

The research program of the Center for Economic Studies (CES) produces a wide range of theoretical and empirical economic analyses that serve to improve the statistical programs of the U.S. Bureau of the Census. Many of these analyses take the form of CES research papers. The papers are intended to make the results of CES research available to economists and other interested parties in order to encourage discussion and obtain suggestions for revision before publication. The papers are unofficial and have not undergone the review accorded official Census Bureau publications. The opinions and conclusions expressed in the papers are those of the authors and do not necessarily represent those of the U.S. Bureau of the Census. Republication in whole or part must be cleared with the authors.

### **Sex Segregation in U.S. Manufacturing**

William J. Carrington  
Department of Economics  
The Johns Hopkins University  
Baltimore, MD 21218  
will@jhu.edu

Kenneth R. Troske  
Center for Economic Studies  
Economic Planning and Coordination Division  
Washington Plaza II, RM 211  
U.S. Bureau of the Census  
Washington, DC 20233-6100  
ktroske@census.gov

CES 96-4 June 1996

All papers are screened to ensure that they do not disclose confidential information. Persons who wish to obtain a copy of the paper, submit comments about the paper, or obtain general

information about the series should contact Sang V. Nguyen, Editor, Discussion Papers, Economic Planning and Coordination Division, Center for Economic Studies, Washington Plaza II, Room 211, Bureau of the Census, Washington, DC 20233-6101, (301-457-1882) or INTERNET address [snguyen@info.census.gov](mailto:snguyen@info.census.gov).

#### Abstract

This paper studies interplant sex segregation in the U.S. manufacturing industry. The study differs from previous work in that we have detailed information on the characteristics of both workers and firms, and because we measure segregation in a new and better way. We report three main findings. First, there is a substantial amount of interplant sex segregation in the U.S. manufacturing industry, although segregation is far from complete. Second, we find that female managers tend to work in the same plants as female supervisees, even once we control for other plant characteristics. And finally, we find that interplant segregation can account for a substantial fraction of the male/female wage gap in the manufacturing industry, particularly among blue-collar workers.

Keywords: Interplant, Gender, Manufacturing, Segregation,

## **I. Introduction**

This paper explores the empirical connection between cross-employer sex segregation and the gap between men's and women's wages. The analysis is motivated by both theoretical and policy issues. The theoretical questions arise from two types of discrimination models. First, Becker's (1957) theory of employer discrimination posits that employers act as if women employees come with an extra psychic cost. If this psychic cost varies across employers, then the market sorts women into those employers who attach the lowest extra cost to their employment. If there are a sufficient number of employers who are indifferent between men and women, then there will be no wage gap; otherwise an equilibrium wage gap exists. Similar implications arise from Becker's theories of customer and coworker discrimination. Second, Lang's (1986) language theory of discrimination posits that communications difficulties between men and women lead to both workplace segregation and reduced wages for women. As with Becker's models, if this theory is to explain the male/female wage gap, then it must be that men and women are segregated into different employers. These facts imply that precise measurements of workplace segregation are crucial to the evaluation of these theories' ability to explain the male/female wage gap.

Our analysis of interfirm gender segregation is also motivated by policy considerations. Most comparable worth policies are designed to

reduce male/female wage differentials within any given employer, but they have no direct impact on wage differences across employers (Johnson and Solon, 1986). Similarly, Title VII of the Civil Rights Act of 1964 is now primarily used to equalize treatment of men and women within a firm; it is rarely used in class action suits based on hiring policies (Donohue and Siegelman, 1991). Thus, the effectiveness of much of the public policy designed to improve women's relative pay is dependent on the extent to which men and women are currently integrated in the workplace.

Early studies of cross-employer segregation (McNulty, 1967; Buckley, 1971; Blau, 1977) found that women were typically segregated into the lowest-paying employers, even within occupations, and these results have been consistently replicated in more recent studies (Pfeffer and Davis-Blake, 1987; Groshen, 1991; Reilly and Wirjanto, 1994; Carrington and Troske, 1995a; Griffin and Trejo, 1995). These results have often been viewed as support for theories such as Becker's and Lang's that emphasize segregation. This paper departs from previous work on cross-employer gender segregation in three ways. First, we study data drawn from relatively large plants in the U.S. manufacturing industry, an important sector that has been understudied in previous work. Second, our data include detailed information on workers' demographic characteristics that was missing in most previous studies. This lets us do a relatively good job of separating a) segregation due

to discrimination from b) segregation due to human capital differences between men and women. And finally, we measure segregation in a different and better way. Standard segregation measures conflate the random allocation of workers to firms with systematic segregation that might be due to discrimination. In contrast, we use new measures of segregation that distinguish between systematic and random components of segregation.

The main results of the paper are summarized as follows. First, there is substantially more interplant gender segregation in U.S. manufacturing than would be expected if workers were allocated randomly to plants. Thus, one of the main predictions of discrimination theory is consistent with the data. However, it is also true that the economy is far from completely segregated, as would be implied by the simplest models of Becker and Lang. Second, there is a positive correlation between the gender of supervisors in a plant on the one hand, and the gender of *non*-supervisors in a plant on the other; women tend to work for women and men tend to work for men. If we assume that female managers are less likely than males to discriminate against female employees, as seems reasonable, then this evidence is consistent with Becker's model. Alternatively, if we assume that communication is easier between managers and employees of the same sex, then the results are also consistent with Lang's model. Finally, we find that, in a purely accounting sense, interplant gender segregation is responsible

for a substantial fraction of the male/female wage gap. Women tend to work in plants where workers, both male and female, tend to be paid less. These results are analogous to the connection between occupational segregation and wages (e.g., Macpherson and Hirsch, 1995), and they imply that even a vigorous comparable worth policy would meet only limited success in reducing the gender wage gap.

The paper proceeds as follows. Section II reviews the major theories of interplant sex segregation. Section III describes our data sources and section IV reviews methods for measuring and interpreting interfirm sex segregation. Section V assesses interplant gender segregation and analyzes why plants vary in their employment of women. Section VI studies the relationship between interplant segregation and the male/female wage gap and, finally, section VII concludes.

## **II. Theories of Interplant Sex Segregation**

This section provides a brief review of the leading theories of interplant sex segregation. Table 1 lists the theories considered here and outlines their implications both for interplant sex segregation and for the male/female wage gap. Becker's (1957) canonical model of employer discrimination posits that employers act as if there is an extra psychic cost attached to the hiring of each female worker. It is typically further assumed that this taste for discrimination varies across employers, as some employers are indifferent to the sex of their

employees, while others may be quite insistent on hiring only men. If men and women are perfect substitutes in production, then this model predicts that women will be completely segregated into those employers with the lowest tastes for discrimination. If women and men are not perfect substitutes, then employers will be more likely to employ both men and women, but there will still be substantial segregation.<sup>1</sup> In either case, women receive lower wages than men if there are too few non-discriminatory employers. Note that competition among women equalizes women's wages across employers in this model. Thus, wages of women and men are independent of the sex composition of coworkers.

The predictions of Becker's (1957) customer discrimination model are similar to those of the employer discrimination model. The main difference is that women are now sorted into those employers with non-discriminatory customers, rather than into those with non-discriminatory managers. Again, the model predicts that there will be a male/female wage gap if there are too few employers with non-discriminatory customers. In contrast, Becker's model of *coworker* discrimination has some unique implications. It too implies segregation, but it also implies that men working in employers with mostly female workers will receive higher wages than similar workers in all-male employers. This prediction does not arise in Becker's models of customer or employer

---

<sup>1</sup>Alternatively, if employers care about the proportion of women hired, then there will be substantial but not complete segregation (Neumark, 1988). Again, women receive lower wages if there are too few non-discriminatory employers.

discrimination. While these models differ in their assumptions and in some of their implications, they all imply that women and men will tend to work for different employers. Thus, workplace segregation is a central implication of all Becker's discrimination models.

