UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 03-15
Economics Series 03
March 2003

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624 98 75

ADDITIVELY DECOMPOSABLE SEGREGATION INDEXES. THE CASE OF GENDER SEGREGATION BY OCCUPATIONS AND HUMAN CAPITAL LEVELS IN SPAIN *

Ricardo Mora[1] and Javier Ruiz-Castillo[2]

Abstract

Building upon the ideas first exposed by Theil and Finizza (1971) and Fuchs (1975), this paper presents an additively decomposable segregation index based on the entropy concept used in information theory. For any pair of classification variables in a given year, the index is decomposed into a *between-group* and a *within-group* term. To analyze intertemporal changes in gender segregation for a given partition, the index is decomposed into two terms that capture, respectively, gender composition effects, and changes in the groups' demographic importance. These decompositions are illustrated with Spanish data on occupations and human capital levels for 1977 and 1992. It is found that, in both years, the higher the educational level, the smaller is gender segregation for all age groups. Moreover, gender segregation decreases with age in all educational categories. However, most gender segregation takes place within, rather than between, age/education categories. Lastly, changes in gender composition across occupations, nearly offset by occupational mix changes, account for a decline of 2% in total gender segregation over this period.

Keywords: additively decomposable indexes; entropy measures, gender segretation.

1 R. Mora, Departamento de Economía, Universidad Carlos III de Madrid; E.mail: ric mora@eco.uc3m.es. Phone: (34) 91 624 9576
2 J. Ruiz-Castillo, Departamento de Economía, Universidad Carlos III de Madrid. E.mail: jrc@eco.uc3m.es. Phone: (34) 91 624 9588

* We wish to acknowledge the excellent work done by the Research Assistant Neus Herranz, who acknowledges financial support from the *Ministerio de Educación y Ciencia*, 2001-2002. Javier Ruiz-Castillo and Ricardo Mora acknowledge financial support from the Spanish *Instituto de la Mujer*, Project Number 16/97. Ricardo Mora acknowledges financial support from DGI, Grant BEC2000-0170. This paper has been presented in the September 2002 meeting of the network *"Living Standards, Inequality, and Taxation"*, in Lubeck, financed by the European Communities (Contract #ERBCHRXCT980248).

I. INTRODUCTION

Gender segregation in the labor market is an important aspect of the way this market works. We can think of gender patterns in labor market outcomes as the result of voluntary choices that reflect differences in individual preferences, as well as technological constraints that favor some gender skills over others in certain economic activities. But gender segregation may also be a mechanism for social enforcement of wage and other forms of gender discrimination. Thus, measuring the extent of this phenomenon and its evolution over time in specific countries is an interesting topic from the point of view of both positive and normative economics.

All previous studies on gender segregation have concentrated on measuring this phenomenon among the employed population. In a few instances, some authors have classified all existing jobs according to two dimensions in order to study different structural aspects of gender segregation in a given moment of time.[1] More often, gender segregation has been studied along a single dimension, usually, occupation. When this is the case, the core of the study is the evolution of gender segregation over time.[2]

Presumably, the distribution of people across occupations is the result of the demand for and the supply of labor. But the interplay between the forces of demand

[1] For instance, the effect of aggregation on the gender segregation induced by occupational choice, or the relative importance of the gender segregation induced by either the occupational or the industrial choice –see, Sections 7.2 to 7.5 in Flückiger and Silber (1999).

[2] See, *inter alia*, Gross (1968), Blau (1977), Blau and Hendricks (1979), Williams (1979), England (1981), Beller (1985), Albelda (1986), Jacobs (1989), Jacobsen (1994), and Blau *et al.* (1998). For a recent treatise on segregation, see Flückiger and Silber (1999).

and supply, at this stage, is conditional on certain productive characteristics of the individuals from both genders. In this paper, individual data on occupations are combined with human capital characteristics -the interaction between age and education levels- to study gender segregation across the resulting groups.

The main objective is to investigate the links between occupational gender segregation and human capital. First, if education widens the career opportunities and occupational choices for females workers, then occupational segregation should differ across human capital categories: the higher the educational level for a given age group, the smaller should be gender segregation. Second, from a human capital perspective, women will choose those occupations where their skills depreciate less if they leave for periods of time because of family obligations. Thus, those females who remain employed in the later part of their life-cycle might be expected to be less segregated by occupation than the group of younger female workers. Finally, the role of human capital factors in gender segregation, that is, by how much is gender segregation reduced when human capital differences are controlled for, is an open question worth investigating.

To investigate these issues, which involve a pair of classification variables, a segregation index with the property of additive decomposability is needed. Unfortunately, the index of gender segregation most frequently used in the literature, the index of dissimilarity of Duncan and Duncan (1955), has not been exploited in this direction.[3] Building upon the ideas first exposed by Theil and Finizza (1971) and

[3] For other limitations of the dissimilarity index, see Zoloth (1976) and Hutchens (1991).

Fuchs (1975), an additively decomposable segregation index is developed which has its origin in the family of income inequality indexes introduced by Theil (1971), based on the entropy concept used in information theory.

The overall measure of gender segregation is decomposed into two components which are closely related to the classification variables: a *between-group* term, which captures the direct gender segregation induced by the first variable, say human capital; plus a weighted sum of *within-group* terms, where each of them captures the gender segregation induced by the second variable, say the occupational choice, within the corresponding human capital category.[4] The index also has a commutative property that will be used in the sequel to study gender segregation induced by human capital differences within each occupation. Finally, for any given partition, the index's structure facilitates the decomposition of the intertemporal change in gender segregation into two terms. The first one accounts for the effect of changes in gender composition across groups, while the second one captures the impact of changes in the subgroups' relative demographic importance.[5]

The interest of the approach is illustrated with an empirical application using Labor Force Survey data for Spain for 1977 and 1992. During this period, the Spanish labor market underwent three important transformations. First, the proportion of females in the employed population grows from 28.6% to 32.9%, as a direct consequence of increasing female labor force participation and slightly decreasing

[4] For an alternative decomposition using the Gini-Segregation Index, see Silber (1989), Deutsch *et al.* (1994), and Sections 7.4 and 7.5 of Flückiger and Silber (1999).

male participation rates. Second, there is a major improvement in educational standards over the period for both male and female workers. Finally, the shares of agricultural and public employment at the beginning of the period were, respectively, well above and below the OECD averages. By 1992, the share of agricultural employment had halved whilst the percentage of workers in the public sector had raised over the OECD average. Thus, the period under study is ideally suited to explore the influence of supply (increasing female labor participation, improvements in education standards), and demand (decreasing agricultural employment coupled with increasing public employment) factors to the evolution of gender segregation.

The rest of the paper contains four Sections and an Appendix. Section II is devoted to the measurement of segregation. Section III studies the pattern of gender segregation in 1977. Section IV deals with the evolution of gender segregation during the 1977-1992 period, while Section V offers some concluding comments. The descriptions of the data as well as the list of 29 occupational categories that are used in the empirical Sections are relegated to the Appendix.

II. THE MEASUREMENT OF SEGREGATION

To explain our measurement approach, we proceed in two steps. We begin with situations in which workers with a given characteristic, say a three-digit occupation, could be classified in terms of a second characteristic, a two-digit occupation, but not vice versa. This case is referred to as "a pair of one-way

[5] Alternative decompositions are found in Fuchs (1975), Blau and Hendricks (1989), and Karmel and

classification variables". In the next step, we confront situations in which individuals can be classified in terms of a first characteristic, say human capital attainment, and/or in terms of a second characteristic, say occupation. This case is referred to as "a pair of two-way classification variables".

II. 1. The Case of a Pair of One-way Classification Variables

II. 1. A. Within-group Segregation

Theil and Finizza (1971) -or TF for short- study racial integration in a city where students of both genders can be classified in J schools and I school districts with $I < J$. In our context, let there be J three-digit occupations, indexed by $j = 1,\ldots, J$, classified into I two-digit occupational groups, indexed by G_i, $i = 1,.., I$. There are a number of procedures for measuring segregation along a single dimension,[6] but not all of them are equally well-suited when one wants to consider two classification variables.

Let F_{ij} and T_{ij} be the number of females and workers of both genders, respectively, in three-digit occupation j within two-digit group i. Let $F_i = \Sigma_{j \in G_i} F_{ij}$ and $T_i = \Sigma_{j \in G_i} T_{ij}$ be the number of females and workers in group i, and let $T = \Sigma_i T_i$ be the total number of workers in the employed population. Let $W_i = F_i / T_i$ be the proportion of females in group i, and let $w_{ij} = F_{ij} / T_{ij}$ be the proportion of females in

Maclahlan (1988).

