Journal of Econometrics 61 (1994) 133–146. North-Holland

Measuring occupational segregation

Summary statistics and the impact of classification errors and aggregation

Joseph Deutsch
Bar-Ilan University, 52900 Ramat-Gan, Israel

Yves Flückiger
Université de Genève, Geneva, Switzerland

Jacques Silber
Bar-Ilan University, 52900 Ramat-Gan, Israel

The study of occupational segregation is shown to be related to that of income inequality, the gender ratio in each occupation playing the role of individual income. Measures of the dispersion, skewness, kurtosis, and concentration of the distribution of this gender ratio are suggested, and an empirical illustration based on Swiss data for the period 1950–1980 is given. Finally bootstrap techniques are used to check the impact of classification errors and aggregation on the measurement of occupational segregation.

Key words: Bootstrap; Duncan; Gini; Occupation; Segregation

JEL classification: C43; J24; J71

1. Introduction

Several studies [Butler (1987), Silber (1989), Hutchens (1991)] have suggested recently that some of the tools applied in the income inequality literature could be applied to the study of occupational segregation. In fact, Duncan and Duncan (1955) were probably the first to observe such a similarity since they derived their concept of the Segregation Curve from the traditional Lorenz Curve which is so popular in the income inequality literature.

Correspondence to: Joseph Deutsch, Department of Economics, Bar-Ilan University, 52900 Ramat-Gan, Israel.

The present paper reports an additional attempt to show the similarity between the study of income inequality and that of occupational segregation by gender. In the next section we derive the exact expressions which give the values of indices measuring respectively the dispersion, the concentration, the asymmetry, and the flatness of the distribution of the ratios of female over males in the various occupations, since this ratio plays the role that income plays when measuring income inequality. An application of the various summary indices to Swiss data during the period 1950–1980 is presented in section 3.

In section 4 an attempt is made, using bootstrap methods, to check the sensitivity of the various summary indices to errors in the classification of the individuals in the various occupations and confidence bounds on the various summary statistics are derived. Section 5 finally takes a look at the impact of the aggregation of occupational categories and stresses the differences which exist in this respect between the analysis of income inequality and that of occupational segregation.

2. Summary measures of the occupational distribution of the sex ratio

The most common measure of Occupational Segregation is certainly Duncan and Duncan's (1955) famous Dissimilarity Index D, which is usually defined as

$$D = 0.5 \sum_{i=1}^{n} \left| \frac{F_i}{F} - \frac{M_i}{M} \right|, \tag{1}$$

where F_i and M_i are, respectively, the number of females and males in occupation i, n is the number of occupations, and F and M are the total number of females and males in the labor force.

Silber (1989b) has shown that the index D could be expressed as

$$D = 0.5 \sum_{i=1}^{n} \frac{M_i}{M} \left| \frac{F_i/M_i}{F/M} - 1 \right|, \tag{2}$$

which shows that the Duncan and Duncan Index is equal to a weighted relative mean deviation of the sex ratios F_i/M_i, the weights being equal to the shares M_i/M.[1]

Silber then proposed, extending earlier work by Butler (1987), to use expression (2) to define an alternative measure of occupational segregation, which would be equal to the weighted Gini Index G_S of the sex ratios F_i/M_i, where

$$G_S = 0.5 \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{M_i}{M} \frac{M_j}{M} \left| \frac{F_i/M_i - F_j/M_j}{F/M} \right|. \tag{3}$$

[1] D may naturally be also expressed as the weighted relative mean deviation of the ratios M_i/F_i, the weights being equal to F_i/F.

It appears therefore, whether one uses expressions (2) or (3), that to study occupational segregation by gender, one has to analyze, in one way or another, the distribution of the sex ratios F_i/M_i. But care should be used in defining the relative frequencies of the ratios F_i/M_i: they have to be equal to the shares M_i/M so that the weighted mean of the distribution is equal to the overall sex ratio F/M. Such an approach leads to the following list of measures summarizing the distribution of sex ratios F_i/M_i.