Lang (1986) presents an alternate model of discrimination. In this model, discrimination arises because male employers and employees find it relatively difficult to communicate with female employees. These communication difficulties lead to transactions costs that employers try to minimize in two ways. First, employers tend to group workers of the same sex together, as this reduces communications costs between employees. Second, employers of either sex prefer to hire workers of their own sex, since this reduces communications costs between employer and employee. Both phenomena tend to segregate men and women in the workforce. In addition, if there are too few female employers, then women workers will receive lower wages in order to induce male employers to hire them. The model is obviously much like Becker's, but the "discrimination" here arises from communications difficulties rather than personal animus.

An alternate theory of racial segregation is based on differences in the education and occupation of men and women. Since employers' skill requirements vary widely (Kremer and Maskin, 1994; Doms, Dunne, and Troske 1995; Troske, 1995), such group variation in skills may induce segregation. For example, high-tech plants that employ many



PhD's will tend to have few female employees, while plants with more traditional pink-collar workforces will tend to employ mostly women. Following Macpherson and Hirsch (1995), we refer to this as the quality-sorting hypothesis. To the extent that we can isolate groups of workers with similar skills, this theory predicts that there should be no systematic within-skill-group segregation and no within-skill-group wage gap. A second alternate theory simply recognizes that the random allocation of workers to employers will typically generate some interplant variation in employee sex composition. This in turn generates some segregation, at least when measured by conventional segregation indices. Of course, random allocation does not cause women to have lower wages than men. Our empirical work assesses the relative importance of these alternate sources of interplant sex segregation.

### **III. Data Sources**

#### **a. The WECD**

This study uses data from the Worker-Establishment Characteristics Database (WECD), a recently developed Census Bureau database that matches information on workers and employers.

The basic design of the WECD is as follows.<sup>2</sup> Manufacturing plants in the Census Bureau's Longitudinal Research Database (LRD) are associated

---

<sup>2</sup>See Troske (forthcoming) for a more complete description of the WECD's development.

with an industry and a geographic block code. Workers responding to the 1990 Census Long Form report the industry and street address of their employer, which has been linked to the block codes used in the LRD. Using this information, workers and firms are matched on the basis of industry and block code. Workers are successfully matched *only* if, in the LRD, there is a unique plant in the appropriate industry and block code. Thus, workers can fail to be matched if a) there is no plant in the LRD in their reported industry and block (perhaps due to reporting error by the worker), or because b) there is more than one plant in the LRD in their reported industry and block, in which case no unique assignment can be made. It is important to note that since the LRD samples only manufacturing establishments, the WECD only contains workers from the manufacturing industry. Nevertheless, the combination of detailed employer and employee data is crucial to the measurement of interfirm segregation and is unmatched by other U.S. data sources.

The design of the WECD implies that we have information on a sample of manufacturing plants, rather than the entire population of plants in manufacturing. The incomplete sample occurs because the LRD contains most but not all manufacturing plants, because plants that share an industry and block with another plant are excluded, and because for some small plants there were no workers who responded to the Census Long Form. Further, within any plant in our sample, we have information on a sample of workers at the plant, rather than the complete

population. The incomplete sampling within establishments arises from two factors. First, the Long Form of the Census is only administered to one out of every six households. Second, some respondents provide incomplete or inaccurate information on their employer's address or industry, which precludes matching such workers to the plant where they work. Together, these factors imply that, among those plants represented in our sample, we have on average approximately 1 out of 12 workers in our sample.<sup>3</sup>

Before moving on, let us emphasize the advantages of using the WECD to investigate patterns of interplant segregation. First, we have as full a set of human capital indicators as can typically be hoped for, including measures of education, age, experience, and occupation.<sup>4</sup> In contrast, most previous segregation studies have used data with relatively crude information about workers' human capital attributes (e.g., Groshen, 1991; Griffin and Trejo, 1995), or data with no worker information at all (Carrington and Troske, 1995a). Second, the WECD provides a useful set of plant characteristics, including sales, value-added, and the capital-labor ratio.<sup>5</sup> In conjunction with the relatively precise human capital measures, these plant characteristics permit an

---

<sup>3</sup>While it would obviously be preferable to have complete information on each plant's workforce, it is still possible to draw useful conclusions from the sample at hand.

<sup>4</sup>This information comes from workers' responses to the Census Long Form.

<sup>5</sup>This information comes from the LRD.

investigation of the characteristics that separate "male" and "female" establishments. And third, the WECD covers the manufacturing sector that has been relatively understudied in previous work.

#### b. Sample Selection Criteria

There are significant differences between men and women in the tendency to work part-time. In 1990, for example, roughly 25% of employed women worked fewer than 35 hours per week, while only 10% of employed men worked part-time by this definition (U.S. Statistical Abstract, 1994). These gender differences in part-time status raise some important sample selection issues. If plants first choose their use of part-time workers and then decide whether or not to hire men or women, including part-time workers in the sample runs the risk of discovering "sex segregation" when what is really at work is the segregation of part-time and full-time workers. In contrast, if some plants choice of part-time vs. full-time workers is in part driven by preferences over the sex of their employees, then excluding part-time workers will lead us to miss an important dimension of interplant segregation.

Since we see no a priori reason to believe that either of these views is strictly correct, we experiment with different sample restrictions. For most of our work, we restricted the sample to workers who a) usually worked more than 30 hours per week in the previous year

and b) worked more than 30 weeks in the previous year, which obviously excludes part-time workers. However, we examine the sensitivity of certain key results to the inclusion of part-time workers in the sample. Other than these restrictions on weeks and hours at work, our only restriction is that a worker's wage not be too much of an outlier. In particular, we required that workers' actual log wage be within five standard deviations of their predicted log wage in a regression of log wages on a wide variety of individual and employer characteristics. This last restriction actually excludes very few people, so it has virtually no effect on our measures of segregation, but it leads to what we believe are slightly more representative measures of wage differences between men and women. The application of these sample restrictions leads to a final sample of 123,183 men and 48,670 women spread across 8308 manufacturing plants.

Table 2 presents selected summary statistics from our WECD sample of workers and firms. The first four rows of table 1 indicate that WECD workers are well-paid relative to the U.S. economy as a whole. This is in part due to the fact that wages are generally higher in the manufacturing industry, but WECD workers' wages are somewhat higher on average than those in the rest of the manufacturing industry (Troske, forthcoming). This is largely because WECD plants are larger than the manufacturing industry average, and because wages are generally higher in large plants (Brown, Medoff, and Hamilton, 1990). Rows 1 through 4

also show that there is a substantial wage gap between men and women in the WECD, and this gap is roughly consistent with that observed in the economy as a whole.

Rows 5 and 6 report of table 2 the average age and potential experience of WECD workers, which are both somewhat higher than the corresponding statistics for all manufacturing workers. Row 7 reports that the average WECD worker works at a very large establishment. This is again in part due to the generally large size of manufacturing establishments, but WECD establishments are large even relative to this industry baseline.<sup>6</sup> Finally, the rows below the numbers 8 and 9 indicate that the restriction to the manufacturing sector has led WECD workers to be concentrated in the Northeast and Midwest and in those occupations with strong representation in manufacturing. Together these facts imply that the WECD is a somewhat select group of workers. There are no obvious reasons why patterns of segregation in the WECD should differ dramatically from that of the aggregate economy, but we must still recognize that these results may not be completely representative of the U.S. economy as a whole.

#### **IV. Measuring Segregation**

This section considers some methodological issues in the

---

<sup>6</sup>The large size of WECD plants arises from the WECD sampling frame. Workers are matched to plants based on industry and location, and unique matches can be made only to plants that are the only one in their industry and location. These unique plants tend to be larger than average (Troske, forthcoming).

measurement of segregation. Economists frequently summarize segregation patterns with an index of segregation, which is simply a statistic that summarizes the extent to which two groups come into contact within a sample. The range of these indices is typically the [0,1] interval. An index of zero corresponds to complete evenness, which occurs when groups are proportionately represented in each plant. If the female population share were 50%, for example, then the sample would be even if every plant employed equal numbers of men and women. In contrast, an index of one corresponds to complete unevenness, which occurs when every plant is either all male or all female. While there are a variety of available segregation indices, we use two particular indices in this study. The first is the popular *dissimilarity index* developed by Duncan and Duncan (1955). If we let  $w_i$  and  $m_i$  equal plant  $i$ 's share of female and male employees in the sample, respectively, then the dissimilarity index is simply

$$D = \sum_i \frac{1}{2} |w_i - m_i|.$$

The dissimilarity index may be interpreted as the share of men (or women) that would have to change plants in order to make the sample completely even. Thus, an index of one implies that all men (or all women) would have to switch plants to achieve integration, whereas an index of zero implies that the population is already completely integrated.

