Wages and Gender Composition: Why Do Women's Jobs Pay Less?

David A. Macpherson, *Florida State University*

Barry T. Hirsch, *Florida State University*

Occupational sex segregation and its relationship with wages during 1973-93 are examined. Wage level and wage change models are estimated using Current Population Survey data matched with measures of occupational skills and job disamenities. Standard analysis confirms that wage levels are substantially lower in predominantly female occupations. Gender composition effects are reduced by about a quarter for women and by over one-half for men following control for skill-related occupational characteristics. Longitudinal analysis indicates that two-thirds or more of the standard gender composition effect is accounted for by occupational characteristics and unmeasured worker skill or taste differences.

I. Introduction

An important contribution to the understanding of gender differences in the labor market has been the finding that wages of both women and men are lower in predominantly female occupations. Such evidence not only enriches our knowledge about the routes through which gender differentials are realized but is useful in evaluating the efficacy of comparable worth and other policies intended to alter the wage structure.[1] The finding

We appreciate the helpful comments and suggestions of Paula England, Erica Groshen, Dan Hamermesh, Ron Oaxaca, and Jay Stewart.

[1] Among the many studies examining sex segregation or the relationship between wages and gender composition are Bergmann (1974), O'Neill (1983), papers in Reskin (1984), Johnson and Solon (1986), Blau and Beller (1988), Sorensen (1989, 1990), Groshen (1991), Fields and Wolff (1991), and England (1992). Sorensen (1990) provides a comprehensive survey and analysis of the literature on wages and gender composition.

[*Journal of Labor Economics*, 1995, vol. 13, no. 3]

0734-306X/95/1303-0006$01.50

that individual wages vary systematically with the gender composition of occupations is seemingly well established in cross-sectional empirical studies For example, Killingsworth (1990, p 24) provides two stylized facts regarding the "femaleness" of occupations (1) both women and men earn less as the proportion female in an occupation increases, and (2) the negative relationship between wages and proportion female is stronger among men than women

Despite a seeming consensus regarding the stylized facts, the magnitude and interpretation of the relationship between wages and gender composition remain in some dispute Most studies examining gender composition have not controlled for a number of the job characteristics that might be expected to affect equilibrium wages Nor is there agreement on the wage-composition relationship using longitudinal analysis, wherein individual wage changes are regressed on changes in gender composition [2] Indeed, Sorensen (1990, pp 76–77) identifies the estimation of wage change models and inclusion of more job characteristics in micro-level wage equations as two of the most important advances needed in this area Absent such evidence, one cannot reject the thesis that the proportion female is a proxy for unmeasured skill and taste differences among workers or of occupational attributes correlated with wages [3]

This article uses the 132 monthly Current Population Survey Outgoing Rotation Group (CPS ORG) files from January 1983 to December 1993,

[2] Studies regressing wage changes on changes in proportion female include England et al (1988) and Gerhart and El Cheikh (1991) Gerhart and El Cheikh examine this issue using the 1983 and 1986 waves of the National Longitudinal Survey of Youth (ages 18–25 in 1983 and ages 21–28 in 1986) Following control for fixed effects, they continue to find a negative and significant relationship for young men (N = 2,460), but a small negative and insignificant effect for young women (N = 2,294) By comparison, England et al find stronger evidence of a negative gender composition effect in their fixed-effects model using the Young Women's and Young Men's cohorts of the original NLS (their sample covers the period 1968–80 for women and 1966–81 for men) Because corresponding estimates from equations without control for fixed effects are not provided, the effect on estimates owing to unmeasured person-specific effects cannot be discerned

[3] Filer (1989) estimates aggregate male and female wage equations at the occupational level, first with gender composition and demographic variables included, and then with the addition of a large number of variables measuring occupational characteristics (for a recent effort along these lines, see England 1992) Filer finds that gender composition has a negative effect on male and female wages in equations excluding occupational characteristics, but no effect when these characteristics are included Filer's results have been subject to much criticism (Smith 1989, Sorensen 1990, Jacobs and Steinberg 1990, but see the reply by Filer 1990) and are not widely accepted In particular, Filer is criticized for not estimating wage equations at the micro level, for the use of an extraordinarily large number of occupational variables in an aggregate wage equation, and for coefficients on several variables at odds with much of the literature Our study is not subject to these same criticisms

plus data from several additional sources, to examine how female and male wages vary with gender composition (proportion female) in a worker's occupation. The article provides recent evidence on changes in the gender wage gap and gender segregation. Principal contributions of the article include the use of several large data sets, the addition of occupational variables not commonly used in wage studies, and the construction of large CPS panels enabling an analysis of individual worker's wage changes and changes in gender composition. The longitudinal analysis enables one to control for unmeasured individual labor quality or taste differences that are correlated with the gender composition of jobs. Thus, we examine whether the wage-composition correlation evident in previous studies is due, at least in part, to occupational characteristics, quality sorting on gender composition, taste differences, or other factors correlated with the proportion female in an occupation.

Section II of this article provides a discussion of the relationship between gender composition and wage rates, with an emphasis on measurement issues and interpretation. Section III describes the construction of the CPS cross-sectional data set and provides descriptive evidence on female and male wages and the gender composition of jobs for the 21-year period from 1973 to 1993. Section IV provides cross-sectional evidence for the years 1983–93 on the effect of gender composition on the wages of women and men, decomposes the gender wage gap into its component parts, compares the effects of gender composition among various worker groups, and examines issues of specification. In Section V, longitudinal evidence from three alternative data sets is examined, with considerable attention given to the issue of measurement error in the change in gender composition variable. A concluding section provides an assessment and interpretation of evidence.

II. Wages and Gender Composition

The relationship between wages and gender composition can be estimated by

$$\ln W_{if} = \sum \beta_{kf} X_{ikf} + \Theta_f \mathrm{FEM}_{if} + e_{if}, \tag{1}$$

$$\ln W_{im} = \sum \beta_{km} X_{ikm} + \Theta_m \mathrm{FEM}_{im} + e_{im}, \tag{2}$$

where subscripts f and m designate female and male, respectively, $\ln W_i$ is the natural log of hourly earnings for individual i, X_k consists of an intercept and variables, indexed by k, measuring personal and/or job characteristics and region, β_k includes a constant and coefficients corresponding to variables in X, FEM is the ratio of female to total employment in the worker's occupation, and Θ is its coefficient, and e is an error term assumed for now to have zero mean and constant variance. A value of $\Theta < 0$ implies that wages decrease with respect to proportion female. If $(\Theta_f - \Theta_m)$ is negative, the logarithmic gender gap widens with respect to FEM; if positive, the gap is narrower in predominantly female jobs.

The logarithmic gender wage gap, $\ln W_m - \ln W_f$, can be decomposed in the following manner:

$$
\begin{aligned}
\overline{\ln W_m} - \overline{\ln W_f} = {} & [\textstyle\sum (p_m\beta_m + p_f\beta_f)(\bar{X}_m - \bar{X}_f)] \\
& + [(p_m\Theta_m + p_f\Theta_f)(\overline{\text{FEM}}_m - \overline{\text{FEM}}_f)] \\
& + [\textstyle\sum (\beta_m - \beta_f)(p_f\bar{X}_m + p_m\bar{X}_f) \\
& + (\Theta_m - \Theta_f)(p_f\overline{\text{FEM}}_m + p_m\overline{\text{FEM}}_f)],
\end{aligned}
\tag{3}
$$

where overbars represent means, and p_m and p_f are the proportion male and female in the sample. The decomposition uses sample proportions to weight the regression coefficients in order to approximate a "nondiscriminatory" or full sample wage structure (see Oaxaca and Ransom [1994] for an analysis of alternative wage decompositions). The first and second terms in brackets (line 1) represent the "explained" portion of the gap, the first being that accounted for by differences in the X's, and the second, that owing to differences in gender density between women and men. The first term can be further disaggregated to examine the separate contribution of selected groups of X's, for example, occupational characteristics. The third term in brackets (line 2) represents the "unexplained" portion of the gender differential, that owing to differences in the coefficients on the X's and FEM.

Interpretation of Θ_f and Θ_m, as well as the decomposition shown in equation (3), depends on the causes of occupational segregation and the routes through which FEM and wage rates are related (for previous discussion of these issues, see, among others, Polachek 1979, Blau 1984, England 1992). Among the nonmutually exclusive explanations for occupational segregation are human capital differences, employer discrimination (based on preferences or statistical discrimination), and premarket differences in family and school inputs and in the socialization process.[4] The most common characterization of the gender composition effect is

[4] An additional explanation is the *devaluation* hypothesis, which argues that employers value less highly work done primarily by women than they value the same work done by men (for a discussion, see England 1992). This explanation has received little attention from economists, because it either is inconsistent with standard theory or reduces to one of the alternative explanations in order to explain wage differentials (e.g., employer discrimination, restrictions on labor mobility, or worker preferences correlated with gender).

that it reflects occupational "crowding" (Bergmann 1974). Women may be crowded into particular occupations, owing either to preferences or to past or present discriminatory barriers to alternative occupations. For example, many women but relatively few men may crowd into occupations with attractive (but costly) job characteristics. In this case the negative effect of FEM would reflect a compensating differential. Crowding lowers the equilibrium wage in these occupations to a level below that for similarly skilled workers in other occupations; interoccupational mobility is insufficient to equalize wages.

The crowding model is useful in explaining $\Theta_f < 0$, but less so in accounting for $\Theta_m < 0$. If men do not face the same barriers as women, why would men accept the lower wages in predominantly female jobs when higher wages are available in male-dominated jobs? If $\Theta_m < 0$, then either predominantly female jobs attract lower-quality (unmeasured) male workers or males in predominantly female jobs have tastes for these jobs and choose to accept lower wages. Stated alternatively, if women face barriers to high-paying occupations, low-paid occupations will attract a disproportionately large number of women and a low proportion of men, hence the negative correlation between FEM and both female and male wages.

The "quality sorting" hypothesis is a related explanation for the wage/gender composition relationship. If women but not men are crowded into low-paying occupations because of discriminatory barriers, then the gender composition of a job becomes an index of labor quality for men and, to a lesser extent, for women. That is, relatively less productive males accept lower-paying jobs in predominantly female occupations. For women, however, the negative correlation between the wage and FEM represents in part the effects of past or present discriminatory barriers. Differences in gender composition across jobs owing to past occupational discrimination by employers or from societal and familial preferences that no longer prevail are likely to have evolved into quality sorting on FEM. Over time, low-paying occupations crowded by women would attract relatively lower-quality males and lose many high-quality females; thus, we observe workers in predominantly female jobs with lower average productivities and wages. Today's labor market equilibrium with sorting on gender composition is a result, in part, of the historical path through which labor markets evolved (Heckman [1991] discusses the difficulty in distinguishing path dependence from worker heterogeneity).

Even in the case of current discrimination, however, FEM serves as a quality index since the probability of a woman being hired into predominantly male jobs is an increasing function of productivity. In short, workers are sorted into occupations based in part on expected productivity, and this productivity may be correlated with gender composition. For example, average and expected tenure may be lower in predominantly female occupations, leading to fewer training investments and a lower equilibrium wage for women and men. If this is a source of the gender composition wage effect, however, it can be measured (i.e., controlled) by inclusion of an appropriate variable measuring average tenure in an occupation.[5]

Taste models of discrimination that posit employer, employee, or consumer prejudices generally lead to the prediction of lower female wages in predominantly female jobs. But if men and women can be rewarded differently within detailed census occupational categories, such models of discrimination lead to the prediction of a weaker (or positive) relationship between wages and FEM for men as compared to women and a larger female-male wage gap in predominantly female jobs. For example, discriminatory employers would pay a premium for men over women, or men preferring not to work with female coworkers would require a wage premium. Employer discrimination can of course be a primary mechanism through which female wages and FEM were initially generated. And the more homogeneous the labor market within occupational cells, the more difficult it is for employers to pay women and men differently. Lower male wages observed currently in such jobs, however, must then reflect the effects of quality sorting on gender composition, worker tastes regarding job characteristics, or immobility or transitional employment among males in predominantly female occupations.[6]

Alternative specifications of cross-sectional models (1) and (2) allow inferences to be made about explanations for the relationship between wages and gender composition. If Θ is sensitive to inclusion of variables measuring occupation-level tenure, training requirements, and work hours, it suggests that human capital differences across occupations, and worker preferences regarding job attachment, help account for the wage-FEM relationship. Sensitivity of Θ to the inclusion of measures of job amenities and disamenities would suggest that the effects of FEM on wages reflect in part compensating differentials and differences in tastes toward these characteristics between workers in predominantly female and male jobs. Equations (1) and (2), however, cannot easily identify unmeasured worker quality and taste differences that may be correlated with FEM (Hwang,

[5] If employers are risk averse, wages will be lower within occupations where productivity is less easily predicted. The issue in this case is not gender differences in productivity but, rather, whether employers can more easily predict productivity for men than for women. Light and Ureta (1992) provide evidence indicating that for recent years (but not previously) tenure can be predicted as accurately for female as for male workers.

[6] For related discussion and analysis, as applied to the *racial* composition of jobs, see Hirsch and Schumacher (1992). Hirsch and Macpherson (1994) conclude that the negative relationship between wages and the proportion black in an occupation is due entirely to the correlation of racial composition with skill-related job characteristics and unmeasured worker-specific quality. For this reason, proportion black is not included here as a control variable.

Reed, and Hubbard [1992] show that bias from unobserved heterogeneity can be large). Estimation of longitudinal wage change models that account for unmeasured individual-specific wage determinants provides a means for examining the importance of these factors.