2.1. Location

The weighted arithmetic mean μ will be equal to

$$\mu = \sum_{i=1}^{n} \frac{M_i}{M} \frac{F_i}{M_i} = \frac{F}{M}. \tag{4}$$

The median $\tilde{\mu}$ will be defined as $G^{-1}(0.5)$, where G^{-1} is the inverse of the distribution function[2] $G(x)$ where x is equal to the sex ratio F_i/M_i.

2.2. Dispersion

Instead of using the weighted standard deviation of the ratios F_i/M_i, we suggest the use of the weighted mean deviation Δ, which is related to the segregation index G_S and is defined as

$$\Delta = 2\mu G_S = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{M_i}{M} \frac{M_j}{M} \left| \frac{F_i}{M_i} - \frac{F_j}{M_j} \right|. \tag{5}$$

2.3. Skewness

One way to measure the skewness of the distribution of the ratios F_i/M_i is to use Pearson's coefficient β_3 [see Kendall and Stuart (1969) for a definition of this index]. The latter however has no clear upper bound so that one may prefer to use the coefficient of asymmetry A_G which was recently proposed by Berrebi and Silber (1987). A_G varies between -1 and $+1$, is related to Gini's Concentration Index and therefore to G_S, and, as shown by Berrebi and Silber (1987), may be expressed in a way very similar to that in which Pearson's β_3 is defined.

The simplest way however to define A_G is to write it as

$$A_G = 0.5 \frac{\Delta_R - \Delta_L}{\Delta}, \tag{6}$$

[2]The distribution function gives evidently the cumulative values of the shares M_i/M, where the occupations are ranked by increasing value of the ratios F_i/M_i.

where Δ_R and Δ_L are the mean deviations of the ratios F_i/M_i which are respectively greater and smaller than the median $\tilde{\mu}$, Δ being equal to the mean deviation of the whole distribution of the ratios F_i/M_i.

2.4. Flatness

To measure the Kurtosis of the distribution of the ratios F_i/M_i, one solution is to adopt Pearson's approach and use his β_4 index [see Kendall and Stuart (1969)]. However, in this case also Berrebi and Silber (1989) have suggested the use of an alternative index, K_G, which is bounded by 0 and 1, is related to Gini's Concentration Index and therefore to G_S, and has been shown to have several features in common with the coefficient β_4.

The simplest way to express this new measure of Flatness is to express it [see Berrebi and Silber (1989)] as

$$K_G = 0.5 \frac{\Delta_R + \Delta_L}{\Delta}, \tag{7}$$

where Δ_R, Δ_L, and Δ have been previously defined.

2.5. Concentration

Here we will naturally use the Segregation Index G_S, which has been defined in expression (2) and is an extension of the Gini Index, to the study of occupational segregation by gender.

Silber (1989b) has shown that this index G_S could be written also as

$$G_S = [\ldots M_i/M \ldots] G [\ldots F_i/F \ldots], \tag{8}$$

where $[\ldots M_i/M \ldots]$ and $[\ldots F_i/F \ldots]$ represent, respectively, row and column vectors of the shares M_i/M and F_i/F, ranked each by decreasing value of the ratios F_i/M_i, while G is a square matrix whose typical element g_{ij} is equal to 0 if $i = j$, -1 if $i < j$, and $+1$ if $i > j$ [see Silber (1989b) for more details on the use of the matrix G].

Since it is known that the Gini Index is equal to twice the area lying between the Lorenz Curve and the diagonal, we can derive from the definition of the Segregation Index G_S in (8) what Duncan and Duncan (1955) called a Segregation Curve. The latter is a plot of the cumulative values of the shares M_i/M (measured on the horizontal axis) and of the shares F_i/F (measured on the vertical axis), both shares being ranked by increasing value of the ratios F_i/M_i. The Segregation Index is then, like the Gini Index, equal to twice the area lying between the diagonal and the Segregation Curve. An illustration of the use of these summary indices and of the Segregation Curve will be given in the following section.

