Long-Term Trends in Occupational Segregation by Sex

Jerry A. Jacobs
University of Pennsylvania

This paper reexamines long-term trends in occupational segregation by sex, using the double-coded 1900 and 1910 Public Use Samples. The analysis addresses the ambiguity in the measurement of long-term trends that arises from using inconsistent or highly aggregated occupational classifications. The revised measures indicate that occupational segregation by sex remained quite constant from 1900 through 1970, although segregation in nonfarm occupations declined slowly. Occupational sex segregation declined between 1970 and 1980 and continued to decline through 1986. An accurate assessment of historical trends is a necessary starting point for theoretical explanations of occupational sex segregation.

Occupational sex segregation has become an increasingly important area of multidisciplinary research. In 1985, over two-thirds of working women were employed in occupations that were 70% or more female. About 25% of the gender gap in wages has been attributed to the sex segregation of occupations. Research interest in this area has been spurred by the rise in women's labor-force participation (Goldin 1983; Smith and Ward 1984), the women's movement (Klein 1984), and most recently by the public debate surrounding comparable-worth proposals (Remick 1984).

In 1968, Gross reported data that documented the remarkable stability of occupational segregation by sex since the turn of the century. This important finding set the stage for much subsequent research. The long-
term stability of sex segregation is frequently the starting point for theoretical discussions in this area (Bielby and Baron 1986; Greenberger and Steinberg 1983; Oppenheimer 1970; Marini and Brinton 1984; Goldin 1983). This finding has raised serious questions for the human capital perspective in economics, as well as for the modernization approach in sociology (Polachek 1979; Smelser 1959). Whereas both these views predicted increasing integration as women’s labor-force participation increased, the overall trends appeared to show no such change. Two lines of reasoning in the human capital framework lead to a prediction of declining sex segregation. One, shared with the modernization perspective in sociology, holds that discrimination (ascriptive criteria) will be eroded by the competitive operation of the marketplace because discrimination is held to cost those who discriminate (Becker 1957). If a component of occupational segregation by sex is due to discrimination, the human capital model predicts its demise. But the human capital model also predicts the decline of voluntary segregation that may result from women’s choices. Increasing wages lead more women into the labor force (Mincer 1962) for longer spells (Smith and Ward 1984), which in turn leads them to invest more heavily in labor-market skills (Polachek 1979), which in turn should reduce the sexual differentiation of occupational choices. The alternative views, which emphasize the continuity of labor-market structures or the interest of men in preserving patriarchy, take the long-term stability of sex segregation as one of the important patterns to be explained (Strober 1984; Milkman 1987; Cohn 1985; Reskin 1988). While no single datum is decisive for the testing of different theories because, for apparently anomalous results, each theory can offer a variety of accounts, evidence on the long-term stability of sex segregation has constituted important grist for many theoretical mills.

Unfortunately, the inconsistency of occupational classifications from census to census has raised questions about the accuracy of this historical picture. This issue has been debated by students of sex segregation without an entirely satisfactory conclusion. In this article, I reexamine the question of long-term trends in occupational sex segregation with new data and assess the implication of my results for different theories of sex segregation.4

MEASURING LONG-TERM TRENDS
Gross (1968) analyzed changes in sex segregation for each decennial census from 1900 to 1960. He reported that in 1900 66.9% of employed

4 For a more extended analysis of these issues and related measures of segregation, see Jacobs (in press).

161
women would have had to change occupations for women to be distributed in the same manner as men. Gross found that this level remained at a nearly identical level (68.4%) in 1960. Thus, despite vast changes in the economy, the technology of production, and the mix of employment and despite the transition from a primarily agricultural society to a postindustrial society, Gross’s data indicated that fundamental differences in work performed by men and women have persisted.

