The Distribution of Wages: A Non-parametric Decomposition

by

Conchita D'Ambrosio*

Working Paper No. 284

October 1999

*Università Bocconi and New York University.

I owe to Wilbert van der Klaauw and Edward Wolff several valuable suggestions and comments. I also thank Roberto Artoni, Francesco Corielli, and Claudio Lucifora for guidance. I am indebted to Markus Jänti and Lars Osberg for useful indications.

Correspondence to: Istituto de Economia Politica, Università Bocconi, Via Gobbi 5, 20136 Milano, Italy.
E-mail: conchita.dambrosio@uni-bocconi.it

ABSTRACT

This paper presents a non-parametric procedure to analyze the effects of different factors on observed movements in any distribution. These effects are estimated by applying kernel density methods to weighted samples in order to obtain counterfactual distributions. The advantage of this approach is that it provides a direct means of investigating if these factors have an impact and where in the density they do so, and it offers a new decomposition method of within and between group components. The approach to the decomposition analysis applied in this paper differs from the classical one of additively decomposable inequality indexes. If the purpose of the analysis is to understand what determined the variation in relative inequality, there is no doubt that the decomposition of the indexes belonging to the generalized entropy family is the best method. If, instead, the aim is to monitor what factors modified the entire distribution, where precisely on the distribution these factors had an effect, and what determined the variation in the level of polarization observed, then that method is useless. The non-parametric method proposed is the one to use, but with one caveat: All the results assume that there are no general equilibrium effects. The paper contains summary statistics of the observed movements and of distance and divergence among the estimated and counterfactual distributions; an original modification of an index of polarization; and an application of the method to the Italian distribution of wages.

INTRODUCTION

The distribution of welfare has always been one of the main issues in economics. On the one hand theorists have been focused on developing measures satisfying appealing properties, on the other applied researchers have used these measures to analyze the welfare of different societies. The article by Lorenz (1905) represents a milestone in this process. By stating to "*plot along one axis cumulated per cents of the population from poorest to richest and along the other the per cent of the total wealth held by these per cents of the population*" Lorenz has offered a criteria to rank distributions. According to this criteria a distribution is preferred to another if it can be obtained by a Pigou-Dalton transfer from the latter. The worst possible distribution is one where there is only one individual who possesses everything, alternatively the best is when the total amount of resources is equally shared among the members of a given society.

Lorenz criteria has, in my opinion, two major drawbacks. First it offers a measure of inequality from the perspective of an impartial observer - an objective measure - which does not indicate how people of a society perceive the level of inequality; on the other hand, it fails to adequately distinguish between convergence to the global mean and clustering around local means.

The latter dissatisfaction has already motivated independent work by Wolfson (1994) and Esteban and Ray (1994) who have conceptualized the notion of polarization.

The goal of this paper is to proceed along those lines and to provide a method able to offer a clear picture of what has happened to any distribution and why. In particular a new decomposition technique of within and between group components is proposed. The approach to the decomposition analysis applied in this paper differs from the classical one of additively decomposable inequality indexes. If the purpose of the analysis is to understand what determined the variation in relative inequality there are no doubts that the decomposition of the indexes belonging to the generalized entropy family is the best one. If, instead, the aim is to monitor what are the factors that modified the entire distribution, where precisely on the distribution these factors had an effect and what determined the variation in the level of polarization observed then the previous method is useless. The non-parametric method proposed is the one to use but it has one caveat: all the results assume that there are no general equilibrium effects as I will explain.

The paper contains an application of the technique to Italian data. I will follow Esteban and Ray and assume that polarization is the result of each individual identifying with people of his own group and feeling alienated towards people of other groups. I will analyze the Italian distribution of wages, provide new summary measures of the observed movements, polarization indexes and measures of distance and divergence among distributions, and try to explain some of the causes of these movements. The technique used is non-parametric. Kernel density estimation methods will allow me to obtain an estimate of the wage distribution and its evolution through time for the whole population and for its subgroups without imposing any assumption about the distribution of the observed data. Counterfactual densities - i.e. what would the density of income have been in one year if workers characteristics - between group component - or the distribution of wages among workers with the same characteristics - within group component - had remained at the level of the previous year - will be estimated by applying the same methods to appropriately weighted samples. Summary statistics of the observed movements and of divergence and distance among the estimated and the counterfactual distributions will conclude. It will be there proposed an original modification of Esteban and Ray index of polarization.

The Italian distribution of wages represents a interesting topic for applying the decomposition. The early 1990's have been a period characterized by significant changes for Italy: the period that goes from 1989 to 1993 was a period of recession for the Italian economy; those years were accompanied by politics aiming to decrease public spending and improve the performance of the Italian economy; some important reforms took place - the main one being the agreements reached among unions, government and industrial employers of 1992-93 - that affected bargaining and determination of wages. Furthermore those years experienced a boom of a particular kind of contract for young workers (*contratti di formazione e lavoro*) implying that young workers, generally more qualified, were payed relatively less. The effects of these changes have been already analyzed in several studies but the main focus has there been on the dispersion and not on the changes on the entire distribution of wages.

The great disparities existing in Italy among its geographic areas are well known. These are characterized by differences in industrial development and composition - less development accompanied by firms with lower technology, smaller dimensions, lower per-worker productivity in the South - in unemployment rate - higher in the South especially among young - in composition of the population in terms of number of family components - more children in the South - in the average age of the population - younger in the South - in the average number of earners - less in the South and lower female participation rate in the South - implying that the effect of the reforms was probably not the same on all the areas. From an accurate look at the estimate of the distribution of the logarithm of wages and its regional decomposition some important facts can indeed be noticed:

1. inequality has increased from 1987 to 1995 as the distribution of wages for the whole country has become more disperse. The Gini coefficient of the distribution of logarithm of real hourly wages increased from 0.099 in 1987 to 0.112 in 1995. The increase in the dispersion of the density was not common to all the areas. Inequality did not increase in the northern area - the Gini index increased from 0.095 to 0.096 - slightly increased in the center - from 0.096 to 0.107 - dramatically increased in the southern area passing from 0.115 in 1987 to 0.142 in 1995;

2. while in 1987 the densities of the three regional areas are quite similar in shape and centered at the same wage level, in 1995 there has been a moving apart of the densities accompanied by a dramatic change in the shape of the distributions.

More than an increase in inequality it is more appropriate to describe the evolution of the wage distribution of those years as subject to an increase in polarization between the northern, the center and the southern areas (Figure 1) as the distributions of the geographic areas moved apart. The Gini coefficient of the whole Italian wage distribution fails to adequately distinguish between convergence to the global mean and clustering around local means.

There is hence the need of extending the economic analysis in order to monitor the evolution and explain what are the factors that caused the observed movements of the entire distribution and not only of some of its moments. This paper then aims to monitor in great detail the effects of

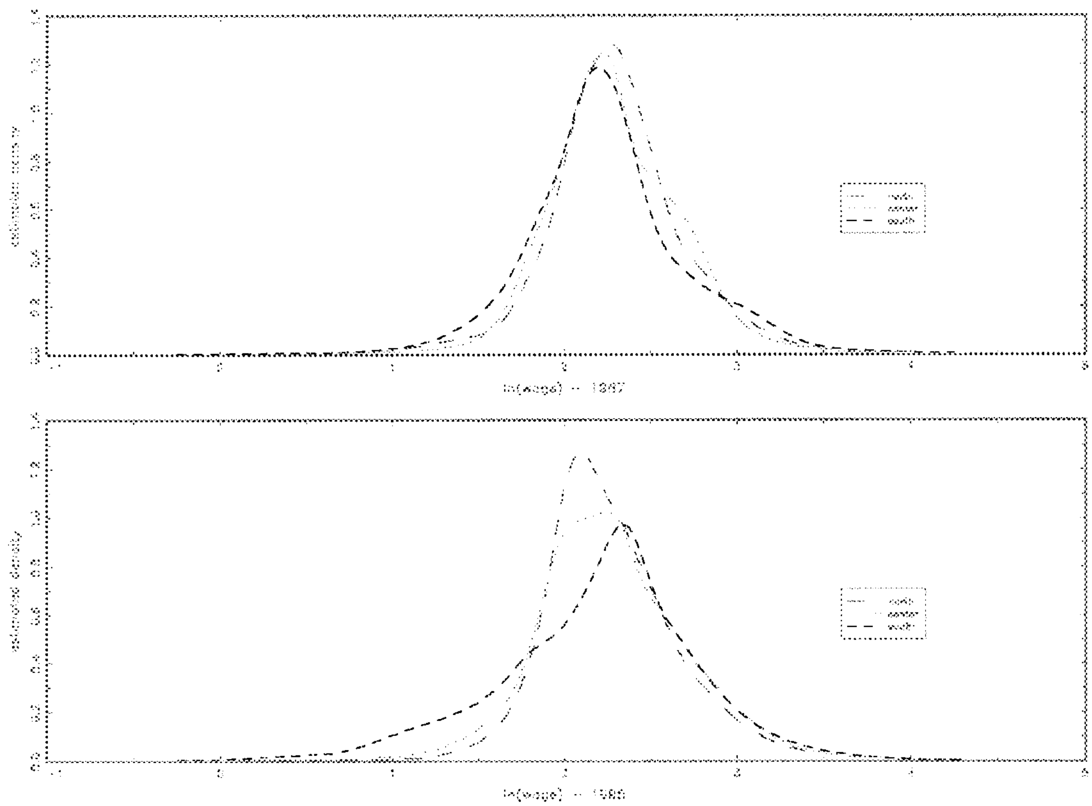

Figure 1: The distribution of the logarithm of real hourly wages.

the changes that occurred in Italy during the early 1990's on the shape of the distribution of hourly wages of the whole country. For the reasons I mentioned I decided to keep the analysis separate among the geographic areas. The next paragraphs contain a description of both the econometric technique used to estimate the densities and the counterfactual ones and of the summary indices of the observed movements, namely measures of distance and divergence and the polarization index proposed by Esteban and Ray and its modification that I propose. The results of the estimation are described in the last paragraph.