Hutchens (1991) criticized the dissimilarity index because it is equally sensitive to mild and extreme departures from evenness. He recommends using instead the *gini coefficient of segregation*, which is more sensitive to the presence of extremely uneven plants. If we first sort the plants on the basis of  $w_i/m_i$ , then the gini coefficient of segregation may be expressed as

$$G = 1 - \sum_{i=1}^T w_i (m_i + 2 \sum_{j=i+1}^T m_j),$$

where T is the number of plants in the sample. As the name suggests, this index is analogous to the gini coefficient of variation widely used in income studies. The only difference is that this index measures interplant variation in gender workforce shares (i.e.  $w_i/m_i$ ) rather than interpersonal variation in income.

To see the difference between the two indices, consider a four plant sample with the following distribution of men and women:

plant 1: 50 women and 50 men  
 plant 2: 50 women and 50 men  
 plant 3: 75 women and 25 men  
 plant 4: 25 women and 75 men.

Both the gini coefficient and the dissimilarity index would characterize this distribution as segregated ( $D=.25, G=.375$ ). Now imagine either a) having plants 1 and 2 trade a man for a woman, or b) having plant 3 trade a man to plant 4 in return for a woman. Either move would result in a sample that was more segregated from the perspective of either



index. However, the gini coefficient would view case b as causing a greater increase in segregation (.39 in case b vs. .38 in case a), whereas the dissimilarity index would treat the two cases symmetrically (.26 in either case). This is an example of the way in which the gini coefficient puts more weight on the tails of the gender share distribution.

The dissimilarity index and the gini coefficient each have their advantages: the dissimilarity is familiar and easy to interpret, while the gini coefficient has more of the behavioral properties that one would like in an ideal segregation index (Hutchens, 1991). Yet a commonly overlooked weakness of both indices is that they are positive when workers are allocated randomly to plants.<sup>7</sup> The problem is that unless the plant is very large, a random draw of employees will not typically reflect the population exactly. To see this clearly, consider a large sample of two-worker plants drawing randomly from a population with equal numbers of men and women. 25% of such plants would have two men, 50% would have one man and one woman, and 25% would have two women. Although such a distribution is completely random, it is far from even, and the gini coefficient and dissimilarity index would be .75 and .50, respectively. While this is an extreme case, it is easy to show that random allocation of workers to plants implies substantial unevenness

---

<sup>7</sup>This critique also applies to other commonly used segregation measures such as Atkinson's index or Theil's Entropy Index.

for plants with as many as fifty or a hundred employees (Carrington and Troske, 1995b).

This observation has three implications for the interpretation of segregation indices. First, it is generally a mistake to conclude from positive and statistically significant segregation indices that there is any systematic sorting of men and women into different employers; such patterns are often equally explicable by chance. Second, random allocation generates far more unevenness among small plants than among large plants. In contrast with the example above, the random allocation of a 50/50 mix of men and women to a large sample of 1000-worker plants would lead to a gini coefficient of .04 and a dissimilarity index of .03. In sum, this implies that it is generally difficult to deduce from such traditional segregation measures whether or not a sample is systematically segregated, and that cross-sample comparisons of segregation are difficult to interpret unless the samples have plants of roughly equal size. Third, it is the number of workers per plant in the sample that dictates the importance of such random segregation. There is no reason to believe that small samples from large plants will be evenly distributed. Thus, since our sample only has on average one twelfth of the workers in an establishment, the fact that we look at relatively large plants does not mean that we are free of this problem.

These concerns lead us to report additional statistics that assist in the interpretation of interplant segregation. Carrington and Troske

(1995b) propose the following modifications of the gini coefficient and the dissimilarity index as a means of distinguishing between systematic and random segregation. We couch the modification in terms of the gini coefficient, but the modified dissimilarity index is completely analogous. Let the *gini coefficient of random segregation*  $G^*$  be the gini coefficient that would occur if a very large number of workers were allocated randomly to employers, taking the sexes' population shares and the size distribution of plants as determined by the sample.<sup>8</sup> The *gini coefficient of systematic segregation* is then defined as

$$\hat{G} = \frac{G-G^*}{1-G^*} \quad \text{if } G-G^* \geq 0$$

and

$$\hat{G} = \frac{G-G^*}{G^*} \quad \text{if } G-G^* < 0.$$

If there is excess unevenness, i.e.  $G > G^*$ , then  $\hat{\alpha} > 0$  is simply the extent to which the sample is more segregated than random allocation would imply ( $G-G^*$ ), expressed as a fraction of the maximum amount of such excess segregation that could possibly occur ( $1-G^*$ ).  $\hat{\alpha} = 1$  is analogous to complete unevenness, as with the standard gini coefficient,

---

<sup>8</sup>The first step in calculating  $G^*$  is to calculate the empirical number of firms in each size class  $s$ ,  $\hat{g}(s)$ . Within any size class  $s$ , random allocation implies that the binomial function  $B(m;s;p)$  is the fraction of firms that will have  $m$  women if  $p$  is the female population share. Thus, random allocation implies that the number of sample units with size  $s$  and  $m$  women should be  $N(m,s;p) = B(m;s;p) \hat{g}(s)$ . Thus, the support of  $N(m,s;p)$  is  $m=0$  to  $s$  for every  $s$  in the support of  $\hat{g}(s)$ . This artificial distribution corresponds to what would be expected if workers were allocated randomly, given  $p$  and  $\hat{g}(s)$ . The gini coefficient of segregation computed from this artificial distribution is what we call the gini coefficient of random segregation.

but  $\alpha = 0$  implies that the sample is equivalent to random allocation. If there is excess evenness, i.e.  $G < G^*$ , then  $\alpha$  is negative and represents excess evenness in the sample  $(G - G^*)$ , expressed as a fraction of the maximum amount of such excess evenness that could possibly occur ( $G^*$ ).

To summarize, our  $\alpha$  is different from the standard gini coefficient in two ways. First, we have set the baseline of 0 to correspond to random allocation rather than complete evenness. Second, we have remapped values of  $G$  that are greater than  $G^*$  into the  $[0,1]$  interval, and remapped values of  $G$  that are less than  $G^*$  into the  $[-1,0]$  interval. In the following analysis, we will examine a similarly modified dissimilarity index. We think that these modified indices provide more useful information than the traditional ones. However, we recognize that some readers are more comfortable with the traditional indices. Thus, in the work that follows we report traditional indices of segregation, indices of random segregation, and the indices of systematic segregation that we developed here. Together, these indices provide a useful summary of segregation patterns.

## **V. Interplant Sex Segregation**

Table 3 examines the extent to which men and women are segregated

across manufacturing establishments in the United States.<sup>9</sup> The six columns each report a different type of segregation index. Columns 1-3 report results for the three versions of the gini coefficient, whereas columns 4-6 report results for the dissimilarity index. While the columns differ by the index reported, the rows of table 3 vary by the sample of workers considered. Row 1 reports results for all workers in the WECD, while others rows report results for samples stratified by broad industries (rows 2 and 3), by selected detailed industries (rows 4-8), by broad occupations (rows 9-14), by selected detailed occupations (rows 15-20), or by schooling (rows 21-24). The numbers without parentheses are the index values, while the numbers in parentheses are bootstrap standard errors.<sup>10</sup>

Row 1 shows that there is substantial sex segregation across U.S. manufacturing plants, as the traditional gini coefficient is .59 and the traditional dissimilarity index is .43. Thus, the distribution of men and women across plants is far from even. However, much of the observed unevenness is attributable to random allocation of workers to plants, as the expected gini coefficient is .24 and the expected dissimilarity index is .16. This shows that, even among these relatively large

---

<sup>9</sup>It is important to remember that, like most previous authors (e.g., McNulty (1967), Buckley (1971), Blau (1977) and Groshen (1991)), this paper studies segregation across plants rather than firms.