Williams (1979) criticized Gross for failing to standardize categories across censuses. Because each census employed a different system of coding occupations, Williams suggested that Gross’s intercensal comparisons were likely to be flawed because of changes in occupational classifications. Williams maintained that increasingly detailed occupational categories in recent censuses tended to inflate the measures of segregation and consequently suppress the pattern of increasing integration. Williams aggregated occupations into larger, temporally consistent categories and found slow but consistent declines in sex segregation over time. England (1981), however, noted that Williams’s approach exaggerates declines over time because it ignores real occupational differentiation that has taken place. Williams’s reliance on aggregated categories ignores the emergence of new, segregated occupations, an especially important consideration over extended time periods. England maintained that while Gross’s approach tends to understate change, Williams’s tends to exaggerate change. She concluded that while there have been changes over the course of the century, the precise magnitude of these changes is hard to estimate. Most students of sex segregation have agreed that the changes during the century have been small, yet the evidence supporting this conclusion has not been entirely satisfactory.

Fortunately, data from the 1900 and 1910 censuses enable me to reexamine this conclusion. Data on samples of individuals in the manuscript censuses have been culled, computerized, and made available for research. The occupations of employed individuals in the 1900 sample have been double-coded, once with the 1900 occupational classification and once with the 1950 census system. The 1910 sample occupations have been double-coded with the 1910 and 1980 occupational categories. These computerized samples allow for a more detailed and accurate reclassification of the occupational data of early censuses than the published data permit. Whereas change in the occupational classification of published data inevitably involves imposing greater aggregation, the availability of the original survey responses allows for the creation of finer distinctions.

The double-coding strategy circumvents the problem of changing classifications. The effect of different occupational classifications can be estimated by calculating indices of segregation for 1900 using both the detailed 1900 and the detailed 1950 census categories. If one obtains es-
sentially the same results by using the more recent, more detailed classification scheme as had been obtained with the contemporaneous occupational codes, then one can be reasonably sure that the results are not artifacts of the coding. I compare temporally distant years with the more recent coding system, providing comparability without sacrificing detail in measurement. I compare 1900 and 1950 by using the 1950 coding system for both years and compare 1910, 1970, 1980, and 1986 by using the 1980 coding system for each year. This procedure allows for the best possible assessment of long-term trends. It addresses Williams's concern by relying on the same categories for all years. It also addresses England's concern by using the most detailed and most recent categories: newly developed occupations simply appear empty in the earlier census.

The following examples of changing occupational classifications highlight the problem of comparability. In 1900 and 1910, barbers and hairdressers were combined in a single category, while in the 1980 classification system, these were distinct occupations. Similarly, in 1900 and 1910, elementary and high school teachers were combined in a single category, whereas in 1980, elementary and secondary teachers were reported separately. In both these examples, the combined category would appear more integrated by sex since the representation of women in one group (hairdressers and elementary teachers) is higher than in the other (barbers and high school teachers). Clearly, these are instances of change in the classification system that are not prompted by changes in the nature of work. One would prefer to compare the two time points with the same occupational classification system.

However, one should note that in other cases the use of an occupational classification system developed for a later period potentially poses a concern about the validity of the measure for the earlier time period. In the 1900 occupational classification system, "broom and brush makers," "corset makers," "straw workers," and "well borers" are listed as separate occupations. In subsequent censuses, these titles were lost as these jobs declined or were subsumed under other titles. As the nature of work changes over time, there may be no ideal way to compare occupational distributions at different time periods. Dropped titles occur mostly in manufacturing employment; there is greater continuity of titles among the professions and craft occupations.

Two points should be noted concerning the issue of the validity of classifications. First, this problem is more serious in knowing the nature of the work than it is in measuring sex segregation. My examination of the occupational titles phased out in later censuses does not suggest that I am significantly misrepresenting the sex composition of occupations in the earlier time period. Second, the use of a variety of classification systems for the same period minimizes this problem; if they all give the same
answer, we can be reasonably assured that the result is not an artifact of coding.\textsuperscript{5}

A clarification of a related concern is in order. I do not claim that the nature of the work is the same in an identically titled occupation in 1910 and 1980. This continuity would be the exception, not the rule. The question I am attempting to answer is whether the sex composition of employment has been stable, for which purpose the data mentioned above will provide the best available answer.