The longitudinal wage change model follows directly from the wage level model. Let the error term e_i from equations (1) and (2) be divided into a person-specific quality or taste component (Φ) fixed over time and a random error term with zero mean and constant variance (e_i'). Adding a time dimension to the earnings equation, the levels formulation in equations (1) and (2) can be rewritten

$$\ln W_{ify} = \sum \beta_{kf} X_{ikfy} + \Theta_f \mathrm{FEM}_{ify} + \Phi_{if} + e'_{ify}, \tag{1'}$$

and

$$\ln W_{imy} = \sum \beta_{km} X_{ikmy} + \Theta_m \mathrm{FEM}_{imy} + \Phi_{im} + e'_{imy}, \tag{2'}$$

where y subscripts year. If the omitted fixed-effect Φ is negatively correlated with FEM, then levels estimates of Θ_f and Θ_m in (1) and (2) are biased downward away from zero.

Letting the change operator Δ represent changes between adjacent years, $[y - (y - 1)]$, $[(y - 1) - (y - 2)]$, and so on, the following longitudinal wage equations are obtained:

$$\Delta \ln W_{ift} = \sum \beta_{kf} \Delta X_{ikft} + \Theta_f \Delta \mathrm{FEM}_{ift} + \Delta e'_{ift}, \tag{4}$$

and

$$\Delta \ln W_{imt} = \sum \beta_{km} \Delta X_{ikmt} + \Theta_m \Delta \mathrm{FEM}_{imt} + \Delta e'_{imt}, \tag{5}$$

where t represents the time periods over which changes are calculated. Person-specific fixed effects owing to unmeasured quality or taste differences fall out, thus (potentially) allowing unbiased estimation of Θ_f and Θ_m. For example, workers with a strong preference for a job characteristic associated with lower wages (e.g., flexibility of hours) are more likely to be observed in both years employed in jobs with a wage lower than that predicted by a wage level regression not controlling fully for that job characteristic.

Levels estimation of equations (1) and (2) in previous studies produces negative values of Θ_f and Θ_m, indicating lower female and male wages in jobs with higher densities of female workers. If this relationship results entirely from a causal effect of gender composition on wages, then longitudinal estimates of Θ_f and Θ_m from equations (4) and (5) should be similar to the levels estimates. If the negative wage/gender composition relationship is due entirely to unmeasured person-specific differences correlated with FEM, longitudinal estimates of Θ_f and Θ_m go to zero. Differences in the levels and longitudinal estimates of Θ_f and Θ_m, therefore, provide evidence as to the relative importance of unobserved worker skills and preferences versus gender-based occupational discrimination (or unmeasured worker and job characteristics not fixed over time).

III. Data and Descriptive Evidence

The primary database for this study is constructed from the 132 monthly CPS ORG files for January 1983–December 1993. In addition, we provide supplemental evidence on gender composition and wages from 1973 forward by using the 1973–78 May CPS public use files and the January 1979–December 1982 CPS ORG files.[7] The CPS ORG files provide unusually large sample sizes. Moreover, because households are included in the CPS in the same month for 2 consecutive years, construction of large 2-year panels of individuals is possible. The appendix provides a detailed description of the construction of our 1983/4–1992/3 panel. Variables measuring occupation and industry characteristics are matched to individuals in the CPS. These variables are calculated by us from the CPS ORG files based on worker occupation and industry codes, calculated from supplementary CPS files containing variables not included in the standard CPS surveys (e.g., company tenure, firm size, and computer use), or obtained from alternative sources such as the *Dictionary of Occupational Titles* (DOT) and matched to individuals on the basis of their recorded occupation.

In the subsequent analysis, we include all female and male workers ages 16 and over, with complete data provided on usual weekly earnings, usual hours worked per week, occupation, race, and other needed variables. Excluded are workers whose principal activity is school (3.6% of the potential sample) or who had either their industry or occupation code allocated by the census (an additional 1.1%). Our principal cross-sectional analysis is for the years 1983–93, while the longitudinal analysis is based on changes

[7] Beginning in January 1983, new occupation and industry codes from the 1980 Census of Population were adopted by the CPS, and union status was asked of the outgoing rotation groups in each monthly survey rather than only in May. Beginning in 1992, the CPS adopted the 1990 Census of Population occupation and industry codes, but changes between the 1980 and 1990 codes are relatively minor. All constructed variables at the detailed occupation and industry level for 1983–93 have been made time consistent. Major industry and occupation dummies used as regression dummies in the table 1 regressions for 1973–93 are time consistent; occupational categories were defined based on the 1970–80 mappings provided in U.S. Department of Commerce (1983). The CPS ORG, or "earnings microdata," files for 1979 forward are made available by the Data Services Group at the Bureau of Labor Statistics (BLS).

for the period 1983/4–1992/3 Wage rates are measured by usual weekly earnings divided by usual hours worked per week, in December 1993 dollars (wages are deflated by the monthly consumer price index, CPI-U) Workers with implied real wage rates less than $1 00 are excluded from the sample (0 1% of the potential sample) The census top coded weekly earnings at $999 in current dollars for the surveys through 1988, after 1988, they were top coded at $1,923 For the years 1989 and forward, we assigned mean earnings above the $1,923 cap based on the assumption that the upper tail of the earnings distribution follows a Pareto distribution The parameters of the Pareto were estimated separately by year and gender For each month in years prior to 1989, mean earnings for those at the $999 cap were assigned based on the mean in 1989 among female or male workers at or above the same real weekly earnings

The total sample size for the 1983–93 wage level analysis is 1,836,541, with 877,070 women and 959,471 men No individuals in the CPS are excluded because of small occupation sample sizes Because estimation is across individuals rather than occupation, those in small occupational cells receive correspondingly little weight in the regression analysis The sample size of the longitudinal data set combining the panels for 1983/4–1992/3 is 459,685, or 25% the size of the full sample The appendix describes fully the construction of the CPS ORG matched panel, as well as a similarly sized March CPS retrospective panel used in our wage change analysis

Gender composition is measured by FFM, the proportion of female workers in the worker's 3-digit census occupation, calculated from the CPS files Descriptive evidence on FFM and the wage-FFM relationship is provided for the period 1973–93 Here, FFM is calculated on an annual basis for the years 1979–93 (we employ 3-year moving averages in subsequent regression analysis for the 1983–93 period) and over 2 years for the 1973–74, 1975–76, and 1977–78 periods All rotation groups in the CPS were asked earnings questions in the May surveys from 1973–78, whereas the annual ORG files for 1979 forward include only the outgoing rotation groups, the quarter sample of the CPS for whom earnings were measured beginning in 1979 Note that measures of FFM from 1973–82 and 1983–93 are not strictly comparable because the earlier years use the 1970 Census of Population occupation codes and the latter a combination of the 1980 and 1990 census codes The modest changes between the 1980 and 1990 codes (the CPS began using the 1990 codes in 1992) permitted us to construct time-consistent categories for 1983–93 (six small occupational categories were merged into larger categories, reducing the number of potential occupations from 503 in 1983–91 to 497 for 1983–93) Despite substantial changes in occupational definitions and categories between the 1973–82 and 1983–93 periods, we do not observe an obvious break in the FFM series between 1982 and 1983

Table 1 provides descriptive evidence for the years 1973–93 on sample sizes, mean female and male real wage rates, the female to male wage ratio, gender composition, the Duncan index of segregation, and the relationship between wages and FEM. As widely recognized, the gender gap changed little during the 1970s, and then narrowed substantially throughout the 1980s and early 1990s. The ratio of female to male hourly earnings, W_F/W_M, increased from .648 in 1973/4 to .669 in 1983, and then to .764 by 1993. Following a real wage decline during 1973–81, wages for women increased by 12.6% during the 1981–93 period, in contrast to a small decrease observed among men.

Mean values of FEM and the Duncan index of segregation by year indicate declining occupational segregation by sex during the 1970s and 1980s, with slow progress evident during the early 1990s. The occupation percentage female among male workers increased from 17.6% in 1973/4 to 28.8% in 1993, the percentage female among women declined from 72.1% to 68.2% during the same period (for previous evidence on changes in occupational segregation see, among others, Blau and Beller [1988], Fields and Wolff [1991] and O'Neill and Polachek [1993]).[8] The Duncan index of segregation, measured by $\frac{1}{2}\sum|m_j - f_j|$, where m and f are the proportions of male and female employment, respectively, in occupation j, varies between zero in the case of an equal occupational distribution and one in the case of complete sex segregation. The Duncan index falls from .685 in 1973/4 to .546 in 1993. Some of the decline appears to result from the change in occupational definitions between 1982 and 1983 (figures for 1973–82 are not directly comparable to those for 1983–93). Indices of occupational segregation are sensitive to the degree of aggregation. For example, at an even more disaggregated level, a large number of jobs are virtually all male or all female (see Groshen 1991).

Although not a central focus of this article, our data set allows us to provide a 1973–93 time series of the relative wage ratio, W_F/W_M, and of the Duncan index of segregation for groups of workers classified by education, age, race, part- or full-time status, private- and public-sector status, production and nonproduction occupational status, and union status. This

[8] Increases in the relative size of the female labor force can produce increases in FEM for both women and men. In table 1, FEM is averaged over *individuals* (this is equivalent to averaging over occupations, weighted by employment). Alternatively, FEM is calculated by averaging over the 497 occupations (unweighted), beginning in 1983 with the new occupational codes. Among occupations that were majority female and majority male in 1983, mean FEM decreased from .746 in 1983 to .724 in 1993 among the female occupations, while increasing from .166 to .195 among the male occupations. If, instead, we calculate changes for occupations that were either two-thirds female or male in 1983, mean FEM decreased from .838 in 1983 to .796 in 1993 among the female occupations, while increasing from .110 to .137 among the male occupations.

Table 1
Mean Wages, the Wage Gap, Gender Composition, and the Wage-Composition Relationship by Year, 1973–93

Year	Female			Male			Female to Male Wage Ratio	Duncan Segregation Index	Female		Male	
	N	Wage	FEM	N	Wage	FEM			θ	θ_{adj}	θ	θ_{adj}
1973/4	30,070	10 26	721	42,842	15 83	176	648	685	−184	−068	−237	−148
1975/6	29,981	10 08	724	40,376	15 45	186	652	675	−193	−087	−234	−161
1977/8	35,415	10 05	712	45,276	15 57	201	645	652	−208	−101	−264	−186
1979	70,930	9 79	711	89,849	15 03	215	651	636	−212	−109	−214	−150
1980	84,271	9 41	708	103,064	14 28	224	659	628	−222	−132	−209	−163
1981	79,779	9 36	707	95,702	14 14	229	662	623	−242	−138	−207	−164
1982	76,254	9 51	708	88,895	14 27	237	666	618	−252	−140	−186	−166
1983	76,159	9 68	699	87,855	14 46	246	669	600	−248	−140	−139	−135
1984	76,659	9 75	698	88,234	14 39	248	678	595	−251	−151	−098	−141
1985	78,753	9 94	694	89,486	14 59	256	681	585	−258	−168	−072	−158
1986	79,318	10 24	694	88,116	14 88	260	688	582	−259	−171	−039	−171
1987	80,687	10 34	694	88,233	14 94	263	692	576	−241	−163	−043	−206
1988	77,356	10 36	689	84,447	14 83	266	699	569	−229	−164	−010	−183
1989	79,495	10 36	687	86,254	14 31	271	721	560	−222	−163	−031	−183
1990	84,021	10 49	686	90,378	14 29	275	734	555	−216	−165	−017	−201
1991	82,152	10 58	687	87,235	14 13	279	749	556	−220	−169	007	−191
1992	81,522	10 66	686	85,330	14 00	284	761	553	−193	−160	009	−184
1993	80,949	10 71	682	83,903	14 02	288	764	546	−181	−174	002	−190

NOTE—Calculations are from the 1973–78 May CPS and the 1979–93 annual CPS ORG files (N = 2,749,246) Wages are measured by usual weekly earnings divided by usual hours worked, in December 1993 dollars Adjustments for top coding are described in the text FEM measure the proportion of females to total employees in workers' detailed occupation, by year The female to male wage ratio is the mean of female real wages to male real wages The Duncan segregation index is calculated by $\frac{1}{2}\Sigma|m_j - f_j|$, where m and f are the proportions of male and female employment in occupation j θ is the log wage regression coefficient on FEM, without controls θ_{adj} is the FEM coefficient, with controls for years schooling completed, potential experience (measured by age minus schooling minus 5) and its square, and dummies for black, other nonwhite married spouse present, ever-married spouse not present, full-time, public sector, large metropolitan area, region (8), industry (13), and occupation (5) In order to insure a time-consisten specification, union status and separate federal, state, and local dummies are not included information is presented in table A1 Although relative wage ratios and the degree of occupational segregation differ significantly among groups, there are increases in the wage ratio and decreases in sex segregation among all groups of workers during the past 20 years

Table 1 also provides estimated slopes of the log wage–FEM relationship, Θ and Θ_{adj}, unadjusted and adjusted for standard worker characteristics, during the 1973–93 period The unadjusted coefficients are obtained from female and male log wage regressions, estimated by year, with only FEM on the right-hand side The adjusted coefficients are obtained by adding a common set of controls over the entire 20-year period—years of schooling, years of potential experience and its square, and dummies for marital status (2), race and Hispanic status (3), part-time status, public-sector status, large metropolitan area, occupation (5), industry (13), and region (8) Excluded are a measure of union status and separate dummies for federal, state, and local employment, because these variables are not available over all the years Estimates in table 1 for 1983 forward differ slightly from results to be shown subsequently (our so-called standard specification) because of the exclusion of union status and detailed public-sector dummies, and because FEM is calculated here on an annual basis rather than as a moving average Sample restrictions are equivalent to those outlined previously

Comparison of the unadjusted and adjusted FEM coefficients illustrates the importance of controlling for worker (and job) characteristics The unadjusted coefficients for women are relatively stable over the 1973–93 period, showing some indication of rising into the mid-1980s and then falling since then By contrast, the male coefficients begin the period larger than the female coefficients and hit a peak in the late 1970s They then decline steadily and are effectively zero in the early 1990s The conventional conclusion that male coefficients exceed those for women is sensitive to time period and choice of control variables Johnson and Solon (1986) use the May 1978 CPS and find larger coefficients for men (they report unadjusted coefficients of −0 343 for men and − 244 for women) As will be seen subsequently, the low unadjusted coefficient for men in later years reflects a nonlinear relationship in which a large number of predominantly male jobs are low skill and low paying Accounting for schooling and a few standard controls increases the magnitude of the male coefficient and makes the wage-FEM relationship more nearly linear For example, with controls, the male coefficient is relatively stable over time The adjusted FEM coefficient for women displays its lowest value in the 1970s, climbs in the mid-1980s, and is stable thereafter

In subsequent regression analysis for the 1983–93 period, FEM and other occupation and industry measures calculated from the CPS ORG files are measured as 3-year moving averages This reduces measurement error that can result from small sample sizes in some occupation and industry cells Although bias resulting from measurement error within small occupation cells will have little effect on wage level equation estimates, measurement error is a concern in the longitudinal analysis. Because absolute and relative changes in gender composition and job characteristics occur slowly over time, little error is introduced from the use of moving averages. Averages for 1983–85 are matched to both 1983 and 1984, and averages for 1991–93 are matched to both 1992 and 1993.

