Does Schooling Contribute to Increase Individuals´ Chances to Access The More Affluent Income Groups?

Héctor R. Gertel
Roberto Giuliodori
Alejandro F. Rodríguez

August 2002

Facultad de Ciencias Económicas
Universidad Nacional de Córdoba

To be presented at the XXXVII Reunión Anual de la AAEP, Tucumán, November 13-15, 2002

Does Schooling Contribute to Increase Individuals´ Chances to Access The More Affluent Income Groups?

Héctor R. Gertel*
Roberto Giuliodori
Alejandro Rodríguez**

1. Introduction

The issue of inequality between groups of economic agents, or relative affluence, its pattern of evolution and the equalizing role attached to education is the object in this research paper. Earlier empirical work has not given precise responses to these research problems of accounting for relative affluence between distributions, although it had long determined the role played by individual and family optimizing behavior of decisions on human capital investments and of circumstances on inequality within each distribution of income. Yet, whether the range and significance of these choices on schooling, training, information and health produced "large" or "small" distances between populations, in the context of studies of education and income distribution is a question that remained largely unsettled; therefore, is worthwhile to research the contribution of schooling to narrow the distance among income recipient units.

Here, we are concerned not with the individual variations within a group, which played a prominent part in the investigations stimulated after Arrow (1973) famous filtering hypothesis and Mincer (1974) seminal work on the decomposition of the variance of the logarithm of income, but with our initial question about the degree of closeness between groups of individuals separated by some statistical attribute, a question of a more general nature introduced in Rao (1952), and investigated more recently in Batacharia and Mahalannobis (1967) and Dagum (1985 and 1997) among others[1] To that end, we specify and estimate a model of the economic distance among income distributions by level of education and gender that will help to measure, in terms of a differential mass of income accruing to each subpopulation, the effect of sorting and grouping economic agents by their educational credentials. This type of model, by allowing a weighted comparison between all binary combinations of subpopulations (having more/less schooling years) produces stronger theoretical results that those obtained through the un-weighted Bhattacharya and Mahalanobis generalized and Euclidian distance methods (Dagum, 2001). The work covers both positive and policy issues, but its central focus is on the role of education in increasing or reducing economic inequality. The paper is organized as follows: the next section introduces the conceptual problems with measuring income distribution differentials with an economic meaning associated with gender and education attached to it. Section three specifies the model of economic distance that measures the degree of differentiation between two subpopulations, while section four presents the data and results of an application of the model that measures the degree of differentiation attributable to education and gender for wage-earners in Greater Córdoba, a typical mid-size metropolitan area in Argentina, during the 1990s. A final section summarizes the main conclusions concerned with the role of education in increasing or reducing gender income differentials and its trend during the decade.

* corresponding author: hgertel@eco.uncor.edu
** Héctor Gertel is with Instituto de Economía y Finanzas - Universidad Nacional de Córdoba; Roberto Giuliodori is with Departamento de Estadística y Matemática – Facultad de Ciencias Económicas –Universidad Nacional de Córdoba; Alejandro Rodríguez is a graduate student from Departamento de Economía – Universidad Nacional de Córdoba. This article was partially funded by the Agencia Córdoba Ciencia (Project ACSE-034002304). The authors are grateful to Paula Auerbach who has collaborated in early stages of this Research Project.

2. Schooling, Gender and Economic Inequality

Hypothesis about the equalizing role of schooling in society abound in the literature. The research reported below is not intended to investigate all the forces and factors affecting the distribution of education and income. Far most modest, the purpose of the work is to make use of the concept of economic distance discussed in Dagum (1985, 1997): to explore how schooling decisions may have affected the gender distribution and structure of personal income by generating an economic distance between groups and how the distances evolved during the past decade in Argentina.

2.1 The Nature of the Problem

Two approaches are currently used to analyze schooling and economic inequality. The first, deals with the specification of scalar measures of the relative degree of inequality in the distribution of income among the unattached members of a population of economic units. A well known measure of this class is the traditional Gini coefficient. It has been observed that the Gini ratio of years of schooling is lower than the Gini ratio of wage-earnings obtained for the same population (Ram, 1984, 1990). The second provides a description of how income distributions between subpopulations differ on the basis of the attached socio-economic and personal characteristics that identify membership to each one of the subpopulations. The decomposition of the Gini coefficient that accounts for the within and between variations, entropy measures and the variance of the log of income belong to this second class. These measures can be linearly decomposed to calculate the relative contribution of one or more attributes to explain between income distribution inequality. Because having invested in education attach an identifying label to individuals, in planning reforms, education policy-makers would gain significant new insights by asking the following questions: Does education labels create membership to exclusive clubs? Are members of a "more education" club more affluent than members of a "less education" club? Is this economic value of membership increasing or decreasing over time? All of these questions may besides be stated by gender.

Under the second approach, the use of alternative decomposition measures to assess the relative contribution of education to the increase or decrease of income inequality expanded considerably after the seminal work by Mincer (Mincer, 1974) stimulated a first round of research on income distribution and inequality by education, experience and sex. Observations on both, the effect of the average level of schooling weighted by its average rate of return on average income of individuals, and of the variance in years of education as partially explaining the variance of income, led to the conclusion that education exerted a net equalizing effect on income distribution. Comparative studies of education and income distribution were conducted in Peru, Brazil and Mexico in the late 1970s. These studies found evidences of an equalizing effect due to increased average schooling in the labor force, despite the fact that overall income distribution in these countries deteriorated during the period (Carnoy, et.al., 1978). Similar conclusions are found for Argentina, in Gertel, De Santis and Pereyra (1987) and Delfino (1998). It is in this sense that, expanding enrollments became a powerful driving force within policies to promote the goals of social and economic mobility in Latin America. Yet, rising average schooling on the population today would not necessarily mean that expectations of improved income and employment would automatically be fulfilled tomorrow for all, because more years of schooling means also higher diversification of opportunities and an increased segmentation in the population. Thus, the positive view about the power of education as an equalizing device that emerged from variance analysis of the Mincer equation, has to be balanced against the screening hypothesis of Arrow, where education contributed, perhaps through the work of self-selection mechanisms, to perpetrate the economic distance between subpopulations in society.

The second round of research on schooling and income distribution focused more on issues associated with identification of the self-selection mechanisms in the process of acquiring more education, casting some doubts on the compensatory power originally attributed to policies of increased schooling. In fact, an increase in the average level of schooling in the population appears to be associated also with a wider variance in income and employment of the individuals due, in part, to the screening or sorting effect of education predicted in Arrow (1974), reinforcing the idea that education can not generate equality alone (Levin and Kelley, 1994, Patrinos, 1995). Clearly, the principal issue in this debate is that by increasing the level of education in the population, educational policy, most probably as an unwanted by-product, may have stimulated the formation of wage-earners clubs with distinct educational credentials. Thus, if education has contributed to create a distance among groups of economic agents, a third set of questions is in place. How is this distance measured? And, how it evolved over time?.

2.2 The Problem of Measuring Distance Between Distributions

The problems of identification and separation in the measurement of the within and between inequality among distributions attracted considerable interest in the past few decades. Popular measures designed to this purpose are of two kinds: the un-weighted class of measures, constructed upon the idea of calculating the overlapping probability area between distributions; and the weighted class of measures, where the overlapping area of the joint distribution is multiplied by a weighting factor in the dimension of income, in order to attach an economic meaning to the measure of distance.