A steep and consistent decline in sex segregation since the turn of the century would lend support to the human capital and modernization perspectives, which predict that the unfettered operation of the market economy undermines discrimination and promotes the ascendancy of universalistic criteria. In contrast, the long-term stability of sex segregation will support structuralist perspectives on sex segregation. Structuralists argue that sex segregation tends to be self-sustaining since, once institutionalized, sex-typed employment patterns are hard to eradicate (Roos and Reskin 1984). While these results will not constitute a decisive test between theories of sex segregation, this analysis is nonetheless essential in providing an accurate starting point for theoretical explanations.

DATA AND METHODS

As I noted above, in order to assess the effect of changing occupational classifications, I examine sex segregation in 1900 with both the 1900 and the 1950 census occupational classification systems. Similarly, I measure the degree of segregation in 1910 with both the 1910 and 1980 census occupational classification systems. In addition, I compare 1910 with 1970, 1980, and 1986, using the 1980 occupational codes in each case.

Data regarding the labor force in 1900 are obtained from the 1900 United States Census Public Use Sample. A sample of 100,438 individuals was drawn from the 1900 manuscript census, including 37,452 individuals with codable occupational information (30,584 men and 6,868 women). Details of the sampling design and occupational coding proce-

\textsuperscript{5} A related problem that has been suggested is that using the classification system for the later period would result in the imposition of artificial distinctions on the earlier period. This issue is of more theoretical than practical concern. Ann Miller, who directed the occupational coding for the 1910 project, indicates that this problem did not arise (personal communication). In any event, coders often do not have sufficient details to make artificial distinctions, even if they were called for. Indeed, the larger problem is that the data do not always allow the coder to make all the distinctions called for. However, this is true for contemporary censuses as well. The problem is handled the same way for contemporary and historical coding; when insufficient information is available to make a distinction between detailed occupations, the individual is placed in the appropriate “not elsewhere classified” category.
dures are described in the 1900 Public Use Sample User's Handbook (Graham 1980). Data for 1910 are obtained from the 1910 Public Use Sample, compiled by a research team at the University of Pennsylvania (Strong 1988). The sample of 366,239 individuals from the 1910 manuscript census includes 120,738 men and 32,886 women with codable occupations. Data for 1980 are obtained from the 1980 census (U.S. Bureau of the Census 1984). A sample from the 1970 census was assigned 1980 occupational codes to facilitate historical comparisons. The data were published as part of the 1980 census volume. And 1986 data are obtained from annual average Current Population Survey data (U.S. Bureau of Labor Statistics 1987).

Special attention needs to be paid to several characteristics of the 1900 labor force. One important consideration in examining occupational segregation at the turn of the century is the size of the agricultural sector. In 1900, 37.5% of the labor force was employed in the agricultural sector, a figure that declined to 2.9% by 1980.

There are several reasons one might choose to exclude agriculture from an analysis of long-term trends. First, there may be more measurement error in the classification of men and women in farm work. The extent of farm women's labor-force participation has been a matter of debate among economic historians for some time (Oppenheimer 1970). Further, occupational distinctions between farm men and women may be artificial since census coders may classify women as unpaid family workers, while men are classified as farmers, no matter how the work is actually divided. Perhaps more important, census takers group most farm employment into a few broad categories, which tends to reduce the measured level of segregation.6

Second, theories about trends in occupational segregation apply most specifically to competition in industrial society and may not apply with equal force to the declining agricultural sector. The degree to which competitive forces drive employers to find the lowest-wage worker for a given job applies more to the industrial economy because the contribution of family members to farm work is often outside the cash nexus of the broader economy.7 The degree of sex segregation on family farms is as

---

6 The coding rule in 1900 was to allocate wives, sons, and daughters of farmers to the category "farm laborers, family workers." Women could be classified as "farm owners" if they were heads of household, or as "farm laborers, wage workers" if they were unrelated to the farm owner. In the 1900 Public Use Sample, 440 women were listed as farm owners, managers, or foremen (vs. 7,169 men); 498 were listed as family workers (vs. 2,458 men); and 383 were listed as farm laborers, wage workers (vs. 2,571 men). The distinctions, of course, do little to clarify the actual sexual division of labor on farms.

