential for the estimation of $\bar{M}$ that the observations to which the Micro equations are fitted be for individual workers so long as they can be clearly classified as either for a group of union workers or for a group of nonunion workers.)

In going from the Micro equation (10) to the Macro equation (11), the distinction between extent of unionism and union status is lost. There is only one unionism variable in (11), extent of unionism $y = U^*$, and its coefficient W_y in (11) picks up the separate effects in the Micro equation (10) of both extent of unionism in its coefficient a_y and union status in its coefficient $\bar{M}$.

Rosen (no. 29) was aware of the importance of distinguishing between union status and extent of unionism in the Macro as well as in the Micro wage equation. The essence of his proposed solution to the problem is this. First rewrite the Micro equation (10) as follows:

$$W = a_{y1}D_1 + a_{y2}D_2 + \ldots + a_{yk}D_k + \bar{M}U + e\,, \qquad (14)$$

where the D's are dummy variables dividing industries into $k + 1$ classes according to their values of U^*, and assume that e is uncorrelated with right-hand variables. (For simplicity I do not show right-hand x and x^* variables.) Now aggregate (14):

$$W^* = a_{y1}D_1 + a_{y2}D_2 + \ldots + a_{yk}D_k + \bar{M}U^* + e^*\,. \qquad (15)$$

Since e is uncorrelated with the D's (and U) in (14), e^* is uncorrelated with the D's in (15). Unfortunately, however, zero correlation between e and U in (14) does *not* imply zero correlation between e^* and U^* in the Macro equation (15). For this reason, when (15) is fitted by employment weighted least squares, there is no assurance that the resulting coefficient W_y for U^* will be equal to $\bar{M}$. Indeed, in the next section I report an experiment in which I fitted both the Macro equation (15) and its Micro counterpart. The coefficent of U^* in the aggregated equation differed considerably from the estimate of $\bar{M}$ in the Micro equation. Nevertheless, the estimates 0.13–0.28 (line 46) of W_y from Rosen's study (no. 29) are considerably lower than his corresponding estimates 0.26–0.34 (line 47) for the same workers for the same year using the same data sources from his study number 30.

IV. Macro versus Micro Estimates: Some Experiments

Weiss-Type Equations

Return to equations (5), (6), and (9):

$$W = a + a_x x + a_y y + \bar{M}U + e; \qquad (5)$$

$$W = c + c_x x + c_y y + e'\,, \quad W_y = c_y\,, \qquad (6)$$

$$W_y = a_y + \bar{M}b_{Uy\cdot x} = J\bar{M}\,, \quad J \equiv \frac{a_y}{\bar{M}} + b_{Uy\cdot x}\,, \qquad (9)$$

where $b_{Uy\cdot x}$ is the partial regression coefficient of union status U on extent of unionism y in a regression that also includes the x variables on the right-hand side. Equation (5) is a Micro wage equation, and (6) is the corresponding Weiss-type Macro equation. Equation (9) relates the coefficient $c_y = W_y$ of y in the Weiss-type equation (6) to the coefficient a_y of y and $\bar{M}$ (the wage gap) of U in the Micro equation (5) when both (5) and (6) are fitted by OLS. Questions: How large is J, and how stable is it across data sets and y variables used?

I am surprised that the literature on union wage effects contains so little information on these questions. The needed data have been available for over a decade and no sophisticated econometrics or computer processing is required. The unpublished Ashenfelter-Taussig paper (study no. 2), which exists only in the form of a short set of notes and tables, contains some relevant information. Their notes show that they fitted a Weiss-type equation (6) for which the coefficient W_y of extent of unionism y was 0.42—see line 4 of table 1. The corresponding wage gap coefficient $\bar{M}$ of union status U in their fitted version of the Micro equation (5) was 0.13. However, in fitting (5) they omitted the extent-of-unionism variable y. Had they not omitted y, I suspect that their estimate of $\bar{M}$ would have been about 0.12 instead of 0.13. Thus, from their results, I estimate that J was about 3.5, a very large ratio.

Table 3 presents the key results from several experiments that I have made in an effort to discover the unionism content of W_y when estimated from Weiss-type equations. The table covers 20 Micro wage equations and 10 matching Weiss-type equations fitted by OLS to the May 1973

Table 3
Comparison of Micro Estimates of $\bar{M}$ with Weiss-Type Estimates of W_y (White Males in Private Sector, May 1973 CPS)
A.

