ASOCIACION ARGENTINA
DE ECONOMIA POLITICA

ANALES | ASOCIACION ARGENTINA DE ECONOMIA POLITICA

XLIII Reunión Anual

Noviembre de 2008

ISSN 1852-0022
ISBN 978-987-99570-6-6

Labor Income Polarization in Greater Buenos Aires

Mariana Viollaz
Sergio Olivieri
Javier Alejo

Labor Income Polarization in Greater Buenos Aires

Mariana Viollaz[*]
Sergio Olivieri
Javier Alejo

CEDLAS [**]
Universidad Nacional de La Plata

Abstract

This paper presents and analyzes a bunch of statistics which characterize the level and evolution of the labor income polarization in Greater Buenos Aires over the past two decades (1986-2006). The empirical evidence reveals two stages throughout those years: the first one distinguished by an increment of all indices and the second one, by shrinkage of them. Inspecting potential factors which could explain those changes, returns to education surge as the main polarization force in the labor market. Hence, an equalizing distribution of the human capital is a possible alternative for a less polarized labor market.

Resumen

Este estudio presenta y discute un conjunto de estadísticas que caracterizan el nivel y la evolución de la polarización laboral en Argentina a lo largo de las últimas dos décadas (1986-2006). Los resultados obtenidos sugieren que el nivel de polarización atravesó una primera etapa de crecimiento para posteriormente iniciar un proceso de reducción. Los retornos a la educación se muestran como la fuerza fundamental para explicar la evolución de la polarización en el mercado laboral. La obtención de una distribución más igualitaria del capital humano parece ser el camino a seguir para alcanzar un mercado laboral menos polarizado.

Keywords: polarization, labor, cohesion, inequality, Argentina, Greater Buenos Aires

JEL codes: I3, D3, D6

---

[*]E-mails: mariana.viollaz@depeco.econo.unlp.edu.ar, solivieri@depeco.econo.unlp.edu.ar and jalejo@depeco.econo.unlp.edu.ar

[**]CEDLAS is Centre for Distributive, Labor and Social Studies at National University of La Plata (Argentina). www.depeco.econo.unlp.edu.ar/cedlas

I. Introduction

Inequality is definitely the income distributional dimension more studied. Economists have contributed to the discussion of social fairness, and have developed a large literature on the measurement of inequality.[1] However, inequality measures consistent with the Dalton-Pigou transfers principle, could not embrace a complete characterization of the income distribution.

Polarization is another dimension of the income distribution which has been studied during the last decade at a fast path. This concept refers to the antagonisms between groups which are internally homogeneous and increasingly different among them. Both polarization and inequality are different although related dimensions of the same distribution. Thus, the analysis of polarization should be considered as complementary to that of inequality.

The motivation for analyzing this phenomenon is the link between polarization and social tensions and instability. Social cohesion is likely to be weak when the dispersion in the socioeconomic characteristics of a population is high. If people have access to substantially different sets of opportunities, and enjoy (or suffer) very different living standards, social tensions are likely to emerge. An economically polarized country is more likely to be socially and politically unstable.[2]

This study documents the characterization of the hourly wages in Greater Buenos Aires from 1986 until 2006, particularly from the economic polarization perspective. The labor market is the main scenario in which inequalities and income poles emerge[3]. The reason could be justified by the lack of capacity of other sources to explain the evolution of income inequality and income polarization. Firstly, household surveys have many deficiencies in capturing capital income, entrepreneur benefits and rents.[4] The Permanent Household Survey (Encuesta Permanente de Hogares or EPH) of Argentina is not an exception. Secondly, in Latin American countries there is evidence of the weakness of their redistributive schemes.[5] Thus, the governmental transfers could not be a feasible source to explain inequality and polarization changes. Lastly, the demographic structure could not seem a relevant explicative factor either.[6] Hence, differences in labor income would be the main sources to explain income inequality and income polarization changes documented in several studies that analyze the evolution of welfare measures such as household per capita income and adult equivalized income.[7]

There are many other factors which could be independent of the labor market and could explain the inequality and polarization movements. In order to take them into account, we present and analyze several labor indicators which depend on demographic and

[1] See Atkinson and Bourguignon (eds.) (2000), Deaton (1997), Cowell (2000) and Lambert (2001).
[2] Of course, the causality can go both directions: socioeconomic fragmentation can be the consequence of social and political instability (Gasparini, *et.al*. 2008).
[3] Gasparini et.al (2008)
[4] Deaton (1997)
[5] Gasparini, et al (2005a)
[6] Haimovich, et. al.(2005)
[7] Gasparini, et. al (2008); Horenstein and Olivieri (2004).

social characteristics of the population and we also perform micro-econometric decompositions.

The rest of the document is organised as follows. In section II we briefly discuss the concept of economic polarization from different perspectives, characteristics and pure income polarization. In section III we present some methodological features as well as a descriptive analysis of the labor market in Greater Buenos Aires, access and employment conditions and polarization by main characteristics of the population. Section IV is focus on the evolution of the pure hourly labor income polarization and the decomposition of those changes. Section V closes with concluding remarks.

II. The measurement of polarization

We rely on the alienation-identification framework proposed by Esteban and Ray (1994): a population is polarized if (i) there are few groups of important size, (ii) in which their members share an attribute and feel some degree of *identification* with members of their own group, and (iii) members of different groups feel *alienated* from each other.

Income polarization measures could be classified into two main sets. Although both sets use income as the variable for alienation, they differ in the nature of identification. While the first uses a discrete variable to provide the relevant grouping of the population, the latter uses income. The first set is known as "polarization by characteristics", whereas the second is called "pure income polarization". For instance, income polarization by the area where the household lives (urban-rural) is part of the first set, while income polarization where individuals identify themselves with those with similar income levels is known as "pure income polarization".

IDENTIFICATION	ALIENATION	TYPE	INDEX
Discrete variable: area, race, educational level, etc	Continous variable: income	Polarization by Characteristics	Gradín Group Polarization (1999) Zhang - Kanbur (2001)
Continous variable: income	Continous variable: income	Pure Income Polarization	Duclos-Esteban-Ray (2004) -EGR - Wolfson

In what follows we provide a brief overview of the polarization measures to be used throughout this paper.[8]

Polarization by characteristics

Although alienation is considered to be into the income space, there might be other population characteristics that create group identity (*e.g*. religion in Northern Ireland, race in USA). As Gradín (2000) states it, "*despite polarization occurring in the income space, groups in the distribution are the result of similarities with respect to a relevant*

[8] See Methodological appendix for further details of polarization indexes.

attribute other than income". Therefore, it is interesting to explore different attributes that could potentially reflect a well-defined social group.

The literature on polarization by characteristics has been recently increasing at a fast pace. Collier and Hoeffler (2001) measure polarization in an empirical analysis of civil war, Reynal-Querol (2001) studies polarization by religion groups and its relationship with the probability of a conflict in sub-Saharan countries, D'Ambrosio (2001) argues that the region of residence accounts for polarization in the Italian distribution of personal income, Gradín (2000) finds that education and socioeconomic conditions are the key variables to explain polarization in the Spanish distribution of income, and Zhang and Kanbur (2001) apply some polarization measures to regional disparities in China.

In this paper we use Gradín (2000) "group polarization", and Zhang and Kanbur (2001) indices. Gradín (2000) makes an extension of the Esteban and Ray (1994) approach to polarization in order to analyze the role of different household characteristics in the formation of groups, and unlike other measures, accounts for both intra-group inequality as well as the overlapping between groups. Zhang and Kanbur (2001) propose an index of polarization which is based on the ratio of the between-group inequality to the within-group inequality – both measured with Theil's Generalized Entropy index, where groups are defined accordingly with an attribute. See the Appendix for more on both indicators of group polarization.

Pure income polarization

To carry out pure income polarization measures we assume that income is a proxy of other relevant characteristics that generate identification among individuals. The first approach to implement a pure income polarization measure is based on the idea of discrete groups, or socioeconomic classes. Following this logic, it is necessary to identify the number and the support interval of each disjoint group. Wolfson (1994), Esteban and Ray (1994) and Esteban, Gradín and Ray (1999) are the main contributions in this approach. Wolfson's (1994) measure assumes two groups of equal size, while the ER (1994) measure allows *n* groups of potentially different sizes. EGR (1999) leaves the determination of the number of groups to the researcher, while implements a methodology to endogenously determine group sizes based on the idea of minimizing income heterogeneity within groups. See the Appendix for further information.

Esteban *et al*. (1999) implement two enhancements on the original ER index (Esteban and Ray, 1994). The first includes a correction to account for intragroup dispersion, and the second, a methodology for selecting group sizes. This approach consists of choosing the *n*-spike distribution that minimizes the income dispersion within all socioeconomic classes (see Appendix).

Although the framework discussed so far follows an intuitive and common way to refer to different socioeconomic strata, the division of the income distributions in a finite number of groups is unnatural, due to the fact that income is a continuous variable. This fact implies some drawbacks: (i) there is a degree of arbitrariness in the choice of the number of income groups, and (ii) continuous changes in polarization are not captured in some cases, given that the population is divided into a finite number of groups.

The Duclos-Esteban-Ray index (DER)[9] sets out to solve these problems. In order to do so, they redefine the axioms that must be satisfied by a polarization index for continuous variables and present a measure of pure income polarization. This new index allows for individuals not to be clustered around discrete income intervals, and lets the area of identification influence be determined by nonparametric kernel techniques, avoiding arbitrary choices. The authors establish that a general polarization measure that respects a basic set of axioms must be proportional to

$$P_\alpha(F) = \int f(y)^\alpha \, g(y) dF(y)$$

where *y* denotes income and *F*(*y*) its distribution. The function *g*(*y*) captures the alienation effect while $f(y)^\alpha$ captures the identification effect. The higher the α parameter, the larger the weight attached to identification in the polarization index.[10] It can be shown that in order to respect the axioms, the parameter α must lie within the interval [0.25, 1]. See the Appendix for details.

It is possible to account changes in polarization through the contribution of alienation, identification and their joint co-movements. Increased alienation is associated with an increase in income distances, while increased identification implies a sharper definition of groups. When taken jointly, these effects may reinforce each other, in the sense that alienation may be highest at the incomes that have experienced an increase in identification, or they may counterbalance each other.