2.2.1 Generalized un-weighted measures

This intuitive idea was introduced by Rao (1952) in his early treatment of clusters and grouping to compare the distance between any two pairs of groups. This issue was also examined by Battacharya and Mahalanobis (1967) in a study of regional disparities in household consumption in India, introducing a decomposition of the Gini ratio to study the relative contribution of the within and the between concentration coefficient. Another early contribution is that of Theil (1967). However, unless the distributions in the comparison are all non-overlapping, an extreme possibility, these class of measures produce an underestimation of the relative contribution computed for the "between groups" component. To avoid this problem, Ebert (1984) and Shorrocks (1984), independently introduced a class of measures that is derived by an axiomatic approach and it is characterized by the properties normally associated with a metric distance function and a set of axioms originally suggested in Rao (1952). Chakravarty and Dutta (1987) observed that the contributions of Ebert and of Shorrocks were of limited use in policy making because metric distance measures are not designed to capture differences in welfare of two distributions and proposed, instead, an axiomatic welfare function that characterizes an ethical distance measure. Nevertheless, the generalized distance of Mahalanobis and a particular measure of it, known as Euclidian distance, are derived from this basic idea but its use would require comparison of normal distributions. Figure I will help to illustrate the general case of normal distributions and equal variance.

The income distributions have to be compared directly, by setting an adequate partition of the total population. In Figures I (a), (b), (c) and (d) the distribution "M", to the right, corresponds to individuals with more education. In figure I (a), to the left of income level Y_1, there is no possibility of error. Any individual with income level $Y<Y_1$ belongs to subpopulation "L", with less education. Yet, to the right of Y_1 there is a possibility of error: individuals with $Y_2>Y_1$ may be part of subpopulation L or M. The area α represents the probability to assign to population L an individual that in fact belongs to population M. A decrease in area α would imply a lower probability of allocation error. In fact, this also would mean that the distance between distributions has increased. Consequently, α has a double character of being a probability (it assumes values in the range 0,1), and a measure of the distance between distributions. And the expression $1 - \alpha$ has a clear meaning: when the expression assumes the value 0, there is a perfect overlapping of distributions while the value of 1 reflects perfect separation of distributions.

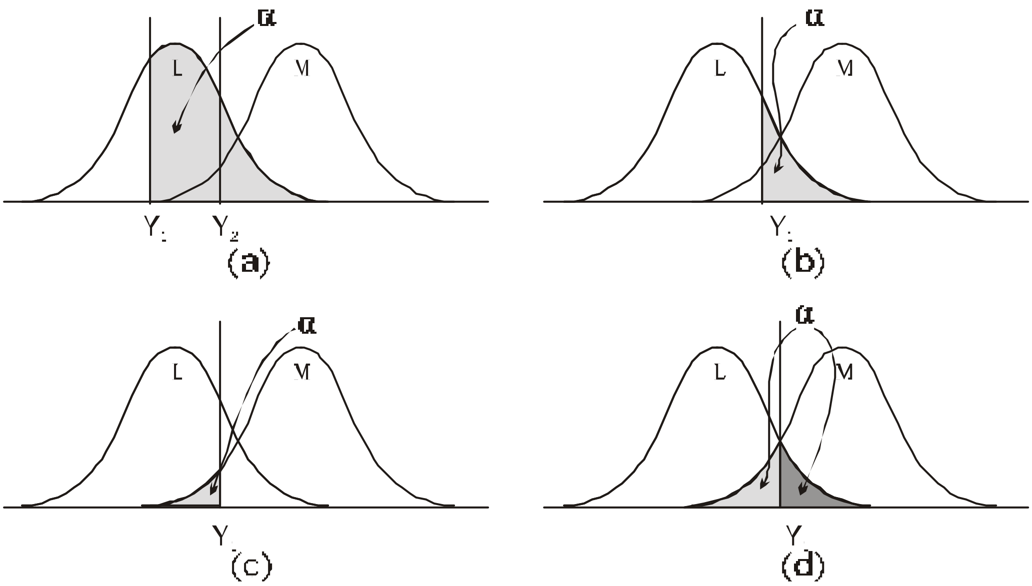

Figure 1

It is easy to observe that α assumes a different value when the probability of error is calculated from distribution "L" , as shown in Figure I (b), or from distribution "M", in Figure I (c), and appears the popular "index-numbers" problem, except when Y_1 is located at the intercepting point of the two distributions, as in figure I (d). It can be demonstrated that under the conditions of normal distributions of equal variance, the value of alpha calculated for Y_1 at the intersecting point is invariant, and can be obtained directly from the Mahalanobis distance in non-orthogonal spaces. This reduces to the Euclidean distance in the case of Orthogonal spaces.

2.2.2 The un-weighted economic measure d_0

The measure d_0 was introduced in Dagum (1985) to address the pervasive effects of the index-numbers problem already discussed. Because income is not normally distributed, a well known fact, relative affluence can not be adequately measured through the class of measures discussed above. Furthermore, it is most likely that two populations of income receivers have different variance. To avoid these problems, Dagum (2001) introduced a new class of metric distance measure, which he called "d_0". It also represents a probability, but it is multidimensional, or bi-dimensional when only two populations are being compared. A simple representation of this last case, for discrete distributions, is

$$d_0 = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} I\left(y_i - x_j\right) \qquad (1)$$

where, y and x are the income of i and j individuals, n and m are the sample sizes, and I summarizes a function that gives indications of whether income y is higher, equal or lower than income x.

The equation (1) for two populations of continuous distributions can be written as:

$$d_0 = P\left[y > x / E_2(y) > E_1(x)\right] = \int_0^\infty \left[\int_0^y f_1(x)dx\right] f_2(y)dy \qquad (2)$$

where fi(y) is the probability density function, that has a CDF Fi(x), and Ei(y) the income expected value of population “i”, with i = 1,2.

A parametric estimation of Fi (x), is obtained, as explained in Gertel, Giuliodori, Auerbach and Rodríguez (2001), from the Dagum three parameters cumulative distribution function,

$$F(x) = \frac{1}{\left(1 + \lambda x^{-\delta}\right)^{\beta}} \qquad (3)$$

where δ and β are equality parameters and λ is a scale parameter.

In equation (1) the expression I(y-x) corresponds to the Indicator Function defined by Dagum, and assumes values as follow:

$$I(y-x) = \begin{cases} 1, \text{ if } & y > x \\ \frac{1}{2}, \text{ if } & y = x \\ 0, \text{ if } & y < x \end{cases} \qquad (4)$$

This definition for d_0 represents the probability that any individual that belongs to the distribution with higher mean income, earns an income that is higher than any other individual who belongs to the distribution with lower average income. A graphic representation for d_0 is provided in Figure 2 below, where x and y values are represented over the V_1 and V_2 axis, respectively, and the frequency distribution is built on the vertical axis. An hypothetical example can be set as follows. Let define a total of five observations for x, and another five for y; then, 25 pair-wise observations are formed, each with a joint probability equal to 0.04 as shown in the vertical axis. Thus, we now have a volume. If d_0 assumes a value of one, it can be interpreted as the total volume of the figure represented in Figure 2, which is associated to the case of completely separated distributions. On the other hand, a value of d_0 in the interval (0,1) is shown as a proportion of the total, to be in this hypothetical example, equal to the volume marked in color. That would indicate the exact degree of overlapping (16 in 25 parts) between the distributions.