[6] See *inter alia*, James and Taeuber (1985), and Siltanen *et al.* (1993). For a recent survey, see Chapters 4 and 5 in Flückiger and Silber (1999).

group *i* and occupation *j*. TF say that the population in group *i* is segregated in occupation *j* whenever w_{ij} differs from W_i. In information theory, the expression

$$I_{ij} = w_{ij} \log (w_{ij} / W_i) + (1 - w_{ij}) \log ((1 - w_{ij}) / (1 - W_i)) \qquad (1)$$

is known as the expected information of the message that transforms the proportions $(W_i, (1 - W_i))$ to a second set of proportions $(w_{ij}, (1 - w_{ij}))$. The value of this expected information is zero when the two sets of proportions are identical; it takes larger and larger positive values when the two sets are more different. For $w_{ij} = 1$ the value of expression (1) is $\log (1 / W_i)$, and for $w_{ij} = 0$ it is $\log (1 / (1 - W_i))$. Thus, for example, when a two-digit group is predominantly male (W_i small), the presence of an all female occupation *j* ($w_{ij} = 1$) implies a large value of I_{ij}. This is intuitively reasonable for a measure of segregation.

TF define the occupational segregation within group *i* as a whole by

$$I_i = \Sigma_{j \in G_i} (T_{ij} / T_i) I_{ij}. \qquad (2)$$

That is to say, I_i is the weighted average of the information expectations in (1), with weights proportional to the number of workers in the occupations within group *i*.

The entropy of a distribution with proportions $(W_i, (1 - W_i))$ is defined as

$$E_i = W_i \log (1 / W_i) + (1 - W_i) \log (1 / (1 - W_i)). \qquad (3)$$

Equation (3) is a measure of the gender mix in group *i*. Notice that E_i takes its minimum value, equal to 0, when $W_i = 0$. Otherwise, E_i is positive and reaches its

maximum value, equal to log 2, when $W_i = 1/2$. To normalize E_i between 0 and 1, from here on it is assumed that all logarithms are in base 2. Analogously, the entropy of a distribution characterized by the proportions $(w_{ij}, (1 - w_{ij}))$ is given by

$$E_{ij} = w_{ij} \log (1 / w_{ij}) + (1 - w_{ij}) \log (1 / (1 - w_{ij})). \quad (4)$$

The average occupational entropy, or the average gender mix of the three-digit occupations in group *i*, is the weighted mean of the E_{ij}'s with weights proportional to the number of workers in the occupations within group *i*: $\mu_i = \Sigma_{j \in G_i} (T_{ij} / T_i) E_{ij}$.

TF establish the relationship between the segregation index in group *i*, I_i, on one hand, and the maximum entropy allowed by the gender composition in that group, E_i, and the average occupational entropy of the three-digit occupations in group *i*, μ_i, on the other. The result is the following:

$$I_i = E_i - \mu_i. \quad (5)$$

Whenever $w_{ij} = W_i$ for all *j*, so that the values of E_{ij} and, hence, of μ_i take their maximum value, E_i, then $I_i = 0$, indicating complete absence of segregation within group *i*. On the contrary, when the w_{ij}'s present the maximum disparity with W_i, so that the values of E_{ij} and μ_i reach their minimum value, equal to 0, then I_i takes its maximum value, E_i, which is bounded above by 1.

II. 1. B. The Decomposition of Overall Segregation

The contribution of TF concludes here. These authors never define a notion of

overall segregation for the population as a whole. In this subsection, two such notions are suggested and are shown to be equivalent.

In the first place, let $F = \Sigma_i F_i$ and $W = F/T$ be the total number and the proportion of females in the employed population, respectively. Consider the expected information of the message that transforms the proportions $(W, (1 - W))$ directly into the proportions $(w_{ij}, (1 - w_{ij}))$:

$$I^{ij} = w_{ij} \log (w_{ij} / W) + (1 - w_{ij}) \log ((1 - w_{ij}) / (1 - W)). \qquad (6)$$

The index I^{ij} provides what is called a *direct* measure of gender segregation in two-digit group i and three-digit occupation j in relation to the entire employed population. Naturally, the greater the discrepancy between the proportion of females in group i and occupation j, w_{ij}, and the proportion of females in the population, W, the greater is the segregation index I^{ij}. The weighted average of the I^{ij}'s, with weights proportional to the number of workers in the three-digit occupation j within two-digit group i, provides a reasonable overall measure of occupational segregation:

$$I = \Sigma_i \Sigma_{j \in G_i} (T_{ij} / T) I^{ij}.$$

Applying the TF result in equation (5), we have that $I = E - \mu$, where $E = W \log (1 / W) + (1 - W) \log (1/(1 - W))$ is the entropy of the distribution characterized by the proportions $(W, (1 - W))$, and $\mu = \Sigma_i \Sigma_{j \in G_i} (T_{ij} / T) E_{ij}$ is the average occupational entropy in the entire population. Therefore, the index I can take values in the interval

[0, E], and E in turn is normalized in the unit interval.[7]

In the second place, the expected information of the message that transforms the entire population proportions (W, $(1 - W)$) into group proportions (W_i, $(1 - W_i)$) is given by

$$I^i = W_i \log (W_i / W) + (1 - W_i) \log ((1 - W_i) / (1 - W)). \qquad (7)$$

Consider the weighted average of the I^is with weights proportional to the number of workers in each group, that is,

$$I^B = \Sigma_i (T_i / T) I^i. \qquad (8)$$

Equation (8) can be interpreted as the *between-group* (direct) gender segregation induced at the two-digit occupational level. On the other hand, given that I_i is the gender segregation in two-digit group i induced by three-digit occupations (see equation (2)), the overall *within-group* gender segregation in the partition by two-digit occupational groups can be defined as

$$I^W = \Sigma_i (T_i / T) I_i. \qquad (9)$$

Hence, the sum of I^B and I^W provides a second reasonable measure of overall occupational segregation.

Notice, however, that applying the TF result in equation (8), we obtain that $I^B = E - \Sigma_i (T_i / T) E_i$. Recall also that, according to equation (5), $I_i = E_i - \mu_i$. Therefore,

[7] For the case of a single variable, Fuchs (1975) also suggests this formula. For the properties of this entropy segregation measure, see Chapter 5 in Flückiger and Silber (1999).

$$I^B + I^W = E - \Sigma_i (T_i / T) E_i + \Sigma_i (T_i / T) (E_i - \mu_i) = E - \Sigma_i (T_i / T) \mu_i = E - \mu = I. \quad (10)$$

Thus, the two measures of overall segregation are equivalent. The direct measure of occupational segregation I is decomposable into a *between-group* term, I^B, which measures the gender segregation at the level of two-digit occupational groups, plus a *within-group* term, I^W, which measures the gender segregation induced by three-digit occupations within each of the two-digit groups.[8]

II. 2. The Case of a Pair of Two-way Classification Variables

So far we have only considered a situation in which workers with a given characteristic could be classified in terms of a second characteristic, but not *vice versa.* In this Subsection, we study situations in which workers can be classified in terms of a first characteristic indexed by $i = 1,\ldots, I$, say human capital, and/or in terms of a second characteristic indexed by $j = 1,\ldots, J$, say occupation. Thus, F_{ij} is the number of females with human capital level i in occupation *j.*

In this case, there are two possible decompositions as the terms in equation (10) must now specify the partition sequel to which they refer. For example, if the population is first partitioned according to human capital, then the *between-group* segregation measure is $I^B_{(i)} = \Sigma_i (T_i / T) I^i$, where I^i was defined in equation (7). On

[8] This is a useful result which has been applied in Herranz *et al.* (2003) to evaluate the impact of aggregation on the measurement of gender segregation. In contrast, in the decomposition based in the Gini-Segregation index, the overall segregation is decomposed into three terms: a between-group term, a within-group term and a third interaction term –see the references in note 3.

the other hand, the term I_{ij} in equation (1) now measures the gender segregation in the group consisting of individuals with human capital *i* in occupation *j*; I_i in equation (2) measures the segregation induced by occupational choices within the group of individuals with human capital *i*; and the *within-group* measure of gender segregation in the partition by human capital, defined in equation (9), must be also indexed: $I^{Wj}_{(i)}$. Therefore, the previous result on the decomposition of the overall segregation index will be written as follows: $I = I^{B}_{(i)} + I^{Wj}_{(i)}$.