THE ESTIMATION METHOD

The main idea of the non-parametric methods for estimating the density function of wages is to let the data speak for itself. The estimate of the density function, $\widehat{f}(y)$, is determined directly from the data of the sample, y_1, y_2, ..., y_N, without assuming as a priori its functional form. The only assumption made is that there exists a density function $f(y)$ from which the sample is extracted.

The non-parametric method used in this work is optimally derived from a generalization of the kernel density estimator to take into account the sample weights attached to each observation, namely from the *adaptive or variable kernel.*

The adaptive kernel is built with a two stage procedure: a density is determined in the first stage in order to obtain the optimal bandwidth parameter; in the second stage the final density is computed. In detail the procedure is as follows:

1. Find a *pilot estimate*, $\widetilde{f}(y)$ such that $\widetilde{f}(y_i) > 0$ $\forall i$ defined as:

$$\widetilde{f}(y_i) = \tfrac{1}{Nh_N} \textstyle\sum_{i=1}^{N} K\left(\frac{y_j - y_i}{h_N}\right) \quad \forall y_j \tag{1}$$

where N is the number of observations of the sample, h_N is the bandwidth parameter and $K(.)$ is the kernel function. In this paper the kernel function that has been used is the normal.

It has been proven[1] that the final estimate is insensitive to the fine detail of the pilot estimate.

2. Define a local bandwidth factor[2] $\lambda(y_i)$:

$$\lambda(y_i) = \left\{ \frac{\widetilde{f}(y_i)}{g} \right\}^{-\frac{1}{2}} \tag{2}$$

where g is the geometric mean of $\widetilde{f}(y_i)$:

$$\log g = \frac{\sum_{i=1}^{N} \log \widetilde{f}(y_i)}{N} \tag{3}$$

The local bandwidth parameter for all y_i depends on the estimated density at y_i.

3. The final estimate is given by:

$$\widehat{f_a}(y_j) = \tfrac{1}{N} \textstyle\sum_{i=1}^{N} \frac{1}{h_N \lambda(y_i)} K\left(\frac{y_j - y_i}{h_N \lambda(y_i)}\right) \quad \forall y_j \tag{4}$$

where in addition to a global bandwidth parameter h_N a local one is included in the estimating procedure $\lambda(y_i)$.

[1] Silverman (1986) pag.101.

[2] As an alternative to the method applied here λ_i can be defined as $\lambda_i = \left\{ \frac{\widetilde{f}(y_i)}{g} \right\}^{-\alpha}$ where $\alpha \in [0, 1]$ is a sensitivity parameter generally set to $\frac{1}{d}$ and d is the dimension of the space where the density is estimated.

The adaptive kernel has been modified in this paper in order to take into account the sample weights, θ_i, associated to each observation. As a consequence every observation is weighted by θ_i and not by $\frac{1}{N}$ implying that the expressions used in (1) is:

$$\widehat{f}(y_j) = \sum_{i=1}^{N} \frac{\theta_i}{h_N} K\left(\frac{y_j - y_i}{h_N}\right) \quad \forall y_j \tag{5}$$

while in (4) is:

$$\widehat{f_a}(y_j) = \sum_{i=1}^{N} \frac{\theta_i}{h_N \lambda(y_i)} K\left(\frac{y_j - y_i}{h_N \lambda(y_i)}\right) \quad \forall y_j \tag{6}$$

where the sample weights are normalized in order to sum to one, $\sum_i \theta_i = 1$.

I estimate the density functions of the logarithm of wages for two different reasons:

1. the kernel estimator has some difficulties in dealing with densities that have a high degree of asymmetry. It is possible to show that the smallest $MISE$ depends on f through $R(f'') = \int f''(y)^2 \, dy$, which is a measure of the total curvature of f. The magnitude of this quantity gives an indication of how well f can be estimated even when h_N is chosen optimally. Hence for a density with high skewness, kurtosis, several modes $\left|f''(y)\right|$ will assume relatively high values implying a larger value of $R(f'')$. It has been shown[3] that the density $Beta(4,4)$ is the easiest to estimate and that the order among some densities is

[3] For the proof see Wand and Jones (1995).

the following:

$Beta(4,4)$
$Normal$
$Gamma(3)$
$Lognormal$

Densities close to normality appear to be easiest for the kernel estimator to estimate. Hence as the density of the level of wages resembles to a *Lognormal* its logarithm will be similar to a *Normal*.

2. I am interested in the movements over time of the distributions. These can be more easily detected by shrinking the tails present in the distribution of the level of wages.

The counterfactual densities are obtained by applying the kernel method to appropriately weighted samples. This technique has been derived from the one proposed by DiNardo, Fortin and Lemieux (1996).

Each observation is actually a vector (y, z, t) - composed of wage y, a vector z of workers characteristics and a date t - belonging to a joint distribution $F(y, z, t)$. The marginal density of wages at one point in time, $f^t(y)$ can be obtained by integrating the density of wages conditional on a set of workers characteristics and on a date t_y, $f(y \mid z, t_y)$, over the distribution of workers characteristics $F(z \mid t_z)$ at the date t_z:

$$\begin{aligned} f^t(y) &= \int_{z\in\Omega_z} dF(y, z \mid t_{y,z} = t) \\ &= \int_{z\in\Omega_z} f(y \mid z, t_y = t)\, dF(z \mid t_z = t) \\ &\equiv f(y \mid t_y = t, t_z = t) \end{aligned} \tag{7}$$

where Ω_z is the domain of definition of workers characteristics.

Two different counterfactual densities can be obtained form (7): the counterfactual density of wages at date t_1 and characteristics at date t_2, represented by $f\left(y \mid t_y = t_1, t_z = t_2\right)$:

$$\begin{aligned} & f\left(y \mid t_y = t_1, t_z = t_2\right) \\ & = \int_{z \in \Omega_z} dF\left(y, z \mid t_y = t_1, t_z = t_2\right) \\ & = \int_{z \in \Omega_z} f\left(y \mid z, t_y = t_1, t_z = t_2\right) dF\left(z \mid t_y = t_1, t_z = t_2\right) \end{aligned} \tag{8}$$

and analogously the counterfactual density of wages at date t_2 and characteristics at date t_1.

Under the assumption that the structure of wages conditional on the distribution of workers characteristics does not depend on the time of the workers characteristics:

$$f\left(y \mid z, t_y = t_1, t_z = t_2\right) = f\left(y \mid z, t_y = t_1, t_z = t_1\right) \tag{9}$$

the counterfactual density of wages at date t_1 and characteristics at date t_2 is:

$$f\left(y \mid t_y = t_1, t_z = t_2\right) \;\; = \int_{z \in \Omega_z} f\left(y \mid z, t_y = t_1\right) dF\left(z \mid t_z = t_2\right) \tag{10}$$

This counterfactual density represents the distribution of wages that would have prevailed in year t_1 if the distribution of workers characteristics had remained at the level of the year t_2. It would be more appropriate to say that the counterfactual density indicate the density that would have prevailed if workers characteristics had remained at their t_2 level and the workers wage distribution would have been the same as observed in t_1 for workers with the same characteristic since general equilibrium effects are excluded from the analysis as the effects of changes in the distribution of z on the structure of wages are not taken into account. What I estimate is indeed the effect of the movements between groups on the total density of wages assuming that the distributions within each group do not change over time.

While assuming that:

$$f(y \mid z, t_y = t_2, t_z = t_1) = f(y \mid z, t_y = t_2, t_z = t_2) \tag{11}$$

the counterfactual density of wages at date t_2 and characteristics at date t_1 is:

$$f(y \mid t_y = t_2, t_z = t_1) \quad = \int_{z \in \Omega_z} f(y \mid z, t_y = t_2)\, dF(z \mid t_z = t_1) \tag{12}$$

This counterfactual density picks the within group component of the observed movements by estimating the effect of changes in the distribution of wages among individuals with the same characteristic on the distribution of wages for the whole population assuming that workers characteristic would not change over time.

The difference between the actual density and the counterfactual one represents the effects on one side of the changes in the distribution of the characteristics of the workers - between group component - and on the other side of changes of the wage structure of workers with given characteristics - within group component.

It is clear from equations (10) and (12) that the counterfactual densities can be obtained by estimating[4], non-parametrically, the component densities:

- $f(y \mid z, t_y = t_i)$ is estimated by applying the kernel method to the appropriate sample in year t_i;

- $F(z \mid t_z = t_i)$ is estimated non parametrically as proportion of workers with given characteristics in year t_i;

[4]An alternative estimation method for the counterfactual density of wage at date t_1 and characteristics at date t_2 is proposed by DiNardo et al. (1996).

THE DECOMPOSITION METHOD

For simplicity of the notation in what follows I will rewrite (7) for y being a discrete random variable:

$$\begin{aligned} f^t(y) &= \int_{z\in\Omega_z} dF(y,z \mid t_{y,z}=t) \\ &= \int_{z\in\Omega_z} f(y \mid z, t_y = t)\, dF(z \mid t_z = t) \\ &= \sum_z \alpha_z^t(y)\, f_z^t(y) \end{aligned} \tag{13}$$

where $\alpha_z^t(y) = f(z \mid t_z = t)$ - the proportion of workers in each group - and $f_z^t(y) = f(y \mid z, t_y = t)$ - the density of wages within each group. The total density of wages, $f^t(y)$, can change over time both because there is a movement of workers between the groups - $\alpha_z^t(y)$'s change - and because the structure of wages within each group changes - $f_z^t(y)$'s vary. The variation in $f(y)$ going from t_1 to t_2 is:

$$f^{t_2} - f^{t_1} \simeq df(t)\mid_{t=t_1} = f'(t)\,dt\mid_{t=t_1} \tag{14}$$

From (8):

$$f'(t) = \sum_z \alpha_z'(t)\, f_z(t) + \sum_z \alpha_z(t)\, f_z'(t) \tag{15}$$

hence (14) is given by:

$$f'(t)\,dt\mid_{t=t_1} = \sum_z \alpha_z'(t)\, f_z(t)\,dt\mid_{t=t_1} + \sum_z \alpha_z(t)\, f_z'(t)\,dt\mid_{t=t_1} \tag{16}$$

I can approximate the following components of (16), under the assumption that both $\alpha_z(t)$ and $f_z(t)$ are linear in $[t_1, t_2]$, by:

$$\begin{aligned} \alpha_z'(t)\,dt &\simeq \alpha_z(t_2) - \alpha_z(t_1) \\ f_z'(t)\,dt &\simeq f_z(t_2) - f_z(t_1) \end{aligned} \tag{17}$$

Hence the variation in f is approximately given by:

$$
\begin{aligned}
& f^{t_2} - f^{t_1} \\
\simeq & \textstyle\sum_z \left[\alpha_z(t_2) - \alpha_z(t_1)\right] f_z(t) \mid_{t=t_1} + \sum_z \alpha_z(t)\left[f_z(t_2) - f_z(t_1)\right] \mid_{t=t_1} \\
= & \underbrace{\left\{\sum_z \left[\alpha_z(t_2) f_z(t_1)\right] - \sum_z \left[\alpha_z(t_1) f_z(t_1)\right]\right\}}_{\text{between group}} + \\
& \underbrace{\left\{\sum_z \left[\alpha_z(t_1) f_z(t_2)\right] - \sum_z \left[\alpha_z(t_1) f_z(t_1)\right]\right\}}_{\text{within group}}
\end{aligned}
\tag{18}
$$

Each component of (18) can be estimated with the non-parametric method as explained in the previous paragraph.