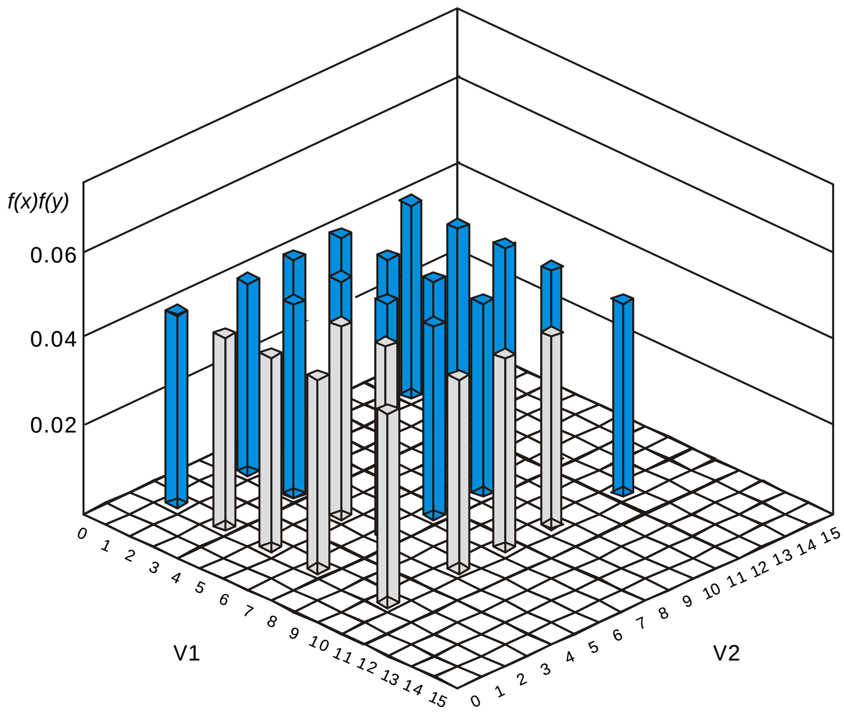

Figure 2

This distribution-free approach based on joint distribution analysis solves the overlapping bias problem associated with the Battacharya and Mahalanobis exercise. However, as Dagum noted, the use of metric distance measures with pure statistical meaning, still is of little help to assess affluence among subpopulations because the dimension of income is missing in the comparison. The next section introduces a new distance measure, d_1, which has the dimension of income, thus it provides with the possibility to explore the role played by attached characteristics of economic agents in explaining differences in affluence among populations.

3. The Model of Income-weighted Measure D_1

A new distance measure d_1 is defined in Dagum (1985) as the mathematical expectation of individual income differences between individuals pertaining to the distribution with higher mean income relative to the population with lower mean income. This measure, that is expressed in monetary units, defines the net economic affluence of the population with higher mean income with respect to the population with the lower mean income, whereas mean income is the criteria for directional economic distance (Dagum, 1997: 524).

By applying the expected value d_1, the author construct an inter-income inequality ratio D_1, that measures the ratio between the net economic affluence and its maximum possible value. Thus, the relative economic affluence ratio D_1 is a normalized measure of dimension zero in the income variable that ranges between zero and one and is invariant with respect to proportional changes of incomes. The usual problems associated with interpreting overlapping in measuring inequality between distributions, and how this new measure helps to find an economic meaning to the comparison among income distributions, are illustrated below. The development of the theoretical model and a numerical illustration of it are presented in this section[2], while the results and discussion of an empirical application to Greater Cordoba are included in the next two sections.

Let consider the economic affluence of subpopulation 2 with respect to 1, the measure d_1 of inter-income distributions inequality developed in Dagum (1985) and expanded in Dagum (2001) is

$$d_1 = \int_0^\infty dF_2(y) \int_0^y (y-x) dF_1(x) \tag{5}$$

where $F_i(y)$ that corresponds to (3), is the cumulative probability density function of population "i", with i=1,2.

For discrete distributions, the value of d_1 is obtained by

$$d_1 = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} (y_i - x_j) I(y_i - x_j) = E[(y-x)I(y-x) / E_2(y) > E_1(x)] \tag{6}$$

where d_1 is the weighted sum of the income differences (y-x) for all y>x, given *$E_2(y)$* > *$E_1(x)$*, and where the weighting factor is the joint density $f_1(x)f_2(y)$.

A solution of (5) is,

$$d_1 = E_2[yF_1(y)] + E_1[yF_2(y)] - E_1(y) \tag{7}$$

As it was said, d_1 has the same dimension as income, and does not have upper limit. Therefore, a normalized and dimensionless measure built on the basis of d_1 proposed by Dagum is

$$D_1 = \frac{d_1 - d_1^*}{\Delta_1 - d_1^*} \tag{8}$$

with d_1* and Δ_1 as the minimum and maximum values, respectively, that d1 can assume. Calculation for d_1* and Δ_1 for continuous distributions are:

$$d_1{*} = \int_0^\infty dF_1(y) \int_0^y (y-x) dF_1(x) = 2E_1[yF_1(y)] - E_1(y) \tag{9}$$

and

$$\begin{aligned} \Delta_1 &= E(|y-x|) = \int_0^\infty dF_2(y) \int_0^\infty (|y-x|) dF_1(x) \\ &= 2d_1 + E_1(y) - E_2(y) \end{aligned} \tag{10}$$

and for discrete distributions are

$$d_1{*} = \frac{1}{nm} \sum_{i=}^{n} \sum_{j=1}^{m} (x_i - x_j) I(x_i - x_j) \tag{11}$$

$$\Delta_1 = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} |y_i - x_j| \tag{12}$$

In short, the ratio D_1 is a novel measure of the degree of closeness between two distributions, with a clear economic meaning because in its construction intervenes d_1, the expected value of the differential mass of income separating the two distributions. Thus, D_1 marks the degree of affluence of one group of individuals over the other in percentage terms of a maximum possible distance between distributions, a result that is obtained by applying (8). Finally, D_1 has the following properties: (i) It is dimension-less, (ii) It assumes values in the interval [0,1], (iii) Its value is zero when the two distributions are identical, and (iv) It assumes the value one when the two distributions are completely separated.

In Table 1, an example that considers three hypothetical situations are considered: scenario I presents identical distributions, scenario II shows that population M is more affluent that L with some overlapping, while scenario III is one of completely separated distributions, with M more affluent that L. The relevant information on income of individuals pertaining to both subpopulations is summarized in each column of Table 1. It also presents results for the average income and the "More Education-Less Education" income differential ratio D_1.