7 This is not to deny the competitive forces working in agriculture. But one might argue that the force of tradition is stronger in agriculture.
much a matter of the division of household labor as it is an indicator of the degree of occupational sex segregation (Rosenfeld 1985).

Third, even if farms were theoretically interesting in this context, their very size and rapid decline would raise the possibility of testing a compositional effect on long-term trends. If nonfarm sex segregation were to decline over time, then we could conclude that this decline in sex segregation was masked by the rapid decline of the relatively highly integrated farm sector.

A second feature of the 1900 labor force that should be noted is the presence of child labor. In 1900, child labor remained a part of the American labor force, with occupational statistics reported for individuals aged 10 and above. An analysis of subsequent periods must adjust for the decline in child labor. While I present results that include children and agricultural workers, I also report figures for individuals above the age of 16 in nonfarm occupations.

Most research on occupational segregation by sex has employed one measure of segregation, the index of dissimilarity ($D$) (for other approaches, see Massey and Denton 1989; and Jacobs 1985, 1986, in press). This measure taps the main dimension of segregation, namely, the degree to which two groups are unevenly distributed over a set of categories. The index of dissimilarity has a convenient interpretation: it represents the proportion of women who would have to change occupations in order to be distributed in the same manner as men. The measure $D$ is symmetrical; that is, the same proportion of men would have to change occupations to be distributed in the same manner as women.

When trends are examined over time, the index of dissimilarity is often supplemented with a size-standardized index of dissimilarity. This index indicates what the degree of segregation would be if all occupations were of equal size. By holding the mix of occupations constant, this measure indicates what changes over time would have occurred if there had been no change in the relative sizes of the occupations.

8 The formula for $D$ is:

$$D = \frac{1}{2} \sum_{i=1}^{n} [(W_i/W) - (M_i/M)] \times 100,$$

where $W_i$ is the number of women in occupation $i$, $W$ is the total number of women, $M_i$ is the number of men in occupation $i$, $M$ is the total number of men, and $n$ is the number of occupations. The measure $D$ varies from 0 to 100, with 0 representing a perfectly even distribution and 100 representing perfect segregation.

9 The formula for the size-standardized index of dissimilarity is:

$$D_{\text{standardized}} = \frac{1}{2} \sum_{i=1}^{n} \left[ \frac{(W_i/T_i)/\Sigma(W_i/T_i)}{\Sigma(M_i/T_i)} - \frac{(M_i/T_i)/\Sigma(M_i/T_i)}{\Sigma(M_i/T_i)} \right] \times 100,$$
RESULTS: SEX SEGREGATION IN THE UNITED STATES SINCE 1900

Table 1 reports comparisons of occupational segregation by sex for 1900 and 1950. For 1900, results using both the 1950 occupational classification and the 1900 classification are reported. In addition, table 1 compares results that include farm and child labor with those that exclude these groups.

Gross (1968) reported that the level of occupational segregation (D) between men and women in 1900 was 66.9. The comparable figure for the 1900 Public Use Sample, when the 1900 coding scheme is used, is 64.9. The difference between these figures may be due to sampling variability and may also reflect small differences in the way the 1900 classification rules were applied by the Public Use Sample coders. When one employs the 1950 coding scheme, the level of sex segregation is 65.0. The difference between the 1900 and 1950 coding schemes in the portrayal of sex segregation in 1900 is a trivial one-tenth of one percent, 64.9 versus 65.0. The 1950 figure was 67.3, indicating virtually no change over this long period. Thus, the historical trends Gross reported were almost exactly correct despite the occupational classification problem. This analysis indicates that available detailed and comparable occupational categories confirm the long-term stability documented by Gross.