			Micro equations			
		Weiss-type Equation	With Uy		Without Uy	
Line No. (1)	y Variable (2)	W_y (3)	$\bar{M}$ (4)	a_y (5)	$\bar{M}$ (6)	a_y (7)
1	U^* by industry	.365	.180	.234	.172	.224
2	BLS-PW by industry	.342	.183	.254	.166	.249
3	BLS-AW by industry	.323	.181	.206	.175	.201
4	CPS-PW by industry	.437	.184	.332	.164	.313
5	CPS-AW by industry	.395	.183	.268	.173	.246
6	U^* by occupation	.124	.172	−.054	.194	−.031
7	CPS by occupation	.159	.172	−.011	.192	.011
8	U^* by SMSA	.100	.172	−.012	.172	−.012
9	CPS-PW by SMSA	.103	.173	.010	.172	.010
10	CPS-AW by SMSA	.089	.174	−.037	.173	−.040

B.

	$W_y - \bar{M}$		$W_y/\bar{M}$		$W_y/(a_y + \bar{M})$	
Line No. (1)	Col. 3 − Col. 4 (8)	Col. 3 − Col. 6 (9)	Col. 3/ Col. 4 (10)	Col. 3/ Col. 6 (11)	Col. 3/ Col. 4 + Col. 5 (12)	Col. 3/ Col. 6 + Col. 7 (13)
1	.185	.193	2.03	2.12	.883	.922
2	.159	.176	1.87	2.05	.782	.823
3	.142	.148	1.78	1.85	.835	.860
4	.253	.273	2.38	2.67	.847	.916
5	.212	.222	2.15	2.28	.875	.942
6	−.048	−.070	.72	.64	1.049	.758
7	−.012	−.033	.93	.83	.990	.782
8	−.072	−.072	.58	.58	.621	.623
9	−.071	−.069	.59	.60	.560	.568
10	−.085	−.084	.51	.52	.649	.671

Current Population Survey (CPS) data file. The workers covered are white male wage and salary workers, at least 15 years of age, employed in the private sector, with needed data, but excluding farm and private household workers. Lines 1–5 also exclude workers in CPS detailed industries with fewer than 20 covered workers, lines 6 and 7 exclude workers in CPS detailed occupations with fewer than 20 covered workers, and lines 8, 9, and 10 exclude workers not residing in a Standard Metropolitan Statistical Area (SMSA). The number of observations (workers) for each of lines 1–5 is 17,546; for each of lines 6 and 7 is 17,758; and for lines 8, 9, and 10 is 12,647.

In all of the equations covered in table 3, the dependent variable is the natural logarithm of a worker's usual hourly earnings—his "usual weekly earnings" divided by his "usual weekly hours." All of the regressions have in common the following right-hand variables: years of school completed, age, one marital status dummy variable, two city-size dummies, a dummy variable for part-time work (usual weekly hours less than 35), five major occupation dummies, three region dummies, four major industry dummies, and an extent-of-unionism y variable that is different on each line of the table. These are the only variables included in the Weiss-type equations. The Micro equations, two for each line of the table, also include the union status dummy variable U, and one of the two Micro equations also includes the interaction variable Uy as indicated in the column headings.

In lines 1–5 the y variable is by industry (154 CPS detailed industries), in lines 6 and 7 by occupation (91 CPS detailed occupations), and in lines 8, 9, and 10 by SMSA (98 listed SMSAs and one catchall category for all other SMSAs). U^* is the fraction of *covered* workers in each industry (line 1), or occupation (line 6), or SMSA (line 8), who are union members as reported in the May 1973 CPS where *covered* means "covered in the fitted wage equations." The remaining seven y variables were estimated by Freeman and Medoff (1979) from U.S. Bureau of Labor Statistics (BLS) establishment data for 1968–72 on collective bargaining coverage, separately for production workers (PW) and all workers (AW) by industry (lines 2 and 3) and from May 1973, 1974, and 1975 CPS data on union membership by occupation (line 7) and separately for PW and AW by industry (lines 4 and 5) and by SMSA (lines 9 and 10).

The coefficient W_y of the extent-of-unionism y variable in each of the 10 Weiss-type wage equations is given in column 3 of table 3. The range of these 10 coefficients is quite wide, from 0.09 to 0.44, indicating great sensitivity of W_y to the choice of the extent-of-unionism y variable used in the regressions. The estimates of a_y (the partial derivative of W with respect to y evaluated in col. 5 at the mean of U among covered workers) from the 20 Micro equations, given in columns 5 and 7, vary with the chosen y variable in a manner similar to that of W_y in column 3.