Scenario I		Scenario II		Scenario III	
Income ($)		Income ($)		Income ($)	
Population L	Population M	Population L	Population M	Population L	Population M
1	1	1	4	1	11
3	3	3	8	3	15
5	5	5	10	5	19
7	7	7	15	7	25
10	10	10	20	10	30
Means		Means		Means	
5.2	5.2	5.2	11.4	5.2	20
$D_1 = 0$		$D_1 = 0.91$		$D_1 = 1$	

Number of cases: M=5, L=5

Table 1

A direct estimation of the dimension-free ratio D_1 is obtained by replacing (6), (11) and (12) into (8). As expected, D_1=0 in scenario I and D_1=1 in scenario III. In scenario II, D_1 should assume a value between zero and one. Calculations corresponding to this scenario of partial overlapping are the following:

$$d_1 = \frac{1}{5\times5}\left[(4-1)+(4-3)+\ldots+(20-1)+(20-3)+(20-5)+(20-7)+(20-10)\right]=6.68$$

$$d_1{*} = \frac{1}{5\times5}\left[(3-1)+(5-1)+(5-3)+(7-1)+(7-3)+(7-5)+(10-1)+(10-3)+(10-5)+(10-7)\right]=1.76$$

$$\Delta_1 = \frac{1}{5\times5}\left[|4-1|+|4-3|+|4-5|+|4-7|+|4-10|+\ldots+|20-1|+|20-3|+|20-5|+|20-7|+|20-10|\right]=7.16$$

and,

$$D_1 = \frac{6.68-1.76}{7.16-1.76} = 0.9111$$

Recalling the economic meaning of d_1 and the range of variation for D_1 $\left(0 \le D_1 \le 1\right)$; the value of 0,91 obtained for scenario II reflects almost complete separation between those individuals who belong to the more education and the less education clubs, respectively.

This measure does not say anything about the inequality within each group which is captured by the Gini ratio. The combined analysis of D_1 and Gini ratio will help to understand two separated phenomena: the significance of the separation between the subpopulations and the inequality within each subpopulation.

4. Income Distribution by Gender and Education in Córdoba

4.1 Data and Methodology

The case considered here is the Greater Córdoba individuals labor-earnings distribution in the 1992-2000 period, based on the Permanent Household Survey (EPH) provided by INDEC (The National Institute of Statistics and the Census) by gender and education. That survey is run twice a year, in May and October. The results reported in this paper are based on the May survey. However, no significant differences have been found when the October survey was processed instead. In the period considered here, the national currency, the peso, was linked to the dollar by a one to one relationship because of the Convertibility Law of 1991.The stability of the peso and low inflation, running below 1 percent per year on average, would facilitate inter-temporal comparability of results.

The parameters in (3) were estimated using STATA 7 (Gertel, et.al., 2001). The estimation of the parameters for (i) total population by gender, (ii) total population by education level (the distribution of individuals having elementary education and those with secondary and post-secondary schooling, respectively), and (iii) total population by gender and education, are included in Appendix I. The corresponding Gini ratios were estimated from the parameters of the model applying the Mathcad 8 software. The ratio D_1 was calculated by applying equation (8) and the statistical hypothesis test were applied according the description in Appendix II.

The observed economic units are individuals with labor income by gender and education. The educational categories selected here are:

-Elementary schooling, or “ES”, (0 to 11 years)
-Completed secondary and Post-secondary, or “SPS”, (12 years, and more)

A more disaggregated analysis is always possible, but this proved to be enough to illustrate the general behavior of the population under study and the trend in inequality over time.

4.2 The Relative Economic Affluence Ratio D_1 and Gini ratios by Education and Gender

Tables 2 and 3 present the Relative Economic Distance Ratio or between-income distribution measure of inequality D_1 and the within Gini ratios for subpopulations of wage-earners by gender and by education, respectively; and Table 4 presents the Relative Economic Distance Ratios and the Gini ratios, for the subpopulations with less and more schooling, discriminated by gender. All tables cover the same period. Figures 3 and 4 show the evolution of the D_1 ratios and the Gini ratios, respectively.

4.2.1 Total Population of Wage-earners by gender

The values assumed by D_1 in Table 2 indicate that the distributions are located at an intermediate point between a total separation and complete overlapping. The female-male economic distance D_1, despite a jump in 1996, decreased from 0.66 in 1992 to 0.58 in 2000, meaning that the masses of income distributions at the end of the period are closer than they were at the beginning.

Gini Ratio and male-female Economic Distance Ratio D_1
Total population

Year	Gini Ratio		D_1
	Female	Male	
1992	*0.3736*	*0.4515*	*0.6633*
1993	*0.3635*	*0.4315*	*0.6313*
1994	*0.3959*	*0.4385*	*0.6052*
1995	*0.3707*	*0.4002*	*0.5436*
1996	*0.3530*	*0.4067*	*0.6274*
1997	*0.4031*	*0.3811*	*0.4546*
1998	*0.3802*	*0.3717*	*0.4686*
1999	*0.3946*	*0.4015*	*0.4296*
2000	*0.4309*	*0.4464*	*0.5804*

χ^2 for D_1s are significant at the 0.01 level.
(See appendix II)
Source: EPH, INDEC

Table 2

The Gini ratio for male had a lightly overall decreasing trend that can be decomposed in two parts: a heavy decrease from 0.45 to 0.37 between 1992-1998 and an increasing trend by the end of the period; whereas the female ratio shows a heavy deterioration, which might be explained by increasing job opportunities for women relative to men.

4.2.2 Total Population of Wage-earners: the groups Secondary and More Schooling vs Elementary Schooling

Let consider now D_1 for the groups Secondary and More Schooling vs Elementary Schooling. The values assumed by D_1 in Table 3 indicate that the distributions are close to complete separation along the period, except for 1998. The economic distance D_1 between these groups remained relatively stable around a 0.81 level.

Gini Ratio and SPS-ES Economic Distance Ratio D_1
Total population

Year	Gini Ratio		D_1
	ES	SPS	
1992	*0.3231*	*0.4628*	*0.8455*
1993	*0.3351*	*0.4475*	*0.8248*
1994	*0.3306*	*0.4516*	*0.7885*
1995	*0.3241*	*0.4563*	*0.8115*
1996	*0.3168*	*0.4623*	*0.8233*
1997	*0.3222*	*0.4225*	*0.7830*
1998	*0.3447*	*0.4139*	*0.6902*
1999	*0.3331*	*0.4412*	*0.8214*
2000	*0.3533*	*0.4742*	*0.8517*

ES: Elementary (0-11 schooling years);
SPS: Secondary and Post-secondary
χ^2 for D_1s are significant at the 0.01 level.
(See appendix II)
Source: EPH, INDEC

Table 3

The Gini ratio for the Secondary and Post-secondary group had a steady trend in the range of 0.41 (1998) and 0.47 (2000) and the Gini ratio for the Elementary Schooling group shows a similar behavior, but at the 0.33 level. This means that the degree of income inequality is higher in the SPS group.

4.2.3. Male and Female Wage-earners with Elementary Schooling

The female-male economic distance for the Elementary Education group is relatively high and stable during the period 1992-1995, at about 0.75 level, followed by heavy fluctuations until 2000, as it is shown in Table 4. The large discrepancy between the male and the female within the low educational credentials group may be revealing that gender discrimination in the unskilled and semi-unskilled labor market had remained high, in Córdoba during the period under study. The heavy fluctuations observed in the last part of the period may have resulted from post-Tequila restructuring effects on larger industrial employers of the metal-mechanics sector, where male employment predominates, and relative stability of employment in the two occupational categories leaded by woman: clerical, and service.