3. The occupational distribution of the gender ratio in Switzerland

In Switzerland, as in many other developed countries, the labor force participation rate of women has increased during the past thirty years, from 51.0% in 1960 to 58.5% in 1990 [OECD (1991)]. Despite this increase there are still large differences betwen men and women. For instance, in 1990 out of 426 potential professional training programs, 92% of the female apprentices were concentrated in only nine professions and 50% of them choose to be either a 'business clerk' or a 'salesman'. At the university level, if we look at the distribution of women across disciplines, we observe that, despite the considerable increase of the proportion of female students over the last twenty years, 43.8% of female students are in social sciences and only 14.6% in the technical sciences. Finally, we may notice also that 73.7% of the part-time jobs are done by women.

When looking at inequality by gender in Switzerland, one has to stress that it is only in 1981 that the Federal Constitution has been modified to introduce the principle of equality between men and women (more precisely the 'equal pay for equal work' principle; art. 4, al. 2 of the Constitution states also that the legislation should provide equality between the genders at work, in education, and in the family). Moreover, different from many other countries (such as the United States, the Scandinavian countries, France, or the United Kingdom), Switzerland has not yet adopted 'equal employment opportunity' measures, even though there have been attempts to abolish some special protective laws for women such as the prohibition of night work for women which might restrict their employment opportunities. Thus, we may expect the employment segregation in Switzerland still to be very high, but it should be interesting to analyze more carefully the evolution of occupational segregation during the last forty years.

In a recent paper, Flückiger and Silber (1990) analyzed employment segregation by gender in Switzerland for the year 1980, attempting to breakdown the overall segregation level into two components, one arising at the time boys and girls choose their field of study and another one which would be the consequence of discrimination in the labor market. The result of this empirical investigation, which was limited to the individuals who attended universities and worked full-time, indicated that occupational segregation by gender in Switzerland, for individuals with a high level of education, seemed to be mainly a consequence of the choice of field of study. To the best of our knowledge, there is no other research dealing with employment segregation by gender in Switzerland, so that a study of the evolution of occupational segregation by gender over time in Switzerland should contribute to a better understanding of the functioning of the labor market in Switzerland.

Table 1 presents the summary indices defined in section 2 for Switzerland, for the Census Years 1950, 1960, 1970, and 1980. The estimates are based on a three-digit classification of the occupations. Unfortunately, the number of occupations varies from one Census to the other, but for the last three Censuses

Table 1
Summary measures of the occupational distribution of the sex ratios in Switzerland 1950–1980.

	Census year			
	1950	1960	1970	1980
Number of occupations	402	530	504	539
Females	21339	21593	25543	28009
Males	50524	50204	50120	49854
Total	71863	71797	75663	77863
Weighted average, $\mu = F/M$	0.422	0.430	0.510	0.562
Weighted mean difference, Δ	0.719	0.725	0.825	0.898
Coefficient of asymmetry, A_G	0.825	0.814	0.769	0.759
Coefficient of flatness, K_G	0.843	0.825	0.788	0.783
Gini's concentration index, G_S	0.852	0.843	0.809	0.799

differences in the number of occupations are small (the number varies from 503 to 539). Despite this problem, the trends are quite clear. Firstly, the overall ratio of female to male workers increased, especially between 1960 and 1980 (the female share in the total labor force increased from 30.1% in 1960 to 36.0% in 1980). Such a trend evidently reflects the increase in the labor force participation rate of women in Switzerland. At the same time however, we also observe an increase in the dispersion of the gender ratios F_i/M_i, as measured by the (weighted) mean difference Δ of these ratios (an increase from 0.73 in 1960 to 0.90 in 1980).

Since the weighted mean difference Δ may also be expressed as

$$\Delta = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{F}{M} \left| \frac{F_i}{F} \frac{M_j}{M} - \frac{F_j}{F} \frac{M_i}{M} \right|,$$

it is clear that, for a given set of shares F_k/F and M_k/M, Δ will increase with F/M. Since we just noticed that F/M increased significantly between 1950 and 1980, we should not be surprised to observe a parallel increase in Δ. This is why it is desirable to standardize Δ when measuring occupational segregation, because this is the only way to make a distinction between trends in labor force participation rates and trends in occupational segregation. The index G_S which was defined in (3) is a standardized measure of Δ and, as indicated earlier, will be the measure of occupational segregation by gender which will be used in the present study. Table 1 indicates that G_S decreased regularly between 1950 and 1960, so that despite the variation in the number of occupations distinguished, one may quite safely conclude that occupational segregation decreased over time.