SUMMARY INDICES

The coefficients needed to summarize the observed movements are of two kinds. First an index is needed to summarize how much any two given densities are different between them: coefficients of distance and divergence; second a class of indexes has to register the moving apart of some densities classified according to the geographic area where the worker resides: the polarization index.

Several coefficients have been suggested in the statistical literature for measuring distance and divergence between probability distributions. The approach chosen in this work follows Ali and Silvey (1966).

Two probability distributions F_1 and F_2 on the real line, with corresponding densities f_1 and f_2, are given, being absolutely continuous with respect to Lebesgue measure and with respect to each other. The measures computed belong to a general class based on the ratio of the densities:

$$\phi(y) = \frac{f_2(y)}{f_1(y)} \tag{19}$$

If F_1 and F_2 are the same then $\phi(y) \equiv 1$. As F_1 and F_2 move apart $\phi(y)$ takes larger values on a set of decreasing F_1-probability and increasing F_2-probability and smaller values on a set of increasing F_1-probability and decreasing F_2-probability. By looking at the expectation of $\phi(y)$ with respect to F_1 - $E_1(\phi)$ - it can be noticed that $E_1(\phi) = 1$ for all F_1 and F_2 hence the coefficient of the F_1-dispersion of ϕ could be a measure of divergence of F_2 from F_1 as it would increase as F_1 and F_2 move apart. The form of the coefficient of divergence that is proposed is based on these intuitions. Ali and Silvey, indeed, state four properties that a coefficient of divergence should satisfy and prove that those are met by any coefficient of the form[5]:

$$E\{C(\phi)\} \tag{20}$$

where C is a continuous convex function on $(0, \infty)$. Notice that the expectation of a convex function of a real random variable measures its dispersion to a greater or lesser extent depending on the nature of this function. Hence depending on the specification of the convex function different measures are obtained.

1. When $E\{C(\phi)\} = E\{(\phi - 1)\log\phi\}$ the measure is the Jeffreys measure of divergence:

$$J(1,2) = \int (f_2(y) - f_1(y)) \log \frac{f_2(y)}{f_1(y)} dy \tag{21}$$

[5]The expectation that is considered in Ali and Silvey is really a generalized expectation, E^*, that is defined even if $\phi = \infty$. For simplicity I avoid this notation but it is worthwile noticing that everything holds even in the case where $\phi = \infty$.

2. For $E\{C(\phi)\} = E\{-\log\phi\}$ and $E\{C(\phi)\} = E\{\phi\log\phi\}$ the measures are the Kullback and Leibler measures of discriminatory information $I(1,2)$ and $I(2,1)$ respectively:

$$\begin{aligned} I(1,2) &= \int f_1(y)\log\tfrac{f_1(y)}{f_2(y)}dy \\ I(2,1) &= \int f_2(y)\log\tfrac{f_2(y)}{f_1(y)}dy \end{aligned} \tag{22}$$

   Jeffrey and Kullback and Leibler measures are based on the Shannon-Wiener definition of information: two populations differ more or less according to how difficult it is to discriminate between them with the best test. The next measures analyzed, the Kolmogorov ones, are measures of distance and differ from the measures of divergence due to symmetry. Indeed:

3. When $\{C(\phi)\} = \frac{1}{2}E\left(\sqrt{\phi}-1\right)^2$ the measure is the Kolmogorov measure of distance, namely:

$$Ko = \frac{1}{2}\int\left(\sqrt{f_2(y)}-\sqrt{f_1(y)}\right)^2 dy \tag{23}$$

4. For $\{C(\phi)\} = \frac{1}{2}E\left|\phi-1\right|$ the measure is the Kolmogorov measure of variation distance:

$$Kov = \frac{1}{2}\int\left|f_2(y)-f_1(y)\right|dy \tag{24}$$

As far as the second class of summary indices is concerned the index of polarization[6] computed is the one suggested by Esteban and Ray (1994) and a modification that I propose.

[6]I could not apply Wolfson's measure of polarization as it is a measure of bipolarization and I was interested in monitoring the movements of the distributions of three groups composed by the regional areas of Italy.

Esteban and Ray introduce a model of individual attitudes in a society and place four axioms to narrow down the set of allowable measures. The notation is the following: $(\eta, y) \equiv (\eta_1, ..., \eta_N; y_1, ..., y_N)$ is a distribution for any positive integer N if $y \in \mathbf{R}^N$, $y_i \neq y_j$ $\forall i, j$ and $\eta > 0$. The total population associated with (η, y) is given by $\sum_{i=1}^{N} \eta_i$. Φ is the space of all distributions. A polarization measure is a mapping $\mathbf{ER} : \Phi \rightarrow \mathbf{R}_+$. In particular Esteban and Ray suppose that each individual is subject to two forces: he identifies with those he considers to be members of his own group - $I : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ represents the identification function - on the contrary, he feels alienated with those he considers to be members of other groups - $a : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ is the alienation function and the individual with wage[7] y feels alienation $a(\delta(y, y'))$ with an individual with wage y'. $\delta(y, y')$ is a measure of distance among the two wages and for Esteban and Ray it is simply the absolute distance $|y - y'|$. The joint effect of the two forces is given by the effective antagonism function, $T(I, a)$ and total polarization in the society is postulated to be the sum of all the effective antagonisms:

$$\mathbf{ER}(\boldsymbol{\eta}, \mathbf{y}) = \sum_{i=1}^{N} \sum_{j=1}^{N} \eta_i^{1+\alpha} \eta_j T\left(I\left(\eta_i\right), a\left(\delta\left(y, y'\right)\right)\right) \tag{25}$$

The measure that satisfies the axioms placed by Esteban and Ray has the following expression:

$$\mathbf{ER}(\boldsymbol{\eta}, \mathbf{y}) = K \sum_{i=1}^{N} \sum_{j=1}^{N} \eta_i^{1+\alpha} \eta_j \delta\left(y, y'\right) = K \sum_{i=1}^{N} \sum_{j=1}^{N} \eta_i^{1+\alpha} \eta_j \left|y_i - y_j\right| \tag{26}$$

for some constants $K > 0$ and $\alpha [1, 1.6]$ that indicates the degree of sensitivity to polarization[8].

[7] Esteban and Ray original index is for the distribution of income.

[8] With $\alpha = 0$ Esteban and Ray index of polarization is proportional to the Gini coefficient normalized using the logarithm of income and not the mean.

This index of polarization is computed empirically as follows:

$$\mathbf{ER}(\alpha) = \sum_{i=1}^{N}\sum_{j=1}^{N} \pi_i^{1+\alpha}\pi_j \left|\mu_i - \mu_j\right| \tag{27}$$

π_i and μ_i represent respectively the relative frequency[9] and the conditional mean in group i for a density of the logarithm of wages $f(y)$, namely:

$$\begin{aligned} \pi_i &= \int_{y_{i-1}}^{y_i} f(y)\,dy \\ \mu_i &= \frac{1}{\pi_i}\int_{y_{i-1}}^{y_i} yf(y)\,dy \end{aligned} \tag{28}$$

In other words what is computed empirically is the degree of polarization in a society where it is assumed that everybody in each given group earns a wage equal to the mean of the group.

I propose a modification[10] of **ER** to compute the level of polarization of a given society without assuming that everybody in each group has a wage equal to the mean of the group and by looking at another characteristic, other than wages, that forms the groups, e.g. region of residence, age, education, industry.

[9]The population weight η_i, $i = 1, ..., N$ are replaced by the population frequencies. The constant K is hence set to $K = \left[\sum_{i=1}^{N}\eta_i\right]^{-(2+\alpha)}$.:

[10]Esteban, Gradin and Ray (1998) and Gradin (1999) have already proposed a modification of **ER** (**P**) to take into account the error of not having included in the analysis the inequality within each group and the overlapping of the groups that has the effect of overestimating the level of observed polarization. In particular:

$$\mathbf{P}(\alpha,\beta) = \mathbf{ER}(\alpha) - \beta\varepsilon \tag{29}$$

where:

$$\varepsilon = G(f) - G(\mu) \tag{30}$$

the difference between the Gini coefficient computed on the ungrouped, $G(f)$, and grouped data, $G(\mu)$. β is the parameter that indicates the importance given to the approximation error.