III. Greater Buenos Aires Labor Market

III.1 Methodological features

This document is based on microdata from the Permanent Household Survey (Encuesta Permanente de Hogares or EPH) carried out by the Instituto Nacional de Estadísticas y Censos (INDEC) since 1974. The database used here is part of the Socio-Economic Database for Latin America and the Caribbean (SEDLAC). This is a large database of household surveys from 21 countries assembled by CEDLAS and the World Bank.

We employ data for individuals living in the Greater Buenos Aires area for the period 1986-2006. We use the October wave from 1986 to 2002 (EPH – Puntual) and the second semester information since 2003 (EPH – Continua). The survey covers only urban population.

For simplicity several sections of this study are focused on years of relative macroeconomic stability separated by equal intervals: 1986, 1992, 1998 and 2004. Also, we include 2006 into the analysis in order to consider the last available information. The first period from 1986 to 1992 was characterized by a drastic fall in GDP and unprecedented rates of inflation. The 1992–98 interval was one of relatively fast growth and structural reforms which were followed by significant changes in the sectoral

[9] Duclos, Esteban and Ray (2004).
[10] When α=0 identification within groups is ignored by the index. In that case, the polarization index coincides with the Gini coefficient.

structure of the economy. The last one, 1998-2004, includes a stage of stagnation and crisis (up to early 2000s) and the beginning of the subsequent recovery.

The variable we use to compute polarization indices and inequality measures is the hourly wage income in the main occupation. The population of interest is made up of people with 15 years old or more.

III.2 Access and employment conditions

This section presents the characterization of the labor market outcomes during the periods previously defined. We analyze the employment access according to different population's groups: gender, education and age; and the evolution of hourly wage and hours of work of employed people. We consider three age groups: between 15 and 24 years old, 25 to 64 and more than 65 years old. The educational level categories were defined by years of education: up to 8 years of education (unskilled), between 9 and 13 years of education (semi-skilled) and more than 13 (skilled).

Employment access

The Greater Buenos Aires employment rate has not changed significantly over the whole period; averaging 52%. Given the sustained increase of the participation rate the conclusion is straightforward: the unemployment rate has experienced a consistent pattern of increase. This phenomenon is particularly noticeable over the growth period of early and mid-90's. For this reason, 1992-98 could be characterized as one of the weakest employment generation periods.

Figure III.1 illustrates changes patterns in the employment, unemployment, and participation rates for groups of individuals defined according to gender, age and educational level.

Grouping by *gender*, we find a sustained increment in women participation rate. In fact, the gender participation gap has been fallen during the whole period. The same happened to the employment gap which has been reduced because of the reduction of the male employment rate up to 1998 and the sustained increase of the women employment rate, particularly, in the last period. In spite of these facts, the gender unemployment gap has been growing since 1998. This could be explained by the more intense impact of the unemployment rate increase over the women. In summary, women have been joining into the labor market, mainly during 1992-98, but not all of them had been absorbed. One possible explanation is a slow reaction of the labor market to the rise of the women labor force participation. Alternatively, women job search could last more because they are looking for specific job attributes like flexible schedules which allow them to balance their productive and reproductive roles.

The classification of the population by age shows the expected employment rate profile: younger people delay their entry into the labor market because they are still investing in human capital. People in the middle age group are in their more active stage of the life cycle while those who have more than 64 are in the retirement period.

Figure III. 2 illustrates this relation for all the years considered. The employment rate of the younger group has deteriorated during the whole period while it has grown up for the others two groups. The less variable employment rate of the older group could be related to the coverage of the pension system[11].

Unemployment is higher among the youth as expected. Their unemployment rate has been growing up to 2004. Notice that while the unemployment rate of the middle and older age groups were falling in the 1998-2004 period, it raised for the younger group. Although a lower employment rate for the youth could be explained by higher levels of human capital investments, the growing unemployment among them bring down that explanation. The conclusion is that the younger are becoming increasingly excluded from the labor market and this could be a source of conflicts and a weaker social cohesion[12].

Figure III.3 reports a positive correlation between educational levels and employment rate. The relation is not so clear when we consider the unemployment rate. In the last years of the sample, the unemployment rate of semi-skilled is higher than for unskilled ones. This could be due to difficulties and longer periods of time necessaries to do the matching in the labor market. For instance, if job requisites are demanding and individuals are wealthy enough, the more skilled-workers dedicate more time to the searching process.

The pattern of change of the unemployment rate was unbalanced among groups of different educational level. When unemployment skyrocketed in 1992-98, the increment was higher for unskilled people than for others groups. Even in a context of strong growth, as consequence of structural reforms at national level; the labor market could not absorb the unskilled labor force. In other words, part of the labor force was becoming increasingly less attractive for the labor market; hence they had fewer chances to find a decent job, and to be integrated into the market economy. In this context social tensions are more likely to emerge. It is interesting to notice that in the 1998-2004 period the unemployment rate of semi-skilled and skilled people has grown while that of unskilled people has been falling. Furthermore, the declining in the last period considered (2004-2006) was more intense for the unskilled group. In spite of this behavior and the economic recovery, the level of the unemployment rate of the unskilled remain higher than in the previous decade.

Employment conditions

This section analyzes hours of work and wage gaps grouping the labor force by gender, age, educational level and informality condition. The reason for considering this last characteristic has to do with the high proportion which represents of the Greater Buenos Aires' labor market.

[11] In a comparative study, Gasparini *et al.* (2006) find that the low employment rates of older people in Argentina, Brazil, Costa Rica and Chile are associated to the high coverage of the pension systems while in countries like Bolivia, Ecuador, Peru, Guatemala, Honduras and Nicaragua the high employment rates of that group are related to weak pension systems.
[12] Even though the unemployment rate fell in the 2004-2006 period among all age groups, the falling was smaller for the youth.

We use two different definitions of labor informality which do not correspond to competing views about informality. Instead, they refer to different phenomena in the labor market.[13] The "productive" definition pictures informal workers as those in low-productivity, unskilled, marginal jobs, while the "legalistic" definition stresses the lack of labor protection and social security benefits.[14]

Figure III.4 and Table III.1 report labor gaps for different groups of workers. The mean hourly wage of men exceeds that of women during the whole period. The gap has shown a volatile path with a growing pattern between 1992 and 2004 and a decline in 2006. In spite of this behavior the gender wage gap remained superior to that registered in 1992: in 2006, a man earns 12% more than a woman while this difference was about 7% in 1992. Women wages are still lower than their male counterparts when controlling for observable characteristics (Table III.2). The gap of hours of work is wider respect to the wage gap and it has been growing up since 1986 with a slight decline in 2006.

People in middle age group earn more than workers in the younger group as expected given their greater experience and seniority. The hours of work gap shows that younger people work less in average than workers in central labor age. This result brings a new sign of the exclusion of the youth from the labor market together with their wage loses because of the growing wage gap.

Considering groups by educational levels, we find that the wage gap between skilled and the rest has significantly widened over the whole period under analysis with a slight decline in 2006. In particular, the wage gap experienced a dramatic increase during 1992-98. That was a period of reforms that were followed by significant changes in the sectoral structure of the economy, and maybe more important, changes in the ways of production used throughout the economy. Notice that the economic changes affected the unskilled and the semi-skilled in roughly the same way: while the wage gap between skilled workers and the rest was growing up, that of semi-skilled and unskilled remained quite stable. As a result, these two groups became increasingly alike, in comparison with the skilled. Table III.2 illustrates changes in the gap skilled/unskilled by showing the coefficients for a college dummy in a Mincer equation. Hence, the distance in terms of hourly wage is the factor that gets aside the group of semi-skilled and unskilled people from those with formal education. It seems that skilled people have taken advantage of the new economic environment and the rest of the workers have struggled with the new economic conditions. A pattern of unbalanced growth of opportunities and outcomes in the labor market may weaken social cohesion and lead to social instability.

Finally, we analyze labor gaps grouping workers by their informality condition. According to both definitions, an informal worker earns less than a formal one and the effect is more intense using the "legalistic" definition. The gap has shown a volatile path using the "legalistic" definition while the pattern was one of consistent increase when we use the "productive" definition with a slight recovery in the last year considered. A similar

[13] Gasparini and Tornarolli (2007)
[14] To implement this classification we consider an individual is an informal worker if he is a salaried worker in a small private firm (up to 5 employees), an unskilled self-employed or zero-income worker.

behaviour is shown by the hours of work gap: informal workers tend to work less respect to workers that do not satisfy that condition. This result is robust to both informality definitions.

The access and employment conditions analysis provide us with relevant evidence to interpret the patterns of changes of labor income polarization indices.

- The women participation rate increased in a sustained manner which translated into a reduction in the gender employment gap. Although, the unemployment gap has grown since 1992. The hours of work gap increased during the whole period analyzed with a recovery during 2006 and the same behaviour was shown by the gender wage gap since 1992.
- There are some signs of exclusion of the youth from the labor market: younger people report the lower employment rate and a higher incidence of unemployment with respect to other age groups. Similarly, there is a gap in terms of hours of work and wage of younger workers with respect to people in central labor age.
- There is a positive correlation between the employment rate and educational levels. The increase in the unemployment rate was more intense for the unskilled in the 1992-1998 period. This result suggests the exclusion of the unskilled from the labor market. Furthermore, skilled workers are moving further away from the rest in terms of hourly wage.
- The wage gap in terms of informality condition is getting wider since 1992 with a recovery in 2006.

III.3 Polarization by characteristics

In this section we try to identify those variables that are more relevant to characterize the labor force population of Greater Buenos Aires into homogeneous groups that antagonize each other in terms of income in the labor market. We consider six alternative groupings of the population according to gender, age, educational level, productive sector, labor relationship[15] and informality condition. Table III.3 presents the Gradín Group Polarization (GGP) and the Zhang and Kanbur (ZK) indices computed for each year of the sample.

For both indicators education is the most relevant variable for income polarization, followed by productive sector for the GGP index and productive sector and informality condition (legalistic definition) depending on the year for the ZK index.

This result suggests that when divided by education, people in each group look more alike, and differences across groups are larger than when dividing by other characteristic. In particular, the classification by gender looks almost irrelevant for polarization in spite of the gender wage gap that was reported in section III.2.