Table 2 presents descriptive statistics for 1993, with the sample segmented into four occupational categories based on gender composition (the break-points for FEM are .25, .50, and .75). As evident from table 2, wages rates for both women and men are substantially lower in predominantly female jobs (FEM ≥ .75). For men, wages (and schooling) are low in occupations that are almost entirely male, while decreasing with respect to gender composition at higher levels of FEM. Women display a similar wage pattern. Notable in table 2 are the substantial differences in occupational characteristics between predominantly female and male jobs. This is evident in the occupational means of job tenure, proportion part-time, occupational training requirements (SVP), computer use, strength, hazards, and physical and environmental conditions. Because there are substantial differences in wages, worker characteristics, and job characteristics among workers in predominantly female and male jobs, we turn below to an analysis of how estimates of Θ vary with specification.

IV. Wages and Gender Composition: Cross-Sectional Evidence

A. Standard Estimates

In this section we present estimates of Θ_f and Θ_m, based first on the estimation of annual cross-sectional log wage equations for the years 1983–93. Results are presented for what is referred to as the "standard" model, regressions including variables measuring individual characteristics, location, and broad occupation and industry of employment (i.e., years of schooling, potential experience and its square, and dummies for union coverage, black, other nonwhite, Hispanic, married with spouse present, ever married without spouse present, part-time, federal worker, state worker, local worker, large metropolitan area, region [8], industry [13], and occupation [5]). These results are representative both of standard estimation techniques and of variables available in most data sets. We subsequently provide wage level results from what is referred to as the "expanded" model, which includes a variety of occupation and industry characteristics measures not routinely included in wage equations.

Results from the standard model, presented in table 3, indicate that the gender composition effect is substantial and similar in magnitude for women and men. Estimates indicate an increase in the absolute value of Θ since the early 1980s. A Θ of, say, −.18 indicates that wages are about

Table 2
Means of Selected Variables by Gender Composition, 1993

Variable	Value of FEM				
	0– 25	25– 50	50– 75	75–1 0	All
Means for females					
Wage (December 1993 $)	10 823	11 867	11 127	10 040	10 709
Schooling	12 628	13 047	13 501	13 006	13 096
Experience	18 551	19 984	19 179	19 708	19 595
Part-time	177	197	231	290	251
Federal	036	058	023	022	031
State	054	046	066	052	054
Local	063	080	132	146	124
Black	123	108	101	109	108
Union	194	128	161	166	158
DOT-GED	2 827	3 452	3 526	3 187	3 295
DOT-SVP	1 815	2 966	2 545	1 240	1 920
Occupation-tenure	7 156	7 112	6 460	5 559	6 175
Occupation-part-time	131	154	214	295	238
Occupation-OJT	366	412	451	403	413
Occupation-computer	373	506	537	577	542
DOT-environment	416	231	130	086	145
DOT-hazards	182	067	054	035	054
DOT-physical	1 665	1 218	1 165	1 623	1 442
DOT-strength	2 469	1 960	1 850	1 772	1 868
Industry-union	186	168	172	170	171
Industry-big firm	452	467	456	412	435
N	4,514	17,757	16,721	41,957	80,949
Means for males					
Wage (December 1993 $)	12 842	16 060	14 751	11 380	14 018
Schooling	12 342	13 501	14 097	13 449	12 995
Experience	19 914	19 528	18 644	16 519	19 422
Part-time	074	090	108	209	092
Federal	027	055	034	036	037
State	033	045	076	073	045
Local	071	051	157	131	078
Black	082	080	085	119	084
Union	248	153	205	210	209
DOT-GED	2 801	3 466	3 584	3 037	3 124
DOT-SVP	1 970	2 980	2 779	1 040	2 334
Occupation-tenure	7 337	7 231	6 771	5 525	7 120
Occupation-part-time	092	139	186	304	132
Occupation-OJT	358	416	465	391	392
Occupation-computer	282	515	572	531	408
DOT-environment	524	229	114	093	353
DOT-hazards	287	073	042	034	173
DOT-physical	2 167	1 242	1 138	1 484	1 702
DOT-strength	2 703	1 967	1 801	1 806	2 301
Industry-union	206	175	200	183	194
Industry-big firm	377	458	455	463	418
N	41,440	27,290	9,791	5,382	83,903

NOTE —All means are calculated across individuals in the 1993 CPS ORG Occupation-tenure is calculated from the May 1983 and 1988 CPS Pension Supplements and the January 1983, 1987, and 1991 CPS, occupation-OJT from the January 1983 and 1991 CPS, occupation-computer from the October 1984 and 1989 CPS, industry-big firm from the May 1983 CPS Pension Supplement and the March 1989–92 CPS, and occupation-part-time and industry-union from the CPS ORG files DOT measures are constructed from data in the *National Occupational Information Coordinating Committee Crosswalk* This provides data for approximately 12,000 occupations from the 1986 revision of the fourth edition of the *Dictionary of Occupational Titles*, plus a "crosswalk" code identifying the 1980 Census of Population occupation code We construct DOT values for each census occupation code by aggregating and taking the unweighted average (employment weights are not available) among all detailed occupations within each census occupation

Table 3
FEM Coefficients by Specification, Gender, and Year from Wage Levels Equations, 1983–93

Specification	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
Females											
Standard	– 1473	– 1583	– 1703	– 1764	– 1693	– 1683	– 1661	– 1669	– 1718	– 1646	– 1762
	(0064)	(0064)	(0065)	(0065)	(0066)	(0069)	(0066)	(0064)	(0066)	(0067)	(0068)
Expanded	– 1001	– 1095	– 1131	– 1327	– 1214	– 1194	– 1242	– 1155	– 1267	– 1133	– 1389
	(0085)	(0086)	(0086)	(0086)	(0087)	(0090)	(0085)	(0083)	(0085)	(0086)	(0087)
N	76,159	76,659	78,753	79,318	80,687	77,356	79,495	84,021	82,152	81,522	80,949
Males											
Standard	– 1502	– 1551	– 1714	– 1818	– 2194	– 1943	– 1956	– 2134	– 1989	– 1985	– 1951
	(0081)	(0081)	(0081)	(0081)	(0083)	(0086)	(0082)	(0080)	(0081)	(0082)	(0084)
Expanded	– 0887	– 1009	– 1029	– 1038	– 1364	– 0989	– 0920	– 1100	– 0877	– 1063	– 0899
	(0101)	(0100)	(0100)	(0100)	(0101)	(0106)	(0101)	(0098)	(0099)	(0100)	(0102)
N	87,855	88,234	89,486	88,116	88,233	84,447	86,254	90,378	87,235	85,330	83,903

NOTE.—The “standard” specification includes variables measured at the individual level years of schooling completed and potential experience and its square, dummies for union coverage, part-time, race (2), marital status (2), public sector (3), large metropolitan area, region (8), industry (13), and occupation (5) The “expanded” specification adds variables measuring means at the occupation and industry levels Occupation variables included are years of tenure, proportion part-time, proportion receiving on-the-job training, proportion using computers at their job, a 1–6 index of general educational development (GED), years required for occupational proficiency or specific vocational preparation (SVP), number of work environment disamenities from zero to five, the proportion in hazardous jobs, number of physical demands from zero to four, strength required measured by a one to five index from low to high Industry variables included are proportion union and proportion in firms with ≥1,000 employees Standard errors are in parentheses

7% lower for both women and men in a typical "female" occupation (FEM $\approx$ .68) than in a "typical" male occupation (FEM $\approx$ .29). Stated alternatively, a movement toward equality of gender composition across occupations (a change to a FEM of .48, the mean in the combined female and male sample for 1993) would be associated with a 3.6% increase in a typical female worker's wage and a 3.4% decrease for a typical male. Changes in Θ_f and Θ_m between 1983 and 1993 (table 3, standard specification), coupled with changes in FEM during this period (table 1), imply that changes in the level and impact of gender composition had a negligible effect in narrowing the gender gap. The difference in the effects of FEM for 1993 and 1983, calculated by

$$(\Theta_f FEM_f - \Theta_m FEM_m)_{93} - (\Theta_f FEM_f - \Theta_m FEM_m)_{83} \\ = [-.0640] - [-.0660] = .0020, \tag{6}$$

indicates the logarithmic gender gap narrowed by only .002 log points owing to changes in gender composition and its wage effects. This is less than 2% of the total .124 narrowing of the gap over the 1983–93 period (see the top line of table 7). Note that the above calculation assumes a causal relationship from gender composition to wages; it is this interpretation that we investigate below.[9]

B. Gender Composition and the Role of Job Characteristics

One explanation for the wage-composition relationship is that it partly reflects compensating wage differentials resulting from differences in job characteristics between predominantly female and male occupations. Moreover, quality sorting by workers may be not on the basis of gender composition per se but, rather, on the basis of job characteristics correlated with gender composition. The job characteristics hypothesis is tested estimating "expanded" wage regressions that include measures of occupation and industry characteristics. Measured at the occupation level are an index of general education requirements (GED), mean years of necessary occupational training (SVP), proportion of workers who report receiving train-

[9] If we use coefficients from our "expanded" model, as presented in table 3 and discussed below, the effect of the change in gender composition implied by equation (6) is .005. Although changes in gender composition had small effects on the gender wage gap during the 1980s, it did not operate in isolation from other structural changes. Blau and Kahn (1992) show that changes in industry coefficients during the 1980s benefited women relative to men, while O'Neill and Polachek (1993) find that changes in occupational returns favored males. Even and Macpherson (1993) provide estimates of the effects of declining unionization on the gender gap. Blau and Beller (1988) show that changes in occupational segregation also had relatively little effect on female-male wage trends in the 1970s.

ing on the job (OJT), mean tenure with the current firm, the proportion of part-time workers, the proportion using a computer on the job, the proportion of jobs facing hazards, and indices of physical demands, environmental conditions, and strength. Measured at the industry level are the proportions of workers in firms with at least 1,000 employees and of workers covered by union contracts.[10]

Results from the expanded model are shown in table 3. The rather clear-cut result is that the magnitude of the relationship between female or male wages and gender composition, following control for job characteristics, is reduced by roughly one-quarter for women and one-half for men. In 1993, for example, the addition of job characteristics variables causes coefficients on FEM to decline in magnitude from −.176 to −.139 among women and from −.195 to −.090 among men. At least some of the negative relationship between wages and the proportion female in an occupation is the result not of gender composition per se but of differences in job characteristics correlated with FEM. The magnitude of the wage-FEM relationship appears to be as strong or stronger in 1993 as in 1983. While we would expect FEM coefficients to be sensitive to changes in the economy's valuation of unmeasured skill and job types with which FEM is correlated, we do not probe this issue directly.

In contrast to Johnson and Solon and other studies (and results from our standard specification), the relationship between wages and gender composition is not found to be systematically stronger among men than among women, once we control for detailed job characteristics. The magnitude of the gender coefficient for women in the expanded model exceeds that for men in 10 of the 11 years from 1983 to 1993. The sharper reduction in Θ_m than in Θ_f following control for measurable skill-related job characteristics supports the thesis that occupational crowding and mobility barriers may account for much of the negative wage-FEM relationship among women, whereas measurable differences in job skills account for relatively more of the negative wage-FEM relationship among men. Subsequent longitudinal analysis will provide estimates of the extent to which FEM is a proxy for *unmeasured* worker skills and preferences.