The idea is a direct application of the method described in the previous paragraphs. The total density of wages, $f^t(y)$, at any point in time, is given by the sum of the densities of each group, weighted by the relative frequency of each group:

$$\begin{aligned} f^t(y) &= \int_{z \in \Omega_z} dF(y, z \mid t_{y,z} = t) \\ &= \int_{z \in \Omega_z} f(y \mid z, t_y = t)\, dF(z \mid t_z = t) \end{aligned} \tag{31}$$

The polarization index has to register the moving apart of the densities classified according to some characteristics of the workers that forms the groups and changes in the frequencies between the groups. Each individual identifies with those of his own group and feels alienated with those he considers to be members of other groups, as Esteban and Ray noted, but now the groups are made by characteristics and not levels of wage. Hence the index of polarization that Esteban and Ray proposed will be modified in order to take into account the distance among distributions of wages between each group. I propose to use as measure of distance among two distributions the Kolmogorov measure of variation distance:

$$Kov_{ij} = \frac{1}{2} \int |f_i(y) - f_j(y)|\, dy \tag{32}$$

and compute the following polarization index obtained from (26) :

$$\mathbf{PK}(\boldsymbol{\alpha}) = \sum_{i=1}^{N} \sum_{j=1}^{N} \pi_i^{1+\alpha} \pi_j Kov_{ij} \tag{33}$$

THE RESULTS

The estimation of the distribution of the logarithm of hourly wages is obtained applying the non-parametric method described earlier to the survey collected by the Bank of Italy, SHIW, of the years 1987 and 1995. In the

1987 survey 6816 workers were interviewed with 3341, 1354 and 2121 of them residing respectively in the North, Center and South of Italy; the correspondent values for 1995 are 6448 workers, 3074 residing in the North, 1360 in the Center and 2014 in the South areas. The definition of wage analyzed is hourly wages computed from yearly net wages and non monetary integrations, hours worked in a week and weeks worked in a year in thousands of lira. Cannari and Gavosto (1994) and Brandolini and Cannari (1994) analyzed the quality of these data and reported that this is the same as the corresponding surveys in other countries. All wages are expressed in real terms by correcting for inflation using CPI (base 1990).

The choice of the period of analysis is determined by the willing of analyzing the effects of the changes that occurred in Italy during the early 1990's[11]on the shape of the distribution of hourly wages of the whole country. During the first half of 1990's Italy has indeed undergone a period of significant changes: the period that goes from 1989 to 1993 was a period of recession for the Italian economy; those years were accompanied by politics aiming to decrease public spending and improve the performance of the Italian economy; some important reforms took place - the main ones being the agreements reached among unions, government and industrial employers of 1992-93 - that affected bargaining and determination of wages. In particular the agreements brought to thc abolishment of the wage indexation mechanism. Since then the settings of wages had to take into account the expected inflation rate as set by the government[12]. Furthermore those years

[11]The last year available of SHIW is 1995, hence the chioce of the ending period.

[12]The consequence of this indexation system was that in those years the forecasts were always smaller than the actual inflation rates causing some workers to experience a real loss in terms of their wage.

experienced a boom of a particular kind of contract for young workers (*contratti di formazione e lavoro*) implying that young workers, generally more qualified - at least high school degree - were payed relatively less. The effects of these changes have been already analyzed in several studies but the main focus has there been on the dispersion of wages and not on the changes on the entire distribution of wages. For the great disparities existing among Italian geographical areas the analysis has been conducted separately for the North, the Center and the South of Italy.

To understand the causes of the observed movement in the aggregate distribution of wages I decompose the total working population into different subgroups according to the following classification[13]:

1. education: the worker has no schooling degree, elementary, junior high, high school degree, laurea or more;

2. age: the worker is between 14 and 20 years old, between 21 and 40, between 41 and 50, between 51 and 65, >65.

3. industry: the worker works in agriculture, industry, services, public administration;

4. occupation: the worker is blue-collar, white collar or teacher, manager or judge or university teacher or elected member.

The distribution of the logarithm of wages for the North of Italy from 1987 to 1995 is characterized by a clear shift of mass towards lower levels of wages as shown in Figure 2. In the lower graph the difference among

[13]Information on industry and occupation is not available for all employees. Hence the sample for these classifications is smaller.

the two distribution is plotted: the density shifted from the interval 2.2 - 3 to 1.6 - 2.2. Only in the lower end of the left tail the density of 1995 is again lower than those of 1987. The mode of the distribution moved left, the mean decreased from 2.297 to 2.271 while inequality as measured by the Gini coefficient did not change - the values being 0.095 in 1987 and 0.096 in 1995.

North	**Jeffreys divergence**	**Kullback-Leibler divergence**
actual	0.0798	0.0375
education between	0.0932 (+16.714)	0.0441 (+17.561)
education within	0.0229 (-71.296)	0.0115 (-69.400)
age between	0.0678 (-15.058)	0.0316 (-15.881)
age within	0.0009 (-98.823)	0.0004 (-98.803)
industry between	0.1328 (+66.438)	0.0646 (+72.020)
industry within	0.0226 (-71.742)	0.0107 (-71.366)
occupation between	0.0726 (-8.997)	0.0337 (-10.236)
occupation within	0.0652 (-18.266)	0.0295 (-21.528)

table 34: measures of divergence between the actual distribution of 1987 and the actual/counterfactual distribution of 1995. North of Italy.
In brackets is the percentage of the change of estimated value with respect to the one computed on the actual densities.

(34)

North	Kolmogorov distance	Kolmogorov variation distance
actual	4.4087	0.0952
education between	5.0461 (+14.458)	0.1063 (+11.636)
education within	0.7826 (-82.248)	0.0507 (-46.767)
age between	3.4363 (-22.056)	0.0844 (-11.295)
age within	0.0197 (-99.553)	0.0082 (-91.363)
industry between	7.3008 (+65.600)	0.1326 (+39.275)
industry within	1.0556 (-76.055)	0.0547 (-42.487)
occupation between	3.3422 (-24.190)	0.0826 (-13.268)
occupation within	0.8379 (-80.994)	0.0556 (-41.6181)

table 35: measures of distance between the actual distribution of 1987 and the actual/counterfactual distribution of 1995. North of Italy. In brackets is the percentage of the change of estimated value with respect to the one computed on the actual densities.

(35)

The decomposition of the observed movement among between group and within group are described in Fig. 3. According to the decomposition based on education and age the effect of the changes within groups, i.e. the modification in the way groups are payed, is able to explain almost all the variation occurred in the density during those years. The counterfactual density obtained by substituting the conditional densities of 1987 for the different age groups into the density of 1995 is almost coincident with the actual density of 1987, as evident from graph 4. In this case the difference among the counterfactual and the actual density - lowest graph on the left - shrinks to zero. For the grouping based on education the difference among the counterfactual obtained by considering the within group effect for 1995 and the actual distribution of 1987 decreases but to a lower extent - graph 2. For industry and occupation the evidence is more mixed: the variation in the density of wages is due to changes within groups but movements between groups have more impact. Regarding education the groups that count the most in determining the total density are those of workers with junior high and high school degree as in both years they represents around 80% of the population[14]. Going from 1987 to 1995 there has been a convergence of the distributions of all groups towards the same level of wages but the one composed of workers with laurea or more who constitute a separate group. This convergence implied a loss in terms of high wages particularly for workers with junior high and high school degree that in 1987 were able to earn relatively more. As far as age is concerned the groups that exert the biggest impact on the total density of wages are those composed of workers between 21 - 40 years old that represents around 56% of the total

[14]The graphs of the densities of the groups classified according to education and age are here omitted but are available upon request.

population. The central mass of the density of this group moved in 1995 completely towards lower levels of wages, with the same shift observed in the aggregate distribution.

The values of the measures of divergence and distance between the distribution of 1987 on one side and the distribution of 1995 - actual and counterfactual - on the other are reported in tables (34) and (35) where in brackets is the percentage of the change of estimated value with respect to the one computed on the actual densities. All measures report a reduction in distance and divergence between the distribution of 1987 and the counterfactual of 1995 according to the decomposition based on education, age, industry and occupation - movements within groups - on age and occupation - movements between groups. All measures reports the greatest reduction - 98.8% according to the measures of divergence, 99.5% for Kolmogorov measure of distance and 91.4% according to Kolmogorov measure of variation distance - among the counterfactual density obtained by substituting the conditional densities of 1987 for the different age groups into the density of 1995, i.e. the within group effect, and the actual density of 1987, confirming what was previously noticed. The measures of divergence attribute a reduction of 71% due to the effect of changes within industries while the second biggest effect is attributed by the measures of distance to the effect of changes within educational groups - 82.3% according to Kolmogorov measure of distance and 46.8% according to Kolmogorov measure of variation distance.

Center	**Jeffreys divergence**	**Kullback-Leibler divergence**
actual	0.0627	0.0310
education between	0.0838 (+33.616)	0.0428 (+37.943)
education within	0.0009 (-98.491)	0.0003 (-98.963)
age between	0.0754 (+20.194)	0.0381 (+22.965)
age within	0.0031 (-95.128)	0.0014 (-95.423)
industry between	0.1443 (+130.011)	0.0765 (+146.603)
industry within	0.0326 (-48.103)	0.0158 (-49.188)
occupation between	0.0637 (+1.578)	0.0334 (+7.543)
occupation within	0.0573 (-8.750)	0.0281 (-9.293)

table 36: measures of divergence between the actual distribution of 1987 and the actual/counterfactual distribution of 1995. Center of Italy. In brackets is the percentage of the change of estimated value with respect to the one computed on the actual densities.

(36)

Center	**Kolmogorov distance**	**Kolmogorov variation distance**
actual	1.7115	0.0788
education between	2.0437 (+19.413)	0.0951 (+20.687)
education within	0.0432 (-97.478)	0.0104 (-86.778)
age between	1.7864 (+4.378)	0.0846 (+7.401)
age within	0.0353 (-97.936)	0.0127 (-83.888)
industry between	5.0672 (+196.066)	0.1342 (+70.350)
industry within	1.4620 (-14.578)	0.0592 (-24.817)
occupation between	1.7407 (+1.706)	0.0828 (+5.123)
occupation within	1.2198 (-28.727)	0.0709 (-9.967)

table 37: measures of distance between the actual distribution of 1987 and the actual/counterfactual distribution of 1995. Center of Italy. In brackets is the percentage of the change of estimated value with respect to the one computed on the actual densities.