[15] Labor relationship variable includes four categories: employer, salaried worker, self-employed and zero-income workers.

Table III.4 shows the sign of the change in the ZK index and its components (between and within). The analysis of the results is made for those variables that are more relevant to explain polarization in the labor market.

The sign of the change in the ZK index is the same when we take into account the educational level or productive sector as variables that determined the grouping of working population. Polarization by those characteristics has shown a volatile pattern through time. When people are group by informality condition (legal definition) polarization has shown a consistent growing pattern since 1992.

The ZK index evolution by educational level could be analyzed taking into account the results obtained in section III.2. The dramatic increase in the wage gap between skilled workers and the rest during the 1992-1998 range explained the higher between component that prevails over the less homogenous definition of the groups. In the 1998-2004 period the wage gap between unskilled and semi-skilled people fell and translated into a more precise definition of the groups. But the less distance between them in the income space determined a reduction in the ZK index. The same applies to polarization by informality condition (legal definition). When the wage gap between informal and no informal workers increased the same thing occurred with the between component. That means that the groups were getting aside in the income space. The higher distance between groups translated into a higher ZK index in spite of the less precise definition of the groups.

IV. Pure Labor Income Polarization

In this section we turn to the analysis of pure labor income polarization. In addition to documenting the level and changes in polarization, this section studies what are the empirical differences between inequality and polarization, and inspect which the sources of these changes are. In order to do so, we divide the section into two subsections; the first one focus on the first two topics and report various indices of pure income polarization (Wolfson, EGR and DER for several parameters) for all years in our sample as well as the Gini inequality index for the hourly wage distribution across. In the second subsection, we use micro-decomposition techniques in order to examine some factors which may explain the changes in labor income polarization.

IV.1 Pure income polarization: levels and changes

Table IV.1 and Figure IV.1 show the evolution of pure income polarization indices for hourly wages in Greater Buenos Aires over the past two decades. The indices were estimated considering different values for the identification parameter. We also present the Gini inequality index in order to compare it with polarization.

If we consider the EGR index and a particular value for the identification parameter, we observe that the bipolarization index is always below the level of the three-spike index. Recall that this approach consists of choosing the *n*-spike distribution that minimizes the

income dispersion within all socioeconomic classes. Following this rule, the distribution of hourly wages would be characterized by three income poles throughout the time.

The distribution of this statistics is unknown, so in order to validate them statistically we construct confidence intervals using bootstrapping technique. Table IV.2 and Figure IV.2 present the results for the 95% confidence intervals for the Gini and DER indices using two alternative values for the α-parameter. This exercise was made with 500 replications. We observe that polarization indices are very precise because of the narrowed range of the interval which not exceeds 0.019. However, the opposite occurs with the inequality index which presents a relatively wider range of variability (0.046).

Figure IV.2 also show the methodological change in the Household Survey in 2003. This change did not generate a significant modification in Gini and DER evolution due to the overlapping between confidence intervals of each household survey.

Table IV.3 presents the sign of the change in pure income polarization and the inequality index for hourly wages for the intervals of time previously defined. Simultaneously, we test the statistic significance of those sings using the re-sampling technique with 500 replications.

The first bracket of time (1986-92) do not shows significance changes not only for inequality but also for polarization indices. The EGR with three-spikes is the exception. This index presents a reduction in polarization at a 10% significance level and at a 5% when we consider an $\alpha-$parameter equal to 1.

The evolution of the indices is totally different during the second period (1992-98). Table VI.3 presents that both inequality and polarization raised at a 1% significance level. Other studies show this dramatic increment in inequality during this stage and they point out the difficulty to find another recent period with significance change.[16]

The last interval (1998-2004) is interesting from the labor income distribution point of view because it compares the last pre-crisis year with the recovery stage which has begun since mid-2002. As said before, inequality and polarization are related concepts but they could show different behaviors and this is what happened during this period. While inequality remains the same between 1998 and 2004, polarization decreases strongly according to almost every index with 5% and 1% significance levels depending on the identification parameter level (Table IV.3).

Finally, during the 2004-06 range there are no significance changes in inequality and polarization measures for the hourly wages in Greater Buenos Aires. This result is reasonable in a sense that the bracket is relatively short. It is a well acknowledge fact that income distribution shows significance changes in its different attributes in mid or long terms.

[16] Gasparini, *et.al.*, (2001)

IV.2 Micro-decomposition

To further inquire as to the source of these changes we performed a micro-decomposition of the hourly labor income. Before entering into the specification of the technique, we wish to underscore the fact that we follow the methodology developed by Gasparini, Marchionni, and Sosa Escudero (2005) which should be consulted by the interested reader.

Methodology and estimation strategy

The micro-decomposition technique is based upon the computation of different distributions; the actual distribution for year *t*, and that resulting from simulating the hourly wages of each individual in year *t* by fixing some argument of their income-determination function at the level of another year, *t'*. Let w_{it} be the individual *i's* hourly wages at time *t* which can be written as a function F of the vector X_{it} of individual observable characteristics that affect wages and employment, the vector ε_{it} of unobservable characteristics and the vector β_t of parameters that determine market hourly wages.

$$w_{it} = F(X_{it}, \varepsilon_{it}, \beta_t) \qquad i = 1, \ldots, N \tag{4.1}$$

where *N* is total population. The distribution of individual hourly wages can be represented as follows:

$$W_t = \{ w_{1t}, \ldots, w_{Nt} \} \tag{4.2}$$

We can simulate individual hourly incomes by changing one or some arguments in equation 4.1. For instance, the following expression represents hourly wage that the individual *i's* would have obtained in time *t* if the parameters had been those of time *t'*, keeping all other things constant:

$$w_{it}(\beta_{t'}) = F(X_{it}, \varepsilon_{it}, \beta_{t'}) \qquad i = 1, \ldots, N \tag{4.3}$$

Hence, the simulated distribution will be:

$$W_t(\beta_{t'}) = \{ w_{1t}(\beta_{t'}), \ldots, w_{Nt}(\beta_{t'}) \} \tag{4.4}$$

The contribution to the overall change in the distribution of a change in the parameters vector or any other *k*-argument of the hourly wage function F, between *t* and *t'*, ceteris paribus, can be obtained by comparing the equations 4.2 and 4.4. Even though we can compare the whole distributions, we are interested in evaluate polarization indexes $PI(W)$. Therefore, the effect of a change in a *k*-argument on the hourly wage distribution is given by,

$$PI[W_t(k_{t'})] - PI[W_t(k_t)] \tag{4.5}$$

The decomposition of the change in the DER polarization index was performed for the years 1986, 1992, 1998, 2004 and 2006 for values of the alpha parameter 0.5 and 0.75, changing parametric estimates of returns to education (β^{ed}), gender gap (β^{g}), returns to experience (β^{ex}), residuals, employment and education levels of the population.

In order to analyze the effects of *changes in parameters* on polarization, the econometric specification of the model corresponds to the reduced form of the labor decisions model

originally proposed by Heckman (1974).[17] Leaving technical details aside and under general conditions, it is possible to derive a reduced form for the equilibrium relations in which wages and hours of work are expressed as functions of the variables taken as exogenous. In this way, the model has two equations –one for wages and one for the number of hours of work- and both are function of factors taken as given that affect wages and hours, which may or may not have elements in common. The error terms represent unobservable factors that affect the determination of endogenous variables.

Considering that we observe positive values of wages and hours of work for a particular individual if and only if the individual actually works, the reduced form model for these two variables is specified as follows[18]:

$$w_{it}^{*} = X_{it}^{1}\,\beta_t + \varepsilon_{it}^{1} \qquad i = 1,\ldots,N \qquad (4.6)$$

$$L_{it}^{*} = X_{it}^{2}\,\delta_t + \varepsilon_{it}^{2} \qquad i = 1,\ldots,N \qquad (4.7)$$

with

$$w_{it} = w_{it}^{*} \quad if\ L_{it}^{*} > 0$$
$$w_{it} = 0 \quad if\ L_{it}^{*} \le 0$$
$$L_{it} = L_{it}^{*} \quad if\ L_{it}^{*} > 0$$
$$L_{it} = 0 \quad if\ L_{it}^{*} \le 0$$

where w_{it} and L_{it} are the observed wages and hours of work respectively. For estimation purposes we assume that ε_{it}^{1} and ε_{it}^{2} have a bivariate normal distribution with $E(\varepsilon_{it}^{1}) = E(\varepsilon_{it}^{2}) = 0$, variances σ_t^{1} and σ_t^{2}, and correlation coefficient ρ. This particular specification corresponds to the Tobit type III model in Amemiya's (1985) classification.[19]

To study the effect of the *random term* on polarization indexes changes we follow an approximate solution of the *rank-preserving transformation* operation.[20] It consists of assuming that both distributions of residuals terms at *t* and *t'* are the same up to a proportional transformation. In order to do so, we supposed previously that the error terms of the hourly wage equations are normally distributed, with zero mean and their variance is estimated as an extra parameter in the Heckman procedure. The rank-preserving transformation is then equivalent to multiplying the residual observed at time *t* by the ratio of standard deviations at time *t'* and *t*. Then,

$$w_{it}\left(\hat{\varepsilon}_{it'}^{1}\right) = X_{it}\,\hat{\beta}_t + \hat{\varepsilon}_{it}^{1}\left(\hat{\sigma}_{t'}^{1} \,/\, \hat{\sigma}_{t}^{1}\right) \qquad i = 1,\ldots,N$$

With the purpose to examine the *employment* effects on labor polarization, the decomposition methodology requires simulating hourly wages for individuals who do not work. Because we do not observe wages, we cannot use the previous equations (4.6 and 4.7) to estimate the residual terms. For each individual in that situation, we assigned as an "error term" a random draw from the bivariate normal distribution implicit in the

---

[17] Gasparini, Marchionni, and Sosa Escudero (2005) pags 64-69.

[18] Even though we estimate the δ parameters, they are not relevant in our analysis. Equation (4.7) is an estimation tool for ε_{it}^{2}.