Let $F_j = \Sigma_i F_{ij}$ and $T_j = \Sigma_i T_{ij}$ be the number of females and workers in occupation *j*, respectively, and let $W_j = F_j / T_j$ be the proportion of females in that occupation. The index of gender segregation induced by human capital within occupation *j* can be defined as: $I_j = \Sigma_i (T_{ij} / T_j) I_{ji}$, where $I_{ji} = w_{ij} \log (w_{ij} / W_j) + (1 - w_{ij}) \log ((1 - w_{ij}) / (1 - W_j))$ is the gender segregation index of workers with human capital *i* for the workers in occupation *j*. Similarly, let $I^j = W_j \log (W_j / W) + (1 - W_j) \log ((1 - W_j) / (1 - W))$ be the index of direct segregation in occupation *j* relative to the employed population as a whole. Following the same argument as in the previous case, it can be shown that the overall index I can be decomposed into the sum of two terms: a *between-group* term, $I^{B}_{(j)} = \Sigma_j (T_j / T) I^j$, which measures the gender segregation induced directly by occupation, and a *within-group* term, $I^{Wi}_{(j)} = \Sigma_j (T_j / T) I_j$, which captures the gender segregation induced by human capital in the

partition by occupation. Therefore,

$$I = I^{B}_{(i)} + I^{Wj}_{(i)} = I^{B}_{(j)} + I^{Wi}_{(j)}. \quad (11)$$

Thus, given a pair of two classification variables, the overall segregation index has a commutative property, i. e. admits two alternative decompositions. In the first one, the term $I^{Wj}_{(i)}$ measures the role of occupation on gender segregation, the impact of human capital being kept constant in $I^{B}_{(i)}$. Similarly, the term $I^{Wi}_{(j)}$ measures the contribution of human capital to overall gender segregation, the impact of occupation being kept constant in $I^{B}_{(j)}$.

Finally, it can be seen that

$$I = \Sigma_i \, (T_i/T) \, I(i) = \Sigma_j \, (T_j/T) \, I(j), \quad (12)$$

where $I(i) = I^{i} + I_i$, and $I(j) = I^{j} + I_j$.

III. THE GENDER SEGREGATION OF THE EMPLOYED POPULATION IN 1977

III. 1. Descriptive Statistics

As explained in the Appendix, the data used comes from the Spanish *Encuesta de Población Activa* (EPA), a labor force survey representative of the household population living in residential housing. The first year of study is 1977, the first time for which microeconomic data is available in electronic support. The target population in 1977 consists of 71,864 individuals, representative of 12,148,346 employed people, of which only 28.6% are females. Individuals are classified

according to two variables. On one hand, human capital categories result from the combination of readily available variables, namely, age and education. This combination gives rise to 11 age/education categories. On the other hand, using an algorithm based in the bootstrap, there is a list of 29 occupations.[9]

We begin by considering the usual case studied in the literature, namely, the partition of the employed population by occupation. The 29 available occupations can be conveniently classified into three main categories: 14 *male occupations*, where the female proportion rate goes from 0 -in the *Armed forces*- to 17.8%; 11 *female occupations*, where the proportion of females goes from 45.9% to 93.9%; and 4 *integrated occupations*, where the proportion of females goes from 22% to 38.9%. In turn, each of these categories can be further divided into a maximum of four groups, depending on whether they contain agricultural, blue collar, white collar, or professional and managerial occupations.

The first 4 columns in Table 1 contain some descriptive statistics for the 29 occupations in 1977. Approximately, 48.8%, 22.8% and 29.5% of the population are employed in male, integrated and female occupations, respectively (see column 3). From another perspective, 20.7% of the population has a job in the agricultural sector, 39.1%, 27.1%, and 12.3% in blue collar, white collar, and professional and managerial occupations, respectively, while the remaining 0.8% is in the military. Naturally, the proportion of female workers increases as we move from male to integrated and female occupations (see column 4).

[9] See the Appendix for a brief explanation of the data and the full description of the 29 occupational

Table 1 around here

Regarding the age/education partition in 1977, some descriptive statistics are presented in the first 4 columns of Table 2. Notice that still in 1977 as much as 16.8% of the Spanish population had a low education (either illiterate or without studies), while only 19.8% had a secondary or a College education (see column 3). The high percentage of workers with only primary education is due to the fact that as late as 1970 compulsory education in Spain had only reached up to that level. Columns 1 and 2 in Table 2 show that the percentages of males and females at different educational levels are surprisingly similar. Nevertheless, in different age brackets the educational experience by gender varies considerably: the percentage of young females (16-30 years old) with a primary or, above all, a secondary education, is larger than the corresponding percentages of young males, while the opposite is the case among workers of more than 30 years of age.

Finally, it is interesting to notice that at all educational levels, except the lowest one, the lower the age bracket, the greater is the proportion of female workers (column 4). This reflects the fact that female labor participation rates for younger females with at least a primary education are above the population average.

Table 2 around here

III. 2. The Role of Occupations and Human Capital Characteristics in Gender Segregation in 1977

As shown in Section II, the gender segregation index in the employed

---

categories and 11 human capital levels.

population, I, can be decomposed into two terms which measure the direct gender segregation in all occupations, $I^{B}_{(j)}$, and the gender segregation induced by age/education characteristics within the partition by occupations, $I^{Wi}_{(j)}$. Similarly, the index I can be decomposed into the direct gender segregation attributable to age/education characteristics, $I^{B}_{(i)}$, and the gender segregation that takes place due to occupational choices within each human capital group, $I^{Wj}_{(i)}$, (see equation 11). Following equation (12), $I = \Sigma_j (T_j / T)\, I(j) = \Sigma_i (T_i / T)\, I(i)$, where $I(j) = I^{j} + I_j$; and $I(i) = I^{i} + I_i$. Columns 5, 6, and 7 in Table 1 (and Table 2) present detailed information for the 29 occupations (and the 11 age/education categories), on the following statistics: I^{j}, I_j, I (j) (and I^{i}, I_i, I(i)). To facilitate the reading, all indexes have been multiplied by 100.

It turns out that $I^{B}_{(j)} = 27.01$, $I^{Wi}_{(j)} = 3.04$, $I^{B}_{(i)} = 1.76$ and $I^{Wj}_{(i)} = 28.29$. Thus, the degree of overall gender segregation is given by:

$$I = 27.01 + 3.04 = 1.76 + 28.29 = 30.05$$

The first conclusion is inescapable: workers' educational choices, even combined with age differences, induce a very low degree of direct gender segregation (1.76 index points, or 5.9% of the total). Alternatively, given the occupational choices, the gender segregation attributable to differences in human capital characteristics within occupations is very low (3.04 index points, or 10.1% of the total). Thus, most of the

gender segregation observed in Spain in 1977 takes place within age/education subgroups or, in other words, between occupations.

This is not surprising in view of the fact that the female proportion across human capital categories differs much less from the overall proportion than the female proportion across occupations (see column 4 in Tables 2 and 1, respectively). Consequently, the range of variation in the index of total gender segregation across age/education categories goes from 21.04 for older workers with a secondary education, to 36.79 for younger workers with a primary education. Instead, this range goes from 2.69 in occupation 15, an integrated agricultural occupation consisting of *Employees in livestock production,* to 143.08 in occupation 26, a white collar female occupation consisting of *Domestic service personnel, typists and other operators* (see column 7 in Tables 2 and 1).[10] As a final symptom of the lesser role of age/education characteristics in gender segregation, only young workers with a secondary or a College education have a direct gender segregation index above 7 points (see column 5 in Table 2), while only within occupations 18, 19 and 27 do the age/education characteristics induce a gender segregation value above 8 index points (see column 6 in Table 1).

This important point does not preclude the detailed analysis of what happens inside each partition. For this purpose, starting with the partition by occupations, recall that the direct segregation index for any occupation, I^{j}, results from the

[10] Recall that while weighted gender segregation indexes are bounded between 0 and 100, each unweighted direct segregation index is bounded only from below.

discrepancy between the proportion of females in the employed population, W = 28.6%, and the proportion of females in that occupation, W_j (see column 4 in Table 1). Naturally, the direct segregation indexes reach high values in the male and female occupations, and low values in the integrated occupations (see column 5 in Table 1). However, the non-linearity of the log function in the index formula implies that whenever the absolute difference between W and W_j is the same for a male and a female occupation, the index for the latter is larger than for the former. Moreover, although small, the gender segregation induced by age/education characteristics is typically larger in female than in male occupations (see column 6 in Table 1). Consequently, total gender segregation tends to be larger in female occupations.

The last column in Table 1 includes the ratio

$$\alpha_j = [((T_j / T)\, \mathrm{I(j)}) / \mathrm{I}\,] / (T_j / T) = \mathrm{I(j)} / \mathrm{I}.$$

The numerator in this expression is the *j*-th occupation relative contribution to the total gender segregation I, while the denominator is this occupation's demographic importance within the employed population. Therefore, when $\alpha_j > 1$ (< 1), this ratio indicates that occupation *j* is contributing to total gender segregation above (below) what could be expected from its demographic weight. In particular, female white collar occupations 23, 27, 25, and 26, as well as the blue collar occupation 21, contribute to total gender segregation from 80% to 380% more than what could be expected from their demographic importance (see column 8 in Table 1). Among male occupations, only blue collar occupations 3 and 6 and the *Armed forces* (occupation 14) contribute between 50% and 60% more than what could be expected from their

demographic weight. Finally, not surprisingly, all integrated occupations contribute to total gender segregation well below their demographic weight.