(37)

The results of the estimation for the Center of the Italy are described in Fig. 4 and Fig. 5. The actual density of wages undergone a process of flattening down from 1987 to 1995 by loosing mass from the center towards the two tails. This shift of density was not symmetric: more mass went on the left tail than on the right one, as evident from the lowest graph in Fig. 4. Both the mode and the mean of the distribution did not change while inequality as measured by the Gini coefficient increased from 0.096 to 0.107. The results of the decomposition among the contribution to the movements observed in the aggregate density of the changes that occurred in the distributions of wages within each group - the within group component - and the shifts of workers between groups - the between group component - are very similar to what was reported regarding the North of the country. For the groups based on education and age all the effect can be attributed to the within group movements but for this geographic area the effect of the movements within groups based on education is higher than in the North. For the Center, indeed, both counterfactual densities obtained by substituting the conditional densities of 1987 for the different age and educational groups into the density of 1995 are surprisingly coincident with the actual density of 1987, as evident from graph 2 for education and from graph 4 for age in Fig. 5. In both cases the difference among the counterfactual and the actual density - lowest graph on the left of both figures - shrinks to zero. Once again for industry and occupation the evidence is more mixed: the changes can be attributed to both within and between groups movements even if the within components are able to reduce the difference among the distributions to a greater extent - graph 6, Fig. 5, for industry and graph 8, Fig. 5, for occupation. For education the groups that count the most in determining the total density are those of workers with junior high and high school degree as in both years they represents around 79% of the population[15]. The evolution of the densities of the different educational groups from 1987 to 1995 in the Center resembles to that of the correspondent ones of the North but the convergence of the densities - all but the one composed of worker with laurea or more - towards the same levels of wage is less evident. There is now greater dispersions within the density of each group especially among those with junior high and high school degree. This dispersion increases in 1995 causing the observed flattening down of the aggregate density. For the grouping based on age, the contribution to the total density is around 51% for the group 21-40 and around 26% for 41-50 one. Going from 1987 to 1995 the latter moved towards higher level of wages while the former became less dispersed and concentrate towards lower level of wages. The movement in the aggregate densities for this decomposition are hence attributed to the increase polarization occurred among those two age groups.

The measures of divergence and distance are reported in tables (36) and (37). All measures agree in attributing some impact to education, age, industry and occupation - within group movements. The measures of divergence attribute the greatest effect to education - within group - which reduces the divergence among the counterfactual distributions of 1995 and the actual one of 1987 of 98% according to Jeffreys measure of divergence and 99% according to Kullback-Leibler measure of divergence. The second biggest effect is due to changes in age within group component with an effect of 95% for both measures. The measures of distance, instead, to not agree in the magnitude of the effects: according to Kolmogorov measure of distance education within-group component shrinks the distance of 97% while

[15]The graphs of the densities of the groups classified according to education and age are here omitted but are available upon request.

this value increases to 98% for age within group; The Kolmogorov measure of variation distance, instead, attributes 87% of the changes to education within group component while only 84% to age within group shifts.

South	Jeffreys divergence	Kullback-Leibler divergence
actual	0.2303	0.1343
education between	0.2339 (+1.541)	0.1376 (+2.453)
education within	0.0061 (-97.348)	0.0023 (-98.262)
age between	0.2301 (-0.104)	0.1351 (+0.583)
age within	0.0020 (-99.119)	0.0008 (-99.429)
industry between	0.3977 (+72.692)	0.2403 (+78.966)
industry within	0.0700 (-70.048)	0.0332 (-75.298)
occupation between	0.2362 (+2.540)	0.1316 (-2.009)
occupation within	0.1359 (-40.974)	0.0596 (-55.637)

table 38: measures of divergence between the actual distribution of 1987 and the actual/counterfactual distribution of 1995. South of Italy.
In brackets is the percentage of the change of estimated value with respect to the one computed on the actual densities.

(38)

South	Kolmogorov distance	Kolmogorov variation distance
actual	7.8866	0.1467
education between	7.5098 (-4.778)	0.1478 (+0.752)
education within	0.2516 (-96.810)	0.0283 (-80.720)
age between	7.5468 (-4.308)	0.1448 (-1.266)
age within	0.0321 (-99.593)	0.0109 (-92.592)
industry between	8.3773 (+6.2219)	0.1803 (+22.884)
industry within	2.7684 (-64.897)	0.0908 (-38.099)
occupation between	9.1847 (+16.459)	0.1584 (+7.998)
occupation within	2.3508 (-70.192)	0.0911 (-37.920)

table 39: measures of distance between the actual distribution of 1987 and the actual/counterfactual distribution of 1995. South of Italy. In brackets is the percentage of the change of estimated value with respect to the one computed on the actual densities.

(39)

The distribution of the logarithm of wages for the South of Italy from 1987 to 1995 is characterized by a clear increase in the mass on the left tail, increasing the skweness of the distribution and by an increase of the density towards higher levels of wages, as shown in Fig. 6. In the lower graph the difference among the two distribution is plotted: the density shifted from the interval 1.8 - 2.3 towards 0 - 1.8 and 2.3 - 3. The mode of the distribution increased, the mean decreased from 2.24 to 2.20 and inequality, as measured by the Gini coefficient, increased from 0.115 to 0.142, the biggest increase among the geographic areas analyzed.

The decomposition of the observed movement among between group and within group are described in Fig. 7. Once again the decomposition based on education and age attribute almost all the effects to the movements within groups. The counterfactual density obtained by substituting the conditional densities of 1987 for the different educational and age groups into the density of 1995 almost reproduce the actual density of 1987, as evident from graph 2 and graph 4. The difference among the counterfactual and the actual density - lowest graph on the left - shrinks, once again, to zero. For the grouping based on education the difference among the counterfactual obtained by considering the within group effect for 1995 and the actual distribution of 1987 decreases but to a lower extent - graph 2 - while in the case of age the difference is almost null. For industry and occupation the evidence is more mixed: the variation in the density of wages is due to changes within groups but movements between groups have more impact, as it was the case for all the other areas. For education the groups that count the most in determining the total density are those of workers with junior high and high school degree as in both years they represents around 63% of the population[16]. The evolution of the densities of the different educational groups from 1987 to 1995 in the South differs dramatically from what was observed in the other areas. Indeed during the years of analysis the densities ordered on the wage scale in an increasing order depending on the diploma attained, the opposite of the convergence phenomena observed elsewhere. The densities of the groups were more similar in shape 1987 than in 1995. The movement observed in the aggregate density is the joint effect of the transformations of the density of workers with junior high and high school degree that were both subject to an increase in the mode and a dramatic increase in the mass left of it especially for those with junior high degree. For the grouping based on age, the contribution to the total density is around 50% for the group 21-40 and around 26% for 41-50 one. Going from 1987 to 1995 both distributions became almost bimodal, the latter moved towards higher level while the distribution of the former towards lower levels of wages. The movement in the aggregate densities for this decomposition are hence attributed to the increase in the inner dispersion occurred among those age groups and to the increase in the level polarization among them.

All measures of divergence and distance - tables (38) and (39) - report a reduction in distance and divergence between the distribution of 1987 and the counterfactual of 1995 according to the decomposition based on education, age, industry and occupation - movements within group. They all agree in attributing the greatest reduction - 97% and 98% according respectively to the Jeffreys and Kullback-Leibler measures of divergence, 99.5% for Kolmogorov measure of distance and 92.6% according to Kolmogorov measure of variation distance - among the counterfactual density obtained

[16]The graphs of the densities of the groups classified according to education and age are here omitted but are available upon request.

by substituting the conditional densities of 1987 for the different age groups into the density of 1995, i.e. the within group effect, and the actual density of 1987, confirming what was previously noticed. The modification of the aggregate density due to the effect of movements within educational groups is 97%, 98%, 96% and 80% according respectively to Jeffreys, Kullback-Leibler measures of divergence, Kolmogorov measure of distance and Kolmogorov measure of variation distance.

Mean	**Total**	**North**	**Center**	**South**
estimated 1987	2.280	2.297	2.279	2.245
estimated 1995	2.252	2.271	2.279	2.202
education between	2.251	2.274	2.256	2.213
education within	2.308	2.343	2.293	2.264
age between	2.257	2.288	2.267	2.202
age within	2.308	2.304	2.290	2.247
industry between	2.181	2.237	2.207	2.075
industry within	2.286	2.307	2.295	2.245
occupation between	2.279	2.305	2.272	2.241
occupation within	2.225	2.259	2.235	2.164

(40)

table 40: means of the actual distributions of 1987 and the actual/counterfactual distribution of 1995.

Gini	**Total**	**North**	**Center**	**South**
estimated 1987	0.099	0.095	0.096	0.115
estimated 1995	0.112	0.096	0.107	0.142
education between	0.118	0.099	0.111	0.150
education within	0.102	0.096	0.095	0.112
age between	0.116	0.098	0.111	0.146
age within	0.102	0.094	0.097	0.115
industry between	0.120	0.098	0.112	0.158
industry within	0.090	0.087	0.086	0.098
occupation between	0.114	0.098	0.108	0.143
occupation within	0.090	0.084	0.089	0.098

(41)

table 41: Gini coefficients of the actual distributions of 1987 and the actual/counterfactual distribution of 1995.

In tables (40) and (41) are reported the means and the Gini coefficients of the estimated distributions.

The mean of the wage distribution of the whole country decrease from 2.28 to 2.25. This reduction is confirmed in the reductions of the mean of two of the regional densities, the North and the South. As far as inequality is concerned, the Gini coefficient of the distribution of logarithm of real hourly wages increased from 0.099 in 1987 to 0.112 in 1995. The increase in the dispersion of the density was not common to all the areas. Inequality did not increase in the northern area - the Gini index increased from 0.095 to 0.096 - slightly increased in the center - from 0.096 to 0.107 - dramatically increased in the southern area passing from 0.115 in 1987 to 0.142 in 1995. The Gini coefficients of the counterfactual densities of all the areas obtained by taking into account the effects of the movements within educational and age groups are very close to the actual of 1987 while those obtained with the decompositions between groups, for every grouping, are higher than the Gini coefficients of the estimated distributions in 1995, implying an additional increase in inequality.

ER	alfa=1	alfa=1.3	alfa=1.6
estimated 1987	0.0089	0.0066	0.0050
estimated 1995	0.0101	0.0075	0.0056
education between	0.0099	0.0074	0.0056
education within	0.0138	0.0104	0.0079
age between	0.0138	0.0103	0.0077
age within	0.0091	0.0068	0.0051
industry between	0.0253	0.0189	0.0142
industry within	0.0097	0.0072	0.0054
occupation between	0.0109	0.0082	0.0062
occupation within	0.0151	0.0113	0.0085

table 42: Esteban and Ray polarization index among the actual distributions of 1987 and the actual/counterfactual distribution of 1995.