Female-Male Economic Distance Ratio (D_1) and Female and Male Gini Ratio by Education

Year	ES			SPS		
	Gini Ratio		D_1	Gini Ratio		D_1
	Female	Male		Female	Male	
1992	*0.2870*	*0.3247*	*0.7754*	*0.4120*	*0.4388*	*0.5867*
1993	*0.2686*	*0.3334*	*0.7582*	*0.3788*	*0.4559*	*0.6504*
1994	*0.2798*	*0.3384*	*0.7542*	*0.4355*	*0.4463*	*0.6113*
1995	*0.2865*	*0.3143*	*0.7452*	*0.4007*	*0.4736*	*0.5870*
1996	*0.2645*	*0.3035*	*0.8229*	*0.3687*	*0.4820*	*0.6962*
1997	*0.3424*	*0.2905*	*0.6873*	*0.4018*	*0.4112*	*0.4667*
1998	*0.2686*	*0.3314*	*0.8496*	*0.3926*	*0.4197*	*0.4406*
1999	*0.3128*	*0.3203*	*0.6657*	*0.3823*	*0.4970*	*0.6447*
2000	*0.3431*	*0.3213*	*0.7675*	*0.4019*	*0.5007*	*0.7019*

χ^2 for D_1s are significant at the 0.01 level. (See appendix II)
Source: EPH, INDEC

Table 4

The male and female Gini ratios show some contrasting characteristics. In 1992, the Gini ratio for the male-elementary schooling group is 13.2 % higher than the Gini ratio for the female-elementary schooling group; the difference presents a downward trend, and its sign reversed by the end of the period, with Gini ratio for the female-elementary schooling group being above the Gini ratio for the male-elementary schooling group by 6.8%, mainly because of the upward Gini ratio for the female-elementary schooling group trend.

The Gini ratio for the male wage-earners with elementary schooling remained stable along the period at a low value of 0.32. The Gini ratio for the female wage-earners with elementary schooling starts below that of men in 1992, at 0.28, and increased steadily until it reached 0.34 by the end of the period. One important reason that could account for this increasing inequality is that rapid technological progress in the nineties permeated the service and trade sectors, reducing work opportunities for unskilled and semi-unskilled workers, that in these sectors are predominantly female.

4.2.4. Male and Female Wage-earners with Secondary and Post-secondary Schooling

The male-female income differential ratio D_1 for the wage-earners with Secondary and Post-secondary schooling (Table 4) is located at mid-range values, below the values for the Elementary Schooling group D_1. It started with 0.58, followed an up-rising trend in 1992-1996 and thereafter fluctuated heavily until 2000. By the end of the period under study, the male-female income differential had increased by 20 %.

The male and female Gini ratios present wide fluctuations with a strong upward trend for male and a stable trend for female. Furthermore, the male Gini for the subpopulation with more schooling investments embodied is higher than the corresponding female Gini. For this reason, the small initial difference of 0.027 (a 6.5% higher Gini for males in 1992), has increased to 0.10 in 2000.

Finally, a visual description of the trend in D_1 and the Gini ratios respectively, is provided through Figures 3 and 4. Figure 3 indicates the evolution of the Economic Distance Ratio D_1 for the subpopulations of (i) female-male total populations, (ii) Secondary and Postsecondary-Elementary Schooling total populations, (iii) female-male Elementary Schooling groups, (iv) female-male Secondary and Postsecondary Schooling groups.

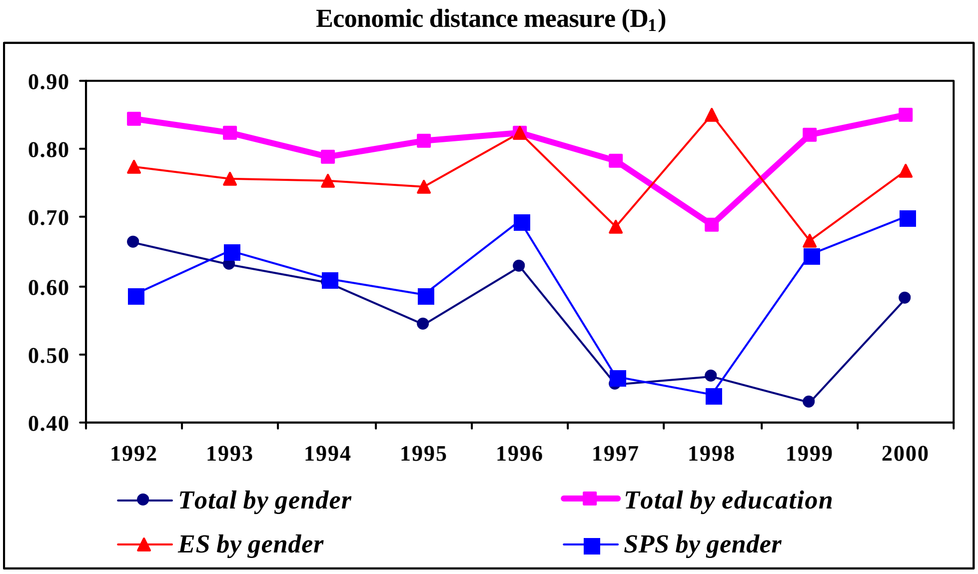

Figure 3

The most striking finding is that of separation between distributions by gender and by education: while male-female D_1 adopts values in the mid-range, the more-less education is clearly placed above it, an closer to total separation of subpopulations. It is also observed that gender discrimination is lower when the more education group is considered. These findings are suggesting that, on contrary to what is generally expected, education rather than gender discriminates more in society.

Figure 4 indicates the evolution of the Gini Ratio for the female and male groups among the Elementary Schooling subpopulation, and the female and male groups among the Secondary and Postsecondary subpopulation.

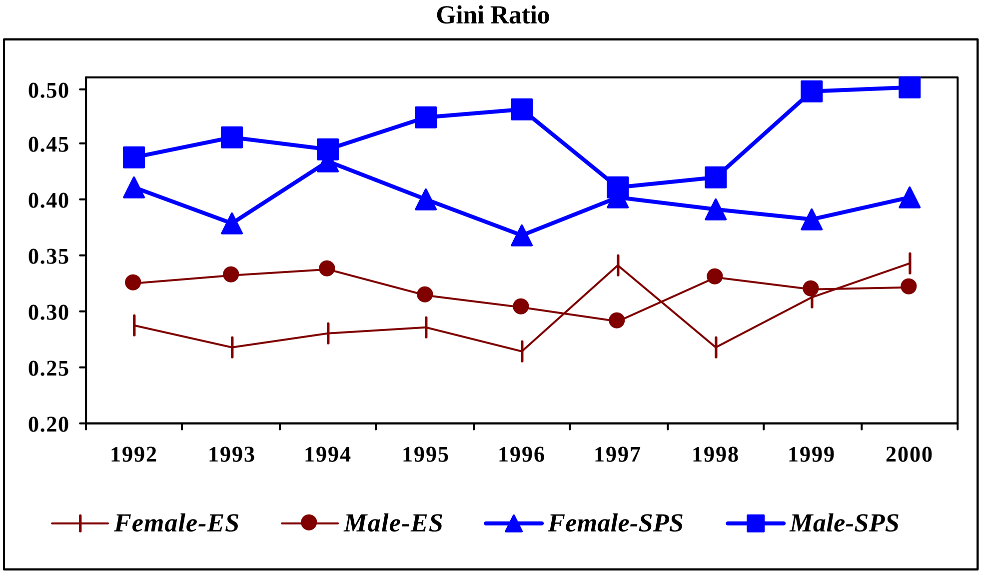

Figure 4

The Elementary Schooling subpopulation exhibits a lower degree of inequality than the Secondary and Postsecondary subpopulation. In both of them, the male group has a higher level of inequality confirming that more education increases job opportunities and this produces more inequality.