[19] Idem

[20] This method consists in replacing the residual in the n^{th} percentile (of residuals) at time *t* by the residuals in the n^{th} percentile at time *t'*, for all *n*. For further details see Bourguignon and Ferreira (2005).

wage-labor supply model whose parameters are consistently estimated by the Heckman procedure.[21]

Lastly for the estimation of the *education level* effect, we used a rough nonparametric method. We divide the adult population in homogeneous groups by gender and age and then replicate the educative structure of a given cell in year *t'* into the corresponding cell in year *t*. [22]

Results

Before we present the results, three observations must be pointed out. First, the preceding decomposition has a restrictive property which is *path dependence.*[23] Table IV.4 reports the average of using alternatively *t* and *t'* as the base year.[24] Second, a positive number of a *k*-argument change reveals two results: i) the sign of it indicates that the *k*-argument effect increased polarization, and ii) the magnitude of it compared with other changes, implies that the change in the *k*-argument was a significant factor affecting polarization in the distribution of the hourly wages. Third, since the results are not sensitive to the choice of the identification coefficient (alpha), we will conduct the analysis considering alpha = 0.5.

Table IV.8 shows that changes in *returns to education* had a cohesion effect on the hourly wage income distribution in the beginning (1986-92) and end (1998-04) of the six-year intervals. Although, they suffer an antagonize effect over the six years in between the previous periods (1992-98). The outcomes are similar when we consider a higher level of identification. In each period, changes in the returns to education represented a relevant factor for explaining polarizations' movements.

Changes in *gender* parameters of the wage equation divide the whole period into two sub-periods. The first (1986-1998) is characterized by the shrunk in the gender wage gap which generates a polarized-diminishing effect. The second (1998-2004) describes a polarized-increasing effect as a consequence of the expansion in this gap, as said in section III.2. Notice, the magnitude of this effect is insignificant relative to the previous one.

The *returns to experience* (age) have a polarized effect on the hourly wage income distribution in the two first periods which may be explained by the increase in the wage gap between middle-age and young-age groups. However, in the latter periods this effect changes its tendency implying cohesion between age groups.

In general, changes in *endowments and returns to unobservables factors* have implied polarizing changes in hourly wages distributions. These effects were particularly strong in 1992-98. The results suggest that an increase in the dispersion of unobservables was,

[21] See Gasparini, Marchionni and Sosa Escudero (2005) pag. 69.
[22] Idem.
[23] This property means that changing the conditional income distribution from the one observed in *t* to that observed in *t'* does not have the same effect on the distribution when this is done with the distribution of characteristics X observed in *t*, as when X is observed in *t'* Bourguignon and Ferreira (2005).
[24] Notice, to address the problem of the path dependence, Shorrocks (1999) provides a formal definition of the appropriate "averaging" concept on the basis of the Shapley values. Detail results were omitted and are disposable by request.

after returns to education, the main factors affecting hourly wages polarization over the period under analysis.

Even the unemployment rates skyrocketed in the mid-90's and they have remained high during the whole period, the *employment effect* on the hourly wage income distribution was negative and negligible (Table IV.6). The reason which contributes to reduce the effect of the great increase in unemployment on hourly wage polarization is that the employment rate did not change significantly during the whole period as stated in section III.2. Hence, there is a minor change in the number of individuals without hourly wage income.

Argentina, as many developing countries, has witnessed a dramatic change in the *educational structure* of its population during the 80's and 90's.[25] The results show that, in Greater Buenos Aires, that change has a polarized-increasing effect on hourly wages distributions over the whole period and particularly, during the 90's. A rough explanation could be the contrast between a higher identification due to more educated population and a higher alienation as a result of educational level groups with relatively high dispersion. Hence, the second effect would compensate the first one implying an increased in antagonism.[26]

The last row in Table IV.6 is estimated as a residual. It includes the effects of *interaction terms and of many other factors* not considered in the analysis. These terms are not as small as we expect, implying either that there are other relevant factors not considered in the analysis, such us, institutional or sector effects, or that they do not tend to compensate each other.

[25] Gasparini, Marchionni and Sosa Escudero (2005)
[26] It is important to notice that in this rough explanation we are not taking into account the correlation between alienation and identification.

V. Concluding remarks

The empirical evidence reveals two stages over the last two decades (1986-2006): the first one characterized by an increment of all polarization indices and the second one, by shrinkage of them.

We identify two vulnerable groups which have been excluded from the labor market, during 1992-98: the youth who joined the labor force and suffered a dramatically increased in their unemployment rates and those unskilled and semi-skilled individuals who experienced a strong separation from the skilled ones.

Inspecting potential factors which could explain those changes, *returns to education* surge as the main polarization force in the labor market. They had a cohesion effect on the hourly wage income distribution in the beginning (1986-92) and end (1998-04) of the six-year intervals. Although, they suffer an antagonize effect over the six years in between the previous periods (1992-98). Other relevant forces are *endowments and returns to unobservables factors* and *other factors* such as interaction terms between unobservables and observables and many other variables e.g. institutional or sector effects.

Finally, an equalizing distribution of the human capital is a possible alternative for a less polarized labor market. A society with high levels of economic polarization would have fewer chances to achieve that objective than those which ensure to their individuals' equal opportunities to obtain economic results. At the same time, education is a key tool for building common values which reinforces social cohesion.

Bibliography

Alesina, and Spolaore (1997), On the number and size of nations *Quarterly Journal of Economics*, 113, pp. 1027-1056.

Atkinson, A. and Bourguignon, F. (2000). Income distribution and economics. *Handbook of Income Distribution*. Elsevier Science B.V. Pg. 1-5, 41-50.

Bourguignon, F and Ferreira, F, (2005) Decomposing changes in the Distribution of household incomes: Methodological aspects, in Bourguignon, Ferreira and Lustig, *The microeconomics of income distribution dynamics (in East Asia and Latin America)*, Oxford University Press, pp 17-46.

Collier, P. and Hoeffler, A.,(2001) Greed and grievance in civil war, World Bank (Washington D.C).

Cowell, F. (2000). Measuring inequality. LSE Handbooks in Economic Series, Prentice Hall/Harvester Wheatsheaf

D' Ambrosio, C (2001) Household characteristics and the distribution of income in Italy: an application of social distance measures, *Review of Income and Wealth*, (Series 47, Number 1, March 2001)

D' Ambrosio, C and Wolf, E. (2001) Is Wealth Becoming More Polarized in The United States?, Working paper 330, Levy Economics Institute, Bard College.

Deaton, A. (1997). *The analysis of household surveys. Microeconomic analysis for development policy*. Washington D.C.: The World Bank.

Duclos, JY, Esteban, J. and Ray, D. (2004). Polarization: Concepts, Measurements, Estimation. *Econometrica* 72 (6), November, pp. 1737-1772.

Esteban, J. (2004) Economic integration and cross-country inequality: the European experience, International Policy Workshop (September, Berlin)

Esteban, J. and Ray, D. (1994), On the Measurement of Polarization. *Econometrica* 62, (4), November, pp. 819-852.

Esteban, J., Gradín, C. and Ray, D. (1999) Extensions of the measure of Polarization, with an application to the income distribution of five OECD countries, mimeo, Instituto de Análisis Económico.

Foster, J. and Wolfson, M. (1992), Polarization and the Decline of the Middle Class: Canada and the US, mimeo, Vanderbilt University.

Ganuza, E., Paes de Barros, R., Taylor, L. and Vos, R. (2001). Liberalización, desigualdad y pobreza: América Latina y el Caribe en los 90. Eudeba, PNUD, CEPAL.

Gasparini, L (2004 a) Different lives: inequality in Latin America and the Caribbean, Working paper CEDLAS. Chapter 2, 2004 World Bank LAC Flagship Report.

Gasparini, L (2004 b) Argentina's Distributional Failure. The role of Integration and Public Policies. Marquez, G. (ed.) *Labor Markets and Globalization in LAC*. Forthcoming, 2005.

Gasparini, L., Haimovich, F. and Pacheco, A. (2005a). Mercados Laborales y Cohesión Social en América Latina. Working paper CIEPLAN, Chile.

Gasparini,L., Horestein, M, Molina, E and Olivieri, S (2008), Polarización económica, instituciones y conflictos. Dinámicas de la cohesión social latinoamericana, PNUD-CIEPLAN, Chile.

Gasparini, L., Marchionni, M. and Sosa Escudero, W. (2001) "Distribución del Ingreso en la Argentina: Perspectivas y Efectos Sobre el Bienestar", Premio Fulvio Salvador Pagani.

Gasparini, L, Marchionni, M and Sosa Escudero, W (2005b), Characterization of inequality changes through microeconometric decompositions: the case of Greater Buenos Aires, in Bourguignon, Ferreira and Lustig, *The microeconomics of income distribution dynamics (in East Asia and Latin America)*, Oxford University Press, pp 47-83.

Gasparini, L. and Tornarolli, L., "Labor informality in Latin America and the Caribbean: patterns and trends from household microdata survey", Working paper N46, CEDLAS–UNLP.

Gradín, C (2003) Polarization and inequality in Spain: 1973-1991, Universidade de Vigo (Vigo, Spain)

Gradín, C. (2000) Polarization by sub-populations in Spain, 1973-91, *Review of Income and Wealth*, (Series 46, Number 4, December)

Haimovich, F., Gasparini, L. and Winkler, H. (2005). Distribución del Ingreso en América Latina: Explorando las Diferencias entre Países. XL Reunión de la AAEP, La Plata.

Horenstein, M. and Olivieri, S. (2004), Income polarization in Argentina: Pure Income Polarization, Theory and Applications, *Económica*, Nº 1-2, Año L, (La Plata, Argentina, Enero-Diciembre).

IADB (1998). *América Latina frente a la desigualdad*. Banco Interamericano de Desarrollo, Washington, D.C.

Lambert, P. (2001). *The distribution and redistribution of income*. Manchester University Press.

Morley, S. (2000). *La distribución del ingreso en América latina y el Caribe*. Fondo de Cultura Económica.

Reynal-Querol, M. (2001) Ethnicity, political system and civil wars, Institut d'Analisis Economic (IAE), CSIC (Barcelona, Spain).

Ricardo, D. (1817). *Principles of Political Economy*. Dent, London.

Sánchez-Páramo, C. and Schady, N. (2003). Off and running? Technology, trade and rising demand for skilled workers in Latin America. Closing the gap in education and technology in Latin America, background paper.