In order to assess the appropriateness of our specification, we examine the signs and magnitudes of the coefficients on the job characteristics variables other than FEM. For ease of presentation, we analyze these (and

[10] A number of previous studies have included DOT occupational measures, and Johnson and Solon (1986) additionally include percentage part-time in their May 1978 wage-level equation. None has included measures of occupation-level tenure, computer use, or on-the-job training. Johnson and Solon report larger estimates of Θ_f and Θ_m; differences between our studies are the inclusion here of more control variables, substantially larger sample sizes, a more recent time period, and estimation of longitudinal models. We examine below the sensitivity of estimates to specification and functional form.

other) issues using a data set pooled over the 1983–93 period, with year dummies included Table A2 provides full regression results for the female and male expanded log wage equations, with occupation and industry characteristics, this corresponds to the results shown in table 4, column 1 Coefficients on earnings function variables measured at the individual level do not warrant comment, apart from the coefficients on years of schooling completed Their relatively low value (043 for women and 045 for men) is due to the inclusion of occupation dummies, GED and SVP Coefficients on the occupation and industry characteristics variables are generally consistent with theory and expectations Variables measuring GED, SVP, OJT, computer use, hazards, and proportion in large firms within the worker's industry are positively related to both male and female wages Results counter to expectations in the female wage equation are a zero coefficient on proportion part-time and small negative coefficients on environmental disamenities and proportion union Contrary to expectations in the male wage equation are small negative coefficients on mean tenure and the strength index and a zero coefficient on physical demands Although our results are largely consistent with expectations, predictions of signs are not unambiguous owing to heterogeneous tastes and sorting and because characteristics such as strength may be correlated with other unmeasured determinants of productivity (in the *change* equation, the only "wrong" sign is on an insignificant DOT-environment coefficient in the male equation)

Table 4
Gender Composition Coefficients from Linear and Dummy Variable Models, Pooled Data Set, 1983–93

Specification	Model 1 FEM	Model 2 FEM25–49	Model 2 FEM50–74	Model 2 FEM75+
Females				
No controls	− 2305	0754	0013	− 0971
	(0021)	(0026)	(0026)	(0024)
Standard	− 1651	− 0538	− 1035	− 1387
	(0020)	(0022)	(0022)	(0021)
Expanded	− 1173	− 0253	− 0653	− 0813
	(0026)	(0022)	(0022)	(0024)
N	877,070		877,070	
Males				
No controls	− 0375	1476	0224	− 1951
	(0025)	(0013)	(0020)	(0026)
Standard	− 1858	− 0030	− 0888	− 1412
	(0025)	(0013)	(0017)	(0022)
Expanded	− 0986	0198	− 0426	− 0608
	(0030)	(0013)	(0018)	(0025)
N	959,471		959,471	

NOTE —Model 1 includes FEM, while model 2 includes the dummies FEM25–49, FEM50–74, and FEM75+ The reference group is FEM < 25 Standard and expanded specifications are described in the note to table 3 All pooled models include year dummies Standard errors are in parentheses

C. Further Results

In this section, we examine the issues of functional form, specification, differences in the effects of gender composition among alternative groups of workers, and estimation of Θ from a data set containing individual data on tenure and firm size.[11] We first examine the linearity of the log wage FEM relationship by estimating specifications that replace FEM with the dummy variables FEM25–49, FEM50–74, and FEM75+, corresponding to the designated range of FEM (the omitted base category is FEM < .25). In table 4, we provide estimates for a model with no control variables, for the standard model, and the expanded model including job characteristics. The results with no controls indicate a U-shaped relationship between log wages and FEM, with wages lowest in occupations with low and high proportions of women. The low average wages in predominantly male occupations reflect the low skill requirements in many of these jobs. Once individual characteristics are included, the relationship becomes closer to linear, with coefficients on FEM25–49, FEM50–74, and FEM75+ of approximately −.05, −.10, and −.14 among women and −.00, −.09, and −.14 among men. When job-level characteristics are added, the coefficients on the categorical dummies change to −.03, −.07, and −.08 among women and .02, −.04, and −.06 among men. In subsequent work, we restrict our analysis to the linear specification.

Table 5 presents the coefficients on FEM obtained from alternative specifications. We include first the equation with no controls, followed by a "base" specification (line 2) including all variables measured at the individual level, with the exception of broad occupation and industry dummy variables. We then add industry and occupation dummies (separately and jointly) to the base model and obtain the "standard" specification (line 5) shown previously. We then add to the standard specification, separately and jointly, the occupational variables GED, SVP, mean tenure, proportion part-time, proportion OJT, and proportion with computer use (lines 6–12). We also add to the standard model all DOT occupation measures other than SVP and GED (line 13) and the industry measures of firm size and union density (line 14). We then present the expanded model (line 15).

[11] An additional issue examined is whether an effective wage floor associated with minimum wage laws or binding reservation wages flattens the wage-FEM gradient, in particular for women. If a significant number of workers are in jobs with wages close to the minimum, there can be relatively little negative effect on wages from increases in the proportion female (for a similar analysis with respect to racial composition, see Hirsch and Macpherson 1994). To address this issue, we created a new sample restricted to workers with wages at least 1.2 times the minimum wage ($3.35 through March 1990, $3.80 from April 1990 through March 1991, and $4.25 after April 1991). This sample restriction resulted in the deletion of 8.8% of the men and 16.7% of the women. Regression results for workers above the cutoff were highly similar to those shown in the article.

Table 5
FEM Coefficient Sensitivity to Specification, Pooled Data Set, 1983–93

Specifications	Females	Males
1 No controls	– 2305	– 0375
2 Base (individual characteristics only)	– 1719	– 1387
3 Base + 13 industry dummies	– 1775	– 0512
4 Base + 5 occupation dummies	– 1543	– 2584
5 Standard model (base model + 5 occupation, 13 industry dummies)	– 1651	– 1858
6 Standard + GED	– 1356	– 1031
7 Standard + SVP	– 0886	– 0935
8 Standard + job tenure	– 1242	– 1579
9 Standard + part-time	– 1059	– 0930
10 Standard + OJT	– 1476	– 1489
11 Standard + computer	– 1620	– 1924
12 Standard + SVP, GED, tenure, part-time, OJT, computer	– 1192	– 0824
13 Standard + DOT environment, hazards, physical, strength	– 1910	– 2326
14 Standard + industry firm size, union	– 1561	– 1997
15 Expanded (standard + all job characteristics)	– 1173	– 0986
16 Expanded minus physical	– 0814	– 0986
17 Expanded, with job characteristics measured gender specific	– 1279	– 1135
18 Expanded, with 49 rather than 13 industry dummies	– 1170	– 1064
N	877,070	959,471

NOTE —Shown are the regression coefficients Θ_f and Θ_m The standard and expanded specifications are described in the text and the note to table 3 Specifications 5 and 15 correspond to standard and expanded models shown in table 4 All models include year dummies Standard errors are approximately 0020– 0025

and the expanded model minus the DOT index of physical requirements (lines 16) Additional specifications shown are one in which all occupation-level variables (other than FEM and the DOT variables) have been calculated on a gender-specific basis (line 17) and one including 49 rather than 13 industry dummies in the expanded specification (line 18)

The principal conclusion to be reached from the results in table 5 is that estimates of the wage-FEM relationship are sensitive to specification Several results are notable For women, the inclusion of occupation and/or industry dummies has virtually no effect on the FEM coefficient By contrast, among males inclusion of industry dummies sharply decreases the magnitude of the FEM coefficient, while inclusion of only five occupation dummies sharply increases its magnitude (lines 3 and 4) These results indicate that the negative wage-FEM correlation for males occurs primarily *within* broad occupation groups, that mean FEM and wages are positively correlated across these broad occupational groups (the simple correlation of mean wages and FEM among our six occupational categories is 293 for men), and that much of the negative FEM-wage correlation among men can be accounted for by industry differences. As pointed out by Johnson and Solon (1986), industry wage differentials are likely to be little affected by comparable worth policies, making such policies relatively less effective. We find virtually identical results, regardless of specification, when a more detailed set of 49 industry dummies is substituted for the 13 industry dummies (i.e., compare lines 15 and 18).

The separate inclusion of human capital and job attachment variables measured at the occupational level is found to reduce substantially estimates of Θ. The addition of SVP, measuring average years training required for occupational proficiency, has a large effect on the magnitude of the FEM coefficients, changing Θ_f from −.17 to −.09 and Θ_m from −.19 to −.09.[12] The effect of proportion part-time is to reduce the FEM coefficient to −.11 for women and −.09 for men. This variable is likely to provide a measure of job attachment and average work hours, which in turn are positively related to training and earnings.[13] Inclusion of mean tenure reduces Θ to −.12 and −.16 for women and men, while inclusion of GED leads to coefficients of −.14 and −.10 for women and men. The addition of OJT has a more moderate effect, while addition of the proportion using computers has little effect on either estimate of Θ (women have moderately higher computer use on the job than do men). The joint addition of SVP, OJT, GED, computer, job tenure, and part-time to the standard model (line 12) produces estimates of $\Theta_f = -.12$ and $\Theta_m = -.08$. These results indicate that a sizable portion of the negative wage-FEM relationship, roughly a quarter among women and half among men, is due to occupational differences in skill requirements and job attachment.

The inclusion of union density and the proportion of workers in large firms, both measured at the industry level, has relatively little effect on estimates of Θ_f and Θ_m (line 14). Inclusion of the DOT occupational characteristics *other than* SVP and GED—environment, hazards, physical, and strength—*increases* the magnitude of Θ for both women and men, to −.19 for women and −.23 for men when added to the standard model (line 13). These results indicate rather clearly that it is *not* differences in occupational working conditions that lead to a negative wage-FEM relationship. In fact,

[12] Some authors have included a SVP *index* measuring variable ranges of years, rather than using the DOT conversion of this index to the number of years training. When included as an index, SVP has a smaller impact on wages and the FEM coefficient.

[13] Rebitzer and Taylor (1991) provide theory and evidence along these lines. In their model, uncertain demand leads firms to choose a mix of primary and contingent workers to perform the same jobs. Firms hire into the primary jobs workers who prefer long hours and with strong job attachment. They predict that jobs with a high percentage of part-time workers will have a large number of contingent workers and, as in our analysis, that the wages of full-time workers will vary inversely with the proportion part-time.

a surprising finding is the sensitivity of Θ_f to the inclusion of the index of physical job attributes (stooping, reaching, seeing, and climbing). Deletion of the physical index from the expanded model *reduces* the magnitude of Θ_f from $-.12$ to $-.08$ (line 16). The reason for this seemingly anomalous result is partly evident in table 2, where it is shown that predominantly female occupations have a relatively high index value of physical demands but relatively low pay. The fact that the physical index enters the female wage equation with a positive and highly significant value (table A2) indicates that it in fact belongs in a wage equation estimating the impact of FEM. More detailed analysis with components of the physical index (these results are not shown) indicates that it is the reaching and seeing components that are driving this result. Estimation of the model with categorical dummies substituted for FEM indicates that exclusion of the physical index has little effect on FEM25–49 and FEM50–74, but sharply reduces the magnitude of FEM75+. In contrast to the results for women, exclusion of the physical index in the male equation has no effect on estimates of Θ_m (lines 15 vs. 16).

Finally, we provide estimates of Θ_f and Θ_m from a specification where all occupation- and industry-level variables other than FEM and the DOT measures are calculated from the CPS separately by gender (line 17). Inclusion of gender-specific measures may be preferable where an included job characteristic variable is highly correlated with FEM but may be a poor proxy for the job attribute expected to provide a compensating differential. Inclusion of gender-specific measures increases the magnitude of the FEM coefficients by about .01–.02 log points; in the full models with gender-specific measures we obtain $\Theta_f = -.13$ and $\Theta_m = -.11$. We attach relatively less weight to these estimates because the gender-specific job characteristics are measured with greater error (sample sizes by gender are less than with the combined samples) and because it may be more appropriate to treat the occupational category as a unified labor market (as assumed in the case of FEM) than as two distinct markets.

A potentially important issue given little attention in the literature on gender composition is the possibility of differences in gender composition effects across sectors, demographic groups, and types of workers. Knowledge of such differences might allow us to ascertain the generality of our results and has the potential to enhance understanding of why occupations with relatively larger numbers of women have lower female and male wage rates. Table 6 provides coefficient estimates (standard errors are omitted) from both the standard and expanded models, disaggregated on the basis of age, education, race, private- and public-sector status, union status, part- and full-time status, and production-nonproduction status. These results provide less insight than expected. For both sexes, the effect of FEM on wages is smallest among relatively senior workers (ages 45+). This is consistent with the hypothesis that occupational crowding has been more

Table 6
Gender Composition Coefficients among Alternative Worker Groups. Wage-Level Equations, Pooled for 1983–93

	Females			Males		
Group	*N*	Standard	Expanded	*N*	Standard	Expanded
All workers	877,071	− 1651	− 1173	959,471	− 1858	− 0986
Age						
16–29	271,800	− 1596	− 1174	293,665	− 1989	1524
30–44	353,996	− 1842	− 1398	387,446	− 2117	− 1074
45–99	251,275	1117	− 0550	278,360	− 1417	− 0161
Education (in years)						
0–11	100,758	− 1591	− 0308	148,138	− 1305	− 0856
12	364,501	− 1459	− 0977	359,693	− 1314	− 0736
13–15	219,245	− 1182	− 0752	215,393	1757	0785
16	122,836	− 2172	− 1376	138,650	− 2910	− 1381
>16	69,731	2272	− 0587	97,597	− 1010	0702
Race						
White	749,327	− 1610	− 1178	844,855	− 1828	− 0942
Black	95,104	− 1658	− 0781	79,429	− 1242	− 0782
Other race	32,640	− 2098	− 1329	35,187	3071	1979
Class						
Private	698,093	− 1678	− 1284	804,611	− 1527	0792
Public	178,978	− 1252	− 0462	154,860	− 2305	− 0793
Union status						
Nonunion	735,434	− 1633	− 1193	732,407	− 1660	− 0879
Union	141,637	− 1244	− 0594	227,064	− 1884	− 0966
Hours status						
Part-time	222,074	0039	− 0605	79,135	− 1791	− 1415
Full-time	654,997	− 2091	− 1340	880,336	− 1808	− 0884
Production status						
Nonproduction	777,075	− 1462	1198	545,499	− 1921	− 0620
Production	99,996	− 2122	− 0924	413,972	− 0994	− 1032

NOTE —The standard and expanded specifications are described in the note to table 3 All models include year dummies Standard errors range from about 0020 to 0100

severe for younger cohorts among whom female labor force participation is highest, or that occupational choice among younger workers is more heavily influenced by unmeasured occupational characteristics correlated with FEM If gender discrimination and occupational barriers were the primary explanation for the wage-FEM relationship, however, we might have expected a stronger gender composition effect among more senior workers Differences in FEM coefficients with respect to education level appear erratic among women, the notable finding among men is the *positive* coefficient in the expanded model among workers with postgraduate education

Among women (but not men), wage penalties associated with gender composition are lower in the public than in the private sector and in the union as compared to the nonunion sector Implementation of pay equity policies has been most extensive in these sectors, although it is not clear the extent to which these differences are related to explicit (or implicit) pay policies designed to lessen gender wage differences. Gender composition effects on wages for women are found exclusively among full-time workers, whereas among males, coefficients on FEM are at least as large for part-time as for full-time workers. Part-time female workers receive low wages (for a given set of measured characteristics) independent of the level of FEM (i.e., Θ_f is close to zero for part-time women), whereas Θ_m is similar among part- and full-time males. Differences in Θ between production and nonproduction workers do not display a consistent pattern across specifications or gender.