As indicated above, long-term comparisons should properly focus on the nonfarm labor force above the age of 16. When the agricultural sector and child labor are removed from the analysis, the results indicate that occupational segregation by sex has in fact been declining slowly. In 1900, D was 75.9 for individuals in nonfarm occupations, a figure that declined to 66.9 in 1950 for the nonfarm occupations. While Gross's estimate of no significant change through 1950 is correct, the analysis of nonfarm occupations gives a somewhat more optimistic picture of the long-term trends in this area.

Gross reported that the size-standardized measure of segregation declines gradually over time. The size-standardized index of dissimilarity declined from 78.5 in 1900 to 61.3 in 1950. This statistic indicates that the more integrated occupations declined in size in this period relative to the more segregated occupations. In other words, had the size distribution of occupations in 1900 persisted, sex segregation would have declined by 1950. That trends in the sizes of occupations inhibited the decline in sex segregation is clearly evident from these results and from those obtained by Gross.

where \( T_i \) is the total number of men and women in occupation \( i \), and other terms are defined as in n. 8 above.

\(^{10}\) My figure corresponds with the figure Williams (1979) reports for 1950, which is just slightly higher than that reported by Gross (1968).
### TABLE 1

**INDICES OF DISSIMILARITY, 1900 AND 1950**

<table>
<thead>
<tr>
<th>YEAR</th>
<th>DATA</th>
<th>OCCUPATIONAL CODING SYSTEM</th>
<th>ENTIRE LABOR FORCE</th>
<th>NONFARM LABOR FORCE OVER AGE 16</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$D$</td>
<td>$D_{\text{standardized}}$</td>
</tr>
<tr>
<td>1900</td>
<td>Published data*</td>
<td>1900</td>
<td>66.9</td>
<td>70.3</td>
</tr>
<tr>
<td>1900</td>
<td>Public Use Sample</td>
<td>1900</td>
<td>65.0</td>
<td>75.1</td>
</tr>
<tr>
<td>1900</td>
<td>Public Use Sample</td>
<td>1950</td>
<td>64.9</td>
<td>78.5</td>
</tr>
<tr>
<td>1950</td>
<td>Published data†</td>
<td>1950</td>
<td>67.3</td>
<td>61.3</td>
</tr>
</tbody>
</table>

† U.S. Bureau of the Census (1953, table 124).
Occupational Segregation

The present results suggest that much of this compositional effect is due to the fact that one very large and declining occupational group, agriculture, had a relatively low level of segregation in 1900. The decline in the unstandardized index of dissimilarity obtained when farming is excluded from the analysis is about half as large (9 points, from 75.9 to 66.9) as that obtained in a size-standardized analysis of changes over time (17.2 points, from 78.5 to 61.3). (Results for the size-standardized measure for the nonfarm occupations are quite similar to those for the entire labor force. This is true because the removal of a few occupations makes little difference when each occupation is accorded the same weight.)

It should also be noted that the level of sex segregation was quite low among children in the labor force. In 1900, the index of dissimilarity among those under age 16 was 44.0, much lower than that obtained for the labor force as a whole, and perhaps lower than among present-day teenagers (Greenberger and Steinberg 1983). The higher level of segregation found for the nonfarm labor force over the age of 16, then, is partly due to the removal of the relatively less segregated child laborers.

Table 2 presents results for 1910 using three different types of data: published 1910 census data using 1910 occupational codes; 1910 Public Use Sample data using 1910 occupational codes; and 1910 Public Use data using 1980 codes. There is a slightly greater range of results for 1910 than for 1900 (63.0–69.0), but the substantive conclusion that approximately two-thirds of men and women working in 1910 would have had to change occupations to be evenly distributed is corroborated by each of these measures. Further, these measures are all quite close to those obtained for 1900 with two different occupational classification systems.