SEDLAC (2008). Socio-economic database for Latin American and the Caribbean. www.depeco.econo.unlp.edu.ar/cedlas

Sen, A. (2000). Social justice and the distribution of income. *Handbook of Income Distribution*. Elsevier Science B.V.

Shorrocks, A (1999) Decomposition procedures for distributional analysis: A unified framework based on the Shapley values. University of Essex, Essex, U.K. Processed.

Sosa Escudero, W. and Gasparini,L (2000) A Note on the statistical significance of changes in inequality, *Económica*, Nº1, Año XLVI, (La Plata, Argentina).

Viollaz, M. (2008), Polarización de ingresos laborales. Argentina 1992-2006, Working paper N70, CEDLAS-UNDP.

Wodon, Q. *et al*. (2000). Poverty and policy in Latin America and the Caribbean. *World Bank Technical Paper* 467.

Wolfson, M. (1994), When Inequality Diverges, *American Economic Review*, 84, 353-358.

Zhang, X. and Kanbur, R (2001), What difference do polarization measures make? An application to China, *Journal of Development Studies*, 37, pp. 85-98.

Table III. 1: Labor gaps by gender, age groups, educational level and informality condition - Greater Buenos Aires (1986 – 2006)

Ratios by gender

Year	Hourly wages (male/female)	Hours of work (male/female)
1986	1.210	1.194
1992	1.071	1.278
1998	1.111	1.324
2004	1.219	1.412
2006	1.118	1.383

Ratios by age groups

Year	Hourly wages		Hours of work	
	(25-65/15-24)	(25-64)/(+65)	(25-65/15-24)	(25-64)/(+65)
1986	1.589	0.876	1.197	1.146
1992	1.560	1.043	1.086	1.223
1998	1.735	0.796	1.100	1.225
2004	1.761	0.416	1.072	1.220
2006	1.783	0.798	1.079	1.235

Ratios by educational level

Year	Hourly wages		Hours of work	
	high /medium level	low / medium level	high /medium level	low / medium level
1986	1.584	0.694	0.909	1.013
1992	1.734	0.739	0.897	0.982
1998	2.155	0.761	0.876	0.966
2004	2.290	0.712	0.896	0.936
2006	2.099	0.721	0.871	0.940

Ratios by informality labor condition

Year	Hourly wages		Hours of work	
	informal / non informal (productive definition)	informal / non informal (legal definition)	informal / non informal (productive definition)	informal / non informal (legal definition)
1986	0.841	0.715	0.958	0.889
1992	0.813	0.770	0.978	0.932
1998	0.691	0.664	0.888	0.896
2004	0.731	0.643	0.981	0.789
2006	0.671	0.499	0.962	0.830

Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October.

Table III.2: Returns to college education and gender gap in terms of hourly wages Coefficients of a Mincer equation - Greater Buenos Aires

	Gender gap	*Educational level gap*
1986	0.248	0.552
	[0.040]***	[0.055]***
1992	0.403	0.526
	[0.056]***	[0.059]***
1998	0.158	0.745
	[0.037]***	[0.048]***
2004	0.097	0.636
	[0.082]	[0.055]***
2006	0.174	0.777
	[0.045]***	[0.042]***

Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October.

Table III.3: Polarization by characteristics – Greater Buenos Aires

	Gender		Age		Educational level		Sector		Labor relation		Informality (Legal)		Informality 2 (Prodcutivity)	
	ZK	GGP	ZK	GGP	ZK	GGP	ZK	GGP	ZK	GGP	ZK	GGP	ZK	GGP
1986	0.01	0.70	0.05	0.78	0.34	1.20	0.07	0.93	0.01	0.70	0.05	0.756	0.014	0.730
1988	0.00	0.66	0.06	0.79	0.38	1.26	0.05	0.82	0.01	0.63	0.08	0.778	0.048	0.794
1991	0.00	0.62	0.04	0.78	0.30	1.17	0.08	0.91	0.03	0.72	0.05	0.780	0.014	0.732
1992	0.00	0.66	0.05	0.81	0.30	1.15	0.07	0.91	0.02	0.76	0.04	0.772	0.000	0.662
1993	0.00	0.65	0.04	0.78	0.26	1.12	0.09	0.91	0.02	0.73	0.02	0.727	0.002	0.672
1994	0.00	0.66	0.05	0.78	0.30	1.15	0.09	0.92	0.02	0.74	0.03	0.743	0.011	0.722
1995	0.00	0.61	0.05	0.78	0.39	1.24	0.07	0.91	0.09	0.81	0.04	0.751	0.017	0.727
1996	0.00	0.58	0.05	0.77	0.42	1.23	0.08	0.93	0.06	0.80	0.02	0.711	0.011	0.709
1997	0.00	0.61	0.05	0.78	0.35	1.19	0.07	0.91	0.06	0.78	0.04	0.764	0.029	0.763
1998	0.00	0.61	0.05	0.76	0.41	1.25	0.10	0.96	0.05	0.77	0.07	0.786	0.040	0.773
1999	0.00	0.59	0.05	0.77	0.42	1.25	0.14	1.02	0.05	0.76	0.07	0.800	0.037	0.768
2000	0.00	0.58	0.06	0.77	0.37	1.23	0.14	1.02	0.03	0.71	0.06	0.775	0.015	0.699
2001	0.00	0.59	0.05	0.74	0.33	1.22	0.08	0.92	0.04	0.70	0.06	0.770	0.021	0.704
2002	0.00	0.60	0.04	0.72	0.41	1.27	0.09	0.94	0.05	0.71	0.14	0.852	0.027	0.714
2003 (1)	0.00	0.63	0.04	0.72	0.35	1.23	0.14	1.03	0.05	0.69	0.05	0.752	0.045	0.762
2003	0.00	0.56	0.04	0.72	0.24	1.15	0.09	0.93	0.04	0.72	0.11	0.851	0.011	0.649
2004	0.00	0.61	0.04	0.74	0.30	1.17	0.10	0.94	0.03	0.74	0.10	0.839	0.028	0.723
2005	0.00	0.62	0.04	0.74	0.40	1.24	0.13	0.99	0.04	0.70	0.12	0.855	0.045	0.755
2006	0.00	0.62	0.05	0.75	0.38	1.23	0.14	1.02	0.04	0.71	0.20	0.922	0.089	0.799

Note: GGP= Gradín Group Polarisation Index with α=1, β=1
ZK=Zhang and Kanbur index.
(1) EPH puntual
Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October.

Table III.4: Signs of changes in the ZK measure and its components over the 1990s

	1986-1992	1992-1998	1998-2004	2004-2006
Gender				
ZK	(-)	(-)	(+)	(+)
Wth	(-)	(+)	(-)	(+)
Btw	(-)	(-)	(+)	(+)
Age				
ZK	(+)	(-)	(-)	(+)
Wth	(-)	(+)	(-)	(+)
Btw	(-)	(+)	(-)	(+)
Educational Level				
ZK	(-)	(+)	(+)	(+)
Wth	(-)	(+)	(-)	(-)
Btw	(-)	(+)	(-)	(+)
Sector				
ZK	(-)	(+)	(+)	(+)
Wth	(-)	(+)	(-)	(-)
Btw	(-)	(+)	(+)	(+)
Labor relation				
ZK	(+)	(+)	(-)	(+)
Wth	(-)	(+)	(-)	(+)
Btw	(+)	(+)	(-)	(+)
Informality (L)				
ZK	(-)	(+)	(-)	(+)
Wth	(-)	(+)	(-)	(+)
Btw	(-)	(+)	(-)	(+)
Informality 2 (P)				
ZK	(-)	(+)	(+)	(+)
Wth	(-)	(+)	(-)	(-)
Btw	(-)	(+)	(+)	(+)

Note: ZK=Zhang and Kanbur index.
Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October.

Table IV.1: Pure polarization - Hourly labor income - Greater Buenos Aires

	Gini	Wolfson	EGR(2)			EGR(3)			DER			
			α =			α =			α =			
			1	1.3	1.6	1	1.3	1.6	0.25	0.5	0.75	1
1986	0.391	0.319	0.170	0.123	0.087	0.632	0.426	0.290	0.288	0.240	0.213	0.196
1988	0.425	0.373	0.191	0.140	0.100	0.688	0.466	0.319	0.308	0.253	0.221	0.201
1991	0.389	0.313	0.170	0.125	0.089	0.620	0.418	0.284	0.288	0.239	0.212	0.195
1992	0.372	0.315	0.163	0.117	0.081	0.590	0.400	0.272	0.279	0.232	0.205	0.187
1993	0.381	0.334	0.169	0.123	0.086	0.613	0.417	0.286	0.285	0.236	0.208	0.189
1994	0.385	0.315	0.169	0.123	0.086	0.621	0.421	0.288	0.286	0.237	0.208	0.190
1995	0.419	0.351	0.186	0.135	0.096	0.678	0.459	0.314	0.304	0.249	0.218	0.198
1996	0.421	0.347	0.185	0.135	0.096	0.682	0.461	0.315	0.305	0.251	0.221	0.201
1997	0.400	0.342	0.173	0.125	0.087	0.647	0.438	0.298	0.295	0.240	0.208	0.186
1998	0.424	0.368	0.188	0.137	0.096	0.688	0.467	0.320	0.309	0.252	0.219	0.198
1999	0.414	0.373	0.187	0.136	0.096	0.668	0.454	0.311	0.304	0.248	0.215	0.193
2000	0.430	0.392	0.195	0.143	0.101	0.701	0.476	0.327	0.315	0.256	0.222	0.200
2001	0.443	0.402	0.201	0.148	0.105	0.722	0.490	0.336	0.321	0.259	0.223	0.198
2002	0.449	0.429	0.210	0.155	0.112	0.724	0.492	0.338	0.326	0.263	0.226	0.201
2003 (1)	0.442	0.401	0.198	0.144	0.100	0.696	0.472	0.323	0.318	0.255	0.218	0.192
2003	0.441	0.378	0.196	0.143	0.102	0.702	0.475	0.324	0.316	0.255	0.220	0.196
2004	0.410	0.343	0.173	0.124	0.084	0.649	0.440	0.300	0.300	0.241	0.204	0.179
2005	0.411	0.348	0.178	0.128	0.089	0.650	0.442	0.299	0.300	0.241	0.205	0.181
2006	0.411	0.351	0.178	0.128	0.088	0.656	0.444	0.302	0.300	0.242	0.206	0.181

Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October.
Notes: (1) EPH Puntual

Table IV.2: Inequality and Pure Polarization confidence intervals- Hourly labor income - Greater Buenos Aires

	Gini			DER			DER		
				α = 0.5			α = 0.75		
	Obs	Lowest	Highest	Obs	Lowest	Highest	Obs	Lowest	Highest
1986	0.391	0.374	0.408	0.240	0.232	0.247	0.213	0.206	0.220
1988	0.425	0.411	0.439	0.253	0.244	0.258	0.221	0.215	0.228
1991	0.389	0.372	0.404	0.239	0.230	0.247	0.212	0.203	0.219
1992	0.372	0.357	0.388	0.232	0.225	0.239	0.205	0.198	0.212
1993	0.381	0.368	0.395	0.236	0.230	0.241	0.208	0.200	0.214
1994	0.385	0.371	0.400	0.237	0.229	0.243	0.208	0.201	0.215
1995	0.419	0.405	0.437	0.249	0.243	0.256	0.218	0.211	0.224
1996	0.421	0.402	0.439	0.251	0.243	0.259	0.221	0.212	0.229
1997	0.400	0.387	0.415	0.240	0.234	0.246	0.208	0.202	0.213
1998	0.424	0.410	0.438	0.252	0.243	0.258	0.219	0.214	0.225
1999	0.414	0.400	0.429	0.248	0.242	0.253	0.215	0.208	0.220
2000	0.430	0.416	0.443	0.256	0.250	0.262	0.222	0.216	0.228
2001	0.443	0.429	0.459	0.259	0.251	0.265	0.223	0.215	0.229
2002	0.449	0.431	0.468	0.263	0.254	0.271	0.226	0.218	0.234
2003 (1)	0.442	0.420	0.466	0.255	0.245	0.264	0.218	0.208	0.227
2003	0.441	0.424	0.460	0.255	0.247	0.263	0.220	0.211	0.227
2004	0.410	0.398	0.423	0.241	0.236	0.246	0.204	0.198	0.209
2005	0.411	0.399	0.427	0.241	0.234	0.246	0.205	0.199	0.210
2006	0.411	0.398	0.426	0.242	0.234	0.248	0.206	0.201	0.212

Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October.
Notes:
(1) EPH Puntual May 2003.
(2) Note: 500 bootstrap replications. * significant at 10%; ** at 5%; ** at 1%

Table IV.3: Pure Polarization and Inequality changes- Greater Buenos Aires

	1986-1992	1992-1998	1998-2004	2004-2006
Gini	(=)	(+)***	(=)	(=)
Wolfson	(=)	(+)***	(=)	(=)
EGR(2)				
α = 1.0	(=)	(+)***	(-)**	(=)
α = 1.3	(=)	(+)***	(-)**	(=)
α = 1.6	(=)	(+)***	(-)**	(=)
EGR(3)				
α = 1.0	(-)**	(+)***	(-)**	(=)
α = 1.3	(-)*	(+)***	(-)**	(=)
α = 1.6	(-)*	(+)***	(-)**	(=)
DER				
α = 0.25	(=)	(+)***	(=)	(=)
α = 0.50	(=)	(+)***	(-)**	(=)
α = 0.75	(=)	(+)***	(-)***	(=)
α = 1.00	(=)	(+)*	(-)***	(=)

Source: Own estimations based on SEDLAC.
Note: 500 bootstrap replications. * significant at 10%; ** at 5%; ** at 1%

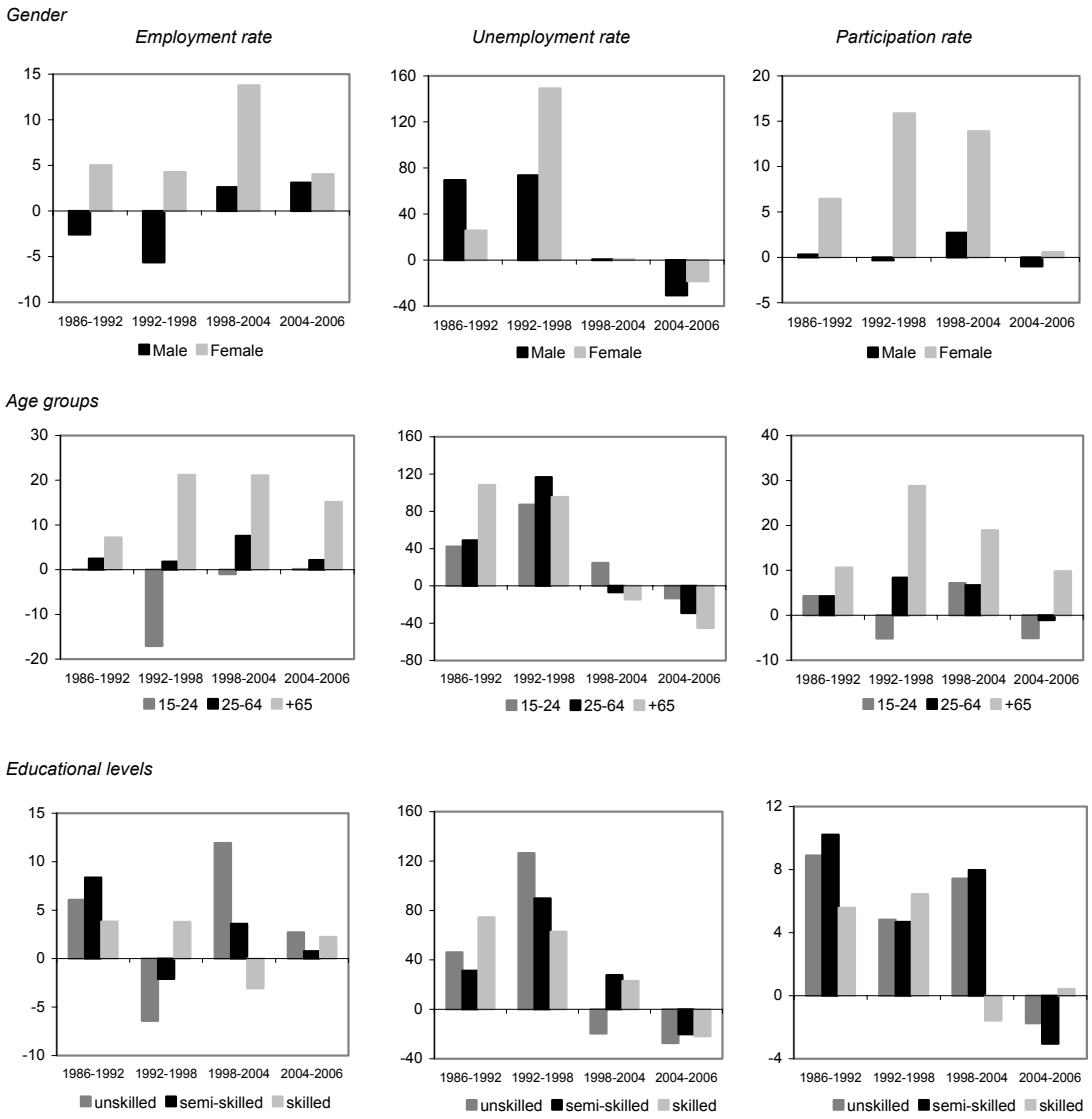
Table IV.8: Decomposition of the change in the DER index: Average results changing the base year in Greater Buenos Aires, selected periods

Indicator	α = 0.50			
	1986 - 92	1992 - 98	1998 - 04	2004- 06
Observed	-0.8	2.0	-1.2	0.1
Effects				
1. Returns to education	-0.6	1.3	-1.3	1.1
2. Gender wage gap	-0.1	-0.1	0.1	0.0
3. Returns to experience	0.1	0.1	-0.2	-0.7
4. Unobservables	0.2	0.8	0.7	-0.5
5. Employment	0.0	-0.1	-0.1	0.0
6. Education structure	0.0	0.4	0.2	0.2
7. Other factors	-0.3	-0.4	-0.5	-0.1

Indicator	α = 0.75			
	1986 - 92	1992 - 98	1998 - 04	2004- 06
Observed	-0.8	1.4	-1.7	0.2
Effects				
1. Returns to education	-0.4	1.0	-1.1	0.9
2. Gender wage gap	-0.1	-0.1	0.1	0.0
3. Returns to experience	0.1	0.0	-0.1	-0.5
4. Unobservables	0.1	0.6	0.5	-0.3
5. Employment	0.0	0.0	-0.1	0.0
6. Education structure	-0.1	0.3	0.1	0.1
7. Other factors	-0.5	-0.3	-1.1	0.0

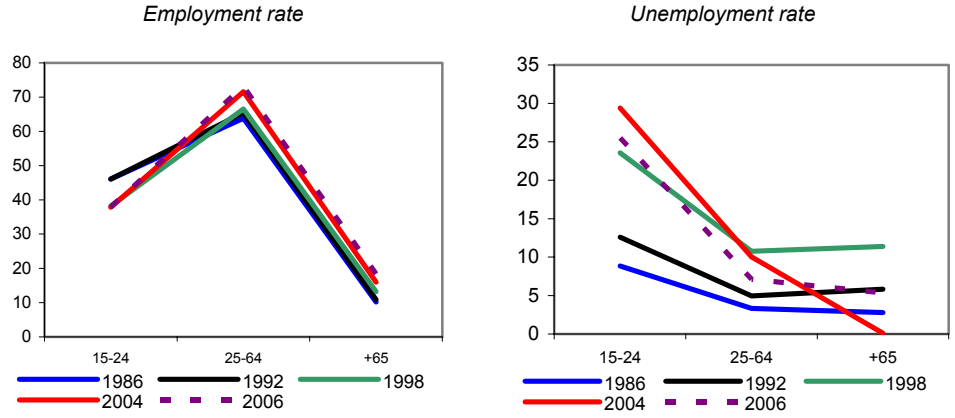
Note: The hourly wages distribution includes those individuals with $w_{it} > 0$, $w_{it}(k_{it'}) > 0$ and we exclude those wages which are greater than 20 times the median.
Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October.