In contrast, the estimates of the wage gap $\bar{M}$ (the partial derivative of

W with respect to union status U, evaluated in col. 4 at the mean of y among covered workers) from the 20 Micro equations, shown in columns 4 and 6, are rather insensitive to which of the 10 y variables enters the equation. The range of $\bar{M}$ in column 4 is 0.172–0.184 and in column 6 is 0.164–0.194.

Panel B of table 3 compares the Weiss-type "union wage effect" estimates W_y with the wage gap estimates $\bar{M}$ from the matching Micro wage equations in three different ways: the differences $W_y - \bar{M}$ in columns 8 and 9, the ratios $W_y/\bar{M}$ in columns 10 and 11, and the ratios $W_y/(a_y + \bar{M})$ in columns 12 and 13. Both the differences $W_y - \bar{M}$ and the ratios $W_y/\bar{M}$ are too unstable across the 10 lines of the table to permit estimation with much precision of the wage gap $\bar{M}$ from knowledge of W_y estimates from Weiss-type equations. There are some regularities, of course, in columns 8–11: all of the Weiss-type W_y's are larger than corresponding wage gap $\bar{M}$'s when the y variable is by industry, and the reverse holds when the y's are by occupation or SMSA.

How closely does W_y approximate the corresponding value of $a_y + \bar{M}$ from the Micro equation? The geometric mean of the ratios in column 12 is 0.79, and the range is 0.56–1.05. The geometric mean of column 13 is 0.78 and the range is 0.57–0.94. The differences $W_y - (a_y + \bar{M})$, calculated from columns 3–5 and not shown in the table, average -0.053 and range from -0.095 to 0.006. Thus these figures suggest that although Weiss-type W_y estimates are somewhat lower than values of $a_y + \bar{M}$ from corresponding Micro wage equations, the interpretation of W_y as roughly the sum of the union/nonunion wage gap $\bar{M}$ and the extent-of-unionism wage differential captured in the coefficient a_y is a valid one.

Table 3 reports results of experiments designed to discover what it is that is estimated by Weiss-type W_y figures. In various other contexts, however, I have fitted by OLS a variety of other Weiss-type equations and matching Micro equations. These are summarized in table 4.

All of the equations covered in table 4 were fitted by OLS to May CPS data for individual workers. Lines 1, 2, and 3 cover all nonfarm, not-

Table 4
Other Comparisons of W_y and $\bar{M}$ (May CPS Data)

LINE NO. (1)	COVERAGE AND DATE* (2)	y (3)	W_y (4)	MICRO EQUATIONS $\bar{M}$ (5)	a_y (6)	$W_y/(a_y + \bar{M})$ (7)
1	All, 1973–75	BLS-AW	.275	.168	.157	.845
2	All, 1973–75	CPS-AW	.349	†	†	†
3	Operatives, 1973–75	BLS-AW	.368	.182	.219	.917
4	White males, 1973	CPS-PW	.390	.176	.299	.821
5	White males, 1973	BLS-PW	.253	.173	.174	.728
6	White males, 1973	CPS-AW	.348	.172	.235	.855
7	White males, 1973	BLS-AW	.269	.167	.164	.812

* See text for details.
† Micro equation not fitted.

private-household wage and salary workers, 16 years of age and older, without missing data, except that line 3 is restricted to operatives (except transport equipment operatives). The worker coverage on lines 4-7 is the same as in table 3 except that in table 4 the observations come from a 20% random subsample of the May 1973 CPS.

The dependent variable in all of the equations is the natural logarithm of usual hourly earnings. All of the equations have an extent-of-unionism y variable by industry on the right-hand side. Column 3 identifies the y variable used on each line where the short-hand identifications have the same meaning in table 4 as in table 3. The estimated coefficients W_y of y in the Weiss-type equations appear in column 4. All of the Micro equations reported in the table include two additional unionism variables: the union membership status dummy variable U and its interaction Uy with y. The estimates of $\bar{M}$ and a_y from the Micro equations are given in columns 5 and 6 and the ratio of W_y to $a_y + \bar{M}$ in column 7.