(42)

What has happened to polarization among the three regional areas according to the index proposed by Esteban and Ray (1994) is reported in table (42). The regional polarization increased in Italy from 1987 to 1995 regardless of the value of degree of sensitivity to polarization - α. The values of the **ER** indexes, indeed, pass from 0.0089 to 0.0101 ($\alpha = 1$), from 0.0066 to 0.0075 ($\alpha = 1.3$) and from 0.0050 to 0.0056 ($\alpha = 1.6$) in the estimated distributions. When the polarization measure is computed on the counterfactual densities the values of **ER** increase for the groupings based education and occupation - within effect - age, industry and occupation - between effect. The reduction observed in polarization is surprising for the grouping based on age - within effect - as the value is almost back to the one of 1987. This result is not true for the decomposition based on education - within effect - as one would have expected from the previous analysis based on the difference among the counterfactual density of 1995 and the estimated one of 1987. This result can be attributed to the peculiarity of the method used to compute this index, which requires to collapse each density on the conditional mean of the group.

PK	**alfa=1**	**alfa=1.3**	**alfa=1.6**
estimated 1987	0.0205	0.0154	0.0116
estimated 1995	0.0308	0.0230	0.0173
education between	0.0310	0.0231	0.0174
education within	0.0219	0.0164	0.0123
age between	0.0284	0.0212	0.0160
age within	0.0205	0.0154	0.0116
industry between	0.0355	0.0265	0.0199
industry within	0.0170	0.0127	0.0095
occupation between	0.0300	0.0224	0.0169
occupation within	0.0233	0.0175	0.0132

table 43: Esteban and Ray polarization index modified by using Kolmogorov measure of variation distance among the actual distributions of 1987 and the actual/counterfactual distribution of 1995.

(43)

I propose a modification of **ER** to compute the level of polarization of a given society without assuming that everybody in each group has income equal to the mean of the group. The index of polarization that Esteban and Ray proposed is here modified in order to take into account the distance among distributions of income between each regional group. I propose to use as measure of distance among two distributions the Kolmogorov measure of variation distance, as previously explained. The results of the measure of polarization computed with the modified index, **PK**, are reported in table (43). Polarization increased regardless of the value of degree of sensitivity to polarization - α. The values of the **PK** indexes, indeed, pass from 0.0205 to 0.0308 ($\alpha = 1$), from 0.0154 to 0.0230 ($\alpha = 1.3$) and from 0.0116 to 0.0173 ($\alpha = 1.6$) confirming what **ER** register. When the index is computed by modifying the density of 1995 the values of **PK** computed among these counterfactual regional distributions increase with respect to the 1995 value for education and industry - between group effect - while it decreases for all the other counterfactuals. Once again is surprising the effect that age - within group - has on the level of polarization as the value obtained among the counterfacual densities is exactly equal to the value computed using the estimated distributions of 1987. The level of polarization is instead almost coincident with the 1987 value if the index is computed using the counterfactual densities obtained by substituting into the 1995 the conditional densities of the groups based on education of 1987. The modified index is hence consistent with the previous analysis: the increase in regional polarization observed in Italy from 1987 to 1995 can be especially attributed to the movements within educational and age group while for the other grouping the evidence is more mixed.

References

[1] Ali, S.M. and S.D. Silvey (1966): "A General Class of Coefficients of Divergence of One Distribution from Another," *Journal of the Royal Statistical Society*. Series A, No.1, 131-142.

[2] Bardone, L., M. Gittleman and M. Keese (1998): "Causes and Consequences of Earnings Inequality in Oecd Coutries," *Lavoro e Relazioni Industriali*, 2, 13-59.

[3] Banca d'Italia (1989): *Supplemento al Bollettino Statistico: i Bilanci delle Famiglie Italiane nell'Anno 1987*, 25 Gennaio 1989, Numero 5, Anno XLII, Roma.

[4] Banca d'Italia (1997): *Supplemento al Bollettino Statistico: i Bilanci delle Famiglie Italiane nell'Anno 1995*, nuova serie, 20 Marzo 1997, Numero 14, Anno VII, Roma.

[5] Brandolini, A. and L. Cannari (1993): "Methodological Appendix: The Bank of Italy's Survey of Household Income and Wealth," in *Savings and the Accumulation of Wealth. Essays on Italian Households and Government Saving Behaviour*, ed. by A. Ando, L. Guiso and I. Visco. Cambridge University Press, Cambridge.

[6] Brunello, G., S. Comi and C. Lucifora (1999): "The returns to Education in Italy: a Review of the Applied Literature," mimeo.

[7] Burkhauser, R.V., A.D. Crews, M.C. Daly and S.P. Jenkins (1996a): "Where in the World is the Middle Class? A Cross-National Comparison of the Vanishing Middle Class Using Kernel Density Estimates," mimeo.

[8] Cannari, L., and A. Gavosto (1994): "L'Indagine della Banca d'Italia sui Bilanci delle Famiglie: una Descrizione dei Dati sul Mercato del Lavoro," *Economia e Lavoro,* 28, 63-79.

[9] Cappellari, L. (1998): *Disuguaglianza, Persistenza e Mobilità: Analisi Longitudinale dei Differenziali Salariali nel Mercato del Lavoro Italiano,* Tesi di Dottorato, Università degli Studi di Pavia.

[10] Casavola, P., A. Gavosto and P. Sestito (1996): "Technical Progress and Wage Dispersion in Italy: Evidence from Firms' Data," *Annales d'Economie et de Statistique,* 41/42, 387-412.

[11] Dell'Aringa, C. and C. Lucifora (1994): "Wage Dispersion and Unionism: Do Unions Protect Low Pay," *International Journal of Manpower,* 15, 2/3, 150-169.

[12] DiNardo, J., N.M. Fortin and T. Lemieux (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1993: a Semiparametric Approach," *Econometrica*, 64, 5, 1001-1044.

[13] Erickson, C.L., and A. Ichino (1993): "Wage Differentials in Italy: Market Forces, Institutions and Inflation," in *Differences and Changes in Wage Structures*, ed. R.B. Freeman and L.F. Katz, The University of Chicago Press, Chicago and London.

[14] Esteban, J.M., and D. Ray (1994): "On the Measurement of Polarization," *Econometrica*, 62, 4, 819-851.

[15] Esteban, J.M., C. Gradin and D. Ray (1997): "Extensions of a Measure of Polarization," mimeo.

[16] Gottschalk, P. and T.M. Smeeding (1997): "Cross-National Comparisons of Earnings and Income Inequality," *Journal of Economic Literature*, 35, 633-687.

[17] Gradin, C. (1999): " The Measurement of Polarization: Analysis by Sub-Populations. The Case of Spain, 1973 - 1991," mimeo.

[18] Härdle, W. (1990): *Applied Nonparametric Regression*, Cambridge University Press, Cambridge.

[19] Jenkins, S.P. (1995): "Did the Middle Class Shrink During the 1980s? UK Evidence from Kernel Density Estimates," *Economic Letters*, 49, 407-413.

[20] Lorenz, M.O. (1905): "Methods for Measuring Concentration of Wealth," *Journal of the American Statistical Association*, (New Series), 70.

[21] Lucifora, C. and F. Origo (1997): "I Differenziali Retributivi 1990-1994: un'Analisi sui Micro-Dati Tratti dagli Archivi INPS," *Rapporto sulle Retribuzioni e sul Costo del Lavoro*, ed. P. Saraceno, Consiglio Nazionale dell'Economia e del Lavoro, Giuffrè Editore, Milano.

[22] Lupi, C. and P. Ordine (1998): "Un'Analisi della Struttura dei Differenziali Salariali in Italia," *Lavoro e Relazioni Industriali*, 2, 69-108.

[23] Manacorda, M. (1997): "Dispersione Salariale e Scolarità," *Rapporto sulle Retribuzioni e sul Costo del Lavoro*, ed. P. Saraceno, Consiglio Nazionale dell'Economia e del Lavoro, Giuffrè Editore, Milano.

[24] Prakasa Rao, B.L.S. (1983): *Nonparametric Functional Estimation*, Academic Press, Orlando, Florida.

[25] Rossi, N. (ed. 1993): *La Crescita Ineguale 1981 - 1991, Primo Rapporto CNEL sulla Distribuzione e Redistribuzione del Reddito in Italia*, Il Mulino, Bologna.

[26] Rossi, N. (ed. 1994): *La Transizione Equa 1992 - 1993, Secondo Rapporto CNEL sulla Distribuzione e Redistribuzione del Reddito in Italia*, Il Mulino, Bologna.

[27] Rossi, N. (ed. 1996): *Competizione e Giustizia Sociale 1994 - 1995, Terzo Rapporto CNEL sulla Distribuzione e Redistribuzione del Reddito in Italia*, Il Mulino, Bologna.

[28] Rossi, N. (ed. 1998): *Il Lavoro e la Sovranita' Sociale 1996 - 1997, Quarto Rapporto CNEL sulla Distribuzione e Redistribuzione del Reddito in Italia*, Il Mulino, Bologna.

[29] Rao, C.R. (1965): *Linear Statistical Inference and Its Applications*, John Wiley & Sons, New York.

[30] Saraceno, P. (ed. 1997): *Rapporto sulle Retribuzioni e sul Costo del Lavoro*, Consiglio Nazionale dell'Economia e del Lavoro, Giuffrè Editore, Milano.

[31] Scott, D.W. (1992): *Multivariate Density Estimation*, John Wiley & Sons, Inc., New York.

[32] Silverman, B.W. (1986): *Density Estimation for Statistics and Data Analysis*,Chapman & Hall, London.

[33] Simonoff, J.S. (1996): *Smoothing Methods in Statistics,* Springer, New York and Heidelberg.

[34] Wand, M.P. and M.C. Jones (1995): *Kernel Smoothing*, Chapman & Hall, London.

[35] Wolfson, M.C. (1994): "When Inequality Diverge," *American Economic Review Papers and Proceedings*, 84, 353-358.

[36] Wolfson, M.C. (1997): "Divergent Inequalities - Theory and Empirical Results," *Review of Income and Wealth*, 4, 401-422.