Figure III.1: Percental change in employment, unemployment and participation rates by gender, age and educational levels – Greater Buenos Aires (1986 – 2006)



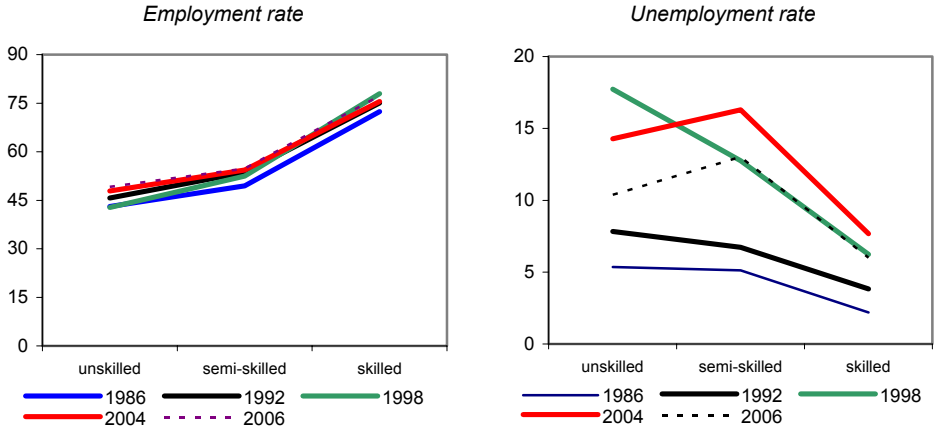
Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October.

Figure III.2: Employment and unemployment rates by age groups – Greater Buenos Aires (1986 – 2006)



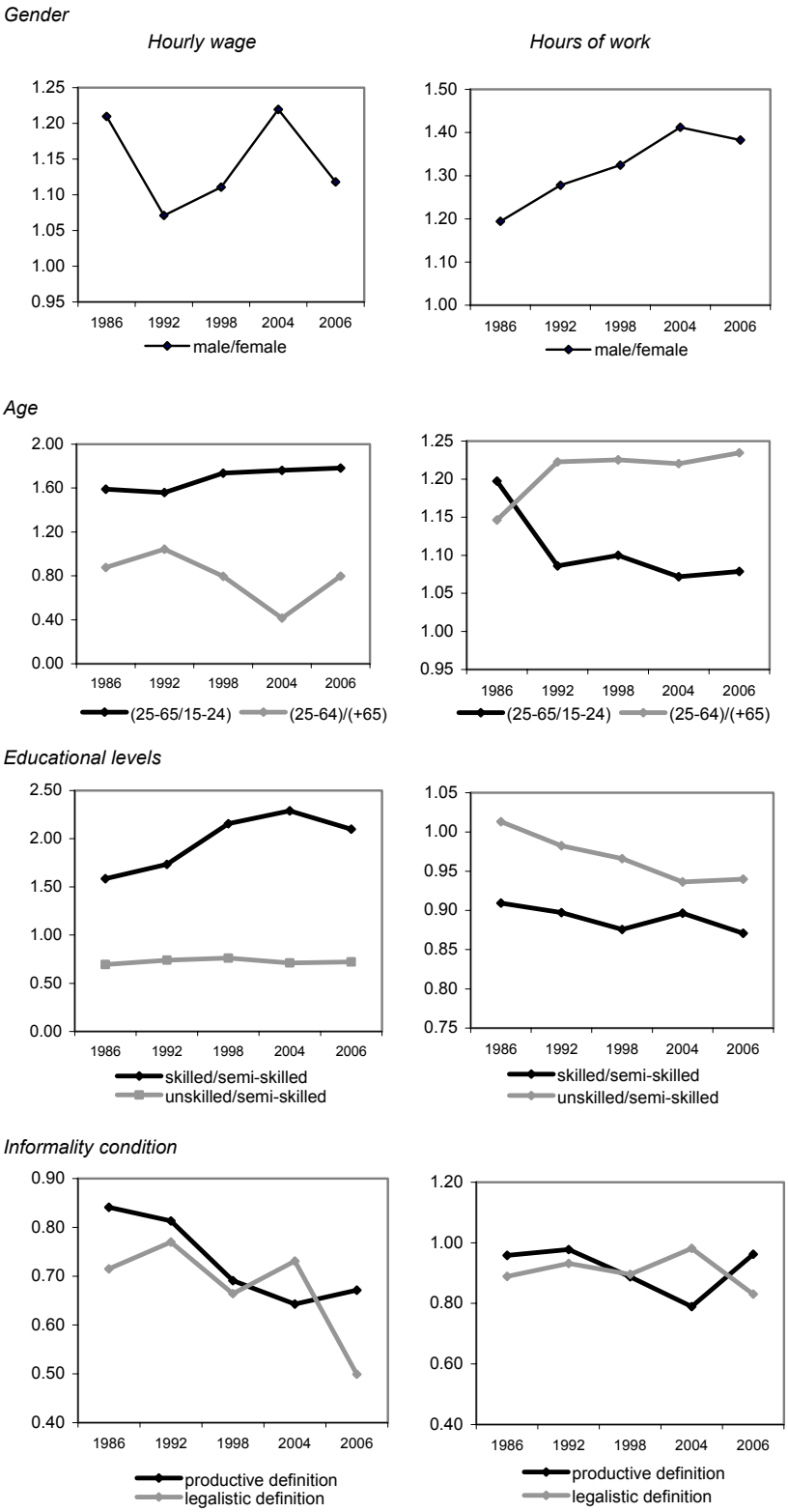
Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October

Figure III.3: Employment and unemployment rates by educational levels – Greater Buenos Aires (1986 – 2006)



Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October

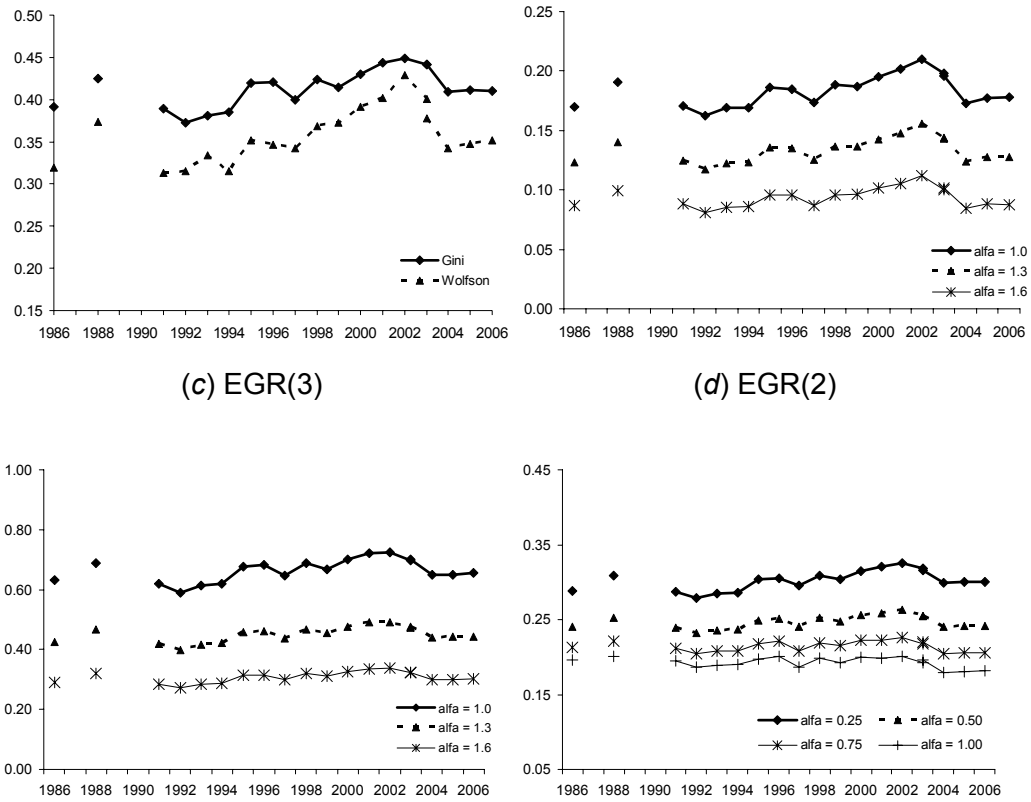
Figure III.4: Labor gaps by gender, age, educational levels and informality condition– Greater Buenos Aires (1986 – 2006)



Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October

Figure IV.1: Wage Inequality and Polarization - Greater Buenos Aires 1986 - 2006

(*a*) *Gini* and *Wolfson* (*b*) EGR(2)

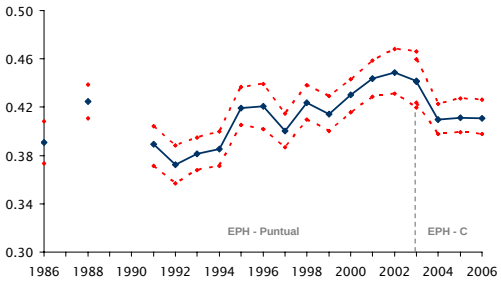


(*c*) EGR(3) (*d*) EGR(2)

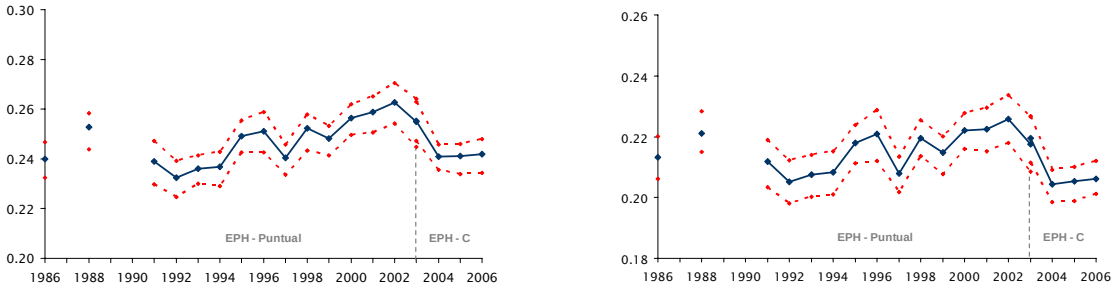
Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October

Figure IV.2: Inequality and polarization confidence intervals- Greater Buenos Aires 1986 - 2006

(*a*) *Gini*



(*b*) DER. α = 0.50 (*b*) DER. α = 0.75



Note: 500 bootstrap replications.
Source: Authors' calculations based on SEDLAC, Greater Buenos Aires, October

Methodological appendix

Gradín Group Polarization Index (GGP)

Gradín (2000) assumes that despite polarization occurring in the income space, groups in the distribution are the result of similarities with respect to a relevant attribute other than income. Thus, he treats the distribution as if it were the aggregate result of more than one stochastic process. In this sense, a population can be divided into "n" groups or sub-populations according with any characteristic (*e.g.* race, region, occupation, etc.). The number of groups depends on the nature of the characteristic. Groups are exogenously conformed according to whether their members share the same category for a given characteristic regardless of their income proximity. Compared to identification by income intervals, we expect higher intra-group dispersion and lower between groups heterogeneity.