The x variables included in the wage equations are as follows:

Lines 1–3: years of school completed and its square, age and its square, the cross-product of age and schooling, two year dummies, two marital status dummies, eight major occupations dummies (omitted, of course, on line 3), 14 major industry dummies, three sex-by-race dummies, seven region-by–rural versus urban dummies, and 98 SMSA dummies.

Lines 4–7: schooling and its square, experience (age minus schooling) and its square, three marital status dummies, eight region dummies, five city-size dummies, eight occupation dummies, 11 industry dummies, and one part-time worker dummy.

The numbers in table 4 closely resemble those by industry on the first five lines of table 3 and lead to the same conclusions.

Aggregate Macro Equations

Return to equations (10) and (11):

$$W = a + a_x x + a_x^* x^* + a_y U^* + \bar{M} U + e \, , \tag{10}$$

$$W^* = a + d_x x^* + W_y U^* + e^* ; \quad d_x \equiv a_x + a_x^* ; \quad W_y \equiv a_y + \bar{M} , \tag{11}$$

where (10) is a Micro wage equation fitted by OLS, (11) is the corresponding aggregate (by industry, city, etc.) Macro equation fitted by employment-weighted least squares, and the asterisks denote means (by industry, city, etc.) of the variables to which they are attached. Thus, when the aggregation is exact, as in (11), the coefficient W_y of the extent-of-unionism variable U^* is exactly equal to the sum $a_y + \bar{M}$ of the coefficient a_y of U^* and the wage gap coefficient $\bar{M}$ of union status U in the matching Micro equation (10). However, if the aggregation is not exact—for example, when the extent-of-unionism variable y in (10) and

(11) is not U^* or when (11) is not fitted with employment weights, then W_y is only approximately equal to $a_y + \bar{M}$, as table 5 below shows.

The table covers 10 Micro equations, all fitted by OLS, and 20 Macro equations of which 10 were fitted with employment-weighted observations and 10 without weighting the observations. The data sources, worker coverage, extent-of-unionism y variables, and the dependent variable W and right-hand x variables in the 10 Micro equations are exactly the same as in table 3. The aggregates W^* of W and the x^*s of the x variables in the Macro equations are all exact. All of the means x^* of the x's, as well as the x's, are included as right-hand variables in the Micro equations along with the extent-of-unionism y variable and the union membership dummy variable U.

Within each column of the table the figures differ by line only because the extent-of-unionism y variable is different on each line. The estimates of the wage gap $\bar{M}$ in column 4 vary little across the 10 lines of the table. Furthermore, the $\bar{M}$ estimates in table 5 are close to their counterparts in table 3 despite the inclusion in table 5 but not in table 3 of the means x^* of the x's as right-hand variables in the Micro equations. In contrast, the estimated coefficients a_y of y in column 3 of table 5 show considerable sensitivity to the choice of y variable, and several of the a_y estimates in table 5 differ by a substantial amount from their table 3 counterparts, indicating sensitivity of estimates of a_y to the inclusion or exclusion from the Micro equation of the means x^*.

The estimates of W_y from the Macro equations in columns 5 and 6 also have much across-lines variability, and some of them are somewhat sen-

Table 5
Comparison of Micro Estimates of a_y and $\bar{M}$ with Aggregate Macro Estimates of W_y
(White Males in Private Sector, May 1973 CPS)

		REGRESSION COEFFICIENTS					
		Micro		Macro: W_y		$W_y/(a_y + \bar{M})$	
LINE NO. (1)	y VARIABLE (2)	a_y (3)	$\bar{M}$ (4)	Weighted (5)	Unweighted (6)	Weighted (7)	Unweighted (8)
1	U^* by industry	.216	.174	.390	.452	1.00	1.16
2	BLS-PW by industry	.157	.177	.257	.269	.77	.81
3	BLS-AW by industry	.154	.179	.287	.312	.86	.94
4	CPS-PW by industry	.236	.173	.384	.448	.94	1.10
5	CPS-AW by industry	.238	.175	.423	.490	1.02	1.19
6	U^* by occupation	−.216	.203	−.013	.061	1.00	a
7	CPS by occupation	−.216	.202	−.016	.058	a	a
8	U^* by SMSA	.242	.168	.410	.366	1.00	.89
9	CPS-PW by SMSA	.204	.169	.327	.237	.88	.64
10	CPS-AW by SMSA	.261	.169	.438	.329	1.02	.76

[a] Not shown because denominator is negative and close to zero.

sitive to weighting. The 20 estimates of W_y range from -0.02 to 0.49, which is fairly similar to the range 0.09–0.44 of the Weiss-type W_y estimates in table 3.