We next consider problems associated with matching aggregate data on occupations to individual worker data (for a discussion of some of these issues, see Kloek 1981, Dickens and Ross 1984, and Moulton 1990). Matching grouped and individual data presents potential problems of several forms. Information from a single occupation is matched to multiple individual workers in that occupation. This has an effect similar to repeating observations, biasing downward standard errors, but not necessarily biasing coefficient estimates. A second problem is measurement error resulting from the heterogeneity of jobs within an occupational category and the resulting imperfect match between individuals' jobs and occupational values. For example, FEM may be measured accurately for the three-digit census occupation for which it is defined but may overstate or understate the proportion of women in an individual worker's more narrowly defined occupation or job. This type of measurement error may bias coefficients on the grouped variables toward zero. An additional distinction worth making is between occupational variables that proxy unmeasured *individual* characteristics and group data measuring relevant *job* characteristics. For example, the variable measuring mean tenure in an occupation is included not only to represent a relevant job characteristic affecting all workers' wages but also as a proxy for missing information on individual workers' tenure.

To provide a check on our results, we reestimate the standard and full models using a two-step estimation strategy suggested by Dickens and Ross (1984). In a first-step wage equation, only variables measured at the individual level or that vary within detailed occupations are included. We then calculate the mean of the equation's error term for each detailed occupation (this is equivalent to including detailed occupational dummies, but is computationally more efficient). Occupational wage differences calculated in the first step then become the dependent variable in a second-step regression, estimated by weighted least squares (WLS), with the square root of occupation sample sizes as weights. Included in the second-step standard equations are FEM and broad occupation dummies, while the second-step expanded model adds all other occupational variables that vary across but not within detailed occupations. We obtain estimates Θ_f (and their standard errors) of $-.1611$ $(.0216)$ and $-.1126$ $(.0246)$ for the standard and expanded models Corresponding estimates of Θ_m are − 1814 (0219) and − 1305 (0249) As expected, WLS point estimates are similar to what were obtained previously using single-step estimation, but standard errors are roughly 10 times larger than those shown previously

Finally, because the CPS does not normally contain measures of tenure on the current job, firm size, or establishment size, we also have estimated identical models using the May 1983 and 1988 CPS Pension Supplements (these results are not shown) Inclusion of individual worker tenure and size variables reduces the unexplained portion of the gender wage gap by a small amount, only 012 in both the standard and expanded specifications Coefficient estimates for Θ_f and Θ_m, however, are highly similar to those presented using the CPS ORG files

D Wage Gap Estimates and Gender Composition

Examined in this section is the sensitivity of gender wage gap estimates to the inclusion of FEM The top line of table 7 provides the total or unadjusted logarithmic gender gap by year, while lines 1a and 1b show the explained and unexplained portions of the gap based on our standard model, absent FEM The unadjusted log wage gap declined throughout the period, from 359 in 1983 to 235 in 1993 Differences in personal and labor market characteristics included in the standard model account for only a fifth of the gap, in 1993, 192 of the 235 total gap remains unexplained Although both the explained and unexplained values (lines 1a and 1b) fell over the 1983–93 period, the fraction of the total gap that is unexplained rose (for a similar finding using the Panel Study of Income Dynamics, see Sorensen 1991)

Lines 2a–2c of table 7 are based on our standard specification, with FEM added to the previous model (specification 1 of table 3) Inclusion of gender composition reduces the unexplained gap by about 05 log points, while accounting for more than half of the explained portion of the gap during most of the period (and 83% by 1993) When job characteristics are added to the model, as shown in the expanded specification on lines 3a–3d the portion accounted for by FEM is reduced by about a third The cross-sectional results throughout the 1983–93 period indicate that gender composition differences between men and women account for roughly 05 log points of the total gender wage gap, the latter declining from 36 in 1983 to 24 in 1993

Although occupational characteristics are jointly not important in explaining the gender wage gap, as opposed to their importance in accounting for the effects of FEM, several are individually important but tend to cancel out each other In line 3c of table 7 we list the contribution of several of the more important variables—SVP, proportion part-time, physical conditions, computer use, and GED—on the explained portion of the gender wage gap Differences between women and men in occupational training

Table 7
Decomposition of Gender Wage Gap, by Specification and Year

Specification	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
Total log gap	359	352	346	338	331	324	295	280	261	245	235
Standard specification without FEM											
1a Unexplained	251	252	251	250	245	243	223	215	205	195	192
1b Total explained	108	099	095	088	086	081	072	065	056	050	043
Standard specification											
2a Unexplained	208	207	202	201	191	194	175	166	158	149	147
2b Total explained	152	145	144	138	140	130	119	114	104	096	089
2c Explained due to FEM	067	070	074	078	083	077	075	078	075	073	074
Expanded specifications											
3a Unexplained	207	207	204	202	192	193	173	166	156	146	145
3b Total explained	152	145	142	137	139	131	122	114	105	098	090
3c Explained due to FEM	042	047	047	052	055	046	045	046	043	044	045
3d Explained due to all job characteristics	015	006	006	000	001	002	005	002	001	− 006	− 010
3e Explained due to selected occupational characteristics											
SVP	015	015	015	012	011	010	012	012	011	008	008
Part-time	008	006	006	008	009	011	015	012	010	006	010
Physical	004	005	004	003	004	004	005	006	005	005	007
Computer	− 011	− 014	− 016	− 017	− 018	− 018	− 015	− 015	− 016	− 016	− 015
GED	− 001	− 002	− 003	− 003	− 004	− 005	− 005	− 006	− 006	− 007	− 008

NOTE —The standard and expanded specifications are described in the note to table 3 The sums of "unexplained" and "total explained" may not sum to the total gap owing to rounding error The calculations are outlined in the text Decompositions are based on the use of a weighted average of female coefficient and male coefficient "multipliers" in the explained portion of the calculation, with weights being the sample proportions of women and men requirements, as measured by SVP, account for about 1.5 percentage points of the gender gap during the early 1980s, but decline to below 1 percentage point by 1993. The proportion part-time can account for roughly 1 percentage point of the gap, while differences in physical requirements accounts for another .5 percentage point. Working in the opposite direction are computer use and GED. Differences in computer use should be associated with about a 1.5 percentage point female wage advantage, and occupational education requirements another .5 point advantage.

V. Longitudinal Analysis of Gender Composition and Wages

A. Results

An important contribution of this study is the estimation of longitudinal wage change models that control for unobserved fixed effects in measuring the relationship between wages and gender composition. If women and men with higher unmeasured skills are more likely to be sorted into predominantly male jobs, and those with lower productivity into predominantly female jobs, then the coefficient on FFM in a longitudinal wage change model should move toward zero (as noted previously, past or present discrimination can produce the sorting pattern described here). That is, workers whose unobserved quality remains constant over a 1-year period would exhibit a relatively small wage change due to a change in the gender composition of a job (this assumes that wage losses associated with firm- and occupation-specific skills are uncorrelated with changes in FFM). Although time in the new job is relatively brief in our panel, it is reasonable to expect most wage effects of gender composition to be attached to the occupation and therefore show up quickly in the new wage.

Using similar reasoning, longitudinal models can account for unobserved taste differences correlated with gender composition and unmeasured job characteristics. For example, workers who place a high weight on jobs with flexible schedules and attractive working conditions are likely to be observed in jobs with lower wage rates. If such workers change occupation (and thus FFM), they are likely to have relatively low wages (conditional on measured characteristics) in both jobs. When the wage equation is estimated in levels form, the coefficient on FFM will be negative if gender composition is correlated with worker preferences of this sort, whereas estimation of a longitudinal wage equation will control for the effects of *unmeasured* job characteristics and tastes, to the extent that they remain fixed among occupational switchers.

Table 8 presents results from both levels and longitudinal wage change equations, estimated using the pooled panel data set constructed from the CPS ORG files and consisting of matched worker pairs for the period 1983/4–1992/3. The levels estimates correspond exactly to the previously estimated standard model (individual characteristics plus broad occupation

Table 8
Panel Data Estimates of FEM and ΔFEM Coefficients for Wage-Level and Wage Change Models, Using the CPS ORG, March CPS, and CPS Displaced Worker Surveys

	Females		Males	
	Levels	Change	Levels	Change
CPS ORG panel				
Standard	− 1632	− 0917	− 1783	− 0852
	(0039)	(0066)	(0049)	(0078)
Expanded	− 1143	− 0549	− 0798	− 0336
	(0050)	(0077)	(0061)	(0088)
N	219,323		240,362	
March CPS panel				
Standard	− 1709	− 1050	− 1617	− 0833
	(0037)	(0112)	(0047)	(0129)
Expanded	− 1223	− 0361	− 0700	− 0223
	(0049)	(0143)	(0058)	(0155)
N	238,299		262,918	
DWS − plant closings and layoffs sample				
Standard	− 2021	− 0913	− 2336	− 0496
	(0309)	(0270)	(0311)	(0279)
Expanded	− 0610	0024	− 0911	− 0326
	(0428)	(0381)	(0381)	(0340)
N	3,883		7,037	
DWS − plant closings only sample				
Standard	− 2289	− 1061	− 1904	− 0312
	(0415)	(0354)	(0443)	(0406)
Expanded	− 0801	− 0483	− 0598	− 0491
	(0582)	(0515)	(0543)	(0488)
N	2,200		3,595	

NOTE—The standard and expanded specifications are described in the note to table 3 All models include year dummies Standard errors are in parentheses Change equations have Δ log(W) as the dependent variable, and coefficients on ΔFEM are presented Change equations from the March CPS files differ in specification from the CPS ORG by the inclusion of changes in region and public-sector status but exclusion of changes in marital status, union status, and in federal, state, and local worker status The DWS results are based on the January 1984, 1986, 1988, 1990, and 1992 CPS Displaced Worker Surveys The sample consists of workers who were age 20 and older and who were displaced from a full-time, private-sector job because of a plant closing, slack work, or a position of shift that was eliminated The sample was further restricted to workers who were reemployed at the survey date in a full-time wage and salary job and excluded those displaced from the construction industry The dependent variable in the change equation is the difference between the log of current weekly earnings and the log of predisplacement weekly earnings In addition to the usual change variables, panel estimates include dummies for year of displacement and survey year

and industry dummies) and expanded model (the standard model plus job characteristics) except that the sample in table 8 consists of matched individuals from the CPS panel during their second year in the survey (1984–93) The panel does not include individuals not employed in adjacent years, those residing in a different households or geographic locations in adjacent years, and those for whom unique matches could not be made (see the appendix) Most likely to be excluded from the panel are young workers Despite these differences, levels results here are highly similar to those obtained previously from the full CPS ORG sample We obtain estimates of Θ_f and Θ_m of − 16 to − 18 in the standard model, while in the expanded model Θ_f is estimated to be $-.11$ and Θ_m to be $-.08$. The similarity in levels estimates between the panel and full samples suggest that the subsequent longitudinal results can be generalized to a representative national sample of female and male workers.

The dependent variable in the longitudinal wage change models is $\Delta \ln W$; longitudinal estimates of Θ_f and Θ_m are based on the coefficients of ΔFEM. For reasons examined below, we present results here based on occupation changers who also report changing industry over the year (more precisely, ΔFEM times a dummy equal to one if both occupation and industry change). Workers recorded as changing occupation but not industry have separate estimates of the ΔFEM variable (not shown in table 8), while workers not recording an occupational change are the reference group.[14] In addition to ΔFEM, whose coefficient is shown in table 8, the standard change model includes control variables measuring changes in experience squared, part-time status, public sector status (federal, state, and local), union coverage status, marital status, broad occupation and industry, and FEM for those changing occupation only (as explained above), plus nine period dummies. The change in experience is one for all workers and thus reflected in the intercept; variables are not included for changes in schooling (because persons whose principal activity was schooling were excluded), race, large metropolitan area, or region (because households moving drop out of the CPS and cannot be matched). The expanded model includes variables measuring changes in the job characteristics.