In the partition by age/education, two points deserve mentioning. In the first place, column 7 in Table 2 shows that the greater the educational level, the smaller is total gender segregation for all age groups (with the sole exception of older workers with a College education that only accounts for 1.3% of total employment). This means that, given the age bracket, the greater the education level, the closer are the proportions of females in the different occupations to the female proportion in the education category in question. This important finding suggests that, as conjectured in the Introduction, more educated female workers encounter fewer barriers to allocate themselves more evenly among the different occupations.

In the second place, interestingly enough, the segregation among the old is smaller than among the previous age brackets in all educational categories. As pointed out before, except for the lower educated, the proportion of females among the employed in 1977 decreases monotonically with age in all educational categories (see column 4 in Table 2). But it would appear that, at every educational level, those females who remain employed in the later part of their life-cycle are less segregated by occupation than at the beginning of their employment career. As highlighted in the Introduction, this fact can be interpreted from a human capital perspective. Given that women must be temporarily absent from their jobs more often than men because of the family obligations they are supposed to attend to, they would tend to

choose those occupations where their skills depreciate less over time.

IV. INTERTEMPORAL COMPARISONS

IV. 1. The Role of Occupations and Human Capital Characteristics in Gender Segregation in 1992

As pointed out in the Appendix, the fundamental changes in the *National Classification of Occupations* and the *National Classification of Industries* that took place in 1993 and 1994, makes it impossible to compare the 1977 data with the period starting in 1993. Therefore, the period of study is 1977-1992. This is an interesting period because, as will be seen presently, there are important changes in male and female behavior relating to labor market participation, investment in human capital through formal education, and occupational choices. The information about the population in 1992 in the partition by age / educational characteristics is in Table 3.

Table 3 around here

The comparison with Table 2 shows the following differences. In the first place, the proportion of females in the employed population has increased by more than 5 percentage points, from 28.6% to 32.9%. In the second place, the employed population is younger in 1992 than in 1977: the presence of workers older than 50 years decreases by 5.1 percentage points. This decrease is somewhat larger among females (5.6 points) than for males (4.7 points). In the third place, there has been a remarkable improvement in educational achievements. As a result, 9.8% of the population has a low education (*versus* 16.8% in 1977), whereas 53.8% has a

secondary or a College education (*versus* 19.8% in 1977).

What are the implications of this upgrading in educational achievements, particularly among the young, for the gender segregation induced by age/education characteristics? In this framework, differences in gender segregation must come from gender differences in the above patterns. The comparison of column 1 in Tables 2 and 3 indicates that the proportion of females with a secondary or a College education has increased, approximately, by a factor of 2.5 and 3, respectively, while the proportion with a low education or, above all, with a primary one, has decreased dramatically. However, judging from the evidence presented in column 2 of these Tables, something similar has also taken place among the males.

Therefore, relative to 1977 the degree of direct gender segregation among the two lowest educational levels has changed very little, while the considerable increase experienced by College graduates below 50 years of age is offset by the decrease among those with a secondary education and older College educated workers who represent a larger proportion of the population. The end result is that $I^{B}_{(i)}$, the direct gender segregation induced by human capital characteristics, takes almost the same low value at the beginning and the end of the period (see column 5 in Tables 3 and 2). In 1992, $I = I^{B}_{(i)} + I^{Wj}_{(i)} = 1.66 + 27.67 = 29.33$, the direct gender segregation in the partition by age/education characteristics amounts to only 5.8% of the total.[11] Thus, the Spanish employed population in 1992 is considerably more

[11] Similarly, the gender segregation attributable to differences in human capital characteristics within occupations is even lower than in 1977 (2 index points, or 6.7% of the total in 1992).

educated than in 1977. Although investment in human capital has been particularly large among females, workers' educational choices in 1992 again induce a very low degree of gender segregation.[12]

As in 1977, column 7 in Table 3 shows two facts: except for the young and the older workers with a College education, the greater the educational level, the smaller is total gender segregation; moreover, the greater the age, the smaller is gender segregation in all educational categories.

IV. 2. Accounting for Changes in Gender Segregation

As we have just seen, overall gender segregation in 1992 is equal to 29.33. Therefore, relative to 1977, there is a slight decrease in total gender segregation of 0.7 points, which represents a 2.4% drop from the 1977 index value.[13] Sampling error can potentially be the source of small changes in gender segregation indexes. In this case, upper (95%) and lower (5%) bootstrap bounds from 5,000 empirical sample replications are equal to 0.14 and -1.46, respectively. Therefore, although the point estimate for the change in gender segregation from 1977 to 1992 implies a decrease in overall gender segregation, this reduction is not statistically significant at the 10%

[12] Albelda (1986) provides indirect evidence about the small role played by educational factors in accounting for changes in gender segregation in the U.S. from 1958 to 1981.

[13] In the only comparable study for the Spanish economy, Sánchez (1991) obtains very similar results using 62 occupations from the same data source for the 1977-1988 period. According to the dissimilarity index, there is a 0.14 per cent decrease in gender segregation; using the Karmel-Maclahlan index, there is a 0.02 per cent increase. On the other hand, the slight decline in gender segregation observed for the Spanish economy in the 1977-1992 period is broadly consistent with the relative stability shown by the dissimilarity index in the U.S. throughout the first half of the twentieth century (see Jacobs 1989, and the discussion in England 1991 of the early papers on the U.S.). This period is characterized by low female labor participation rates comparable to the Spanish ones: in 1960, that rate was 37.7 per cent in the U.S. – see Beller (1981). For a general discussion of the main theories on the persistence of occupational segregation, see, for example, Anker (1997) and Preston and Brandeis (1999).

confidence interval.

The main aim of this Subsection is to identify the sources for this negligible change in overall gender segregation, denoted by $\Delta \equiv I_{92} - I_{77}$. Given the reference demographic weights β_i in the partition by age / education characteristics, $i = 1,\ldots, 11$, the following decomposition will be used in the sequel:

$$\Delta = \Sigma_i \, (T_{i92} / T_{92}) \, I_{i92} - \Sigma_i \, (T_{i77} / T_{77}) \, I_{i77} =$$

$$\Sigma_i \, GC_i + DM_i = \Sigma_i \, TOTAL_i, \qquad (13)$$

where

$$GC_i = [I_{i92} - I_{i77}] \, \beta_i \qquad (14)$$

measures the change in gender segregation induced by gender composition changes in the partition by age / education characteristics and

$$DM_i = [(T_{i92} / T_{92}) - \beta_i) \, I_{i92} + (\beta_i - T_{i77} / T_{77})] \, I_{i77} \qquad (15)$$

shows the change in gender segregation induced by changes in the demographic mix of the age / education categories. Finally,

$$TOTAL_i = GC_i + DM_i.$$

IV. 2. 1. The Partition By Human Capital Characteristics

The relevant information when $\beta_i = T_{i77} / T_{77}$ for all i is presented in Table 4. Two points deserve to be noticed. (i) Changes in the gender composition lead to a positive GC term. This is mainly due to the moderate increase in gender segregation among workers with a low education or with a primary education and more than 30

years of age. (ii) The slight decrease in overall gender segregation during the period is due to the offsetting influence of changes in the mix of age/education categories that lead to a negative value of the DM term, which is equal to – 1.9. The improvement in the employed population's educational standards is reflected in an increase in the proportion of workers in the upper tail of the educational distribution that leads to positive DM terms. This is offset by the reduction in the proportion of workers in the lower tail of the distribution, which is weighted by relatively high values of total gender segregation indexes in 1992, I_{i92} (see equation 15 with $\beta_i = T_{i77} / T_{77}$ for all i). At both tails of the distribution, the DM effect is stronger for the group through which the main change takes place, namely, the young.[14]

Table 4 around here

IV. 2. 2. The Partition By Occupations

The analysis for the partition of occupations is more complex. The relevant information when $\beta_j = T_{j77} / T_{77}$ for all j is presented in Table 5, which is organized as follows. First of all, changes in gender composition across occupations depend on two factors: changes in the frequency distribution of women, denoted by $(\Delta F_j / F)$, and changes in the female proportion within each occupation, denoted by ΔW_j. Therefore, columns 1, 2, and 3 of Table 5 refer to $(\Delta F_j / F)$, ΔW_j and GC_j, respectively.

[14] This result is robust to the choice of weights: using the 1992 population weights, $\beta_i = T_{i92} / T_{92}$ for all i in equations 14 and 15, the GC and DM terms become 0.1 and - 0.8, respectively.

In order to facilitate the discussion, the 29 occupations have been aggregated into 11 categories[15], classified into four groups according to the sign of both the change in the female frequency distribution and the sign of GC.[16] In the second place, changes in the occupational mix depend on changes in the population frequency distribution, denoted by $(\Delta T_j / T)$. Therefore, columns 4 and 5 contain the information on $(\Delta T_j / T)$, and DM_j. Finally, column 6 in this Table captures the sum of the two factors in equation 13, that is, $TOTAL_j = GC_j + DM_j$.