<sup>10</sup>It may seem puzzling that the "Expected" indices of columns 2 and 5 have standard errors. However, recall that these indices are what would be expected conditional on the sample distribution of establishment sizes and gender. Since these distributions vary somewhat across bootstrap samples, the expected indices are themselves stochastic.

plants, a statistically significant segregation index is not always indicative of any systematic allocation of workers to employers. Nevertheless, while positive, the expected segregation indices are substantially less than what is actually observed. For example, the systematic gini coefficient of row 1 indicates that excess unevenness is 45% of the maximum that could possibly be observed. Thus, the distribution of all men and women is substantially more uneven than random allocation would predict: men and women are systematically segregated in U.S. manufacturing.

While these results are consistent with the theories of Becker and Lang, it is important to recognize that the sexes could be segregated by factors other than discrimination. In particular, the quality-sorting hypothesis suggests that men and women have different skills and that this creates segregation independently of discrimination. To distinguish these hypotheses requires that we examine segregation among men and women of similar skills. For this reason, the rest of table 3 examines sex segregation among relatively homogenous groups of workers. Rows 2-8 of table 3 examine the extent to which interindustry segregation can explain the aggregate segregation patterns of row 1. While such interindustry segregation may itself be attributable to discrimination, it might also plausibly be attributed to interindustry variation in required worker skills. Rows 2 and 3 take a crude attack by simply breaking industries out into durables and non-durables. The

indices show that within-industry segregation is very similar to that of the aggregate, so that little aggregate segregation is due to interindustry segregation at this level.

Rows 4-8 of table 3 examine segregation within a set of detailed industries that are well-represented in the WECD.<sup>11</sup> The results show that segregation within-detailed industries is generally substantially less than that within manufacturing as a whole. For example, the systematic gini coefficient for the Meat Products industry is only .28, whereas the corresponding index for nondurable manufacturing as a whole is .45. Similarly, the systematic dissimilarity index is .30 in durable manufacturing as a whole, but it is only .16 in the household appliance industry. These results suggest that much of the aggregate segregation in U.S. manufacturing is due to the sorting of men and women into different detailed industries. In addition, note that systematic segregation is particularly low in the Newspaper Publishing industry, which employs a relatively large number of white-collar workers. These results suggest a white-collar/blue-collar distinction that we will shortly explore further.

Much of the debate on labor market differences between men and women is based on the idea that women are segregated into a few, relatively low-paying occupations. One side of this debate views

---

<sup>11</sup>We chose these industries because they had a large number of plants in the WECD. Interested readers may obtain results for other detailed industries from the authors.

occupational segregation as the result of discrimination (e.g., Bergmann, 1986), while the other side views occupational segregation as the result of men's and women's differential human capital investments, which are themselves driven by traditional divisions of family responsibility (e.g., Mincer and Polachek, 1974). While these views differ, they both suggest that much of the interfirm segregation documented in row 1 may reflect a) interfirm *occupational* segregation, and b) occupational segregation by sex. Rows 9-14 of table 3 address this issue with measures of interfirm segregation within six crude occupational categories. Before addressing the substance of these rows, note first that there is a mechanical reason why the "expected" segregation indices of columns 2 and 5 are much higher for these "within-occupation" rows than they were for the previous analyses. This occurs because the restriction to particular occupations means that we have fewer sample workers per establishment. As the discussion of the previous section pointed out, the random allocation of fewer workers across the same number of employers leads to an increase in random unevenness, and hence to an increase in expected segregation. This illustrates the pitfalls of using traditional segregation indices to compare segregation across different samples. Without accounting for random unevenness, one would conclude that within-occupation segregation was more severe than it was in the entire population, which we believe would be the wrong conclusion.



More substantively, the move to a within-occupation analysis leads to reduced estimates of the systematic component of segregation for some occupations. For example, the systematic gini coefficient for sales and service occupations in our sample is only .21, and the systematic dissimilarity index for professionals, technicians, and managers is only .13. Thus, for these occupations, there is only a limited amount of systematic interfirm segregation within U.S. manufacturing. However, the same is not true of more blue-collar occupations. For example, craftsmen, operatives, and laborers are all much more segregated than random allocation would predict. Thus, some but not all of the interfirm segregation documented in row 1 is attributable to occupational segregation.

Much of the literature on occupational segregation is concerned with segregation across quite narrowly defined occupations. While a within-occupation analysis of all detailed occupations would be unwieldy, rows 15-20 of table 3 analyze segregation within a few detailed occupations, which were chosen because they are well-represented in our samples. Moving to samples defined by such narrow occupations leads to a small average number of workers per plant, which in turn leads to very high expected indices of segregation. Thus, the high traditional segregation indices found in these samples (e.g., .91 for the gini coefficient for mechanics and repairers) is largely due to random unevenness. Nevertheless, there is a substantial amount of

systematic segregation within these occupations. As before, there is an important blue-collar/white-collar distinction between "engineers, architects, and surveyors" on one hand, and all the other occupations on the other.

As a final exercise, rows 21-24 of table 3 analyze segregation within groups of workers stratified by educational attainment. The results again suggest that differences in educational attainment do not explain aggregate interfirm gender segregation. Of course, this is not surprising as the education gap between men and women is not large. More interestingly, the results suggest that there is substantially less systematic segregation amongst workers with at least some post-high school education. For example, the systematic gini coefficient is .23 for workers with a college degree or more, but it is .59 for high school dropouts. These results are consistent with the results for occupation, where it was the blue-collar occupations that were the most segregated. These results may indicate that educated, white-collar men are less resistant to working with women than are their blue-collar counterparts. Alternatively, it may be that affirmative action and civil rights pressures to integrate operate most strongly on white collar workers.

The results of table 3 demonstrate that men and women are substantially more segregated across manufacturing plants than random allocation would imply. There are several theories consistent with this fact. First, it may be that prejudicial attitudes or communications

barriers serve to segregate men and women, as in the models of Becker and Lang. Alternatively, it may simply be that men and women have different skills and occupations, and that plants vary in the types of workers they employ. The fact that there is substantial segregation within occupations and industries suggests that gender differences in skills cannot completely explain the observed segregation patterns. However, it must be recognized that the occupational and industry groups considered in that analysis were quite broad, and it may be that segregation along more refined dimensions is what drives interplant segregation.

One useful though imperfect means of distinguishing between these hypotheses is to examine the correlation between the sex of supervisors and the sex of supervisees. Lang's theory clearly implies that communications difficulties between managers and workers leads female managers, who presumably speak the "language" of female workers, to employ more female workers. In addition, if female managers are less likely to have a taste for discrimination against women, then Becker's theory predicts that women managers are more likely to supervise female employees.<sup>12</sup> In contrast, the quality-sorting hypothesis does not predict that otherwise identical male and female managers will

---

<sup>12</sup>While there is not much direct evidence on whether or not women are less likely to discriminate against women, such evidence as exists suggest that there probably is some same sex preference (Goldberg, 1968; Ferber and Huber, 1975). However, in her case study of a large corporation, Kanter (1977) argues that women managers may suffer from "tokenism" which causes them to be even tougher on women subordinates than are men.

systematically supervise workers of either sex. Thus, subject to some provisos, the correlation between the sex of supervisors and the sex of supervisees provides some clue as to the proper interpretation of our results on interfirm segregation.