The estimates reported in the first two lines of table 8 indicate clearly that the gender composition effects reflect in part unmeasured worker-specific skills and/or preferences correlated with FEM. For women, the estimate of Θ_f in the standard model drops in magnitude from $-.16$ to $-.09$, while for men Θ_m changes from $-.18$ to $-.09$. These results support the hypothesis that more productive women and men sort into or are selected for higher-paid "male jobs," while less able workers are more likely to work in occupations that have a higher proportion of women. Alternatively, the reduction in estimates of Θ_f and Θ_m may reflect worker taste differences regarding unmeasured job amenities and disamenities correlated with FEM. We should emphasize, however, that FEM remains a significant and nontrivial determinant of wage rates, even after controlling for unmeasured person-specific differences.

Once we control for both measured job characteristics and unmeasured worker-specific effects in the expanded longitudinal model, however, the

[14] Because FEM changes over time within occupations, nonswitchers realize very small changes in gender composition. For convenience, we do not include a separate variable equal to a nonswitching dummy times ΔFEM; rather, wage change for this group is reflected in the equation intercept. Results with this variable included are virtually identical to those shown.

gender composition effect becomes rather small, − 05 for women and − 03 for men That is, job characteristics and unmeasured skills and tastes account for roughly two-thirds of the standard gender composition effect among women (i e , the change from 163 to − 055) and roughly four-fifths of the effect among men (the change from − 178 to − 034) A coefficient of, say, − 05 suggests that differences in the gender composition of occupations can account directly for only 02 log points (05 × 40, where 40 is the difference in mean FEM between women and men) of the sizable gender wage gap, which averaged roughly 30 over the 1983–93 period Our panel results indicate that gender composition has a relatively small direct or causal effect on wages Rather, FEM is correlated with differences in job characteristics, worker-specific productivity differences among observationally equivalent workers, and taste differences regarding job characteristics These factors in turn produce labor market sorting such that wages and FEM are negatively correlated

B Measurement Error and Alternative Estimates

An important concern in the longitudinal analysis is possible measurement error in the explanatory change variables Measurement error biases toward zero regression coefficients, and this downward bias is most severe in models where intertemporal variance owing to measurement error is large relative to true variance of a right-hand-side variable (see, e g , Freeman 1984) Measurement error bias may be particularly serious where a substantial number of persons have their occupation misclassified (Mellow and Sider 1983) and where the time period is sufficiently short that there are few true occupational changers

Although we expect bias from measurement error to be present in our analysis, several points are in order Our sample has excluded all worker-year pairs where occupation or industry has been allocated by the census in either the first or second year Second, there exists serial correlation in response error (for related evidence on earnings, see Bound and Krueger, 1991), so that a respondent who reports the incorrect occupational category in year 1 may report the same category in year 2 In this case, "two wrongs make a right" because the person would be classified correctly as having ΔFEM = 0 Third, even among those workers misclassified by occupation, it is likely that they have recorded a closely related occupation whose FEM may not be too different from actual FEM (we present evidence below) This is in contrast to the frequently discussed case of mismeasurement of union status, where workers are assigned values of zero or one Finally, we can gauge the seriousness of measurement error by comparing our results with those from other data sets where occupational change is unlikely to be measured with significant error

We consider several pieces of evidence Our most important control for measurement error in the longitudinal analysis is to present only coefficient estimates on ΔFEM for workers who are recorded as changing both occupation and industry, because workers who report changes both in industry and occupation are less likely to have remained in the same job than those workers recorded with changes only in occupation Thus, estimates of **Θ** should be less affected by measurement error for industry movers than for industry stayers In order to illustrate the importance of this distinction, we provide in table 9 results from longitudinal equations with alternative treatment of ΔFEM In addition to providing the coefficients on ΔFEM for those changing occupation and industry (col 1), we present ΔFEM coefficients for occupation switchers who did not change industry (col 2) and coefficients for all occupation switchers with no distinction based on industry change (col 3) Results are consistent with the expectation that measurement error biases coefficients toward zero Indeed, in the standard model, our preferred coefficients for occupation and industry switchers are roughly three times larger (in absolute value) than corresponding estimates for occupation-only changers

As argued above, those reporting a change in occupation in the CPS are likely to realize a change in gender composition smaller than if their new occupation were randomly misreported Evidence supports this proposition The average absolute value of ΔFEM for those who change both occupation and industry is 23, for those reporting only a change in occupation, the corresponding value is 20 By contrast, if these same persons are randomly assigned a CPS occupation (an equal probability is assigned to each eligible occupation), the mean absolute value of ΔFEM is 36 for both samples Recorded occupations for workers falsely categorized as changing occupation are clearly not selected randomly And the relatively

Table 9
ΔFEM Coefficients Based on Alternative Definitions of Occupational Switching

	Matched CPS ORG, 1983–93		
Specification	Occupation and Industry Changers	Occupation-Only Changers	All Workers
Females			
Standard	− 0917	− 0271	− 0549
	(0066)	(0059)	(0046)
Expanded	− 0549	0142	− 0161
	(0077)	(0071)	(0061)
Males			
Standard	0852	− 0282	0538
	(0078)	(0072)	(0056)
Expanded	− 0336	0234	− 0018
	(0088)	(0082)	(0069)

NOTE —The standard and expanded specifications are described in the note to table 3 Standard errors are in parentheses large absolute value of ΔFEM among designated switchers (.23) suggests that it provides a high ratio of signal to noise.

In order to explore further the degree of measurement error owing to misreported occupational changes, we used supplementary information from the January 1987 CPS public-use survey, which explicitly asked individuals in all rotation groups whether they had changed occupations during the previous year and, if so, their previous occupation and industry. Individuals from rotation groups 5–8 in January 1987 are matched with their previous responses in outrotation groups 1–4 in our January 1986 sample (we use the same sample restrictions as in our prior analysis). We first use the January 1986 and January 1987 occupation and industry codes to calculate (as previously done) the number who change reported occupation only and the number who change both occupation and industry. We then use information from the January 1987 CPS to see which workers explicitly say they changed occupation during the previous year. For purposes of exposition, we refer to this latter group as "verified" changers. Of those measured as changing reported occupation but not industry ($N = 4{,}116$), just 7.2% are found to be verified changers, as measured by the January 1987 CPS. But among those reporting both occupation and industry change between 1986 and 1987 ($N = 2{,}673$), 28.3% are verified changers. These results strongly support our decision to use coefficient estimates only for those who change both occupation and industry in order to reduce measurement error in the ΔFEM variable. But they also suggest that substantial measurement error remains even in this measure.

In order to further assess the generality of the longitudinal results from the constructed CPS ORG panels, we estimate similar wage level and wage change models using two alternative data sets. The first is a data set constructed from the March CPS files for 1983–93. The March CPS records information not only on current occupation, earnings, and other characteristics but also on an individual's occupation, industry, and class of worker in the longest job held during the previous year, annual earnings, weeks worked, and hours worked per week the previous year, and state of residence the previous March.[15] The major advantage of the March CPS files is that occupational change is far less likely to be measured with error. The constructed CPS ORG panel previously used relies on information from two interviews, 1 year apart, possibly conducted by different individuals with different household members, and coded by different census coders. By contrast, the March file relies on information from a single

[15] Current earnings are reported by only a quarter of the March CPS (the outgoing rotation groups). We retain the full March sample by matching each worker's reported earnings from the April, May, or June ORG file to the March record. Similarly, Funkhouser (1993) has matched the April 1983 CPS immigration supplement with the April–July CPS ORG records.

interview with a single household member by a single interviewer and with a single occupation coder (when occupation does not change). Another important advantage of the March files is that they include information on individuals who have changed households or location or who could not be matched from the ORG files from separate years. Disadvantages of the March data set are that the wage change variable is constructed from two different earnings measures (calculated wages from the March retrospective questions are larger than from the ORG earnings supplements) and the previous year's earnings may be determined in part by jobs other than the longest held. Neither of these is likely to seriously bias estimates of Θ—differences in the March and ORG wage measures will be reflected largely in the constant of the wage change equation, while mismeasurement of wages owing to multiple occupations should not bias Θ if the measurement error is uncorrelated with ΔFEM. There will be a bias toward zero in estimates of Θ, however, to the extent that last year's wage reflects the wage on the new occupation (with the new FEM value), lessening the measured wage change associated with ΔFEM. That is, for those changing occupations in the latter half of the previous year rather than in the second year prior to the March survey, there will be bias in the estimate of Θ, on the order of about 15%.[16]

An additional panel data set is constructed using the January 1984, 1986, 1988, 1990, and 1992 CPS Displaced Workers Surveys (DWS). The DWS provide information on whether workers have lost or left a job during the past 5 years because of a plant closing, an employer going out of business, a layoff from which a worker was not recalled, or other similar reason. The DWS has the same advantage as the March surveys in that it is relatively unlikely that there will be a false identification of nonswitchers as switchers. The DWS has an advantage as compared to the March CPS panel in that the prior wage and occupation refer directly to the last job held. The most important differences between the DWS and either the ORG or March panels is that occupational changes extend beyond a year and job changes among displaced workers are more likely to be exogenous. And the analysis can be further limited to those affected by plant closings, because it can

[16] Assume that there is one occupational change and that changes are distributed evenly throughout a year. The mean and median switcher will change occupations during the first week in November, the midpoint between July 1 (the second half of year 1) and March 15 (the approximate date of the March survey). Annual earnings in year 1 will therefore be a weighted average of old and new occupational earnings, with a weight of about 15% on the new occupation (53 out of 365 days). Measured wage change and estimates of Θ likewise will be understated (i.e., biased toward zero) by roughly 15%. Additional disadvantages of the March data are that information is not available for the previous year on the union and marital status variables, and less information about class of worker is available prior to 1989.

be argued that some layoffs may not be completely exogenous (for similar reasoning and use of the DWS, see Gibbons and Katz 1992).

Table 8 presents three alternative sets of estimates from the standard and full specifications for wage levels and wage change models, in addition to those from the CPS ORG panels. We provide comparable estimates from the March CPS files for 1983–93 and the five January CPS DWS files for 1984–92 with separate estimates for all displaced workers and for the subset of displaced workers affected by plant closings. Despite differences in samples and specification, the results from the alternative data sets are broadly similar. In all cases, wage change equation estimates of Θ_f are about −.10 in the standard specification and from zero to −.05 following control for job characteristics. Estimates of Θ_m from the March CPS files are similar to those presented previously from the CPS ORG panel, being −.08 in the standard and −.02 in the expanded models. Estimates of Θ_m from the DWS wage change equations are only about −.04 prior to controls for job characteristics and positive (but not statistically significant) following inclusion of controls. Interestingly, the DWS estimates of Θ_m from the standard wage *level* equations are larger in absolute value than those obtained in the other data sets. Although sample sizes are small and significance levels low, the suggestion from the DWS results is that changes in gender composition owing to exogenous occupational change have little effect on wages, following control for person-specific effects and job characteristics.

Our primary concern with the CPS ORG panel results was the possibility of bias toward zero in estimates of Θ_f and Θ_m owing to measurement error in the ΔFEM variable, or bias due to endogenous occupational change. On the basis of the results in table 8, it appears that bias for these reasons is not serious. Longitudinal estimates of Θ are in fact larger in magnitude using the matched ORG files than the retrospective March CPS, where occupational change is measured far more accurately, and coefficient bias should be relatively minor (roughly 15%). Results from the DWS, which should not have such bias and which reflect exogenous occupational change, suggest that the direct effect of gender composition on wages is quite small.[17]

VI. Interpretation and Conclusions

Previous literature exploring the relationship between the gender composition of occupations and wages has emphasized the negative effect of

[17] In results not reported, we investigated whether there exist symmetric responses to increases and decreases in gender composition. Separate estimates of Θ for workers increasing and decreasing FEM produced opposite results using the March CPS from those using the ORG and DWS panels. Because we have no convincing explanation for why there should be an asymmetry in coefficients or differences across data sets, we did not further explore this issue.

proportion female on the earnings of both women and men. Past estimates based on levels estimation have not accounted for several important dimensions of worker productivity, tastes, and job characteristics. The few longitudinal studies examining this issue have been plagued by relatively small sample sizes. Little attention has been given to issues of specification or linearity of the wage-FEM relationship, bias from measurement error in wage change equations, or demographic and sectoral differences in the effects of gender composition.

This article takes advantage of large representative national samples from the January 1983–December 1993 monthly CPS surveys, as well as data on occupation and industry characteristics constructed from the *Dictionary of Occupational Titles* and various CPS supplements. The database allows us to examine changes over time in the gender composition of jobs for both women and men and its changing effect on wages and the gender wage gap. Most important, we are able to estimate longitudinal wage change models for large samples of worker-year pairs for 1983/4–1992/3. Supplementary analysis is provided using data sets constructed from the 1973–78 May CPS, the 1979–82 CPS ORG files, the March CPS for 1983–93, the five DWS for January 1984–92, the 1983 and 1988 May–June Pension Supplements, and CPS supplements containing information on tenure, firm size, job training, and computer use.