Columns 7 and 8 show the ratios $W_y/(a_y + \bar{M})$. On lines 1, 6, and 8 of column 7, the Macro equation is an exact aggregation in all respects of the matching Micro equation and, of course, the ratio $W_y/(a_y + \bar{M})$ then must be unity. None of the other 14 ratios in columns 7 and 8, however, is unity to at least two decimal places and, though the geometric mean of these 14 ratios is 0.91, which is fairly close to unity, the range of the ratios is wide, from 0.64 to 1.19. The differences $W_y - (a_y + \bar{M})$, not shown in the table, average -0.013 (zero differences are excluded in calculating the average) and range from -0.135 to 0.077. Thus, though on the average the values of $a_y + \bar{M}$ approximate well the values of W_y from matching aggregate Macro equations, the error in a particular approximation may be substantial in magnitude and of ambiguous sign. Furthermore, both the differences $W_y - \bar{M}$ and the ratios $W_y/\bar{M}$ are widely dispersed. The differences range from -0.22 to 0.32 and the ratios from -0.08 to 2.8.

Let us return to equation (15), which incorporated Rosen's suggestion (in study no. 29) for maintaining in the Macro equation the distinction between extent-of-unionism effects and wage gap (or union status) effects. I fitted an employment-weighted Macro equation exactly like that on line 1 of table 5 except for including three dummy variables classifying industries by their values of U^* as in equation (15). The estimate of W_y (equal to the estimated coefficient of U^*) was 0.365, more than twice as large as the wage gap $\bar{M}$ estimate from the matching Micro equation. Thus in this experiment his suggestion failed to work.

In this section I have presented evidence that:

1. Macro estimates of the union wage effect measured by W_y from wage equations fitted by least squares to aggregated cross-section data tend on the average to estimate the sum of two quite distinct coefficients in corresponding micro equations: the union/nonunion wage gap $\bar{M}$ and the partial derivative a_y of W with respect to the extent-of-unionism variable y used in the Macro equations. A similar proposition holds for Weiss-type W_y estimates, with the qualification that Weiss-type W_y figures tend to be somewhat smaller than values of $a_y + \bar{M}$ from matching Micro equations.

2. Ratios of Macro W_y's to corresponding values of $a_y + \bar{M}$ (or $\bar{M}$) from Micro equations as well as differences $W_y - (a_y + \bar{M})$ (or $W_y - \bar{M}$) have substantial dispersion.

3. Estimates of a_y are quite sensitive to the choice of y variable and to other aspects of wage equation specification.

4. Therefore, estimating the wage gap $\bar{M}$ from knowledge of W_y involves much imprecision.

V. The Verdict on the Macro Estimate

Sections III and IV convince me that the interpretation of W_y as estimating the wage gap is incorrect. In my judgment the estimates of W_y summarized in table 1 should be ignored in estimating the mean wage gap in the U.S. work force as a whole and in its parts.

Of course, there are other ways of interpreting W_y than as a wage gap estimate. In particular, in some of the Macro studies $\bar{y}W_y$ is interpreted, I think, as the mean real wage gain attributable to unionism. The wage gain concept is quite different from that of the wage gap. The wage *gap* for a worker, measured in the presence of the existing unionism, is the excess of his wage if unionized over his wage if nonunion. The wage gain, on the other hand, is the excess of his real wage in the presence of unionism over his real wage in the absence of unionism.

Assume for the moment that this wage gain interpretation of W_y is correct. For the white males covered in tables 3 and 5, the mean extent of unionism $\bar{y}$ varied from line to line as follows depending on the y variable used: line 1, 0.32; line 2, 0.46; line 3, 0.36; line 4, 0.37; line 5, 0.29; line 6, 0.32; line 7, 0.29; line 8, 0.33; line 9, 0.37; and line 10, 0.24. In tables 3 and 5 the estimates of $\bar{y}W_y$ ranged from slightly negative to 0.17. Thus the experiments reported in tables 3 and 5 indicate that $\bar{y}W_y$ is quite sensitive to the choice of the extent-of-unionism y variable.