The groups are more homogeneous and distributions are more stable along the period, according to the values of the Gini ratios. The Table also shows that the Secondary and Post-secondary education group has a much more higher degree of inequality than it was observed in Table 2 for the female-male comparison.

5. Summary and Conclusions

Theory provides un-weighted and weighted measures of distances between income distributions of two populations. Dagum developed a measure of the relative economic distance that is computed through a coefficient ranging from 0 to 1, and is based on the expected value of the differences between incomes of pair-wise individuals. This measure was applied to the analysis of education, gender and income distribution of wage-earners in the metropolitan area of Córdoba.

The main results are summarized bellow:

The female-male Relative Economic Distance Ratio, considering the total population, is located at an intermediate point between a total separation and complete overlapping. The female-male economic distance D_1, despite a jump in 1996, decreased from 0.66 in 1992 to 0.58 in 2000, meaning that the masses of income distributions at the end of the period are closer than they were at the beginning.

Separation between the Secondary-Post Secondary and the Elementary Education groups is almost total. The economic distance D_1 between these groups remained relatively stable around a 0.80 level (except for 1998). These values indicate that the distributions are close to complete separation along the period.

The female-male Relative Economic Distance Ratio among the Elementary Schooling group ranges from 0.7 to 0.8 and is higher than the same ratio calculated for the Secondary and Postsecondary population group.

Within each education group, however, both female and male Gini ratios have similar values, being significantly lower in the Elementary Education group (about 0.3 for elementary schooling and 0.4 for the Secondary and Postsecondary group).

The previous results could be interpreted as meaning that having invested little in education does not contribute to reduce gender inequality, while, on the other hand, increasing the level of investment in education seems to contribute to its reduction.

Finally, education is traditionally presented as a powerful tool to abate inequality in general, and gender inequality more specifically. This paper re-examined these issues of education, gender and economic inequality. To that end, it made use of the Relative Economic Distance ratio introduced by Dagum, in which he incorporates an economic meaning to earlier measures of distance proposed in the literature. The paper also presented estimates of the Relative-Economic-Distance ratios between subpopulations of wage-earners in Córdoba as well as the within distributions Gini Ratios for the period 1992-2000. It was found, as expected, that a female-male economic distance exists; however, the estimated value for the economic distance ratio between the Secondary and Further Education group and the Elementary Education group is shown to be much higher.

Notes:

[1] A comprehensive review of early and recent approaches to the decomposition of the Gini Income Inequality is that of Dagum (1997). An influential contribution in this field is that of Bourguignon (1979). As noted in Dagum (1997), Bourguignon demonstrated that a class of additive decomposability of the Gini Ratio can be obtained by applying income-weights or population-weights as developed in Theil, or Bhattacharya and Mahalanobis works, respectively, "although weighting coefficients do not necessarily add to one" (p.902). The un-weighted class of decomposition measures is discussed in Shorrocks (1984) and Yitzhaki (1994). A summary discussion of metric distance measures is found in Sharma (1996), chapter 3.

[2] It can be proved (Dagum, 1997: 524-526) that D_1 is a key element in a new class of decomposition of the Gini ratio where the Gini results from: (i) the Gini inequality within subpopulations, plus (ii) the gross Gini inequality between populations; in turn, (ii) has two parts: a part that measures the income difference weighted overlapping of distributions (D1) and a part reflecting the net contribution of the Gini inequality between subpopulations to the total Gini. However this type of application of D1 remains outside the scope of this presentation.

References:

Arrow, Kenneth (1973) “Higher Education as a Filter” Journal of Public Economics, 2:193-216.

Bourguignon, F. (1979) “Decomposable Income Distribution Measures” Econometrica 47:901-920.

Carnoy, Martin, with Lobo, J., Toledo, A. and Velloso, J. (1974), Can Educational Policy Equalize Income Distribution in Latin America? London, Saxon House.

Chakravarty, S.R. and Dutta, B. (1987) “A note on measures of Distance between Income Distributions” Journal of Economic Theory, 41 (1), 185-188.

Dagum, Camilo, (1985) “Analyses of Income Distribution and Inequality by Education and Sex” Advances in Econometrics, Vol.4:167-227.

Dagum, Camilo (1997) “A new approach to the decomposition of the Gini income inequality ratio” Empirical Economics, 22,4: 515-531.

Dagum, Camilo (2001) “Desigualdad del rédito y bienestar social, descomposición, distancia direccional y distancia métrica entre distribuciones” Estudios de Economía Aplicada, No 17, 2001: 5-52.

Delfino, José A. (1998) Educación, capital humano y crecimiento económico en Argentina, en: Rol de la Educación y la Capacitación en el Crecimiento Económico de la Argentina, Córdoba, Eudecor.

Ebert, U. (1984) “Measures of Distance between Income Distributions” Journal of Economic Theory, 32, 266-274.

Gertel, Héctor R., De Santis, M. and Pereyra, L. (1987) “Educación y distribución de ingresos en Córdoba”, ANALES, XXI Reunión anual de la Asociación Argentina de Economía Política.

Gertel, H., Giuliodori, R., Auerbach, P., Rodríguez, A. (2001) “Unemployment And Income Distribution Analysis, New evidences using a Dagum parametric income distribution model”, ANALES, XXXVI Reunión anual de la Asociación Argentina de Economía Política.

Henry M. Levin and Carolyn Kelley (1994) Can Education Do It Alone? Economics of Education Review, vol. XIII (2) pp.97-108.

Mincer, Jacob (1974) Progress in Human Capital Analysis of the Distribution of Income, National Bureau of Economic Research, Working Paper N 53.

Patrinos, H. A. (1995) “Socioeconomic Background, Schooling, Experience, Ability and Monetary Rewards in Greece” Economics of Education Review, 14: 85-91.

Sharma, S. (1996) Applied Multivariate Techniques, New York, John Willey & Sons. Inc. pp.39-50.

Shorrocks, Anthony F. (1984), “Inequality decomposition by population subgroups”, Econometrica, Vol.53, 5, 1369-1386.

Ram, Rati (1984) ”Population Increase, Economic Growth, Educational Inequality, and Income Distribution: Some Referent Evidence” Journal of Development Economics, 14 (April), 419-428.

Ram, Rati (1990) ”Educational Expansion and Schooling Inequality: International Evidence and some Implications” The Review of Economics and Statistics, Vol.72, 2, (May), 266-274.

Rao, C.Radhakrishna (1952) Advanced Statistical Methods in Biometric Research, New York, John Willey & Sons.

Theil, Henry (1967) Economics and Information Theory, Amsterdam, North-Holland.