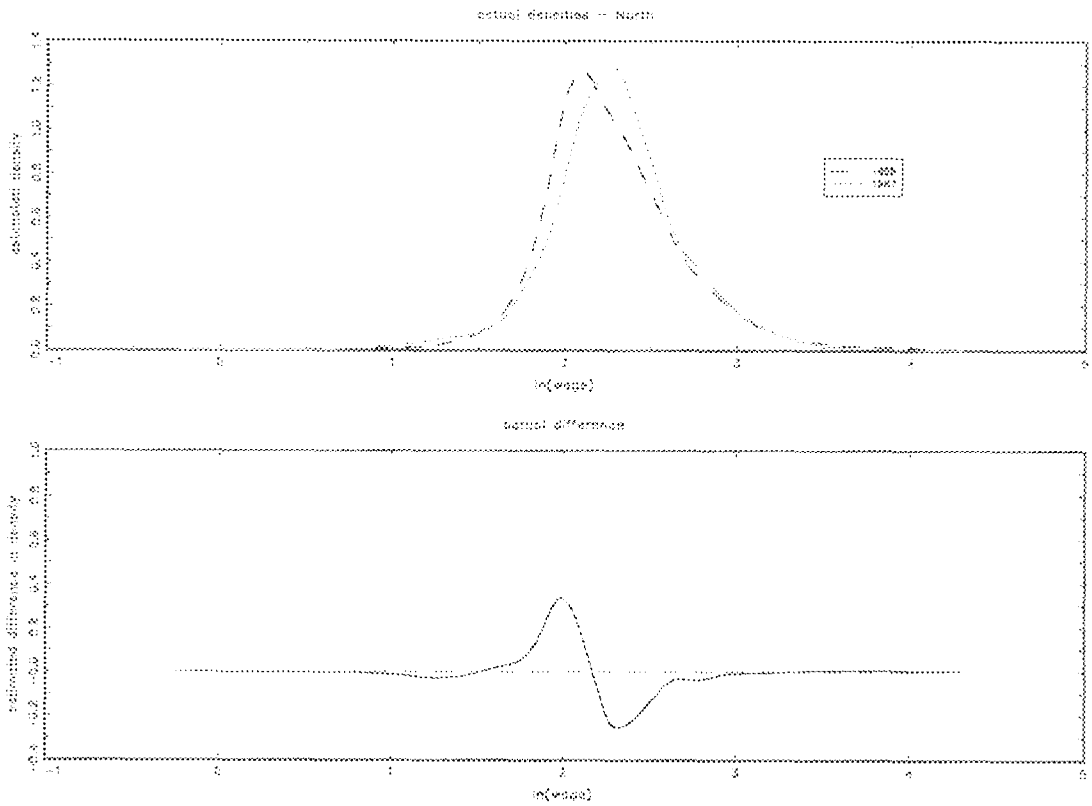

Figure 2: The distribution of the logarithm of real hourly wages. North of Italy.

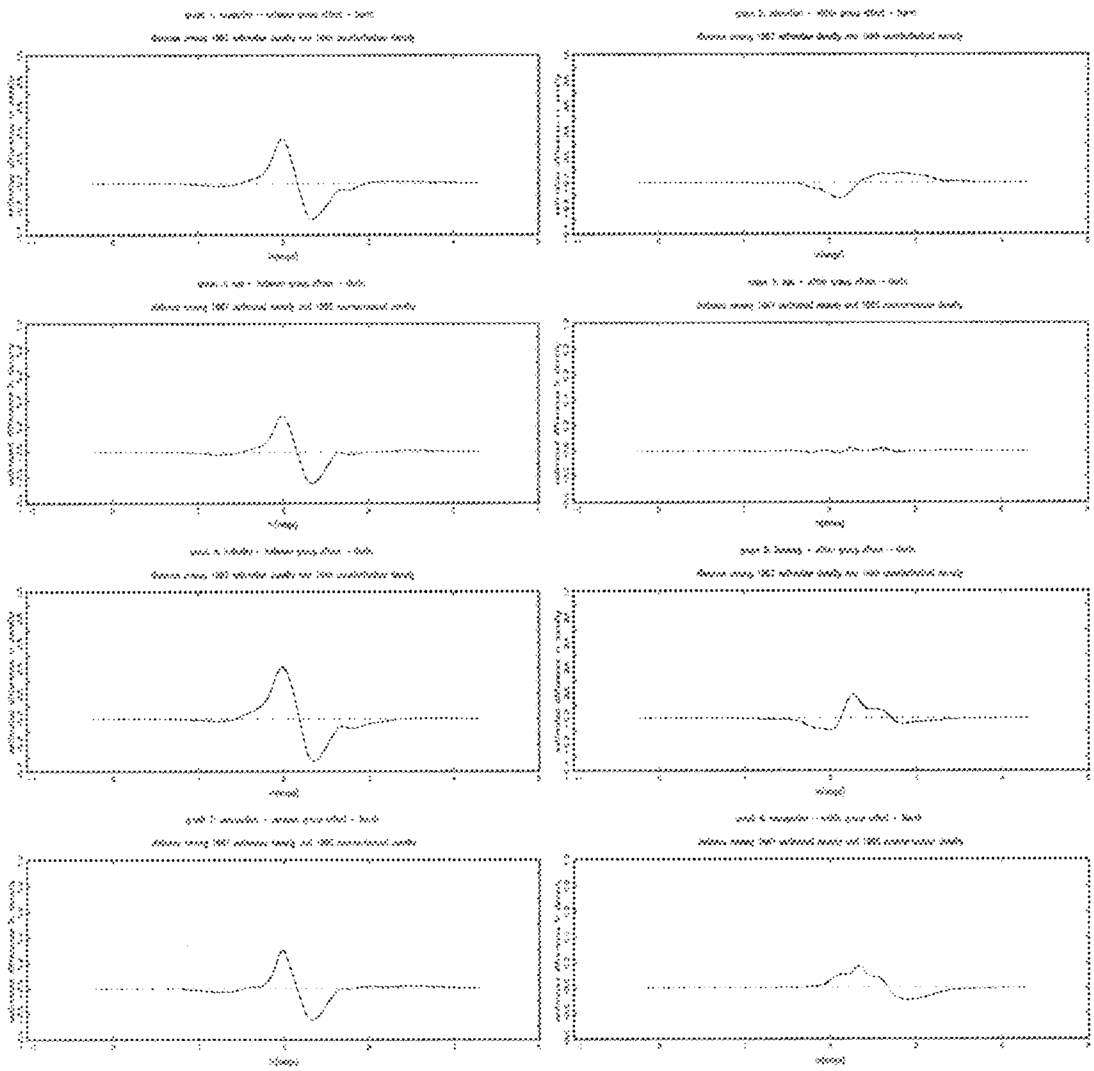

Figure 3: Distances among 1987 estimated density and 1995 counterfactual densities obtained applying the between and within group decomposition - North of Italy.

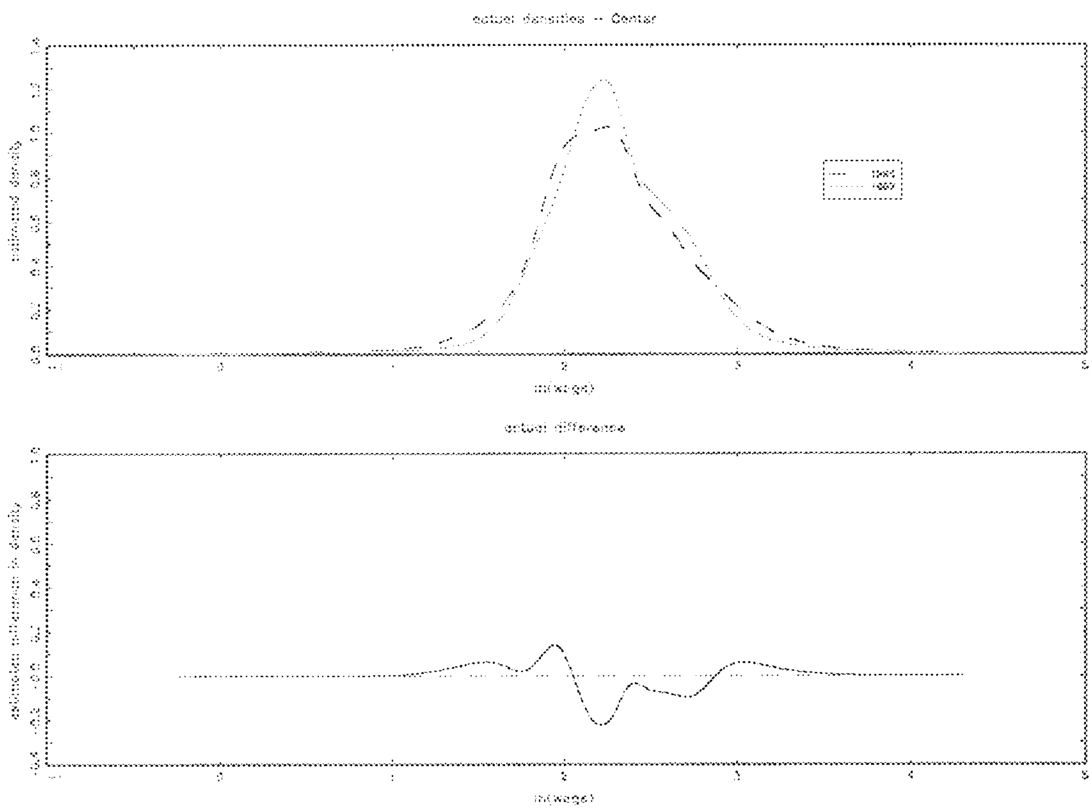

Figure 4: The distribution of the logarithm of real hourly wages. Center of Italy.