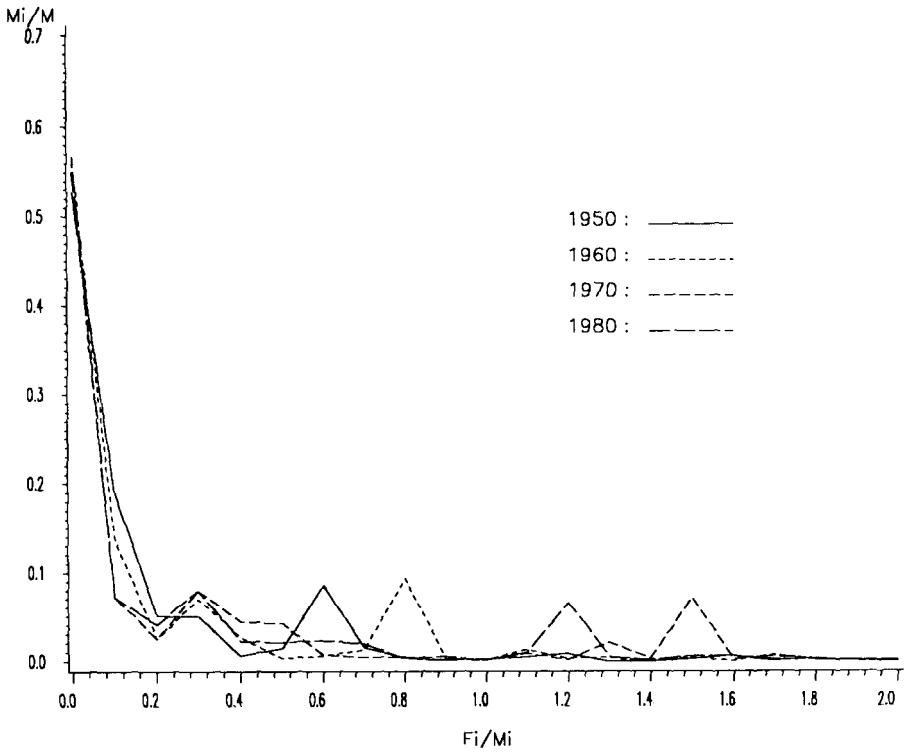

Fig. 1. Frequency distribution of F_i/M_i.

This picture is confirmed by the fact that the coefficient of asymmetry A_G also decreased.[3] As indicated by expression (6), A_G compares the dispersion of the ratios F_i/M_i which are greater than the median $\tilde{\mu}$ with that of the ratios F_i/M_i which are smaller than $\tilde{\mu}$. A decrease in A_G is therefore likely to indicate that over time there was a decrease in the relative importance of typically 'female occupations' (which are represented by a very high ratio F_i/M_i). It is also likely that there was a decrease in the relative importance of typically 'male occupations' since one observes also a decrease over time in the value of K_G (from 0.846 in 1950 to 0.78 in 1980). Such a change indicates a decrease in the pointedness (an increase in the flatness) of the distribution of these F_i/M_i ratios, but since A_G decreased, it is clear that the trend towards less purely 'male' or 'female' occupations was stronger in the latter case.

These observations seem to be confirmed by fig. 1 which gives the frequency distribution of the ratios F_i/M_i for those observations for which F_i/M_i is smaller

[3]One should remember however that the index G_S is a function of the coefficient A_G [see Berrebi and Silber (1987)].