Table 4 addresses this issue with OLS regressions in which the unit of observation is a plant and the dependent variable is the female share of non-supervisory employees. Thus, these regressions try to explain why it is that some plants' supervisees are primarily women while other plants' supervisees are primarily men.<sup>13</sup> The right-hand side variables in the regressions include the log of plant employment, controls for the average age and education of non-supervisory and, separately, supervisory workers, as well as dummies for region and 2-digit industry. More importantly, column 1 includes the percent of supervisors that are female. The regression indicates that a 50% increase in women's share of a plant's supervisors is associated with roughly a 7.5% increase in women's share of the plant's non-supervisors. Thus, column 1 demonstrates that female managers tend to supervise female employees, although the relationship is far from completely systematic.

Column 1 of table 4 is consistent with the discrimination models of Becker and Lang, as female supervisors are more likely to hire and

---

<sup>13</sup>To be in the sample used to estimate these regressions, plants had to have at least one supervisor and one non-supervisor in our sample. This led to the exclusion of some of the smaller plants in the sample.

supervise female workers. However, these regressions can still be interpreted in the human capital framework without too much trouble. It is true that men and women are allocated somewhat differently within the broad class of supervisory occupations. While the inclusion of 2-digit industry dummies means that column 1 is comparing men and women within fairly narrowly defined industries, it remains possible that differences in the within-industry occupational structure of men and women is what drives column 1. As an example, consider the printing and publishing industry in general, and the newspaper industry in particular. Female managers in this industry may be likely to work in editorial or advertising establishments, while men may be more likely to supervise the manufacture and distribution of the papers. To an extent, one might take this as evidence of discrimination of one sort or another, but it might also be attributed to non-discriminatory differences in the type of human capital held by male and female managers and supervisees.

Columns 2-4 of table 4 present three crude ways of trying to address this issue. To the extent that male/female differences in human capital or occupation can be reduced to a scalar measure, as opposed to simple qualitative differences, controlling for the wages and/or productivity of both supervisors and non-supervisors may provide a cleaner measure of the degree to which the correlation between sex of supervisor and sex of supervisees is driven by discrimination. Column 2 repeats the regression of column 1 with the inclusion of the average

wage of non-supervisory employees and, separately, the average wage of supervisors. Row 13 shows that these variables add substantially to the explanatory power of the model, as the r-square increases by roughly a third. The added variables also attenuate the coefficient on percent female supervisors, although it remains moderately large and significant. Column 3 omits the wage variables and instead adds the log of labor productivity, which is defined as the dollar value of plant shipments divided by the number of employees.<sup>14</sup> This variable has much less explanatory power, and it also has less of an effect on the coefficient on percent female supervisors. Finally, column 4 repeats the exercise when both labor productivity and the wage variables are included. The results are very similar to those of column 2. In sum, table 4 shows that, holding the wages and productivity of their supervisees constant, men are more likely to supervise other men and women are more likely to supervise other women. This suggests that to the extent that human capital differences explain interplant gender segregation, it must be along qualitative rather than quantitative dimensions.

These results suggest that the systematic segregation documented in table 3 is at least in part the result of supervisors' systematic choice of supervisees of their own sex. These choices may arise from

---

<sup>14</sup>Similar results were obtained when we defined labor productivity to be value-added/employees.

discrimination due to prejudice as in Becker's model, or from discrimination due to language differences between the sexes. Of course, given the relatively crude occupational classification system considered in the analysis, it remains possible that occupational segregation along finer dimensions is what drives the observed patterns of interfirm segregation. However, together with previous research by Bielby and Baron (1984), Groshen (1991), and others, the results we report here suggest that, even within narrow occupations, there is usually some interfirm gender segregation.<sup>15</sup> Thus, we believe that there is an important systematic component of interfirm gender segregation, particularly among less-educated and blue-collar workers, and that this segregation is not completely explained by sex differences in human capital acquisition.

Before moving on, let us note again that the preceding analysis was conducted on a sample of workers who worked at or near full-time. We focussed on this sample because discrimination against full-time workers is particularly troubling, and because segregation in a broader sample might be partly due to segregation of part-time and full-time workers. However, since women are disproportionately part-time and because part-time/full-time segregation might itself reflect sex discrimination, it is worth comparing the above results to those

---

<sup>15</sup>One problem with these previous studies is that there was usually little effort made to distinguish systematic from random segregation. Thus, some of the patterns of intraoccupational sex segregation found by previous authors could be due to random allocation of workers.

obtained in a sample which includes part-time workers. In short, it makes very little difference in the analysis.<sup>16</sup> There are of course some minor changes in overall measures of segregation, and even some more significant changes in segregation within particular industries or occupations. However, the major results still hold: men and women are systematically segregated, particularly in blue-collar occupations and industries.

## **V. Interplant Segregation and the Male/Female Wage Gap**

This section investigates two aspects of the relationship between interfirm segregation and the male/female wage gap. We first decompose the gender wage gap into between- and within-plant components. This exercise is motivated by public policies such as Title VII and comparable worth that are largely devoted to eliminating within-plant pay differences between equally qualified men and women. This implies that these policies will not be effective unless within-plant pay differences are the primary source of women's low wages.

Our approach is to regress wages on a set of plant fixed effects, either before, after, or at the same time that we control for workers' personal characteristics, and to see how much of the gender wage gap can be explained by the location of men and women in different plants. Let

---

<sup>16</sup>Full results from the sample that includes part-time workers are available from the authors on request.



$Y$  = log hourly wages, let  $X$  = a set of personal characteristics including terms in education, experience, and detailed occupation, and let  $Z$  = a set of plant fixed effects. Columns 1-3 of table 5 then report results from a two-step procedure in which we first estimated  $Y = X'S$ , and then regressed the residuals of this first regression on the plant fixed effects  $Z$ . In essence, this procedure gives personal characteristics first crack at explaining the gender wage gap, and the plant effects are given the opportunity to explain the residual gender wage gap. Column 1 reports results for all workers, while columns 2 and 3 report separate results for blue collar and white collar workers. Row 1 shows that the unadjusted difference in log hourly wages between men and women is .429 for all workers, .439 for white collar workers, and .446 for blue collar workers in our sample.

Row 2 of table 5 reports the residual wage gap left after we control for the effect of personal characteristics. This is approximately .280 for white collar workers and approximately .371 for blue collar workers. Row 3 reports the residual male/female wage gap after netting out the fixed effects estimated in the second step, without first adjusting for personal characteristics, and Row 4 reports the residual wage gap after we net out both the personal characteristics and the plant fixed effects. Two facts stand out from these calculations. First, it is clear that plant fixed effects can statistically account for a substantial fraction of the male/female wage

gap, both as a whole and separately for the blue and white collar samples, even after we control for a large array of other productive characteristics. Second, the role of the plant fixed effects is much more important for blue collar workers than it is for white collar workers. For white collar workers, column (2) shows that human capital characteristics can account for 36% of the gender gap in log hourly wages. Subsequent controls for employer can account for an additional 24% of the gap so that, among white collar workers, employer identity plays an important but secondary role in wage determination. In contrast, column (3) shows that employer identity plays a dominant role in statistically explaining the gender wage gap among blue-collar workers. This can be seen by the fact that while human capital characteristics can explain only 17% of the gender wage gap, plant fixed effects can account for almost 68% of the residual wage gap. Thus, among blue-collar workers, the male-female wage gap is largely accounted for by women's location in plants that generally pay low wages.

Columns 4-6 reverse the order of the decomposition by regressing wages on the plant fixed effects in the first step, and then regressing the residuals on the personal characteristics in the second step. This gives the plant effects first crack at explaining the male/female wage gap. Row 3 of column (4) shows that the male/female wage gap remaining after controlling for plant fixed effects is .230 for all workers, or about 52% of what it was without controlling for these effects. Column

(6) shows that the role of plant fixed effects is particularly strong among blue-collar workers, where they explain over two thirds of the male/female wage gap. Columns 7-9 regress Y on X and Z simultaneously, so that personal characteristics and plant fixed effects are given equal opportunity to explain the gender wage gap. The results are similar to those of the previous columns, as plant fixed effects explain a substantial fraction of the male/female wage gap.