This interpretation of W_y, however, is incorrect. Return to equation (11), which is an exact aggregation of the Micro equation (10):

$$W = a + a_x x + a_x^* x^* + a_y U^* + \bar{M}U + e \,, \tag{10}$$

$$W^* = a + (a_x + a_x^*)x^* + W_y U^* + e^* \,, \quad W_y = a_y + \bar{M} \,. \tag{11}$$

It follows from (10) and (11) that if $\bar{y}W_y$ $(=\bar{U}W_y)$ is the overall mean real wage gain, given x and x^*, then the real wage gain of workers in unorganized sectors ($U = U^* = 0$), given x and x^*, on the average is zero. The notion that workers in unorganized sectors are insulated from spillovers from the rest of the economy is a strange one. Surely in the general equilibrium of the economy in the presence of unionism the real wage gains or losses of (nonunion) workers in unorganized sectors depend upon the extent of unionism in the whole economy and the distribution of workers by union status among the parts of the economy.

Essentially what I am arguing here is that the coefficients in (10) and (11), estimated from data in the presence of unionism, depend upon the distribution of unionism in the economy. Thus, if this distribution is changed, the estimated coefficients in the wage equations also will change. Therefore I reject the interpretation of $\bar{y}W_y$ as an estimate of mean real wage gains.

Appendix

1. Ashenfelter and Johnson 1972.
2. Ashenfelter and Taussig 1971.
3. Cain et al. 1981.
4. Dalton and Ford 1977.
5. Ehrenberg 1973.
6. Ehrenberg and Goldstein 1975.
7. Fottler 1977.
8. Fuchs 1968.
9. Gustman and Clement 1977.
10. Hamermesh 1971.
11. Hamermesh 1975.
12. Hendricks 1977.
13. Hirsch 1981.
14. Hirsch (forthcoming).
15. Hirsch and Rufolo (forthcoming).
16. Kahn 1977.
17. Kahn 1979.
18. Kasper 1970.
19. Killingsworth (forthcoming).
20. McLaughlin 1980.
21. Mitchell 1980.
22. Mixon 1978.
23. Pashigian 1980.
24. Pencavel 1970.
25. Potthoff 1977.
26. Raimon and Stoikov 1969.
27. Rapping 1967.
28. Rosen 1969*a*.
29. Rosen 1969*b*.
30. Rosen 1970.
31. Throop 1968.
32. Victor 1979.
33. Weiss 1966.
34. Wessels 1981.

References

Ashenfelter, O., and Johnson, G. E. "Unionism, Relative Wages, and Labor Quality in U.S. Manufacturing Industries." *International Economic Review* 13 (October 1972): 488–508.

Ashenfelter, O., and Taussig, M. "Alternative Estimates of Union-Nonunion Wage Differentials." Mimeographed. Princeton, N.J.: Princeton University, Industrial Relations Section, May 1971.

Cain, G. G.; Becker, B. E.; McLaughlin, C. G.; and Schwenk, A. E. "The Effect of Unions on Wages in Hospitals." *Research in Labor Economics* 4 (1981): 191–320.

Dalton, J. A., and Ford, E. J., Jr. "Concentration and Labor Earnings in Manufacturing and Utilities." *Industrial and Labor Relations Review* 31 (October 1977): 45–60.

Ehrenberg, R. G. *An Economic Analysis of Local Government Employment and Wages*, Final Report DL 91-25-73-09-1, U.S. Department of Labor, 1973, Chapter IV.

Ehrenberg, R. G., and Goldstein, G. S. "A Model of Public Sector Wage Determination." *Journal of Urban Economics* 2 (July 1975): 222–45.

Fottler, M. D. "The Union Impact on Hospital Wages." *Industrial and Labor Relations Review* 30 (April 1977): 342–55.