Table 5 around here

IV.2.2.A Changes in the Occupational Mix of the Economy

From 1977 to 1992 the employment population increases by only 2 per cent, approximately. By activity sectors, the pattern of change is the following (see column 4 of Table 5): agricultural (category I) and blue collar occupations (categories III, IV, and VIII) decrease by 10.8% and 2.8%, respectively, while all white collar (categories II, V, IX and X) and professional and managerial occupations (categories VI, and VII), increase by 6.1% and 5.1%, respectively. The relative size of the *Armed forces* (XI) remains essentially the same. Thus, the decline of agriculture and industrial activities, and a terciarization of the economy in which the public sector plays a

[15] In particular, the original female white collar occupations include occupations 22 to 27. However, in 1992 the proportion of employment in the public sector in occupations 22, 24 and 25 is only 0.8, 2.0 and 0.4%, while in occupations 23, 26, and 27 this proportion is 31.3, 9.9, and 59.8%, respectively. Therefore, in Table 5 this set of occupations is subdivided into two categories: occupations 22, 24 and 25 are classified as "private", while occupations 23, 26, and 27 are classified as "public".

[16] The information for the full partition of 29 occupations is available on request.

major role characterize this period.[17]

These changes in the pattern of economic activity affect the distribution of the employed population across male, integrated and female occupations. (i) Integrated occupations, which represent 22.8% of the population in 1977, go down to 19.5%. Essentially, this decrease is driven by occupation 15 (*Agricultural workers in livestock production*) whose relative size decreases by 5.5 points. (ii) The proportion of male occupations remains constant. This is because the decrease in agricultural occupations 1 and 2 is offset by the corresponding increase in white collar occupations (7, 8, 9) and the professions grouped in occupations 10 to 13, while male blue collar occupations maintain their relative importance at 25.8% of total employment. (iii) Thus, the decrease in integrated occupations is matched by an increase in female occupations. However, female blue collar occupations closely related to agricultural activities (*Textile and tobacco industries*, 20, and the *Clothing industry*, 21), as well as what has been called the private white collar ones (category II in Table 5), lose ground during this period. On the contrary, female professional and managerial occupations, as well as white collar public occupations (category X in Table 5), increase their relative importance during the period.

As a consequence of these trends we should expect DM negative values for agricultural occupations (category I in Table 5), female blue collar occupations closely related to agricultural activity (category III), as well as the female private

[17] Whereas employment in the private sector actually decreases by 600,000 persons, in the public sector there is an increase of 847,000 jobs. As a consequence, the percentage represented by the public sector increases from 10.8 to 17.4 per cent.

white collar ones (category II). We should also expect positive values for tertiary occupations, namely, the different groups of professional and managerial occupations (categories VI and VII), as well as white collar occupations (categories V and IX), particularly those linked to the public sector (category X). This is, indeed, what is found in column 5 of Table 5.

IV.2.2.B Changes in Gender Composition Across Occupations

In order to study the evolution of gender composition during this period, the first fact to be stressed is the important increase in the overall proportion of female workers from 28.6% to 32.9%. The key to understanding the change in gender segregation indexes across occupations is the connection between this fact and the set of female proportions W_j for every *j*. Columns 1 and 2 in Table 5 inform about the changes in the female frequency distribution and female proportions, while column 3 presents the implications for GC values.

The first four rows of Table 5 consider cases in which there is a decrease in the presence of women, while the next six rows consider occupational categories in which there is an increase.

(i) The greatest decrease, which is parallel to the decline in employed population in the agricultural sector as a whole, has been in agricultural occupations (category I). Except in occupation 15 (*Agricultural workers in livestock production*, an integrated occupation), there are minor increases in female proportions; that is to say, relatively more males have abandoned agricultural activities. These changes give rise to GC small negative values in all but occupation 2 (*Fish, game, and forestry*

workers).

(ii) By construction, women proportions in female occupations are characterized by high values. However, in categories III (female blue collar occupations closely related to agriculture), and II (white collar occupations where the private sector predominates, including *Administrative staff and auxiliary jobs in the service sector*, as well as *Concierges, building supervisors, cleaning and domestic service in all service sectors*), such high female proportions are reduced. This, together with the increase in the female proportion for the population as a whole, *W*, gives rise to GC negative values. In category IV, consisting of integrated blue collar occupations, there is a smaller decrease of female proportions; given the important increase in *W*, this translates into smaller GC negative values.

(iii) In categories V and VI (male white collar, and professional and managerial occupations, respectively), the increase in the proportion of women more than offsets the increase in *W*, giving rise also to GC negative values. In female professional and managerial occupations (grouped in category VII) already characterized in 1977 by a higher female proportion, this pattern generates a smaller GC negative value.

(iv) In the last group of occupations, all GC values become positive. On one hand, the explanation lies in the large increase in female proportions in white collar occupations dominated by the public sector (category X), which were already characterized in 1977 by very high female proportions. On the other hand, the entrance of women in male blue collar and integrated white collar occupations

(category VIII and IX, respectively) would lead us to expect a decrease in gender segregation; however, that increase is offset by the relatively larger increase in *W*, resulting in GC positive values.

In all categories in group 1, gender composition and occupational mix effects reinforce each other giving rise to a total effect equal to – 6.2 index points. In group 2, GC and DM offset each other, while in group 3 positive GC and DM effects yield an increase in total segregation equal to 5.3 index points. This accounts for the decline in gender segregation in 0.7 = - 6.2 + 0.2 + 5.3 index points. However, there is a decline of 3.8 index points due to gender composition effects. This indicates that, had we not had a positive DM effect, equal to 3.1 index points, the reduction in overall gender segregation might have been close to 10 per cent of the 1977 value.[18]

V. CONCLUSIONS

The property of additive decomposability for any partition has been extensively studied in the field of income inequality for quite some time. In the presence of two (or more) classification variables, this property is also essential in the field of gender segregation. Following up on the seminal work by Theil in income distribution theory, this paper has presented an additively decomposable gender segregation index based on the entropy concept used in information theory.

Overall gender segregation in a given year has been decomposed into a

[18] This result is robust to the choice of weights: using the 1992 population weights, the GC and DM terms become – 3.1 and 2.4, respectively. On the other hand, using the Karmel-Maclahlan decomposition,

between-group term that measures, for instance, the direct gender segregation that can be attributed to human capital characteristics; and a *within-group* term, that captures gender segregation caused by occupation within human capital categories. The index has a commutative property that permits a similar decomposition where the role of the two partitions is interchanged. In intertemporal comparisons, the change in total gender segregation has also been decomposed into two terms. The first measures the effect of changes in gender composition in the groups of a given partition, while the second captures the impact of changes in the groups' relative demographic importance.

These two decompositions have been applied to Spanish data in 1977 and 1992. In both years, it has been found that, even when differences in educational achievements across age groups are taken into account, the direct gender segregation induced by human capital characteristics accounts for less than 10% of total gender segregation. Thus, most gender segregation must be attributed to occupational choices within age/education groups. Also, as conjectured in the Introduction, the higher the educational level (the older the group), the smaller is gender segregation in all age groups (educational categories).

Together with the decline in agriculture and blue collar occupations, the most important change in the employment structure during the 1977-1992 period in Spain is the terciarization of the economy. This has been mainly caused by the increase in the size of the public sector. Such changes in the occupational mix caused a 10%

Sánchez (1991) also obtains for the 1977-1988 period that a DM term and a term that includes our GC

increase in gender segregation. This is offset by changes in the gender composition across occupations in a scenario characterized by a considerable increase in the proportion of females, which goes from 28.6% to 32.9% of the employed population. The net result is a small, not statistically significant 2% decrease in gender segregation over this period.

The occupational categories where changes in gender composition induce a reduction in gender segregation are the following: agriculture and closely related female blue collar occupations; female white collar occupations where the private sector plays a larger role than the public sector; and male white collar, professional and managerial occupations. The main occupational categories responsible for an increase in gender segregation are the male blue collar occupations, where the inroads made by women are not enough to offset the increase in the proportion of females in the employed population. In integrated white collar occupations, as well as female white collar occupations where the public sector has a dominant position, the increased presence of women also leads to slight increases in gender segregation.

This last result is intriguing, because in Spain, as in many other countries, openings in certain occupations within the public sector are filled through publicly advertised examinations, open to anyone with the appropriate educational credentials. Therefore, it would appear that in the public sphere there is less room for gender discrimination and we might expect occupational gender segregation in the public sector to be smaller than in the private sector. Hence, gender segregation

---

concept have positive and negative signs, respectively.

in female white collar occupations where the public sector has a dominant position might be smaller than in other occupations. In fact, to properly study differences between the private and the public sector, it would be necessary to look *inside* occupations where the public sector reaches a minimum size, like female white collar occupations. Then it would be possible to study whether gender segregation is larger in the private or the public sector, and whether gender segregation has increased or not in these two sectors. These issues, which are being separately investigated[19], provide a good example of the type of problems that can be analyzed using the additive decomposability property of the gender segregation index presented in this paper.