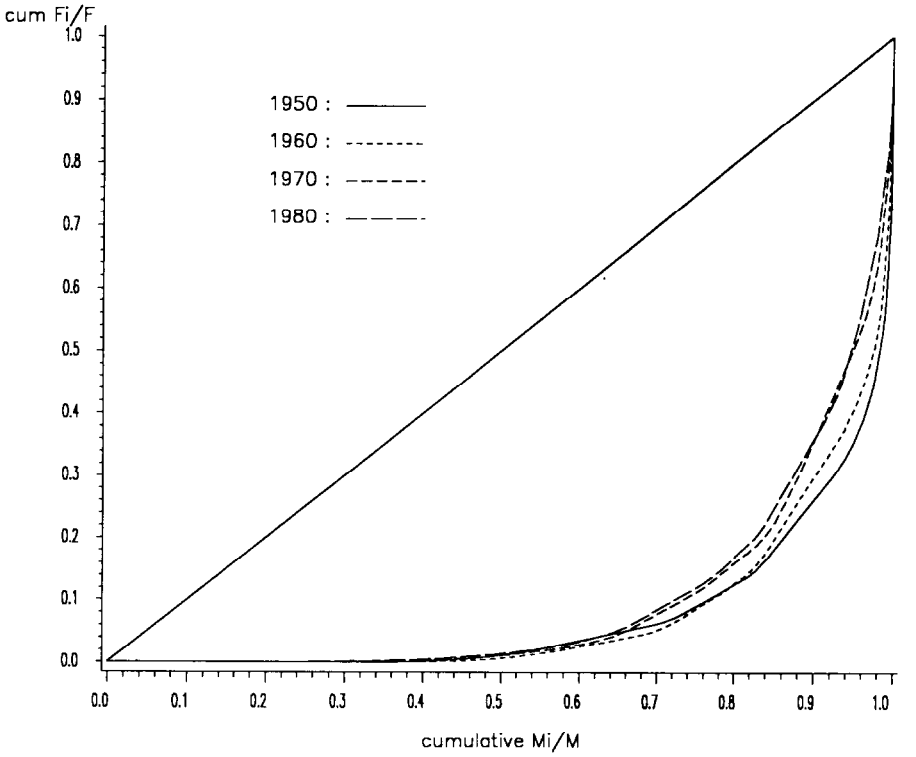

Fig. 2. Segregation curves: 1950–1980.

than 2. It appears clearly that for each Census year there was a secondary mode whose value was greater than the mean F/M, and this secondary mode seems to have shifted towards the right over time (although the secondary mode for 1980 is located on the left of that of 1970, as appears in fig. 1). This is one of the reasons why a decrease in the degree of asymmetry was observed over time (recall however that fig. 1 reports only the ratios which are smaller than 2 and hence gives an incomplete picture).

The decrease in the value of the segregation index G_S is apparent when looking at fig. 2: as a whole the segregation curves moved to the left over time, hence the observed decrease in G_S. One should notice, however, that the four segregation curves cross each other so that no clear-cut conclusions in terms of stochastic dominance may be derived from the data [see Hutchens (1991) for more details on the application to the study of occupational segregation of concepts used in the income inequality literature]. We now turn to study the effect of the classification of the occupations on the amount of segregation by gender.

Table 2
Bootstrap of the summary measures of the occupational distribution of the sex ratios in Switzerland, 1950–1980.

	Census year							
	1950		1960		1970		1980	
	1%	99%	1%	99%	1%	99%	1%	99%
Mean diff., Δ	0.716	0.724	0.722	0.731	0.821	0.832	0.894	0.905
Asymmetry, A_G	0.814	0.835	0.803	0.823	0.751	0.779	0.737	0.767
Flatness, K_G	0.828	0.854	0.811	0.832	0.766	0.796	0.755	0.789
Gini conc., G_S	0.848	0.857	0.840	0.850	0.806	0.817	0.796	0.806

4. The impact of classification errors

When a detailed occupational classification is used, it is likely that errors occur when determining the occupation of the various individuals. It is therefore of interest to check whether such errors will affect the value of the Segregation Index G_S. We have therefore used the bootstrap method to derive confidence bounds on the index G_S and on the summary statistics defined in section 2.