The results of table 5 show that there is an important, and in the case of blue-collar workers dominant, role played by interplant segregation in accounting for the wage gap between men and women. Women work in low-paying plants while men work in plants that pay relatively high wages. Comparable worth policies are designed to even out interoccupational wage differences within firms, and by extension within plants. The important role of interplant pay differences suggests that such policies can have only a limited effect on the male-female wage gap, even if they are effective at evening pay differences within firms. In addition, the results are consistent with discrimination theories which posit that women are crowded into a few non-discriminatory employers.

Our second exercise relates the wages of male and female workers to the gender makeup of their coworkers. This question is motivated by several observations. First, Becker's theory of coworker discrimination (although not his theory of employer discrimination)

posits that workers in integrated plants may receive higher wages. For example, if male workers demand a higher wage in order to work with women, then men in integrated plants will receive higher wages than similar men working in all-male plants. A second motivation is the emphasis of the gender wage gap literature on the interoccupational relationship between average wages and gender composition. While this focus does not have a clear theoretical basis, our interfirm analysis of gender composition and wages is directly analogous.

Our approach here is to estimate an individual-level hourly wage regression with personal and plant characteristics on the right hand side. The personal characteristics include terms in experience and education, sex, marital status, occupation, and race, and the plant characteristics include total employment, and the female share of establishment employment. Table 6 presents the results of several such regressions. Columns 1 and 2 report results for all workers. The columns differ in that column 1 includes only the variables listed above, while column 2 adds a measure of plant-level labor productivity.<sup>17</sup> Both regressions indicate that, holding characteristics including sex constant, workers earn less if they work in plants with largely female workers. For example, row 2 of column 2 indicates that the representative male working in a 50% female plant earns wages that

---

<sup>17</sup>Labor productivity is defined to be the dollar value of shipments divided by employment.

are more than 10% lower than observationally equivalent males working in all-male plants. The combination of rows 2 and 3 show that the relationship between wages and percent females in the establishment is even greater among women. The regression indicates that the representative woman in a 50% female plant earns wages that are over 15% lower than observationally equivalent women working in plants that are nearly all male. A comparison of columns 1 and 2 shows that these relationships are not much changed if we control for (admittedly crude) measures of labor productivity.

The regressions of columns 1 and 2 control for 1-digit occupation and 2-digit industry (within manufacturing of course), so these results are not driven by the fact that men and women are sorted into broadly different occupations and industries. Nevertheless, it is possible that segregation along finer dimensions of occupation and industry is what drives these results. For example, the quality-sorting hypothesis suggests that, within any 2-digit industry, plants with many women tend to employ workers with relatively low skills. Alternatively, it could be that the mere fact of having female coworkers tends to drive down wages. There is no completely satisfactory way of sorting out these alternative interpretations, but the rest of table 6 makes several crude attempts at doing so.

Columns 3-6 of table 6 present wage regressions where the samples are divided along industrial lines. Columns 3 and 4 report regressions

for industries that are relatively female-intensive (i.e. the fraction of women employees in our sample exceeds 33%), while columns 5 and 6 report regressions for industries that are relatively male-intensive. The motivation for these regressions is that there might be differential scope for quality-sorting depending upon whether women are an important component of the industry's overall workforce. The results of these regressions may be summarized as being not too different from those of columns 1 and 2. Namely, workers with female coworkers earn lower wages than observationally similar workers with predominantly male coworkers.

Columns 7-12 take an analogous approach to samples stratified by broad occupation. The results display substantial variation across occupations. Columns 7 and 8 show that, among managers, there is no wage penalty for male managers that work with predominantly female underlings. However, women managers in predominantly female plants do earn substantially less than women managers in more integrated plants. Columns 9 and 10 show that, among sales and clerical occupations, both men and women are penalized for working in largely female plants, but the size of this penalty is substantially less for women. Despite this variation, there is still a common theme. Wages for both men and women tend to be lowest in those plants with a predominantly female workplace.

These results are somewhat hard to explain with any of the discrimination theories. In Becker's theory of employer discrimination and in Lang's language theory, competition forces the wages of workers

within either sex to be the same across plants regardless of the plants' gender composition. In Becker's theory of coworker discrimination, prejudiced men should receive *higher* wages for working with predominantly female coworkers. None of these predictions are consistent with the data. In contrast, the quality sorting hypothesis suggests that highly female plants pay lower wages simply because their employees, both male and female, tend to have lower skills. This is roughly consistent with what we find here. However, the quality sorting hypothesis suggests that there should be no gender wage gap once we control for labor productivity, which is not what we find. Thus, although our labor productivity measure is crude, the quality sorting hypothesis is not completely consistent with the data either.

As with the results on segregation, we examined the sensitivity of these results to the inclusion of part-time workers in the sample. Not surprisingly, there are some minor changes in coefficients in the models that we estimate. However, the general results are quite unaffected by the inclusion of part-time workers.<sup>18</sup>

## **VI. Conclusions**

This paper has examined the extent and causes of interfirm gender segregation in U.S. manufacturing. A primary finding is that there is

---

<sup>18</sup>Full results from the sample including part-time workers are available from the authors upon request.

substantial interfirm gender segregation, which is consistent with earlier results. However, much of the traditionally measured segregation is attributable to the random allocation of workers to plants. Thus, previous studies may have overstated the systematic component of workplace gender segregation. This is particularly true among white-collar, highly educated workers. A secondary finding is that managers and their subordinates tend to be of the same sex, even within industries and occupations. While this may be partly the result of qualitative differences in human capital acquisition, it is also consistent with theories of discrimination based on animus (Becker) or language (Lang).

We also examined the distribution of men's and women's wages across U.S. manufacturing plants. We found that differential pay across establishments can statistically account for a substantial fraction of the overall gender pay gap. Indeed, among blue collar workers plant fixed effects explain more of the gender pay gap than do a full array of traditional human capital measures. Among other things, this finding suggests that comparable worth policies can have only limited success in reducing the gender pay gap. We also examined the relationship between wages and the fraction of female workers in a plant, holding the worker's sex and other characteristics constant. While there is some variation across occupations and industries, the basic finding is that both men and women earn less when they work in plants that are



predominantly staffed by women. We argue that this finding is not explicable by theories of discrimination. The quality-sorting hypothesis of Macpherson and Hirsch (1995) offers one potential explanation of these facts, as perhaps low-skilled men and women are all concentrated in highly female plants. However, the fact that controlling for labor productivity leaves this relationship largely intact suggests that the quality-sorting hypothesis is not a complete explanation either. We are left with the view that subtle combinations of these theories are required to explain the data, and that perhaps we need some altogether new theories of gender differences in the labor market.

## References

- Becker, Gary S. 1957. The Economics of Discrimination. Chicago: The University of Chicago Press.
- Bergmann, Barbara R. 1986. The Economic Emergence of Women. New York: Basic Books.
- Bielby, William, and James Baron. 1984. "A Woman's Place is With Other Women: Sex Segregation Within Organizations." In Sex Segregation in the Workplace: Trends, Explanations, Remedies, ed. Barbara Reskin, 27-55. Washington, D.C.: National Academy Press.
- Blau, Francine D. 1977. Equal Pay in the Office. Lexington, Mass.: D.C. Heath.
- Brown, Charles, James Hamilton, and James Medoff. 1990. Employers Large and Small. Cambridge, MA: Harvard University Press.
- Buckley, John. 1971. "Pay Differences Between Men and Women in the Same Job." Monthly Labor Review 94(11):36-9.
- Carrington, William J., and Troske, Kenneth R. 1995a. "Gender Segregation in Small Firms." Journal of Human Resources, 30(3): 503-33.
- Carrington, William J., and Troske, Kenneth R. 1995b. "On Measuring Segregation in Small Unit Samples." Working paper, The Johns Hopkins University.
- Doms, Mark, Dunne, Timothy, and Troske, Kenneth R. 1995. "Workers, Wages, and Technology." Working paper, Center for Economic Studies, U.S. Census Bureau, Washington, D.C.
- Donohue, John J. III, and Peter Siegelman. 1991. "The Changing Nature of Employment Discrimination Litigation." Stanford Law Review, 43: 983-1033.
- Duncan, Otis, and Beverly Duncan. 1955. "A Methodological Analysis of Segregation Indices." American Sociological Review 20(2):210-217.
- Ferber, Marianne Abeles, and Joan Althaus Huber. 1975. "Sex of Student and Instructor: A Study of Student Bias." American Journal of Sociology 80(4):949-63.