On the other hand, to our knowledge, all previous studies on gender segregation, including this one, refer exclusively to the employed population. However, individual occupational choices are conditional on the labor market participation and human capital investment decisions made prior to the occupational choice. Thus, a possible extension of this paper's approach is to consider, not only the gender segregation of the employed population, but also the gender segregation of the entire non-student population of legal working age.

[19] See Mora and Ruiz-Castillo (2003).

DATA APPENDIX

The Spanish data for this study comes from EPA (*Encuesta de Población Activa*), a labor force survey conducted by the *Instituto Nacional de Estadística*. The EPA consists of about 50,000 household observations per quarter, representative of the Spanish household population living in private residential housing. It investigates the relationship with economic activity and other characteristics of every household member over 14 years of age. The EPA is a rotating panel in which each household is interviewed during 7 consecutive quarters; thus, one eighth of the sample is renewed every quarter. In this paper, data from the second quarter is taken as representative of the year as a whole.

The time period in this paper starts in 1977, the first year for which micro-economic data is available in electronic support. In 1993 and 1994 there are fundamental changes in the National Classification of Occupations (NCO) and in the National Classification of Industries (NCI), making it impossible to compare the 1977 data with the period starting in 1993. Therefore, the period studied is 1977 – 1992.

According to EPA, the employed population in 1977 and 1992 is, approximately, 12,148,346 and 12,361,738 people, respectively. There are 71,864 and 62,332 individual observations in 1977 and 1992, respectively, which can be classified according to the two-digit NCI of 1974 and the two-digit NCO of 1979.[20] It is clear that the use of more detailed categories leads to larger index values, since broader categories mask some of the segregation within them (England, 1981). Consequently, researchers have always sought to work with the largest possible occupation's space.[21] However, the idea that, *ceteris paribus*, the larger the number of occupations the better, has been questioned because of the possible bias due to small cell size (Blau *et al.*, 1998): random allocations of individuals across occupations may generate relatively high levels of gender segregation purely by chance. Moreover, when the number of occupations is very large, results on segregation are difficult to interpret. Finally, in this paper occupations must be large enough in order to be meaningfully partitioned by age/education characteristics. Given that we are limited by a relatively small sample size because our data come from a labor force survey rather than a Census, we need to search for the smallest

[20] Because EPA is a labour force survey rather than a census, there are a relatively low number of two-digit occupations and industries. In Herranz *et al.* (2003) occupations are taken as the basic partition and combined with the list of two-digit industries to obtain a 106 occupational classification.

[21] In empirical studies using Census data, the occupational space typically reaches several hundred categories. For instance, in the U.S. Blau et al. (1998) work with 470 occupations from the 1970, 1980, and 1990 Census.

possible set of occupations.

Herranz *et al.* (2003) explore how far it is possible to aggregate an initial list of occupations without reducing the gender segregation value too much. Using an algorithm based on the bootstrap, that paper shows that an original list of 106 occupations for 1977 and 1992 can be aggregated into a common list of 29 occupational categories, which are fully described below.

Individuals in each occupation must be further partitioned according to productive characteristics. The available sample size limits the number of subgroups that can be considered. In particular, we distinguish three age categories (16-30; 31-50; 51-99), and four educational attainment levels (illiterates and without formal studies or "low education"; with less than 9 years of education or "primary education"; between 9 and 12 years of education or "secondary education"; and College education). Since it might be argued that the educational experience varies considerably by age, a final partition consisting of 11 age/education subgroups (where the low education category had to be combined with a 16-50 age interval) has been constructed.

LIST OF OCCUPATIONS

The 106 initial occupations are listed within the 29 final categories obtained with the bootstrap algorithm.

MALE

Agriculture

1 Independent farm workers, fishermen in farms and other agricultural production.
Farm workers, ranchers, ranch hands in other industries

2 Fish and game workers
Forestry workers

Blue collar

3 Construction workers and bricklayers
Drivers, other transport personnel
Electricians in other industries
Iron and steel workers
Miners and quarry workers.
Machine operators, radio & TV station operators, and sound-system operators
Stonemasons
Chemical laboratory workers in other industries

4 Construction workers in other industries

Foundry workers
Furniture makers and carpenters
Workers not classified in other subgroups (unskilled workers) in services
Graphic arts workers
Wood and paper mill workers
Painters
Furriers and leather workers

5 Mechanics, machinists, watchmakers and other precision mechanics
Shoemakers in repair services

6 Plumbers, welders, sheet metal workers

White collar

7 Personnel in protection and security services
Foremen and overseers
Mailroom workers and office assistants
Engineers, inspectors, and conductors in passenger transport

8 Employees in accounting, cashier, teller positions in other industries
Sculptors, painters, decorators, photographers
Sales assistants, sales representatives in wholesale trade
Stockbrokers, bonds brokers, real estate agents, and insurance brokers
Accountants and bookkeepers
Adding machine operators and data processors

9 Sales personnel and sales representatives

Professional and managerial

10 Companies Directors and managers
Owners or managers of commercial establishments in wholesale trade
Head of sales and head buyers
Inspectors of transport and communication services
Operator of agricultural or fishing enterprises
Directors and managers of commercial establishments
Owners or managers of commercial establishments in other industries
Members of governmental branches

11 Owners or managers of hotel, restaurant services in restaurants
Head clerks and office managers
Directors and managers of hotel in restaurant services

12 Physicians, veterinarians, and pharmacists
Legal professionals
Professional musicians and show business professionals
Statisticians, mathematicians, computer analysts, and other like technicians
Economists
Chemists, physicists, and geologists
Writers and journalists
Biologists and agricultural and forestry specialists
Sports professionals

13 Draftsmen and engineering technicians
Architects and engineers
Pilots and Officers of air and maritime navigation

Armed forces

14 Members of the Armed Forces

INTEGRATED

Agriculture

15 Farm workers, ranchers, and ranch hands in farms
Independent farm workers and fishermen in livestock production

Blue collar

16 Food and drink preparation workers in food and kindred products
Workers not classified in other subgroups (unskilled workers) in agriculture and industry
Cargo handlers in other industries
Cargo handlers in agriculture and mining
Glass and ceramic factory workers
Rubber and plastic manufacturing plant workers
Chemical laboratory workers in chemicals and allied products

17 Electricians in equipment manufacturing
Crafts people and similar not classified in above subgroups
Jewelers and silversmiths
Garment workers: upholsterers

White collar

18 Employees in administrative services in non-classified areas in other services
Employees in administrative services in non-classified areas in agriculture and mining
Employees in administrative services in non-classified areas in wholesale trade
Employees in administrative services in non-classified areas in hotels and restaurants
Supervisors of domestic service personnel

FEMALE

Agriculture

19 Farm workers, ranchers, and ranch hands in livestock production

Blue collar

20 Textile workers
Cargo handlers in manufacturing.
Food and drink preparation workers in other industries
Shoemakers in other industries
Paper and cardboard factory workers
Tobacco production workers

21 Garment workers: other

White collar

22 Sales assistants and sales representatives in retail
Employees in administrative services in non-classified areas in retail
Sales assistants and sales representatives in other industries

23 Concierges, building supervisors, and cleaning service personnel in other services
Hair stylists and beauty treatment personnel
Concierges, building supervisors, and cleaning service personnel in trade and transport,
Chefs, cooks, and food service personnel in other industries
Dry cleaning and laundry service employees
Telephone and telegraph operators
Concierges, building supervisors, and cleaning service personnel in agriculture and mining

24 Chefs, cooks, and food service personnel in hotels, restaurants, and other lodging services
Personnel in other services not classified in other subgroups in education and health
Personnel in other services not classified in other subgroups in other industries

25 Concierges, building supervisors, and cleaning service personnel in personal household

26 Domestic service personnel and other like personnel
Stenographers, typists, and key-punch operators

27 Medical, veterinary, and pharmaceutical assistants and technicians
Employees in accounting, cashier, and teller positions in trade and miscellaneous repair

Professional and managerial

28 Owners or managers of commercial establishments in retail
Owners or managers of hotel, restaurant services in hotels and other lodging services

29 Teachers
Professionals or technicians in non-classified areas

REFERENCES

Albelda, R. (1986), "Occupational Segregation by Race and Gender, 1958-1981", *Industrial and Labor Relations Review*, **39**: 404-411.

Anker, R. (1997), "Theories of Occupational Segregation by Sex: An Overview", *International Labour Review*; 136(3), Autumn 1997, pages 315-39.

Beller, A. (1985), "Changes in the Sex Composition of U.S. Occupations, 1960-1981", *Journal of Human Resources*, **20**: 235-250.

Bergmann, B. (1974), "Occupational Segregation, Wages and Profits When Employers Discriminate by Race or Sex", *Eastern Economic Journal*, **1**: 103-110.

Blau, F. (1977), *Equal Pay in the Office*, Lexington: Heath.