The bootstrap [see Efron (1982) and Efron and Gong (1983) for an introduction to this approach] provides a method for estimating the distribution of the summary statistics previously defined, by resampling with replacement from the data set. More precisely, for each gender we have used the actual shares of the various occupations in the population as the expected proportions of these occupations in the sample we built[4] and did 500 drawings to derive the distribution of the summary statistics, this allowing us finally to construct a 1%–99% confidence interval.

The results of such an experiment are given in table 2 which presents these 1%–99% confidence bounds. To check whether there were significant changes in the various summary statistics between two Censuses, one has to combine the results of tables 1 and 2. One would then observe, for example, that the expected (actual) value (0.843) of the Segregation Index G_S in 1960 lies outside the bounds of G_S for 1950 (0.848; 0.857) as well as for 1970 (0.806; 0.817), so that one may conclude that the value of G_S in 1960 was significantly lower than that observed in 1950 and higher than that given for 1970.

Similar comparisons for all the summary statistics and Census years would show that in most cases there was a significant variation of these indices between two Censuses. The only exceptions are the following ones: the Asymmetry Index A_G in 1960 was not significantly different from that for the year 1950 and the same is true when comparing A_G in 1980 with A_G in 1970. One may observe,

[4]Computer space limitations forced us to use a 10% sample of the original population because otherwise there would have been too many potential combinations.

also, that the actual value of A_G in 1950 lies outside the bounds of A_G for 1960, and the same is true when comparing 1970 with 1980. However, because of the previous findings we still have to conclude that there was no significant difference in the asymmetry of the distribution (of the ratios F_i/M_i) between the years 1950 and 1960 on one hand, 1970 and 1980 on the other.

The other exception concerns the comparison of the indices K_G in 1970 and 1980. It appears that the expected value of K_G in 1970 (0.788) lies within the bounds of K_G for 1980 (0.755; 0.789) and that the expected value of K_G for 1980 (0.783) lies within the bounds of K_G for 1970 (0.766; 0.796). There is therefore no significant difference in the Flatness of the Distribution between the Census years 1970 and 1980.

When looking at the index G_S however, we may safely conclude that there was a significant decrease over time in the level of occupational segregation (compare the value of G_S in table 1 with its bounds as given in table 2).

The present analysis has therefore shown that the existence of possible classification errors in the definition of the occupations of the various individuals will not modify the previous conclusions that occupational segregation by gender in Switzerland decreased between 1950 and 1980. One should be more careful, however, with respect to statements concerning the decrease over time in the asymmetry or the flatness of the distribution of the ratios F_i/M_i. We now turn to the analysis of another problem, that of the possible impact of the aggregation of occupations.

5. The effect of aggregation on occupational segregation

Aggregation is an important issue because it will tell us at which level occupational segregation occurs: does it appear only when a three-digit classification is used or does it already exist at the level of two-digit occupations?

Since the index G_S is used to measure occupational segregation, the analysis of the impact of aggregating the occupational categories will be similar to that of the breakdown of the Gini Index by population subgroups [see Silber (1989b)]. More precisely, let G_T refer to the value of the Segregation Index when a three-digit occupational classification is used and G_B the corresponding value of G_S when a two-digit classification is used. Let I refer to the total number of two-digit occupations i and n_i to the number of three-digit categories within the two-digit occupation i, and let $M_{i,h}$ and $F_{i,h}$ $(h = 1, \ldots, n_i)$ refer respectively to the number of males and females in the subcategory h belonging to occupation i. Finally define $M_{i.}$, $F_{i.}$, M, and F as

$$M_{i.} = \sum_{h=1}^{n_i} M_{ih}, \quad F_{i.} = \sum_{h=1}^{n_i} F_{ih}, \quad M = \sum_{i=1}^{I} M_{i.}, \quad F = \sum_{i=1}^{I} F_{i.}.$$

We may then write

$$G_T = [\ldots M_{ih}/M \ldots]G[\ldots F_{ih}/F \ldots], \tag{9}$$

$$G_B = [\ldots M_{i.}/M \ldots]G[\ldots F_{i.}/F \ldots]. \tag{10}$$

Similarly the within two-digit occupation i Segregation Index G_i will be defined as

$$G_i = [\ldots M_{ih}/M_i \ldots]G[\ldots F_{ih}/F_{i.}, \ldots]. \tag{11}$$

Note that in expressions (9) to (11) the elements of the vectors to the right and left of the matrix G have to be ranked in the same way [by decreasing values of the ratios F_{ih}/M_{ih} in (9) and (11) and $F_{i.}/M_{i.}$ in (10)].