Prior to controlling for detailed job characteristics, the cross-sectional relationship between proportion female in an occupation and wages was found to be highly negative for women and men. Estimates of the effects of proportion female are substantially lower (by roughly half) using longitudinal analysis and our standard specification, indicating that person-specific labor quality or preferences account for much of the previously observed relationship. When variables measuring occupation and industry characteristics are added to wage change equations, the estimated effects of gender composition again are substantially reduced. The remaining effects of gender composition on female or male wages appear rather small. Two-thirds to all of the gender composition effect is accounted for by measured job skills and characteristics and by unmeasured worker-specific skills and preferences.

We conclude that predominantly female jobs pay lower wages to women and men largely because of their skill-related characteristics and quality sorting on FEM. This is in part the result of past and present occupational discrimination that has led to an equilibrium in which the *unmeasured* skills of women and men increase with the proportion male in an occupation. Measured job characteristics matter but are less important than unmeasured worker-specific skills and tastes. The job characteristics that are most important are related to training and job attachment. Wages decrease with respect to FEM because workers in predominantly female occupations generally require less training to acquire proficiency, and these occupations are more likely to have large numbers of part-time workers and a lower level of worker tenure. Other measurable occupation and industry characteristics, reflecting, for example, job amenities and disamenities, computer use, formal OJT, the job environment, industry unionization, and presence of large firms, have a relatively small effect on FEM coefficient estimates.

What are the implications of this study for pay equity or other public policies that aim to adjust relative wages of occupations, on the basis in part of measured worker and job characteristics? Our results indicate that the direct effect of gender composition on wages is rather small, following what we argue are appropriate controls for job characteristics and unmeasured skills and preferences. Policies that eliminate only the remaining relationship of wages and FEM would have little effect on the sizable gender wage gap. Policies that alter wage rates more substantially run the risk of distorting what appear to be legitimate compensating differentials for skills or job attributes, although such policies are likely to narrow gender wage differentials.

Although our conclusion regarding the desirability of comparable worth policies is similar to that of Johnson and Solon (1986), the reasoning differs. They conclude that comparable worth policies, which are typically implemented within firms, would be ineffective in eliminating wage differences correlated with gender composition because these are highly correlated with industry wage differentials not subject to direct change via pay equity policies (we provide similar evidence for males). We conclude that occupational wage differences correlated with gender, following use of appropriate controls, are sufficiently small that they should not be a major focus of public policy. Stated alternatively, after controlling for gender composition, as well as personal and job characteristics, much of the still sizable female-male wage gap remains unexplained. It may be more appropriate for public policy to focus on causes of the gender wage gap per se, rather than on occupational wage differences correlated with gender composition. It is worth recalling that the 1980s and early 1990s have witnessed a substantial narrowing of the gender gap and improvement in relative female earnings, and this narrowing has not been the result of changes in gender composition and its estimated effects.

While gender composition per se may be relatively unimportant as a proximate cause of low wage rates, occupational characteristics and worker skills and preferences correlated with gender composition are important. Narrowing of the gender wage gap will result if there continues to be narrowing of differences in experience, industry structure, and occupation characteristics between predominantly female and male jobs. Far more difficult to evaluate is why a labor market equilibrium has arisen in which predominantly female occupations have become associated with job characteristics and worker endowments leading to lower pay. Among the possible (nonmutually exclusive) explanations are historical patterns of occupational discrimination that led to the current sorting equilibrium, current discriminatory barriers to women in occupations with more job training, longer tenure, and other attributes associated with higher pay, and a sexual division of labor such that relatively many women and few men choose jobs associated with lower-paying characteristics.

Although beyond the scope of this article, limited evidence suggests that all three factors are important. Historical patterns of sex discrimination have had important effects on the gender composition of occupations. Both formal and informal occupational barriers to women were commonplace, although the incidence and survival of these barriers showed some sensitivity to their economic costs.[18] Moreover, current differences in the occupational choices of, and division of labor between, women and men are influenced heavily by the past. Less certain is the extent to which occupational choices by women are constrained currently by barriers in the labor market. O'Neill (1983) estimates equations with gender composition as the dependent variable and concludes that current occupational segregation results to no small degree from preferences expressed at early ages, prior to entry into the labor market. Because preferences are determined in part by perceptions of the present and expectations of what is possible in the future, interpretation of such evidence is not unambiguous.

More direct tests of occupational barriers are less readily available. Gupta (1993) uses information in the NLSY on job aspirations of young women and men in 1979 and broad occupational level achieved in 1982 ("female" occupations plus three broad groups including nonfemale occupations). She concludes that gender differences in occupational attainment result from both differences in preferences and employer selection. Given their occupational choice, women were less likely to be chosen from the queue in professional/managerial and service occupations. Research by Padavic (1992) on the attitudes of female clerical workers following their temporary transfer to "male" production jobs during a strike likewise indicates that both differences in preferences and employment constraints affect the occupational distribution.[19]

[18] See, e.g., Kossoudji and Dresser (1992), who provide an analysis of the rise and fall of female industrial employment during the 1940s, and Goldin (1990), who examines the economic and noneconomic determinants of female wages and employment during much of this century.

[19] Padavic (1992) provides an analysis of a natural experiment reminiscent of the World War II experience. During a union strike at a utility company, nonunion female clerical and administrative workers, in addition to nonunion male workers, were moved into predominantly male blue-collar jobs at eight plants. The production jobs paid significantly higher wages but had less attractive working conditions (less flexibility, less cleanliness, less socializing, and more extensive physical demands). Following the conclusion of the strike, workers were returned to their previous

Finally, our analysis has implications for wage equation specification and interpretation The inclusion of the proportion female variable in cross-sectional female and male wage equations can be justified on statistical grounds Absent detailed controls for job characteristics and person-specific skills and tastes, gender composition is an important correlate of wages It is important, however, that researchers recognize that the direct effect of gender composition on wages is small, rather, FEM serves as a proxy for unmeasured skills, preferences, and job attributes Progress in understanding gender differences in the labor market is unlikely to be enhanced significantly by further emphasis on pay differences between women's and men's jobs, per se More promising will be renewed attention given to how and why the labor market sorts women and men into jobs with different characteristics and productivities and a continuing investigation into the sources of a wage gap that remains sizable, independent of the gender composition of jobs

Appendix

Construction of Longitudinal Samples from the CPS ORG Files and the March CPS

Households are included in the CPS for 8 months—4 consecutive months in the survey, followed by 8 months out, followed by 4 months in Outgoing rotation groups 4 and 8 are asked earnings supplement questions (weekly earnings, hours, union status, etc) The CPS contains household identification numbers (ID) and record line numbers, but not individual identifiers Individuals can be identified potentially for the same month in consecutive years, that is, individuals in rotation 4 in year 1 can be matched to individuals in rotation 8 in year 2

The longitudinal ORG file was created in the following manner Separate data files were created for males and females and for pairs of years (rotation 4/1983 and rotation 8/1984, rotation 4/1984 and rotation 8/1985, etc) Within each file, individuals were sorted as appropriate on the basis of ascending and descending household ID, year, and age To be considered an acceptable matched pair, a rotation 8 individual had to be matched with a rotation 4 individual with identical household ID,

jobs Padavic administered a questionnaire both to the women who had taken the production jobs and to those who had not been transferred, asking if they would like to be transferred to the production job Despite the higher pay, most women did not want to be transferred to the production job Evidence in the study supports the view that individual choice, job characteristics, and childhood activities were important determinants of occupational choice A limitation of this study is that women in clerical jobs have selected those jobs on the basis in part of job characteristics, the ideal experiment would shift a representative sample of female (and male) workers to the production jobs and then administer a similar set of questions to both groups We should note that our interpretation of the Padavic study is not identical to that of the author identical survey month, and an age difference between zero and two (as surveys can occur on different days of the month, age change need not equal one). Several passes were necessary because a single household may contain more than one male or female pair. Checks were provided to insure that only unique matches were selected. For each rotation 8 individual, the search was made through all rotation 4 individuals with the same ID to make sure there was only one possible match; the file was resorted in reverse order, and each selected rotation 4 individual was checked to insure a unique rotation 8 match. As uniquely matched pairs were identified, they were removed from the work file. Incorrect changes in the variables marital status, veteran status, race, and education (e.g., a change in schooling other than zero or one, a change from married to never married, etc.) were used to delete "bad" observations in households where there were multiple observations and ages too close to separate matched pairs. Several passes at the data were made. In households where two pairs of individuals could be separated based on a 1-year change but not the 0–2-year age change, a 1-year criterion was used. If a unique pair could not be identified based on these criteria, they were not included in the data set (e.g., four observations with two identical pairs, or three individuals with two possible matches using the 0–2-year age change criterion).

The match rate in the longitudinal analysis is just under two-thirds of employed wage and salary workers in any year. The principal reasons that matches cannot be made are if a household moves (thus changing the household ID), if an individual moves out of a household, if a worker becomes self-employed, if an individual drops out of the labor market or fails to meet other sample selection criteria, or if the census is unable to reinterview a household and/or receive information on the individual. Note that the match rate reported here is similar to the match rate of 68.8% for the 1987–88 CPS reported by Card (1992) using a broader-based sample and a less stringent probabilistic matching algorithm obtained from the Bureau of Labor Statistics. Peracchi and Welch (in press) analyze attrition rates among matched March CPS files and conclude that age is the most important determinant of a successful match. Other factors that lessen match probabilities are poor health, low schooling, and not a household head, while gender and race are unimportant match predictors following control for other factors.

The sample size of the CPS ORG panel is 25% that of the full CPS ORG. The difference can be approximated as follows. Because the unit of observation in the panel is the pair of observations in adjacent years, the potential sample size is initially cut in half. Because we do not match the half of the 1983 sample that entered the CPS in 1982, or the half of the 1993 sample that exits the CPS following their 1994 interview, the potential longitudinal sample is reduced further by 9% (i.e., 91% of the full 1983–93 sample is in for 2 years). A 100% match rate of workers between adjacent years would produce a sample for each period half as large as the corresponding cross sections. We achieve a match rate of about 64%, owing to individuals who changed households, who entered or exited employment between years, or who could not be reliably matched based on available information Finally, sample sizes are reduced further to roughly half the normal size for the 1984/5 panel and to one-quarter for 1985/6 This is the result of a CPS test sample from July–September 1985 that implemented new population weights Rotation 4 households interviewed in July 1984–September 1985 were not reinterviewed a year later in 1985 and 1986 Hence, the combined multiplier relating the size of the longitudinal sample to the initial full sample is about 25, that is, 50 × 91 × 64 × 875 = 25, where 875 is the ratio (10 − 1 25)/10, with the numerator representing the remaining panels following the "loss" of 1 25 panels for 1984/5 and 1985/6

The March CPS longitudinal file is a retrospective panel All rotation groups in March are asked for information about earnings, weeks worked, hours worked last year, and occupation and industry on the longest job held last year A quarter sample in March (the ORGs) are asked current earnings, hours, and so forth The entire March sample is matched to their earnings supplement records in their outgoing month, in March, April, May, or June These records were matched initially on the basis of household ID and line number, followed by checks on changes in gender and age to insure an accurate match The March retrospective panel is 76 4% the size of a March sample based on the presence of earnings last year (and other typical variables) Losses are due to households moving, individuals leaving the household, changing employment status (i e, leaving the labor force or shifting to self-employment), changing line number, failing to be reinterviewed, and missing hours or weekly earnings in the earnings supplement among employed wage and salary workers who are otherwise matched

Table A1
Female/Male Wage Ratios and Duncan Segregation Index, by Worker Group and Year, 1973–92

Group	1973–74	1975–76	1977–78	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
Female to male wage ratio																		
Age																		
16–29	766	783	766	769	781	785	796	819	812	825	832	839	845	855	881	893	904	919
30–44	618	607	614	626	633	647	654	662	674	673	681	700	702	731	736	754	775	779
45+	597	610	591	595	601	593	596	583	590	595	601	593	608	630	639	657	665	665
Education (in years)																		
<12	617	645	635	646	652	666	659	667	665	664	689	680	700	702	727	754	763	756
12	643	648	643	650	657	662	673	675	682	689	699	701	706	718	728	746	755	757
13–15	675	657	659	671	682	684	695	700	701	710	710	715	731	747	751	762	784	787
16	667	642	634	640	648	639	640	648	658	654	661	672	675	714	735	743	751	746
>16	731	759	737	719	718	720	709	704	721	718	714	723	714	750	739	742	767	771
Race																		
White	642	645	635	641	649	653	657	661	668	672	678	683	690	714	724	738	753	754
Black	717	766	795	771	787	775	790	788	800	810	823	826	819	858	871	886	900	897
Other race	716	705	698	725	715	732	715	706	718	732	726	709	738	779	793	795	782	808
Class																		
Private	613	619	613	652	660	663	666	67	679	683	689	693	700	725	734	750	763	765
Public	745	738	729	681	681	686	702	681	696	659	708	723	710	721	756	758	777	777
Union status																		
Nonunion	648	645	635					658	667	670	674	678	687	711	719	735	749	750
Union	712	746	738					755	762	772	789	796	794	823	848	844	853	860
Hours status																		
Part-time	528	569	606	804	834	835	845	841	877	844	906	870	901	931	912	982	979	893
Full-time	657	667	661	662	669	676	681	685	697	701	704	709	713	739	749	762	775	785
Production status																		
Nonproduction	612	612	615	616	627	627	629	630	632	633	635	635	640	669	677	688	702	702
Production	618	625	619	636	635	647	654	654	654	658	666	670	673	681	675	700	698	700
Duncan segregation index																		