Yitzhaky, S. (1994) “Economic Distance and Overlapping of Distributions”, Journal of Econometrics, 61, 147-159.

Appendix I

Table A 1

Income Distribution Statistics, Female and Male

Year and Gender	Parameters of cdf* **b**	**d**	**l**	Sample Size	F	Mean Estimated	Mean Observed	Median Estimated	Median Observed
1992									
Female	*1.3294*	*2.5443*	*0.0282*	*575*	*90910*	*370.4*	*368.0*	*285.5*	*300.0*
Male	*3.9722*	*1.9418*	*0.0314*	*1008*	*276769*	*605.4*	*547.1*	*395.0*	*400.0*
1993									
Female	*1.4789*	*2.5707*	*0.0374*	*574*	*119752*	*438.0*	*430.8*	*340.2*	*330.0*
Male	*1.9557*	*2.1186*	*0.0863*	*932*	*250241*	*681.0*	*651.6*	*470.6*	*500.0*
1994									
Female	*1.9747*	*2.2919*	*0.0437*	*575*	*152777*	*509.3*	*488.4*	*372.4*	*370.0*
Male	*3.4894*	*2.0047*	*0.0613*	*909*	*214557*	*789.9*	*768.4*	*528.9*	*530.0*
1995									
Female	*1.7449*	*2.4696*	*0.0537*	*459*	*119700*	*536.7*	*525.8*	*409.1*	*400.0*
Male	*2.8571*	*2.2010*	*0.0729*	*768*	*191183*	*764.2*	*732.6*	*547.4*	*500.0*
1996									
Female	*1.1215*	*2.7685*	*0.0515*	*411*	*78906*	*452.5*	*486.9*	*362.6*	*350.0*
Male	*2.3010*	*2.2049*	*0.0740*	*696*	*171324*	*690.2*	*683.6*	*493.3*	*500.0*
1997									
Female	*1.0796*	*2.4459*	*0.0734*	*413*	*132008*	*478.9*	*467.5*	*359.0*	*350.0*
Male	*2.1069*	*2.3592*	*0.0699*	*711*	*143167*	*647.9*	*641.5*	*482.2*	*500.0*
1998									
Female	*0.9020*	*2.6884*	*0.0678*	*364*	*78515*	*443.6*	*452.8*	*348.5*	*350.0*
Male	*1.0072*	*2.6866*	*0.1304*	*601*	*176377*	*597.2*	*608.7*	*470.0*	*500.0*
1999									
Female	*0.7817*	*2.6763*	*0.1035*	*416*	*93449*	*481.9*	*489.7*	*375.2*	*400.0*
Male	*1.4445*	*2.3459*	*0.1035*	*620*	*156374*	*634.6*	*659.7*	*467.3*	*450.0*
2000									
Female	*0.6425*	*2.5663*	*0.1120*	*446*	*137609*	*438.6*	*433.0*	*329.1*	*320.0*
Male	*1.7364*	*2.0770*	*0.0984*	*650*	*168180*	*680.1*	*630.0*	*461.1*	*450.0*

*/ The β and δ parameters are the equality parameters in equation (3); any increase in β values should be read as implying that the number of individuals al the lower deciles in the income distribution is now smaller. Any increase in δ values implies, instead, that the mass of income accruing to the mid-and middle-high income recipient units is now bigger.

Table A 2
Income Distribution Statistics
Elementary Schooling and Secondary-Post secondary Subpopulations

Year and	Parameters of cdf			Sample		Mean		Median	
schooling	**b**	**d**	**l**	Size	F	Estimated	Observed	Estimated	Observed
1992									
ES	*1.0018*	*3.0935*	*0.0214*	*889*	*144596*	*345.2*	*347.2*	*289.1*	*300.0*
SPS	*1.9397*	*1.9878*	*0.1012*	*694*	*218011*	*737.2*	*654.9*	*482.9*	*500.0*
1993									
ES	*0.9049*	*3.0534*	*0.0432*	*852*	*184756*	*410.5*	*412.1*	*341.2*	*350.0*
SPS	*2.1094*	*2.0349*	*0.1198*	*654*	*179261*	*835.6*	*769.9*	*560.3*	*600.0*
1994									
ES	*1.4425*	*2.8256*	*0.0421*	*825*	*144378*	*476.7*	*480.2*	*387.0*	*400.0*
SPS	*1.3471*	*2.1157*	*0.2458*	*658*	*249406*	*910.7*	*885.5*	*620.4*	*600.0*
1995									
ES	*0.8955*	*3.1660*	*0.0690*	*617*	*147201*	*485.6*	*481.4*	*409.1*	*400.0*
SPS	*8.6451*	*1.8798*	*0.0327*	*609*	*161209*	*952.5*	*831.6*	*607.2*	*600.0*
1996									
ES	*0.7093*	*3.4577*	*0.0509*	*535*	*91474*	*422.2*	*421.3*	*365.5*	*350.0*
SPS	*4.4568*	*1.8931*	*0.0522*	*571*	*162610*	*844.1*	*787.9*	*538.8*	*500.0*
1997									
ES	*0.6074*	*3.5744*	*0.0623*	*587*	*123445*	*427.2*	*435.4*	*372.3*	*390.0*
SPS	*1.9332*	*2.1623*	*0.1171*	*527*	*161611*	*779.3*	*740.9*	*547.2*	*550.0*
1998									
ES	*0.7505*	*3.1137*	*0.0633*	*510*	*111481*	*432.1*	*423.4*	*360.8*	*351.0*
SPS	*1.2792*	*2.3180*	*0.1535*	*448*	*183844*	*705.2*	*697.5*	*513.4*	*500.0*
1999									
ES	*0.5263*	*3.6326*	*0.0603*	*515*	*96682*	*401.2*	*409.3*	*350.2*	*350.0*
SPS	*2.2061*	*2.0541*	*0.1048*	*516*	*175822*	*800.3*	*775.8*	*541.7*	*540.0*
2000									
ES	*0.4731*	*3.5493*	*0.0554*	*520*	*115824*	*366.0*	*372.1*	*315.4*	*300.0*
SPS	*1.5922*	*1.9808*	*0.1566*	*568*	*170205*	*821.2*	*715.8*	*532.5*	*500.0*

*/ The β and δ parameters are the equality parameters in equation (3); any increase in β values should be read as implying that the number of individuals al the lower deciles in the income distribution is now smaller. Any increase in δ values implies, instead, that the mass of income accruing to the mid-and middle-high income recipient units is now bigger.