Define a partition $\rho^c = (q_1,...q_n; m_1,...m_n)$, where q_i is the population share in group *i* and $m_1 \leq m_2 \leq .... \leq m_3$ indicate average incomes of the groups. The measure is defined in accordance with the EGR(1999) index as:

$$GP(F;\alpha,\beta,\rho^c) \equiv P(F;\alpha,\beta,\rho^c) - (-\beta) = ER(\alpha,\rho^c) - \beta\left[\varepsilon(F;\rho^c) - 1\right]$$

$$\varepsilon(F;\rho^c) = G(F) - G(\rho^c)$$

The error term is expressed in parallel to EGR(1999) and accounts for both intra-group inequality as well as overlap between groups.[27]

The index is sensitive to the number of categories for which the characteristic is expressed. In particular, the smaller the number, the larger we expect both terms in the index, so the net effect is ambiguous. The most relevant characteristics will be those showing at the same time high polarization between the groups and homogeneity within them.

Zhang and Kanbur Index (ZK)

Zhang and Kanbur (2001) propose an index of polarization based on the ratio of the between-group inequality to the within-group inequality – both measured with Theil's Generalized Entropy index. This polarization index captures the average distance between groups in relation to income differences within groups. As the groups become internally more homogeneous, within-group inequality diminishes, differences across groups are, relatively speaking, magnified and polarization is higher. Similarly, if we leave within-group inequality unchanged as the distance between-group increases, polarization rises.

The measure for polarization suggested by Zhang and Kanbur is:[28]

$$ZK = \frac{T_B}{T_W}$$

[27] For a more detailed treatment of the subject we refer the reader to Gradín (2000) and Esteban, Gradín and Ray (1999)
[28] For a more detailed treatment of the subject we refer the reader to Zhang and Kanbur (2001)

$$ZK = \frac{\sum_{j=1}^{K} \frac{n_j}{N} \frac{\mu_j}{\mu} \ln\left(\frac{\mu_j}{\mu}\right)}{\sum_{j=1}^{K} \frac{n_j}{N} \frac{\mu_j}{\mu} T_j} \text{ where } T_j = \frac{1}{n_j} \sum_{j=1}^{K} \frac{y_j}{\mu_j} \ln\left(\frac{y_j}{\mu_j}\right)$$

K: number of groups; n_j: number of individuals in each group; *N*: total number of individuals; μ_j: mean income of each group; μ: mean income; *yi:* individual income.

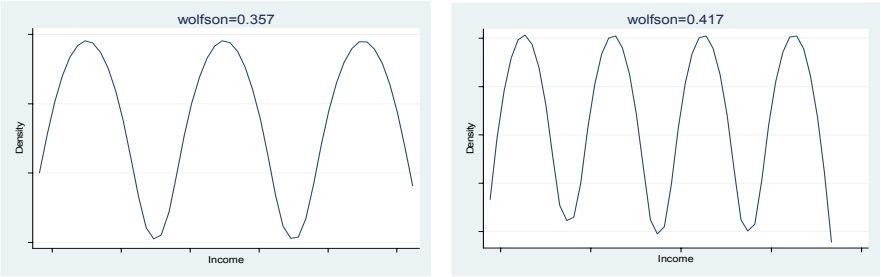
Wolfson(1994)

Wolfson's polarization measure is derived from the Lorenz curve. It is defined as twice the area between the Lorenz curve and the tangent line at the median point (see figure bellow). It can be written as:

$$P^W = \frac{\mu}{m}\left[0.5 - L(0.5) - \frac{G}{2}\right]$$

where μ= mean, m= median, L(0.5) = value of the Lorenz curve at the median income and G = Gini coefficient. Polarization reaches the maximum value when half of the population has zero income and the other half has twice the mean. Wolfson shows that like the Gini this index lies between zero and one.

This measure has problems when there are several income poles. The income distribution in the second panel of the next graph is intuitively less polarised than the income distribution in the first panel, since income masses are less identified. However, the Wolfson index shows the opposite result because it implicitly assumes the existence of two groups of equal size.



ER(1994)

Esteban and Ray (1994) introduce a model of individual attitudes in a society and use four axioms to narrow down the set of possible measures. In particular, they suppose that each individual is subject to two forces. On the one hand she identifies with those she considers to be members of her own group. *I*: $\Re_+ \to \Re_+$ represents the identification function. On the other hand, she feels alienated from those she consider to be members of other groups. *a*: $\Re_+ \to \Re_+$ is the alienation function. An individual with income *y* feels alienation *a*(δ(*y, y'*)) from an individual with income *y'*, where δ(*y, y'*) stands simply for the absolute distance |*y-y'*|. Note that alienation, as well as identification, is perfectly symmetric in this scheme. The joint effect of the two forces is given by the effective

antagonism function, *T(I,a)*. Total polarization in the society is postulated to be the sum of all the effective antagonisms:

$$P(\pi, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_i \pi_j T(I(\pi_i), a(\delta(y, y')))$$

Esteban and Ray demonstrate that the only measure of this family which satisfies the axioms has the following expression.

$$P^*(w, y) = k \sum_{i=1}^{n} \sum_{j=1}^{n} \pi_{i=1}^{1+\alpha} \pi_j \left| y_i - y_j \right|$$

For *k*>0 and $\alpha\varepsilon$[1, 1.6] that indicates the degree of sensitivity to polarization.

EGR(1999)

Esteban, Gradín and Ray (1999) state that the ER (1994) polarization measure for discrete groups or "n-spike representation" should be used only after the population has been regrouped in a way that captures the group identification structure of society. This clustering will lose some of the initial information that concerns the dispersion of the population around the clusters that are treated as single groups: the ER measure needs to be corrected. EGR propose the following polarization measure:

$$P(f; \alpha, \beta) = ER(\alpha, \rho) - \beta\varepsilon(f, \rho)$$

The first term is the ER measure of polarization and the second term is a measurement error or lack of identification weighted by a free parameter β.

Diagrammatically an n-spike representation is equivalent to transforming the original Lorenz Curve into a piecewise linear Lorenz curve (with n pieces) (see figure bellow). In other words each individual in a given group is assumed to have the same income. Hence the minimal error term is obtain through the minimisation of the area between the original Lorenz curve and the piecewise linear representation. It is therefore immediate that:

$$\varepsilon(f, \rho^*) = G(f) - G(\rho^*)$$

where G(.) assigns the Gini coefficient to the distribution variable in its argument. ρ^* is the optimal n-spike representation that best approximates to the real distribution. Combining the previous equations:

$$P(f, \alpha, \beta) = ER(\alpha, \beta) - \beta\left[G(f) - G(\rho^*)\right]$$

Duclos-Esteban-Ray index (DER)

The following axioms that are satisfied by the DER index are based on a density with finite support (kernel), and symmetric reductions in dispersion that concentrate the density around its mean (squeezes).

Axiom 1: if a distribution is made up of a basic density, then a squeeze cannot increase polarization.

Axiom 2: if a symmetric distribution is composed by three basic densities then a squeeze in the outer densities should not reduce polarization.

Axiom 3: if we consider a symmetric distribution made up of four basic densities with disjoint supports, then a move of the center distributions towards their outer neighbours, while keeping the disjoint supports, should increase polarization.

Axiom 4: Given two distributions F and G, if P(F) ≥ P(G), being P(F) and P(G) the respective polarization indexes, it must be that P(αF) ≥ P(αG), where αF and αG represent a rescaled version of F and G.

The authors establish that a general polarization measure that respects the previous axioms must be proportional to:

$$P_\alpha(f) \equiv \iint f(x)^{1+\alpha}\, f(y)|y-x|dydx$$

where *f*(*y*) and *f*(*x*) denote the income (or other well-being measure) density function. The formula can be rewritten as

$$P_\alpha(F) = \int f(y)^\alpha\, g(y)dF(y)$$

where *F*(*y)* denotes the income distribution function, *g*(*y*) captures the "alienation" effect, and *f*(*y*)$^\alpha$ the "identification" effect.

If we have a sample of incomes with independent and identically distributed observations ranked from smallest to highest, the DER operational formula is:

$$P_\alpha(\hat{F}) = n^{-1}\sum_{i=1}^{n} \hat{f}(y_i)^\alpha \left[\hat{\mu} + \left(y_i\left(\overline{w}^{-1}\left(2\sum_{j=1}^{i} w_j - w_i \right) - 1 \right) - \overline{w}^{-1}\left(2\sum_{j=1}^{i-1} w_j y_j - w_i y_i \right)\right)\right]$$

where y_i is the *i*-th individual income, $\hat{\mu}$ is the sample mean, w_i is the weight of individual *i*, and $\overline{w} = \sum_{j=1}^{n} w_j$.

The function $\hat{f}(y_i)$ is a nonparametric kernel estimate of the income density, using a bandwidth that minimizes the mean square error of the estimator h^*. Duclos, Esteban and Ray (2004) provide other formulas that are easier to compute. The first can be used with normal distributions and will not exceed the *h** that minimizes the mean squared error by more than 5%.

$$h^* \cong 4.7 n^{-5} \sigma \alpha^{-1}$$

The second is for distributions with skewness greater than 6:

$$h^* \cong n^{-5} IQ \frac{(3.76 + 14.7\sigma_{\ln})}{(1 + 1.09 * 10^{-4}\sigma_{\ln})^{(7268+15323\alpha)}}$$

where *IQ* is the interquantile range, and σ_{ln} is the variance of log-income.