- Goldberg, P. A. 1968. "Are Women Prejudiced Against Women?" Trans-Action 5:28-30.
- Griffin, Peter, and Stephen Trejo. 1995. "The Measurement of Gender Segregation Using Firm-Level Data." Working paper, California State University, Long Beach.
- Groshen, Erica L. 1991. "The Structure of the Female/Male Wage Differential." Journal of Human Resources 26(3):457-472.
- Hellerstein, Judith, David Neumark, and Kenneth Troske. 1996. "Wages, Productivity, and Worker Characteristics." Working paper, U.S. Census Bureau, Center for Economic Studies.
- Johnson, George, and Gary Solon. 1986. "Estimates of the Direct Effects of Comparable Worth Policy." American Economic Review 76(5):1117-25.
- Kanter, Rosabeth Moss. 1977. Men and Women of the Corporation. New York: Basic Books.
- Kremer, Michael, and Maskin, Eric. 1995. "Segregation by Skill and the Rise in Inequality." Working paper, Massachusetts Institute of Technology, Cambridge, MA.
- Lang, Kevin. 1986. "A Language Theory of Discrimination." Quarterly Journal of Economics.
- LaPiere, Richard T. 1934. "Attitudes vs. Actions." Social Forces 13:230-37.
- McNulty, Donald. 1967. "Differences in Pay Between Men and Women Workers." Monthly Labor Review 90(12):40-3.
- Macpherson, David A., and Barry T. Hirsch. 1995. "Wages and Gender Composition: Why Do Women's Jobs Pay Less?" Journal of Labor Economics 13(3): 426-71.
- Mincer, Jacob, and Solomon Polachek. 1974. "Family Investments in Human Capital: Earnings of Women." Journal of Political Economy 82:S76-S108.
- Neumark, David. 1988. "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination." Journal of Human Resources 23: 279-95.

- Pfeffer, Jeffrey, and Alison Davis-Blake. 1987. "The Effect of the Proportion of Women on Salaries: The Case of College Administrators." Administrative Science Quarterly 32(1):1-24.
- Reilly, Kevin T., and Wirjanto, Tony S. 1995. "Does More Mean Less? The Male/Female Wage Gap and the Proportion of Females at the Establishment Level." Working paper, University of Waterloo.
- Smith, James P., and Finis Welch. 1984. "Affirmative Action and Labor Markets." Journal of Labor Economics 2(2): 269-301.
- Troske, Kenneth R. Forthcoming. "The Worker-Establishment Characteristics Database." In Labor Statistics Measurement Issues, eds. John Haltiwanger, Marilyn Manser, and Robert Topel. Chicago: NBER.

**Table 1**

**Theories of Interplant Sex Segregation**

---

---

<u>Theory</u>	<u>Implications for Interplant Segregation</u>	<u>Implications for the Male/Female Wage Gap</u>
1. Becker's theory of employer discrimination	Employers hire either all women or all men.	Women earn less than men if there are an insufficient number of unprejudiced employers. Wages do not vary by sex of coworkers.
2. Becker's theory of customer discrimination	Employers hire either all women or all men.	Women earn less than men if there are an insufficient number of unprejudiced employers. Wages do not vary by sex of coworkers.
3. Becker's theory of coworker discrimination	Employers hire either all women or all men.	Wages of men are increasing function of female coworker share.
4. Lang's language theory	Employers hire either all women or all men.	Women earn less than men if there are an insufficient number of employers versed in their language. Wages do not vary by sex of coworker.
5. Skill-based segregation	Segregation may arise due to differences in occupational or educational attainment. Within groups of similar skill, however, there should be no systematic segregation.	Skill-based segregation does not lead to a male/female wage gap within groups of workers with similar skills.
6. Random Allocation of workers to employers	Random allocation generates some segregation according to conventional measures.	Random allocation does not lead to a wage gap by itself.

---

---

**Table 2**  
**WECD Sample Means**

<u>Variable</u>	<u>Men</u>	<u>Women</u>
1. Hourly Wage	14.97	9.65
2. Log Hourly Wage	2.59	2.16
3. Annual Earnings	33,436	20,150
4. Log Annual Earnings	10.29	9.79
5. Age	40.5	39.7
6. Potential Labor Market Experience	22.3	21.9
7. Total Employment at Plant	2087	1360
8. Region (%)		
Northeast	26.4	29.3
Midwest	47.6	41.0
South	20.5	24.8
West	6.1	5.5
9. Occupation (%)		
Managers and other professionals	17.7	13.4
Clerical and other non-production workers	8.4	25.3
Sales occupations	5.6	4.9
Operatives and fabricators	36.6	39.8
Precision production, craft, and repair occupations	25.3	10.0
Laborers	6.2	6.6

All data drawn from the Worker-Establishment Characteristics Database.

**Table 3**  
**Indices of Interfirm Sex Segregation**

Sample	Gini Index			Dissimilarity Index		
	(1)	(2)	(3)	(4)	(5)	(6)
	<u>Gini</u>	<u>Expected Gini</u>	<u>Systematic Gini</u>	<u>Dissimilarit y</u>	<u>Expected Dissimilarit y</u>	<u>Systematic Dissimilarit y</u>
1. All workers	.59 (.01)	.24 (.01)	.45 (.01)	.43 (.01)	.16 (.00)	.33 (.01)
<u>Within broad industry</u>						
2. Nondurables	.59 (.01)	.26 (.00)	.45 (.01)	.43 (.01)	.17 (.00)	.31 (.01)
3. Durables	.55 (.01)	.23 (.01)	.41 (.01)	.40 (.01)	.15 (.01)	.30 (.01)
<u>Within selected detailed industry</u>						
4. Meat Products	.43 (.05)	.21 (.02)	.28 (.06)	.30 (.05)	.14 (.01)	.19 (.06)
5. Apparel, excluding knit	.52 (.03)	.35 (.01)	.26 (.04)	.37 (.02)	.23 (.01)	.18 (.03)
6. Newspaper publishing	.40 (.02)	.31 (.01)	.13 (.02)	.29 (.01)	.21 (.01)	.10 (.02)
7. Motor vehicles and equipment	.41 (.03)	.15 (.02)	.31 (.03)	.32 (.02)	.10 (.01)	.24 (.02)
8. Household appliances	.31 (.03)	.12 (.01)	.22 (.03)	.22 (.03)	.08 (.01)	.16 (.03)

Notes: Numbers in parentheses are bootstrap standard errors. See text for description of the samples and indices.