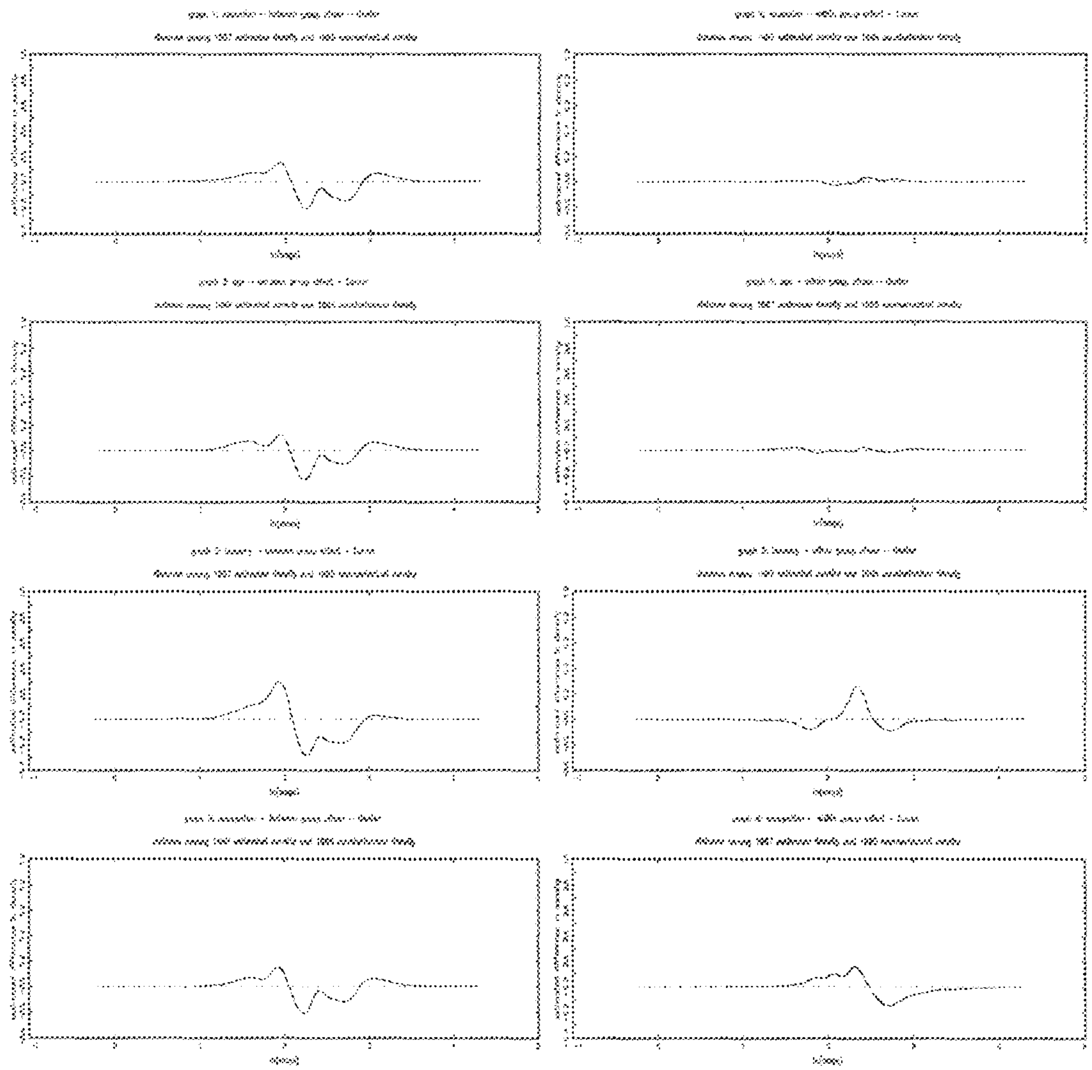

Figure 5: Distances among 1987 estimated density and 1995 counterfactual densities obtained applying the between and within group decomposition - Center of Italy.

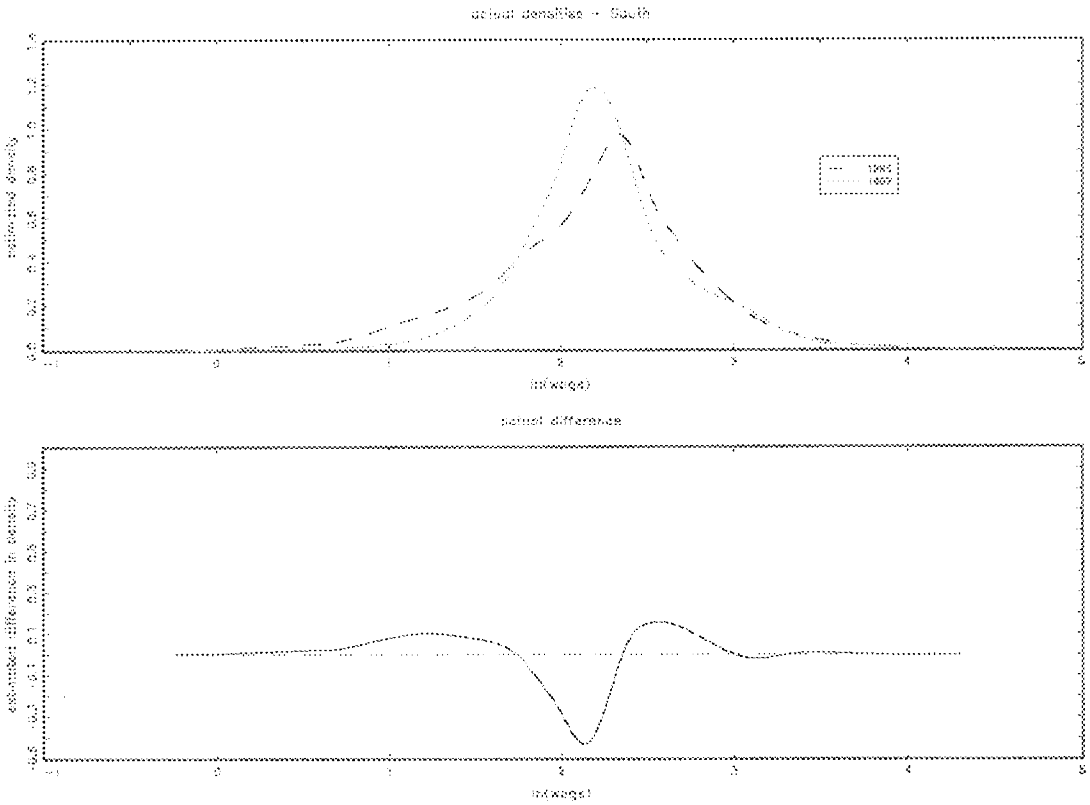

Figure 6: The distribution of the logarithm of real hourly wages. South of Italy.

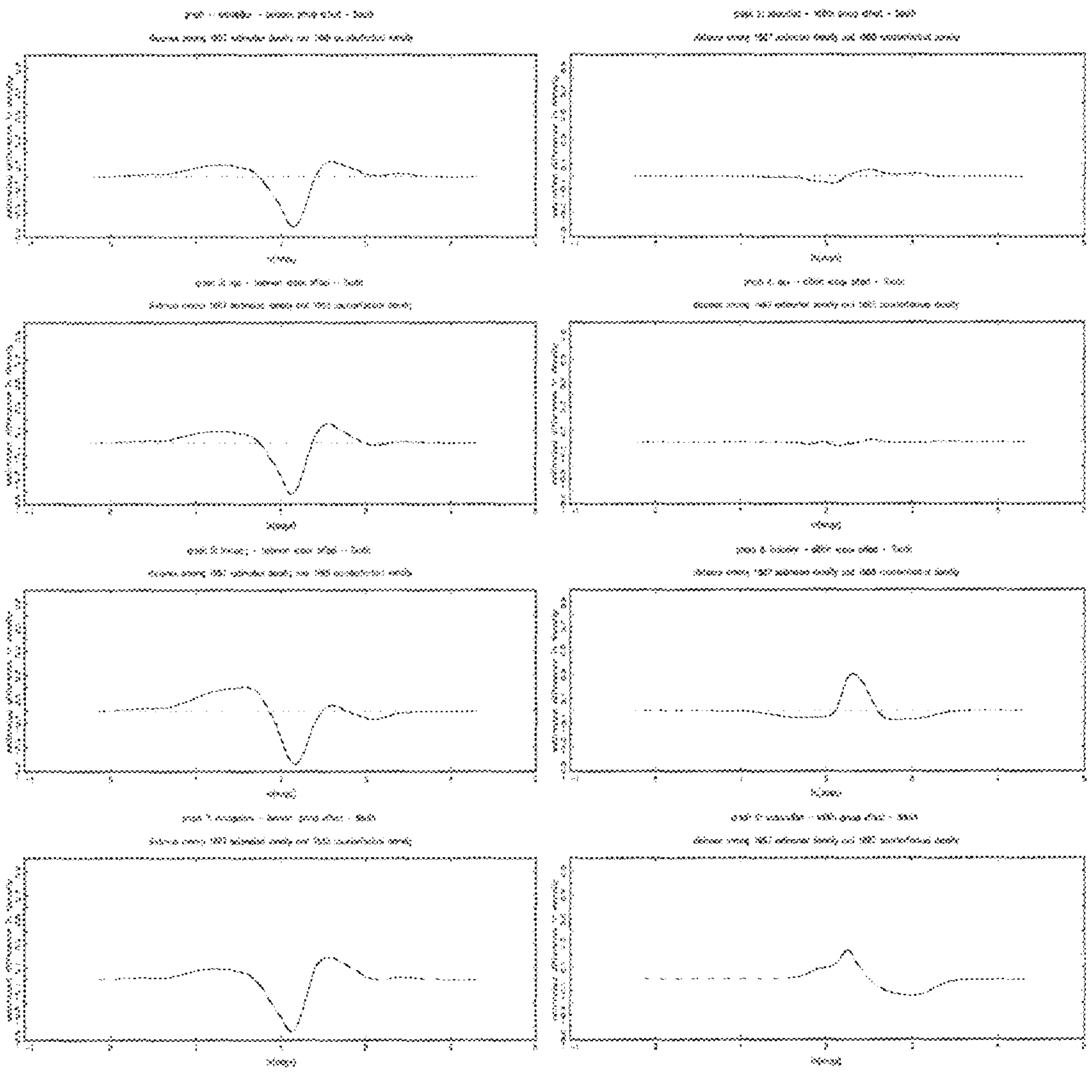

Figure 7: Distances among 1987 estimated density and 1995 counterfactual densities obtained applying the between and within group decomposition - South of Italy.