Table 2 also brings trends in sex segregation up to date, using 1980 occupational codes to compare 1910, 1970, 1980, and 1986. Between 1910 and 1970, occupational segregation by sex remained constant. However, men and women in nonfarm occupations did experience a decline in sex segregation, with the index of dissimilarity declining from 74.3 in 1910 to 67.5 in 1970 to 59.7 in 1980. Sex segregation in nonfarm occupations declined by roughly one point per decade, on average, between 1910 and 1970.

The nonfarm trends moved in the same direction as the size-standardized trends. The size-standardized index of dissimilarity declined from 73.8 in 1910 to 61.3 in 1970, a drop of 12.5 points. The drop in the size-standardized measure is roughly twice as large as that in the nonfarm labor force (6.8 points, from 74.3 to 67.5), as was also the case in the 1900 and 1950 comparison. The long-term decline in agricultural employment

11 They report high levels of sex segregation among teenagers, but do not report a specific index of dissimilarity.
TABLE 2


<table>
<thead>
<tr>
<th>Year</th>
<th>Data</th>
<th>Occupational Coding System</th>
<th>Entire Labor Force</th>
<th>Nonfarm Labor Force over Age 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>1910</td>
<td>Published data*</td>
<td>69.0</td>
<td>68.1</td>
<td>71.4</td>
</tr>
<tr>
<td>1970</td>
<td>Published data*</td>
<td>67.0</td>
<td>70.5</td>
<td>74.9</td>
</tr>
<tr>
<td>1980</td>
<td>Published data†</td>
<td>63.0</td>
<td>73.8</td>
<td>74.3</td>
</tr>
<tr>
<td>1986</td>
<td>Published data‡</td>
<td>67.6</td>
<td>61.3</td>
<td>61.2</td>
</tr>
</tbody>
</table>

inhibited declines in overall sex segregation and, by itself, represents a significant component of the compositional effect documented with the size-standardized measure.

The gradual decline in sex segregation that began in the 1970s continued into the 1980s, with the index of dissimilarity dropping to 57.3 by 1986. This decline thus represents a continuing trend and is not a temporary, artifactual phenomenon caused by a limited set of occupations moving from male-dominated to female-dominated fields (Reskin and Roos, in press).

DISCUSSION

Four different occupational classification systems are employed in this analysis: 1900 and 1950 occupational coding systems are used for measuring segregation in 1900, and 1910 and 1980 systems are used for measuring 1910. All four classification systems produce approximately the same levels of sex segregation for the 1900–1910 period, ranging from a low of 63.0 to a high of 69.0. There can be little doubt about the level of occupational sex segregation at the turn of the century or its continuation for most of the century. Of course, the levels of segregation at the firm or job level cannot be addressed by these data.

My reassessment of occupational trends since the turn of the century reinforces conventional wisdom in some ways and modifies it in others. Gross's (1968) conclusions about the long-term stability in sex segregation are corroborated and are not artifacts of changing classification. However, nonfarm segregation has declined somewhat since 1900. Nonfarm segregation was about 10 points higher than overall segregation in 1910. There was approximately a seven-point decline in segregation between 1910 and 1970 and another seven-point decline since 1970. In other words, there was roughly as much change in nonfarm sex segregation between 1910 and 1970 as there was between 1970 and 1980.

These results do not constitute definite support for any theory of sex segregation. Human capital theorists might point to the recent changes as evidence that the market is responsive to changing investments of women, while structuralists might point to the continuity through 1970 as evidence of the persistence of discriminatory barriers against women.

The present results seem consistent with Kessler-Harris's (1982) view of change in women's working conditions over time. She argues that over the long term, change has been incremental but concludes that there are potentially radical consequences of incremental change. In short, the glacial pace of change in sex segregation through 1970 has given way to modest but consistent change since that time. While the emphasis on continuity has been an important strength, the structuralist perspective
should attempt to identify conditions under which the degree of sex segregation would be expected to change (Baron, Mittman, and Newman 1988).

REFERENCES