Adapting Silber's (1989b) analysis to the case of occupational segregation, we may define the contribution C_W of the segregation within two-digit occupations to the total segregation G_T as

$$C_W = \sum_{i=1}^{I} (F_{i.}/F)(M_{i.}/M)G_i. \tag{12}$$

It can be shown [see Silber (1989b)] that the total segregation G_T is equal to the sum of three terms written as

$$G_T = G_B + C_W + C_O, \tag{13}$$

where C_O is an interaction term equal to the residual term which one obtains by subtracting the sum $G_B + C_W$ from G_T.

This residual term may be given [see Silber (1989b)] an interesting intuitive interpretation since it is equal to the difference existing between G_T and the segregation level one would obtain if the three-digit occupations were, firstly, classified by decreasing value of the ratios $F_{i.}/M_{i.}$ corresponding to the two-digit occupations to which they belong, and secondly within each two-digit category, by decreasing value of the ratios F_{ih}/M_{ih}. The difference between such a classification and that used in the computation of G_T (where the three-digit occupations are immediately classified by decreasing value of the ratios F_{ih}/M_{ih}) is precisely a measure of the overlap[5] which exists between the I distributions of the ratios F_{ih}/M_{ih}, hence the name of the term C_O in eq. (13).

We have applied such a decomposition to the data of the 1980 Census in Switzerland and the results are given in table 3. It appears that most of the occupational segregation is already apparent at the two-digit level and that the

[5]See the paper by Yitzhaki in this volume for a more detailed analysis of the overlap of distributions.

Table 3
Decomposition of the segregation index of the occupational distribution in Switzerland, 1960–1980.

	Census year		
	1960	1970	1980
Total segregation, G_T	0.843	0.809	0.799
Between component, G_B	0.718	0.677	0.655
Within component, G_W	0.031	0.019	0.023
Overlap component, G_O	0.095	0.113	0.121

contribution of the segregation existing within the two-digit categories is very small. Note that this does not imply that there is no important segregation within some two-digit occupations. As expression (12) indicates, C_W depends not only on the various indices G_i but also on the shares $F_{i.}/F$ and $M_{i.}/M$ of these two-digit occupations in the total male and female workforce, so that even if for a given two-digit occupation i, G_i is high, its contribution to C_W may be low if F_i/F or M_i/M are low.

What is most interesting is that the value of C_O is greater than that of C_W, and this proves that when one uses less aggregated data, one may obtain new gender ratios F_{ih}/M_{ih} which may be very different from the aggregated ratio $F_{i.}/M_{i.}$, hence the existence of overlap. It is the existence of such an overlap which precisely prevents us from applying the parametric methods used in income distribution analysis to the study of occupational segregation. When using income data, one often ignores the within income classes inequality, hence the attempt to use parametric forms to derive estimates of the complete income distribution. When studying occupational segregation at a high level of aggregation, one not only ignores the within occupation segregation but also the amount of overlap between the distributions of the ratios F_{ih}/M_{ih} corresponding to the various two-digit occupations i.

Let us now look in more detail at the results of the breakdown of G_T. The results are given in table 3, but only for the Census years 1960, 1970, and 1980 because in 1950 the classification of the occupations was very different from that used in subsequent Census years. The results of table 3 show that occupational segregation by gender decreased over time (between 1960 and 1970) not only when a three-digit classification is used but also when the analysis is based on two-digit occupations (see the variation over time of G_B). Here also we have estimated, using again bootstrap techniques, confidence bounds (at the 1%–99% levels) on the indices G_T, G_B, C_W, and C_O. The results are given in table 4 and allow us to safely conclude that, both at the three-digit level (see the value of G_T) and at the two-digit level (see the value of G_B), occupational segregation decreased between 1960 and 1980.