Blau, F. and W. Hendricks (1979), "Occupational Segregation by Sex: Trends and Prospects", *Journal of Human Resources*, **12**: 197-210.

Blau, F., P. Simson, and D. Anderson (1998), "Continuing Progress? Trends in Occupational Segregation Over the 1970s and 1980s", *Feminist Economics*, **4**: 29-71.

Deutsch, J, Y. Flückiger y J. Silber (1994), "Measuring Occupational Segregation", *Journal of Econometrics*, **61**: 133-146.

England, P. (1981), "Assessing Trends in Occupational Sex Segregation, 1900-1976", en I. Berg, ed., *Sociological Perspectives on Labor Markets*, Academic Press.

Duncan, O. And B. Duncan (1955), "A Methodological Analysis of Segregation Indexes", *American Sociological Review*, **20**: 210-217.

Flückiger, Y. and J. Silber (1999), *The Measurement of Segregation in the Labor Force*, Heidelberg: Physica-Verlag.

Fuchs, V. (1975), "A Note on Sex Segregation in Professional Occupations," *Explorations in Economic Research*, **2**: 105-111.

Gross, E. (1968), "Plus ça change…? The sexual structure of occupations over time", *Social Problems*, **16**: 198-208.

Herranz, N., R. Mora and J. Ruiz-Castillo (2003), "An Algorithm to Reduce the Occupational Space in Gender Segregation Studies", Working Paper No. 03-14,

Economic Series 03, Universidad Carlos III.

Hutchens, R. M. (1991), "Segregation Curves, Lorenz Curves and Inequality in the Distribution of People Across Occupations", *Mathematical Social Sciences*, **21**: 31-51.

Jacobs J. (1989), "Long-Term Trends in Occupational Segregation by Sex", *American Journal of Sociology*, **95**: 160-173.

Jacobsen, J. (1994), "Trends in Work Force Sex Segregation, 1960-1990", *Social Science Quarterly*, **75**: 204-211.

James, D.R. and K.E. Taeuber (1985), "Measures of Segregation", in N.B. Tuma (ed.), *Sociological Methodology*, San Francisco: Jossey-Bass, 1-31.

Karmel, T. And M. Maclahlan (1988), "Occupational Sex Segregation: Increasing or Decreasing?," *Economic Record*, **64**: 187-195.

Mora, R. and J. Ruiz-Castillo (2003), "Gender Segregation By Occupations In the Public and the Private Sectors. The Case of Spain In 1977 and 1992", Working Paper No. 03-16, Economic Series 05, Universidad Carlos III.

Preston, J., and U. Brandeis (1999) "Occupational Gender Segregation: Trends and Explanations", *Quarterly Review of Economics and Finance*; 39(0), Special Issue 1999, 611-24.

Sánchez, M (1991), "La segregación de las mujeres en el mercado de trabajo español: evidencia para el período 1964-1988," *Hacienda Pública Española*, 123-139.

Silber, J. (1989), "Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality," *Review of Economics and Statistics*, **LXXI**: 107-115.

Siltanen, J., J. Jarman, and R. Blackburn (1993), *Gender Inequality in the Labor Market: Occupational Concentration and Segregation, A Manual on Methodology*, Geneva: International Labor Office.

Theil, H. (1971), *Principles of Econometrics*, New York: John Wiley and Sons.

Theil, H. and A. J. Finizza (1971), "A Note on the Measurement of Racial Integration of Schools by Means of Information Concepts," *Journal of Mathematical Sociology*, **1**: 187-194.

Williams, G. (1979), "The Changing U.S. Labor Force and Occupational Differentiation By Sex," *Demography*, **16**: 73-88.

Zoloth, B. (1976), “Alternative Measures of School Segregation”, *Land Economics*, **52**: 278-298.

Table 1. Descriptive Statistics and Gender Segregation Indices in the Partition by Occupations, 1977

OCCUPATION	Female[a]	Male[b]	Total[c]	W_i[d]	I^{ie}	I_i[f]	I(i)[g]	a_i[h]
MALE	**10.17**	**62.91**	**47.81**	**6,09**				
Agriculture	**4.26**	**10.75**	**8.89**	**13,72**				
1.	4.24	9.40	7.92	15,32	7.06	1.03	8.10	26.94
2.	0.02	1.36	0.98	0,71	43.48	0.52	44.01	146.47
Blue Collar	**2.20**	**35.28**	**25.81**	**2,44**				
3.	0.21	15.98	11.47	0,51	44.68	0.12	44.80	149.11
4.	1.74	13.13	9.87	5,05	26.46	1.77	28.23	93.95
5.	0.23	4.15	3.03	2,16	36.48	1.21	37.69	125.44
6.	0.03	2.02	1.45	0,57	44.36	0.76	45.11	150.15
White Collar	**1.46**	**7.04**	**5.45**	**7,69**				
7.	0.26	3.05	2.25	3,32	32.02	1.87	33.89	112.79
8.	1.12	2.47	2.08	15,43	6.94	4.18	11.12	37.01
9.	0.08	1.53	1.12	2,08	36.82	1.28	38.10	126.82
Prof. & Manag.	**2.24**	**8.70**	**6.85**	**9,38**				
10.	0.26	3.38	2.49	3,02	33.07	0.76	33.84	112.62
11.	1.09	2.02	1.75	17,82	4.53	3.60	8.13	27.07
12.	0.74	1.61	1.36	15,52	6.84	2.80	9.64	32.10
13.	0.15	1.69	1.25	3,51	31.33	2.67	34.00	113.17
Armed Forces: 14.	**0**	**1.13**	**0.81**	**0**	48.65	0	48.65	161.92
INTEGRATED	**25.34**	**21.83**	**22.83**	**31,77**				
Agriculture: 15	**12.01**	**9.28**	**10.06**	**34,17**	1.05	1.65	2.69	8.97
Blue Collar	**5.69**	**7.73**	**7.14**	**22,79**				
16.	4.58	6.52	5.97	21,98	1.64	6.40	8.05	26.79
17.	1.11	1.20	1.18	26,89	0.11	4.87	4.98	16.56
White Collar: 18.	**7.64**	**4.82**	**5.63**	38,89	3.50	10.55	14.05	46.76
FEMALE	**64.48**	**15.26**	**29.35**	**62,88**				
Agriculture: 19.	**3.12**	**1.19**	**1.75**	**51,21**	16.20	8.96	25.16	83.72
Blue Collar	**13.86**	**3.06**	**6.15**	**64,48**				
20.	6.90	2.60	3.83	51,56	16.69	3.50	20.19	67.19
21.	6.97	0.46	2.32	85,76	102.64	3.70	106.34	353.91
White Collar	**37.95**	**7.32**	**16.09**	**67,51**				
22.	10.55	2.78	5.00	60,39	31.40	1.47	32.86	109.38
23.	7.68	1.46	3.24	67,80	47.37	5.89	53.25	177.24
24.	5.25	2.30	3.14	47,81	11.81	3.04	14.85	49.42
25.	7.44	0.36	2.39	89,16	116.69	6.32	123.01	409.41
26.	4.87	0.13	1.48	93,88	139.16	3.91	143.08	476.19
27.	2.15	0.30	0.83	74,37	64.58	8.61	73.19	243.58
Prof. & Manag.	**9.55**	**3.68**	**5.36**	**50,98**				
28.	5.15	2.43	3.21	45,87	9.61	5.39	15.00	49.93
29.	4.40	1.25	2.15	58,59	28.03	1.31	29.34	97.66
TOTAL	**100.00**	**100.00**	**100.00**	**28.62**	$\hat{I}^{B}_{(j)}$**=27.00**	$\hat{I}^{W}_{(j)}$**=3.04**	**I=30.04**	

Notes

[a] **Female**: $100(F_j / F)$; [b] **Male: $100(M_j/M)$;** [c] **Total**: $100(T_j / T)$; [d] $W_j = 100(F_j / Tj)$

[e] $I^{B}_{(j)} = (T_j/T)\, I^j$ = Direct gender segregation induced by occupational choices

[f] $I^{wi}_{(j)} = (T_j/T)\, I_j$ = Gender segregation induced by age / education characteristics within occupations

[g] $I = I^{B}_{(j)} + I^{wi}_{(j)}$ = Gender segregation in the employed population

[h] $a_j = I(j)/I$

Table 2. Descriptive Statistics and Gender Segregation Indexes in the Partition by Age/ Education Characteristics, 1977