Table A 3

Income Distribution Statistics, Female and Male Elementary Schooling Population

Year and	Parameters of cdf			Sample		Mean		Median	
schooling	**b**	**d**	**l**	Size	F	Estimated	Observed	Estimated	Observed
1992									
Female	*0.4224*	*4.7210*	*0.0038*	290	31871	244.1	252.3	*227.1*	*210.0*
Male	*1.7368*	*2.8001*	*0.0210*	599	92650	400.8	393.1	*324.7*	*300.0*
1993									
Female	*0.5086*	*4.6643*	*0.0069*	299	50986	294.4	299.0	*273.7*	*280.0*
Male	*0.8856*	*3.0857*	*0.0665*	553	112646	471.0	473.2	*392.9*	*400.0*
1994									
Female	*0.7201*	*3.9182*	*0.0163*	283	46149	345.1	351.8	*309.4*	*300.0*
Male	*2.3883*	*2.6018*	*0.0387*	542	82344	553.9	547.3	*435.3*	*400.0*
1995									
Female	*0.7060*	*3.8469*	*0.0187*	198	39033	348.9	348.2	*311.2*	*300.0*
Male	*1.0863*	*3.1247*	*0.0818*	419	80036	553.6	544.3	*465.2*	*480.0*
1996									
Female	*0.3116*	*6.1208*	*0.0031*	165	12707	284.1	287.6	*275.5*	*300.0*
Male	*0.8260*	*3.4582*	*0.0645*	370	66291	482.5	480.9	*418.0*	*400.0*
1997									
Female	*0.5050*	*3.5810*	*0.0268*	188	48487	310.5	313.9	*269.6*	*280.0*
Male	*0.9938*	*3.4470*	*0.0502*	399	53170	482.9	492.6	*418.4*	*400.0*
1998									
Female	*0.2367*	*7.1847*	*0.0014*	155	18779	271.9	279.8	*268.4*	*280.0*
Male	*0.7165*	*3.2887*	*0.0959*	355	73787	496.7	486.1	*422.1*	*400.0*
1999									
Female	*0.3039*	*5.0914*	*0.0140*	180	27337	301.1	325.0	*282.2*	*300.0*
Male	*0.6108*	*3.5892*	*0.0743*	335	65393	451.0	454.5	*393.6*	*400.0*
2000									
Female	*0.3635*	*4.1375*	*0.0112*	181	26887	247.5	257.1	*221.5*	*200.0*
Male	*0.5652*	*3.6788*	*0.0658*	339	65012	428.1	433.6	*375.8*	*400.0*

*/ The β and δ parameters are the equality parameters in equation (3); any increase in β values should be read as implying that the number of individuals al the lower deciles in the income distribution is now smaller. Any increase in δ values implies, instead, that the mass of income accruing to the mid-and middle-high income recipient units is now bigger.

Table A 4

Income Distribution Statistics, Female and Male
Secondary plus Post-Secondary Population

Year and	Parameters of cdf			Sample		Mean		Median	
schooling	**b**	**d**	**l**	Size	F	Estimated	Observed	Estimated	Observed
1992									
Female	*4.2720*	*2.0977*	*0.0226*	*285*	*53479*	*540.0*	*485.7*	*375.8*	*400.0*
Male	*1.2541*	*2.1971*	*0.2258*	*409*	*152202*	*832.6*	*772.7*	*583.3*	*600.0*
1993									
Female	*1.9734*	*2.3877*	*0.0651*	*275*	*57052*	*610.4*	*574.1*	*457.4*	*450.0*
Male	*1.6390*	*2.0489*	*0.2266*	*379*	*84276*	*990.0*	*911.9*	*662.7*	*600.0*
1994									
Female	*1.5339*	*2.1550*	*0.1169*	*291*	*92240*	*688.4*	*621.1*	*478.4*	*500.0*
Male	*1.9603*	*2.0531*	*0.2344*	*367*	*93905*	*1110.4*	*1094.9*	*748.4*	*750.0*
1995									
Female	*8.6383*	*2.1020*	*0.0204*	*261*	*53651*	*729.9*	*660.5*	*511.3*	*500.0*
Male	*8.4856*	*1.8220*	*0.0437*	*348*	*77066*	*1119.1*	*960.0*	*693.4*	*700.0*
1996									
Female	*2.2313*	*2.4192*	*0.0523*	*246*	*47091*	*592.4*	*620.6*	*448.1*	*420.0*
Male	*3.2035*	*1.8522*	*0.1035*	*325*	*72361*	*1018.5*	*914.5*	*632.7*	*600.0*
1997									
Female	*2.4130*	*2.2206*	*0.0585*	*223*	*56803*	*636.0*	*599.4*	*456.8*	*450.0*
Male	*1.7434*	*2.2414*	*0.1686*	*304*	*83253*	*866.1*	*844.6*	*622.3*	*600.0*
1998									
Female	*1.0544*	*2.5215*	*0.1283*	*206*	*54067*	*598.2*	*585.2*	*455.8*	*450.0*
Male	*1.4228*	*2.2531*	*0.1753*	*242*	*82280*	*791.9*	*793.1*	*567.7*	*600.0*
1999									
Female	*1.0746*	*2.5796*	*0.1366*	*234*	*62792*	*621.8*	*617.6*	*479.9*	*490.0*
Male	*6.0752*	*1.7617*	*0.0494*	*282*	*85762*	*1010.5*	*907.0*	*601.6*	*600.0*
2000									
Female	*0.8145*	*2.6001*	*0.1633*	*263*	*61765*	*579.5*	*555.9*	*444.8*	*450.0*
Male	*1.9563*	*1.8480*	*0.1832*	*305*	*69609*	*1038.4*	*853.6*	*633.9*	*600.0*

*/ The β and δ parameters are the equality parameters in equation (3); any increase in β values should be read as implying that the number of individuals al the lower deciles in the income distribution is now smaller. Any increase in δ values implies, instead, that the mass of income accruing to the mid-and middle-high income recipient units is now bigger.

Appendix II

Statistical hypothesis testing of D_1

To decide whether the estimated income differential ratio D_1 is significantly different from zero, the null hypothesis

$$H_0: \quad F_1(y) \equiv F_2(y) \quad \text{all y,} \qquad \text{(II.1)}$$

Has to be tested against the alternative hypothesis

$$H_1: \quad F_1(y) \neq F_2(y) \qquad \text{(II.2)}$$

Includes all ways, such as change in means, variance and asymmetries, in that distributions can differ.

However, if it considers only shift but without considering changes in variances and symmetries, the alternative hypothesis become

$$H_1: \quad F_1(y) \equiv F_2(y - \boldsymbol{q}), \boldsymbol{q} > 0 \qquad \text{(II.3)}$$

To infer whether D_1 is or not significantly differ from zero, since the second distribution (Q_2) is more affluent than the first distribution (Q_1), it applies the one-tail Kolmogorov-Smirnov test for two subpopulations. Its corresponding statistic is

$$D^+ = \sup\left[F_{1m}(y) - F_{2n}(y)\right] \qquad \text{(II.4)}$$

where "*m*" is the sample size from Q_1 and "*n*" from Q_2. If the ratio D_1 has been obtained from observed data; then, the statistics D^+ is calculated from observed income distribution. On the other hand, if D_1 has been estimated from parametric income distribution, then D^+ is estimated from it.

The statistic $4mn\left(D^+\right)^2 / (m+n)$ has for large-samples, a chi-square distribution with two degrees of freedom. That is:

$$\frac{4mn\left(D^+\right)^2}{m+n} \rightarrow \boldsymbol{c}^2(2) \qquad \text{(II.5)}$$

which is the test applied in this work.

Given an "ε" level of significance, if the statistic $4mn\left(D^+\right)^2 / (m+n)$ is less than the critical values $\boldsymbol{c}_e^2(2)$, the null hypothesis is accepted; hence its interpretation is that the income distance between the two distributions under study is not significantly different from zero.