**Table 3 (Continued)**

**Indices of Interfirm Sex Segregation**

Sample	Gini Index			Dissimilarity Index		
	(1) Gini	(2) Expected Gini	(3) Systematic Gini	(4) Dissimilarit y	(5) Expected Dissimilarit y	(6) Systematic Dissimilarit y
<u>Within broad occupation</u>						
9. Prof./managers	.63 (.02)	.54 (.02)	.19 (.01)	.44 (.01)	.36 (.02)	.13 (.01)
10. Sales/service	.79 (.01)	.73 (.01)	.24 (.03)	.59 (.02)	.52 (.02)	.15 (.02)
11. Clerical	.68 (.01)	.59 (.01)	.21 (.02)	.49 (.01)	.40 (.01)	.14 (.01)
12. Craftsmen	.83 (.01)	.56 (.01)	.62 (.02)	.67 (.01)	.38 (.01)	.46 (.02)
13. Operatives	.80 (.01)	.38 (.01)	.68 (.01)	.62 (.01)	.24 (.01)	.50 (.01)
14. Laborers	.83 (.01)	.60 (.01)	.56 (.02)	.65 (.01)	.42 (.01)	.40 (.02)
<u>Within selected detailed occupation</u>						
15. Engineers, Architects, and Surveyors	.75 (.05)	.71 (.05)	.13 (.03)	.54 (.06)	.50 (.06)	.08 (.03)
16. Mechanics and Repairers	.91 (.01)	.84 (.01)	.45 (.05)	.76 (.67)	.67 (.02)	.27 (.05)
17. Precision Production Occupations	.85 (.01)	.61 (.01)	.60 (.02)	.68 (.01)	.42 (.01)	.45 (.02)
18. Textile, Apparel, and Furnishings Machine Operators	.84 (.02)	.50 (.01)	.67 (.03)	.64 (.02)	.34 (.01)	.46 (.03)
19. Machine Operators, Assorted Materials	.83 (.01)	.56 (.01)	.62 (.01)	.66 (.01)	.38 (.01)	.45 (.01)
20. Fabricators, Assemblers and Hand Working Occupations	.84 (.01)	.48 (.01)	.69 (.01)	.67 (.01)	.31 (.01)	.52 (.02)
<u>Within schooling group</u>						
21. < 12 years	.80 (.01)	.51 (.01)	.59 (.01)	.62 (.01)	.34 (.01)	.42 (.01)
22. 12 years	.69 (.01)	.35 (.01)	.53 (.01)	.52 (.01)	.23 (.00)	.38 (.01)
23. 13-15 years	.63 (.01)	.45 (.01)	.32 (.01)	.45 (.01)	.29 (.01)	.23 (.01)
24. > 16 years	.65 (.02)	.55 (.02)	.23 (.02)	.47 (.02)	.37 (.02)	.16 (.01)

Notes: Numbers in parentheses are bootstrap standard errors. See text for description of the samples and indices.



**Table 4**

**Plant-Level OLS Models of Employee Sex Composition**

<u>Independent Variable</u>	Dependent Variable = Female Share of Non-Supervisory Employment			
	(1)	(2)	(3)	(4)
1. Percent female supervisors	.152 (.011)	.124 (.010)	.144 (.011)	.120 (.010)
2. Log of establishment employment	.005 (.003)	.021 (.003)	.005 (.003)	.020 (.003)
3. Percent of non-supervisors with a college degree	-.161 (.028)	-.028 (.026)	-.158 (.029)	-.029 (.027)
4. Percent of establishment workers in managerial occupations	-.039 (.048)	.257 (.045)	-.041 (.048)	.249 (.046)
5. Percent of establishment workers in sales occupations	.103 (.044)	.191 (.041)	.114 (.044)	.193 (.041)
6. Percent of establishment workers in craft occupations	-.131 (.041)	-.049 (.037)	-.150 (.041)	-.053 (.038)
7. Log of average hourly wages of non-supervisory employees	-	-.347 (.012)	-	-.342 (.012)
8. Log of average hourly wages of supervisory employees	-	.035 (.010)	-	.035 (.010)
9. Log of labor productivity (x 100)	-	-	-.010 (.002)	-.003 (.002)
10. 2-digit industry dummies	yes	yes	yes	yes
11. Region dummies	yes	yes	yes	yes
12. Number of plants in sample	4465	4464	4358	4358
13. R-square	.337	.441	.346	.443

Notes:

1. All data drawn from the Worker-Establishment Characteristics Database.
2. All regressions included controls for the average age and education of supervisors, as well as the average age of non-supervisory employees.
3. To be in the sample for this table, establishments had to both have at least one supervisor and one non-supervisor in the WECD.

**Table 5**

**Decomposing the Male/Female Wage Gap into Within- and Between-Plant Components**

	Order of the Decomposition								
	Step 1: $Y = X'\beta + u_1$ Step 2: $Y - X'\beta = Z'\gamma + u_2$			Step 1: $Y = Z'\gamma + u_1$ Step 2: $Y - Z'\gamma = X'\beta + u_2$			Step 1: $Y = X'\beta + Z'\gamma + u_2$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	<u>Total</u>	<u>White Collar</u>	<u>Blue Collar</u>	<u>Total</u>	<u>White Collar</u>	<u>Blue Collar</u>	<u>Total</u>	<u>White Collar</u>	<u>Blue Collar</u>
1. Unadjusted Male/Female Wage Gap ( $Y_m - Y_w$ )	.429	.439	.446	.429	.439	.446	.429	.439	.446
2. Male/Female Wage Gap Adjusted for Personal Characteristics ( $Y_m - X_m'\beta$ ) - ( $Y_w - X_w'\beta$ )	.317	.280	.371	.366	.344	.416	.345	.302	.400
3. Male/Female Wage Gap Adjusted for Plant Fixed Effects ( $Y_m - Z_m'\hat{\gamma}$ ) - ( $Y_w - Z_w'\hat{\gamma}$ )	.342	.332	.194	.230	.285	.143	.275	.352	.197
4. Male/Female Wage Gap Adjusted for Personal Characteristics and Plant Fixed Effects ( $Y_m - X_m'\beta - Z_m'\hat{\gamma}$ ) - ( $Y_w - X_w'\beta - Z_w'\hat{\gamma}$ )	.230	.173	.119	.167	.190	.123	.191	.214	.151

Notes:

1. Y = hourly wages
2. X = worker characteristics including flexible terms in experience and education, race, marital status, 10 occupation dummies, 4-digit industry dummies, MSA, region, MSA x region.
3. Z = a set of plant fixed effects.
4. In addition to our previous restrictions, we required that each individual be in a plant with at least three people in our sample. This reduced the sample size by approximately 10%.

**Table 6**  
**Individual-Level Models of Hourly Wage Determination**

<u>Independent Variable</u>	Sample											
	<u>By Industry</u>						<u>By Occupation</u>					
	<u>All Workers</u>		<u>Female-Intensive</u>		<u>Male-Intensive</u>		Managerial		Sales		Laborers	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1. Female	-.083 (.010)	-.086 (.010)	-.090 (.016)	-.091 (.016)	-.083 (.015)	-.087 (.015)	-.072 (.017)	-.074 (.017)	-.111 (.014)	-.109 (.015)	-.050 (.015)	-.054 (.015)
2. Female share of establishment employment	-.214 (.021)	-.190 (.021)	-.254 (.028)	-.218 (.028)	-.188 (.030)	-.169 (.029)	.025 (.026)	.031 (.025)	-.178 (.026)	-.159 (.026)	-.297 (.025)	-.268 (.025)
3. Female x female share of establishment employment	-.091 (.022)	-.090 (.021)	-.047 (.030)	-.053 (.029)	-.113 (.039)	-.108 (.037)	-.164 (.036)	-.162 (.036)	.071 (.027)	.064 (.026)	-.128 (.029)	-.127 (.029)
4. Log of establishment employment	.068 (.004)	.065 (.004)	.062 (.006)	.060 (.006)	.071 (.005)	.068 (.005)	.035 (.003)	.034 (.003)	.056 (.003)	.054 (.003)	.083 (.005)	.079 (.005)
5. Labor productivity (x 1000)	-	.258 (.028)	-	.312 (.040)	-	.233 (.030)	-	.105 (.028)	-	.219 (.029)	-	.301 (.033)
6. R-square	.497	.504	.469	.475	.448	.457	.412	.414	.427	.433	.477	.488
7. Number of observations	169,099		63,950		105,149		27,512		31,470		110,117	

Notes:

1. The dependent variable is the log of the hourly wage, which is defined as annual earnings divided by annual hours of work.
2. Each regression also included as regressors the following variables: a quartic in experience, five dummy variables for educational attainment, interactions of all education and experience terms, and dummy variables for race, marital status, 1-digit occupation, region, msa residence, and 2-digit industry.
3. Standard errors have been corrected for heteroscedasticity and for the clustered sample design.
4. All data drawn from the Worker-Establishment Characteristics Database (WECD).