AGE/EDUCATION	Female[a]	Male[b]	Total[c]	W_i[d]	I^i[e]	I_i[f]	I(i)[g]	a_i[h]
LOW EDUCATION	**17.35**	**16.54**	**16.77**					
1. 16 - 50	8.76	9.32	9.16	27.38	0.06	31.23	31.29	104.13
2. More than 50	8.58	7.22	7.61	32.29	0.46	26.35	26.81	89.24
PRIMARY EDUCATION	**60.22**	**64.79**	**63.48**					
3. 16 - 30	25.57	18.92	20.82	35.15	1.45	35.34	36.78	122.42
4. 31 - 50	21.93	29.40	27.26	23.03	1.15	28.65	29.81	99.20
5. More than 50	12.71	16.47	15.40	23.64	0.91	21.99	22.90	76.22
SECONDARY EDUCATION	**16.84**	**13.36**	**14.36**					
6. 16 - 30	12.81	6.68	8.43	43.48	7.19	27.80	34.99	116.46
7. 31 - 50	3.02	4.79	4.29	20.19	2.69	24.19	26.88	89.47
8. More than 50	1.00	1.89	1.64	17.51	4.81	16.23	21.04	70.02
COLLEGE EDUCATION	**5.60**	**5.31**	**5.39**					
9. 16 - 30	2.34	1.17	1.50	44.64	8.32	22.77	31.09	103.48
10. 31 - 50	2.28	2.70	2.58	25.35	0.39	23.73	24.11	80.26
11. More than 50	0.97	1.45	1.31	21.21	2.06	20.22	22.28	74.15
TOTAL	**100.00**	**100.00**	**100.00**	**W=28.62**	**$I^B_{(i)}$ =1.77**	**$I^{Wj}_{(i)}$=28.27**	**I=30.04**	

Notes

[a]**Female**: $100(F_i/F)$; [b]**Male**: $100(M_i/M)$; [c]**Total**: $100(T_i/T)$; [d]$W_i = 100(F_i/Ti)$

[e]$I^B_{(i)} = (T_i/T)\, I^i$ = Direct gender segregation induced by age/education characteristics

[f]$I^{Wj}_{(i)} = (T_i/T)\, I_i$ = Gender segregation induced by occupational choices within age/education characteristics

[g]$I = I^B_{(i)} + I^{Wj}_{(i)}$ = Gender segregation in the employed population

Table 3. Descriptive Statistics and Gender Segregation Indexes in the Partitions by Age/ Education Characteristics, 1992

AGE/EDUCATION	Female[a]	Male[b]	Total[c]	W_i[d]	I^{ie}	I_i[f]	I(i)[g]	a_i[h]
LOW EDUCATION	**10.06**	**9.67**	**9.80**					
1. 16 - 50	4.38	4.33	4.35	33.18	0.00	34.81	34.81	118,71
2. More than 50	5.68	5.34	5.45	34.28	0.06	28.12	28.18	96,09
PRIMARY EDUCATION	**30.25**	**39.48**	**36.44**					
3. 16 - 30	4.45	5.53	5.18	28.27	0.72	35.51	36.23	123,53
4. 31 - 50	17.10	21.90	20.32	27.68	0.92	31.24	32.15	109,64
5. More than 50	8.70	12.05	10.95	26.16	1.55	25.32	26.87	91,63
SECONDARY EDUCATION	**42.09**	**40.28**	**40.88**					
6. 16 - 30	25.66	20.77	22.38	37.72	0.74	29.15	29.89	101,94
7. 31 - 50	14.76	16.54	15.95	30.45	0.20	26.01	26.21	89,36
8. More than 50	1.67	2.98	2.55	21.59	4.50	19.29	23.80	81,14
COLLEGE EDUCATION	**17.60**	**10.57**	**12.88**					
9. 16 - 30	6.42	2.22	3.60	58.67	20.07	15.91	35.98	122,70
10. 31 - 50	9.58	6.36	7.42	42.50	2.89	19.84	22.73	77,52
11. More than 50	1.59	1.99	1.86	28.16	0.75	24.38	25.13	85,71
TOTAL	**100.00**	**100.00**	**100.00**	**W=32.90**	**$I^{B}_{(i)}$ =1.66**	**$I^{wj}_{(i)}$=27.67**	**I=29.33**	

Notes

[a]**Female**: $100(F_i/F)$; [b]**Male**: $100(M_i/M)$; [c]**Total**: $100(T_i/T)$; [d]$W_i = 100(F_i/Ti)$

[e]$I^{B}_{(i)} = (T_i/T)\, I^i$ = Direct gender segregation induced by age / education characteristics

[f]$I^{wj}_{(i)} = (T_i/T)\, I_i$ = Gender segregation induced by occupational choices within age / education characteristics

[g]$I = I^{B}_{(i)} + I^{wj}_{(i)}$ = Gender segregation in the employed population

[h]$a_i = I(i)/I$

Table 4. 1977 *Versus* 1992: Gender Composition and Demographic Mix Effects In the Partition by Age/education (1977 demographic weights, i. e. $b_i = T_{i77}/T_{77}$)

	GC[a]	DM[b]	TOTAL[c]
LOW EDUCATION			
1. 16 - 50	0.3	- 1.7	- 1.4
2. More than 50	0.1	- 0.6	- 0.5
PRIMARY EDUCATION			
3. 16 - 30	- 0.1	- 5.7	- 5.7
4. 31 - 50	0.6	- 2.2	- 1.6
5. More than 50	0.6	- 1.2	- 0.6
SECONDARY EDUCATION			
6. 16 - 30	- 0.4	4.2	3.7
7. 31 - 50	0.0	3.1	3.0
8. More than 50	0.0	0.2	0.3
COLLEGE EDUCATION			
9. 16 - 30	0.1	0.8	0.8
10. 31 - 50	0.0	1.1	1.1
11. More than 50	0.0	0.1	0.2
TOTAL	**1.2**	**- 1.9**	**- 0.7**

Notes

[a]**GC =** Change in gender segregation induced by gender composition changes in the partition by age/education characteristics

[b]**DM =** Change in gender segregation induced by changes in age/education categories' demographic mix

[c]**TOTAL = GC + DM**

Table 5. 1977 *Versus* 1992: Gender Composition and Demographic Mix Effects In the Partition by Occupations. Selected Occupational Categories (1977 demographic weights, i. e. $b_j = T_{j77}/T_{77}$)

OCCUPATION	D $(F_j/F)^a$	D $W_j{}^b$	GC[c]	D $(T_j/T)^d$	DM[e]	TOTAL[f]
GROUP 1	**- 26.5**	**(-)**	**- 3.8**	**- 14.3**	**- 2.4**	**- 6.2**
I. Agriculture (1, 2, 15, 19)	-11.1	(+)	- 0.2	- 10.8	- 0.7	- 0.9
II. Female, WC[h], Private (22, 24, 25)	- 7.3	(-)	- 2.3	- 0.7	- 0.8	- 3.1
III. Female, BC[g] (20, 21)	- 6.9	(-)	- 1.1	- 2.2	- 0.8	- 1.9
IV Integrated, BC[g] (16, 17)	- 1.2	(-)	- 0.2	- 0.5	0.0	- 0.2
GROUP 2	**9.8**	**(+)**	**- 1.1**	**7.5**	**1.3**	**0.2**
V. Male, WC[h] (7, 8, 9)	2.5	(+)	- 0.5	2.4	0.4	- 0.1
VI. Male, PM[l] (10, 11, 12, 13)	3.1	(+)	- 0.4	2.3	0.2	- 0.1
VII. Female, PM[l] (28, 29)	4.1	(+)	- 0.2	2.8	0.6	0.4
GROUP 3	**16.8**	**(+)**	**1.1**	**6.9**	**4.2**	**5.3**
VIII. Male, BC[g] (3, 4, 5, 6)	0.5	(+)	0.7	- 0.1	0.1	0.8
IX. Integrated, WC[h] (18)	6.0	(+)	0.2	2.6	0.5	0.7
X. Female, WC[h], Public (23, 26, 27)	10.2	(+)	0.1	4.4	3.7	3.7
XI. Armed Forces: (14)	0.0	(≈)	0.1	- 0.1	- 0.1	0.0
TOTAL	**0.0**	**(+)**	**- 3.8**	**0.0**	**3.1**	**- 0.7**

Notes

[a] $\mathbf{D\ (F_j/F) = (F_{j92}/F_{92}) - (F_{j77}/F_{77})}$ = Change in the female frequency distribution across occupations

[b] $\mathbf{D\ W_j = W_{j92} - W_{j77}}$ = Change in the female proportion across occupations

[c] $\mathbf{GC_j = [I_{j92} - I_{j77}]\ (T_{j77}/T_{77})}$ = Change in gender segregation induced by gender composition changes in the partition by occupations

[d] $\mathbf{D\ (T_j/T) = (T_{j92}/T_{92}) - (T_{j77}/T_{77})}$ = Change in the population frequency distribution across occupations

[e] $\mathbf{DM_j = [(T_{j92}/T_{92}) - (j_{i77}/T_{77}))]\ I_{j92}}$ = Change in gender segregation induced by changes in occupations' demographic mix

[f] $\mathbf{TOTAL_j = GC_j + DM_j}$

[g] **BC** = Blue Collar; [h] **WC** = White Collar; [l] **PM** = Professional and Managerial