Freeman, Richard B., and Medoff, James L. "New Estimates of Private

Sector Unionism in the United States." *Industrial and Labor Relations Review* 32 (January 1979):143–74.
Fuchs, V. R. *The Service Economy.* New York: Columbia University Press (for NBER), 1968.
Gustman, A. L., and Clement, M. O. "Teachers' Salary Differentials and Equality of Educational Opportunity." *Industrial and Labor Relations Review* 31 (October 1977): 61–70.
Hamermesh, D. S. "White-Collar Unions, Blue-Collar Unions, and Wages in Manufacturing." *Industrial and Labor Relations Review* 24 (January 1971): 159–70.
———. "The Effects of Government Ownership on Union Wages." In *Labor in the Public and Nonprofit Sectors,* ed. Daniel S. Hamermesh. Princeton, N.J.: Princeton University Press, 1975.
Hendricks, W. "Regulation and Labor Earnings." *Bell Journal of Economics* 8 (Autumn 1977): 483–96.
Hirsch, B. T. "The Inter-Industry Structure of Earnings and Unionization." Mimeographed. Greensboro: University of North Carolina at Greensboro, July 1981.
———. "The Inter-Industry Structure of Unionism, Earnings, and Earnings Dispersion." *Industrial and Labor Relations Review* 36 (October 1982): 22–39.
Hirsch, W. Z., and Rufolo, A. M. "Effects of Prevailing Wage Laws on Municipal Government Wages." *Journal of Urban Economics* (1983), in press.
Kahn, L. M. "Union Impact: A Reduced Form Approach." *Review of Economics and Statistics* 49 (November 1977): 503–7.
———. "Unionism and Relative Wages: Direct and Indirect Effects." *Industrial and Labor Relations Review* 32 (July 1979): 520–32.
Kasper, H. "The Effects of Collective Bargaining on Public School Teachers' Salaries." *Industrial and Labor Relations Review* 24 (October 1970): 57–72.
Killingsworth, M. R. "Union-Nonunion Wage Differentials: Estimates from a New Industry Cross-Section." Mimeographed. New Brunswick, N.J: Rutgers University, 1981.
Lewis, H. Gregg. *Unionism and Relative Wages in the United States.* Chicago: University of Chicago Press, 1963.
McLaughlin, C. G. "The Impact of Unions on Hospital Wages." Unpublished Ph.D. dissertation, University of Wisconsin—Madison, 1980.
Mincer, Jacob. *Union Effects: Wages, Turnover, and Job Training.* New York: Columbia University Press (for NBER), 1981.
Mitchell, D. J. B. "Some Empirical Observations of Relevance to the Analysis of Union Wage Determination." *Journal of Labor Research* 1 (Fall 1980): 193–215.
Mixon, J. W., Jr. "Earnings Differentials in U.S. Manufacturing, 1972." *Quarterly Review of Economics and Business* 18 (Spring 1978): 75–82.
Pashigian, B. P. "Has Occupational Licensing Reduced Geographical Mobility and Raised Earnings?" In *Occupational Licensure and Reg-*

ulation, edited by Simon Rottenberg. Washington, D.C.: American Enterprise Institute, 1980.
Pencavel, J. H. *An Analysis of the Quit Rate in American Manufacturing Industry.* Princeton, N.J.: Princeton University, Industrial Relations Section, 1970.
Potthoff, R. F. "A Study of the Determinants of Geographic Earnings Differentials." Mimeographed. August 1977.
Raimon, R. L., and Stoikov, V. "The Effect of Blue-Collar Unionism on White-Collar Earnings." *Industrial and Labor Relations Review* 22 (April 1969): 358–74.
Rapping, L. A. "Monopoly Rents, Wage Rates, and Union Wage Effectiveness." *Quarterly Review of Economics and Business* 7 (Spring 1967): 31–47.
Rosen, S. "On the Interindustry Wage and Hours Structure." *Journal of Political Economy* 77 (March/April 1969): 249–73. (*a*)
———. "Trade Union Power, Threat Effects, and the Extent of Organization." *Review of Economic Studies* 36 (April 1969): 185–96. (*b*)
———. "Unionism and the Occupational Wage Structure in the United States." *International Economic Review* 11 (June 1970): 269–86.
Throop, A. W. "The Union-Nonunion Wage Differential and Cost-Push Inflation." *American Economic Review* 58 (March 1968): 79–99.
U.S. Bureau of the Census. *Statistical Abstract of the United States: 1980.* Washington, D.C.: Bureau of the Census, 1980.
Victor, R. B. "Municipal Unions, Wages and Employment." Mimeographed. November 1979. (A portion of this is contained in Victor's "Municipal Unions and Wage Patterns" in Industrial Relations Research Association, *Proceedings of the Thirty-second Annual Meeting* [Madison: University of Wisconsin, Industrial Relations Research Assn., 1979], pp. 294–99.)
Weiss, L. W. "Concentration and Labor Earnings." *American Economic Review* 56 (March 1966): 96–113.
Wessels, W. J. "Wages, Unionization, Strike Frequency and Union Political Power." Mimeographed. Northern California State University, 1981.