Such a conclusion is certainly robust for the 1970–1980 comparison since the number of both the two-and three-digit occupations was greater in 1980 than in

Table 4
Bootstrap of the segregation index components of the occupational distribution in Switzerland, 1960–1980.

	Census year					
	1960		1970		1980	
	1%	99%	1%	99%	1%	99%
Total segregation, G_T	0.840	0.850	0.806	0.817	0.796	0.806
Between component, G_B	0.711	0.724	0.670	0.683	0.649	0.663
Within component, G_W	0.030	0.032	0.018	0.019	0.022	0.024
Overlap component, G_O	0.092	0.101	0.110	0.119	0.117	0.127

1970, a result which per se should have led ceteris paribus to higher levels of occupational segregation. But the fact that the bootstrap analysis here also attempted to take into account possible classification errors, should lead us to conclude that the observed decrease in occupational segregation between 1960 and 1970 was also significant, even though the number of three-digit occupations was higher in 1960 than in 1970 (530 versus 504).

The contribution of the within two-digit occupations segregation, though small (see the values of C_W in tables 3 and 4), was statistically significantly different from one Census year to the other. Such a conclusion holds also for the index C_O of overlap. However, whereas the trend in C_W is not very clear (it first decreased, then increased), the overlap component C_O shows a significant upward trend over time. Note also that the bootstrap analysis shows that even with classification errors C_O was always much higher than the contribution of the within two-digit occupational segregation C_W.

6. Conclusion

The present study has attempted to take a new look at the way occupational segregation by gender should be analyzed. We suggested the application of techniques which are familiar in the income inequality literature by showing that the gender ratio F_i/M_i could play the role that income plays in the inequality literature. Formulations were then derived, giving the value of the Segregation Index G_S which is in fact Gini's Concentration Index of the ratios F_i/M_i as well as of indices measuring the asymmetry and the flatness of the distribution of these ratios. We first estimated the values of these summary indices on the basis of Swiss Census data for the years 1950, 1960, 1970, and 1980. Then we checked whether the results were sensitive to the existence of classification errors. Finally, we studied the impact of aggregation on the level of segregation and gave an illustration of this problem using data from the 1980 Census.

References

Berrebi, Z.M. and J. Silber, 1987, Dispersion, asymmetry and the Gini index of inequality, International Economic Review 28, 331–338.

Berrebi, Z.M. and J. Silber, 1989, Dispersion, the Gini index of inequality and the flatness of an income distribution, Mathematical Social Sciences 18, 229–237.

Butler, R.J., 1987, New indices of segregation, Economics Letters 24, 359–362.

Duncan, O.D. and B. Duncan, 1955, A methodological analysis of segregation indices, American Sociological Review 20, 210–217.

Efron, B., 1982, The jackknife, the bootstrap and other resampling plans (Society for Industrial and Applied Mathematics, Philadelphia, PA).

Efron, B. and G. Gong, 1983, A leisurely look at the bootstrap, the jackknife and cross validation, American Statistician 37, 36–48.

Flückiger, Y. and J. Silber, 1990, Career choice and the measurement of segregation by sex in employment: The case of students in Switzerland, in: Modelling the labor market, Proceedings of the meeting of the Applied Econometrics Association, Strasbourg.

Hutchens, R.M., 1991, Segregation curves, Lorenz curves and inequality in the distribution of people across occupations, Mathematical Social Sciences 21, 31–51.

Kendall, M.G. and A. Stuart, 1969, The advanced theory of statistics (Charles Griffin, London).

OECD, 1991, Historical statistics (OECD, Paris).

Silber, J., 1989a, Factor components, population subgroups and the computation of the Gini index of inequality, Review of Economics and Statistics LXXI, 107–115.

Silber, J., 1989b, On the measurement of employment segregation, Economics Letters 30, 237–243.

Yitzhaki, S., 1994, Economic distances and overlapping of distributions, Journal of Econometrics, this issue.