Age																		
16–29	682	670	643	627	622	620	610	608	604	594	589	582	575	565	563	566	560	550
30-40	699	683	667	640	635	619	615	594	586	583	577	572	563	559	558	556	556	548
45+	700	704	685	675	663	661	664	643	644	625	628	623	616	605	596	599	594	591
Education (in years)																		
<12	692	685	668	652	647	653	656	635	630	626	620	612	604	605	588	596	595	583
12	728	718	705	691	687	684	677	660	659	651	652	645	638	637	629	630	625	620
13–15	706	692	673	646	652	638	635	624	617	608	604	608	603	588	592	594	590	583
16	671	661	606	592	571	560	546	536	527	520	519	509	495	477	476	478	473	464
>16	533	558	520	490	473	474	486	472	463	448	450	444	444	431	437	431	442	451
Race																		
White	688	680	657	641	634	629	621	606	601	591	588	585	574	564	560	561	558	553
Black	699	666	657	636	615	618	621	597	591	588	582	566	566	569	559	571	568	537
Other race	751	697	604	613	599	586	597	572	554	550	563	536	544	554	541	519	509	519
Class																		
Private	697	685	665	639	632	627	621	604	599	590	586	581	574	566	561	562	558	551
Public	636	637	608	564	553	537	551	505	502	504	492	472	470	443	453	456	431	429
Union status																		
Nonunion	688	677	662					599	594	583	580	572	564	554	549	550	546	538
Union	673	672	640					625	622	613	615	608	613	605	606	610	609	604
Hours status																		
Part-time	665	650	628	598	612	595	592	584	589	586	580	574	571	549	542	536	529	530
Full-time	677	667	640	626	617	611	605	587	581	569	569	564	555	545	541	546	543	537
Production status																		
Nonproduction	668	652	628	607	593	588	583	564	561	546	544	536	529	518	511	510	505	500
Production	670	663	652	633	625	617	610	603	595	588	587	587	582	586	578	575	570	567

NOTE —Figures were calculated from the 1973–78 May CPS and the 1979–93 CPS ORG files (N = 2,749, 246) The sample includes wage and salary workers ages 16 and over, with the exclusion of earners whose principal activity is school Further description of the sample is in the text The female to male wage ratio is the mean of female real wages to male real wages Wage rates are defined as usual weekly earnings (1993 \$) divided by usual hours worked per week The Duncan segregation index is calculated by $\frac{1}{2}\Sigma|m_j - f_j|$, where m and f are the proportions of male and female employment is cooupation j

Table A2
Regression Coefficients, Expanded Wage Model, 1983–93 CPS ORG

Variables	Females	Males
FEM	− 1173 (0026)	− 0986 (0030)
Individual characteristics		
Schooling	0432 (0002)	0455 (0002)
Experience	0151 (0001)	0257 (0001)
Experience^2/100	− 0249 (0002)	− 0398 (0002)
Union	1485 (0013)	1570 (0012)
Part-time	− 0967 (0011)	− 1600 (0017)
Married with spouse	0344 (0012)	1223 (0013)
Other ever married	0321 (0014)	0682 (0018)
Black	− 0441 (0014)	− 1113 (0016)
Other race	− 0302 (0022)	− 0791 (0023)
Hispanic	− 0543 (0019)	− 1106 (0019)
Federal	0329 (0028)	0119 (0027)
State	− 0433 (0023)	− 0681 (0025)
Local	− 0668 (0018)	− 0841 (0021)
Large metropolitan area	1271 (0010)	1280 (0011)
Job characteristics		
DOT-GED	0531 (0013)	0225 (0013)
DOT-SVP	0135 (0007)	0297 (0006)
Occupation-tenure	0058 (0003)	− 0035 (0003)
Occupation-part-time	0099 (0047)	− 1610 (0055)
Occupation-OJT	3006 (0049)	1325 (0048)
Occupation-computer	1485 (0031)	1809 (0034)
DOT-environment	− 0509 (0027)	0213 (0015)
DOT-hazards	0686 (0046)	0188 (0030)
DOT-physical	0408 (0009)	− 0000 (0009)
DOT-strength	0181 (0015)	− 0251 (0015)
Industry-union	− 0456 (0043)	0475 (0042)
Industry-big firm	1772 (0024)	1858 (0027)
Region (8)	yes	yes
Industry (13)	yes	yes
Occupation (5)	yes	yes
Year (10)	yes	yes
R^2	4495	4894
N	877,070	959,471

NOTE —The mean of the log wage is 2 1884 for women and 2 4941 for men Standard errors are in parentheses The omitted reference group is full time, non-union, white, non Hispanic, never married, private-sector worker in the Northeast and not in a large metropolitan area, professional or managerial occupation, agricultural sector, in 1983 Variables preceded by "occupation" and "industry" are means of variables in workers' designated occupation and industry Occupation-tenure is calculated from the 1983 and 1988 May CPS Pension Supplements and the January 1983, 1987, and 1991 CPS surveys, occupation OJT from the January 1983 and 1991 CPS, and occupation-computer from the October 1984 and 1989 CPS Industry-big firm (proportion in firms with 1,000+ workers) is calculated from the 1983 May CPS Pension Supplement and the March 1989–92 CPS surveys All other "occupation" and "industry" variables are calculated from the 1983–93 CPS ORG files DOT measures are taken from the *Dictionary of Occupational Titles,* matched at the census occupation level DOT-SVP is years required for occupational proficiency or specific vocational preparation, DOT-GED is a 1-6 index of general educational development, DOT-environment is the number of work environment disamenities from 0-5, DOT-hazards is the proportion in hazardous jobs, DOT-strength is measured by a 1–5 index from low to high strength required, and DOT-physical is the number of physical demands from 0–4

References

Bergmann, Barbara R "Occupational Segregation, Wages and Profits When Employers Discriminate by Race or Sex" *Eastern Economic Journal* 1 (April 1974) 103–10

Blau, Francine D "Occupational Segregation and Labor Market Discrimination" In *Sex Segregation in the Workplace Trends, Explanations, Remedies,* edited by Barbara F Reskin Washington, DC National Academy Press, 1984

Blau, Francine D, and Beller, Andrea H "Trends in Earnings Differentials by Gender, 1971–1981" *Industrial and Labor Relations Review* 41 (July 1988) 513–29

Blau, Francine D, and Kahn, Lawrence M "Race and Gender Pay Differentials" In *Research Frontiers in Industrial Relations and Human Resources,* edited by David Lewin, Olivia S Mitchell, and Peter D Sherer, pp 381–416 Madison, WI Industrial Relations Research Association, 1992

Bound, John, and Krueger, Alan B "The Extent of Measurement Error in Longitudinal Earnings Data Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9 (January 1991) 1–24

Card, David "The Effect of Unions on the Distribution of Wages Redistribution or Relabeling?" Working Paper no 4195 Cambridge, MA National Bureau of Economic Research, October 1992

Dickens, William T, and Ross, Brian A "Consistent Estimation Using Data from More than One Sample" Technical Working Paper no 33 Cambridge, MA National Bureau of Economic Research, 1984

England, Paula *Comparable Worth Theories and Evidence* New York Aldine de Gruyter, 1992

England, Paula, Farkas, George, Kilbourne, Barbara Stanek, and Dou, Thomas "Explaining Occupational Sex Segregation and Wages Findings from a Model with Fixed Effects" *American Sociological Review* 53 (August 1988) 544–58

Even, William E, and Macpherson, David A "The Decline of Private-Sector Unionism and the Gender Wage Gap" *Journal of Human Resources* 28 (Spring 1993) 279–96

Fields, Judith, and Wolff, Edward N "The Decline of Sex Segregation and the Wage Gap, 1970–80" *Journal of Human Resources* 26 (Fall 1991) 608–22

Filer, Randall K "Occupational Segregation, Compensating Differentials, and Comparable Worth" In *Pay Equity Empirical Inquiries,* edited by R T Michael, H I Hartmann, and B O'Farrell, pp 153–70 Washington, DC National Academy Press, 1989

——— "Compensating Differentials and the Male-Female Wage Gap A Comment" *Social Forces* 69 (December 1990) 469–73

Freeman, Richard B "Longitudinal Analyses of the Effects of Trade Unions" *Journal of Labor Economics* 2 (January 1984) 1–26

Funkhouser, Edward "Do Immigrants Have Lower Unionization Propensities than Natives?" *Industrial Relations* 32 (Spring 1993) 248–61

Gerhart, Barry, and El Cheikh, Nabil "Earnings and Percent Female Longitudinal Study" *Industrial Relations* (Winter 1991) 62-78
Gibbons, Robert, and Katz, Lawrence F "Does Unmeasured Ability Explain Inter-industry Wage Differentials?" *Review of Economic Studies* 59 (July 1992) 515-35
Goldin, Claudia D *Understanding the Gender Gap An Economic History of American Women* New York Oxford University Press, 1990
Groshen, Erica L "The Structure of the Female/Male Wage Differential Is It Who You Are, What You Do, or Where You Work?" *Journal of Human Resources* 26 (Summer 1991) 457-72
Gupta, Nabanita Datta "Probabilities of Job Choice and Employer Selection and Male-Female Occupational Differences" *American Economic Review Papers and Proceedings* 83 (May 1993) 57-61
Heckman, James J "Identifying the Hand of the Past Distinguishing State Dependence from Heterogeneity" *American Economic Review Papers and Proceedings* 81 (May 1991) 75-79
Hirsch, Barry T, and Macpherson, David A "Wages, Racial Composition, and Quality Sorting in Labor Markets" Discussion Paper no 1038-94 Madison, WI Institute for Research on Poverty, July 1994
Hirsch, Barry T, and Schumacher, Edward J "Labor Earnings, Discrimination, and the Racial Composition of Jobs" *Journal of Human Resources* 27 (Fall 1992) 602-28
Hwang, Hae-shin, Reed, W Robert, and Hubbard, Carlton "Compensating Wage Differentials and Unobserved Productivity" *Journal of Political Economy* 100 (August 1992) 835-58
Jacobs, Jerry A, and Steinberg, Ronnie J "Compensating Differentials and the Male-Female Wage Gap Evidence from the New York State Comparable Worth Study" *Social Forces* 69 (December 1990) 439-68
Johnson, George, and Solon, Gary "Estimates of the Direct Effects of Comparable Worth Policy" *American Economic Review* 76 (December 1986) 1117-25
Killingsworth, Mark R *The Economics of Comparable Worth* Kalamazoo, MI Upjohn Institute for Employment Research, 1990
Kloek, T "OLS Estimation in a Model Where a Microvariable Is Explained by Aggregates and Contemporaneous Disturbances Are Equicorrelated" *Econometrica* 49 (January 1981) 205-7
Kossoudji, Sherri A, and Dresser, Laura J "Working Class Rosies Women Industrial Workers during World War II" *Journal of Economic History* 52 (June 1992) 431-46
Light, Audrey, and Ureta, Manuelita "Panel Estimates of Male and Female Job Turnover Behavior Can Female Nonquitters Be Identified?" *Journal of Labor Economics* 10 (April 1992) 156-81
Mellow, Wesley, and Sider, Hal "Accuracy of Response in Labor Market Surveys Evidence and Implications" *Journal of Labor Economics* 1 (October 1983) 331-44
Moulton, Brent R "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units" *Review of Economics and Statistics* 72 (May 1990) 334-38

Oaxaca, Ronald L., and Ransom, Michael R. "On Discrimination and the Decomposition of Wage Differentials." *Journal of Econometrics* 61 (March 1994): 5–21.

O'Neill, June. "The Determinants and Wage Effects of Occupational Segregation." Working paper. Washington, DC: Urban Institute, March 1983.

O'Neill, June, and Polachek, Solomon. "Why the Gender Gap in Wages Narrowed in the 1980s." *Journal of Labor Economics* 11, pt. 1 (January 1993): 205–28.

Padavic, Irene. "White-Collar Work Values and Women's Interest in Blue-Collar Jobs." *Gender and Society* 6 (June 1992): 215–30.

Peracchi, Franco, and Welch, Finis. "How Representative Are Matched Cross Sections? Evidence from the Current Population Survey." *Journal of Econometrics* (in press).

Polachek, Solomon W. "Occupational Segregation among Women: Theory, Evidence and a Prognosis." In *Women in the Labor Market*, edited by Cynthia B. Lloyd, Emily Andrews, and Curtis Gilroy. New York: Columbia University Press, 1979.

Rebitzer, James B., and Taylor, Lowell J. "Work Incentives and the Demand for Primary and Contingent Labor." Working Paper no. 3647. Cambridge, MA: National Bureau of Economic Research, March 1991.

Reskin, Barbara F., ed. *Sex Segregation in the Workplace: Trends, Explanations, Remedies.* Washington, DC: National Academy Press, 1984.

Smith, James P. "Commentary" (Comments on Filer, 1989). In *Pay Equity: Empirical Inquiries*, edited by R. T. Michael, H. I. Hartmann, and B. O'Farrell, pp. 171–75. Washington, DC: National Academy Press, 1989.

Sorensen, Elaine. "Measuring the Effect of Occupational Sex and Race Composition on Earnings." In *Pay Equity: Empirical Inquiries*, edited by R. T. Michael, H. I. Hartmann, and B. O'Farrell, pp. 49–69. Washington, DC: National Academy Press, 1989.

———. "The Crowding Hypothesis and Comparable Worth Issue: A Survey and New Results." *Journal of Human Resources* 25 (Winter 1990): 55–89.

———. "Why the Gender Gap Declined in the 1980s." Mimeographed. Washington, DC: Urban Institute, November 1991.

U.S. Department of Commerce, Bureau of the Census. *CPS Occupation Categories: Detailed 1970 and Major 1980 Classifications.* Washington, DC: U.S. Government Printing Office, 1983.