J Econ Inequal (2009) 7:333–350
DOI 10.1007/s10888-008-9078-z

A dwindling middle class? Italian evidence in the 2000s

Riccardo Massari · Maria Grazia Pittau · Roberto Zelli

Received: 18 December 2006 / Accepted: 25 January 2008 / Published online: 29 March 2008

Abstract In recent years, increasing attention has been paid to household impoverishment in Italy. Most of previous analyses dealing with this issue are based on summary statistics, which may not capture the whole income distribution. This paper employs a non-parametric tool, the "relative distribution", to describe patterns of changes in the entire Italian household income distribution over the period 2000–2004. This approach also allows for a decomposition of the relative density to isolate changes due to differences in location from changes due to differences in shape, thus enabling deeper analysis of income polarization. During the 2000s there was a significant location effect, and also increased income polarization, which has particularly affected incomes below the median. Analyses by social groups, according to the employment status of the household head, show significant re-distribution effects within groups.

Keywords Income distribution · Relative distribution · Polarization

JEL Classification D31 · C14

R. Massari
DCNAPS, Sapienza Università di Roma, Rome, Italy
e-mail: riccardo.massari@uniroma1.it

M. G. Pittau (✉)
Statistics Department, Columbia University, New York, USA
e-mail: grazia.pittau@uniroma1.it

R. Zelli
DSPSA, Sapienza Università di Roma, Rome, Italy
e-mail: roberto.zelli@uniroma1.it

1 Introduction

> Rather than revelling in the *dolce vita*, Italians are battling with the *carovita* (the high cost of living), ...Newspaper headlines warn that the "Middle class has gone to hell" and "Italians don't know how they will make it to the end of the month" (The Guardian, Tuesday, December 28, 2004).

In recent years, so-called "household impoverishment" in Italy, especially with regards to the middle income class, has led researchers and policy makers to turn their attention to distributional issues. However, while there is a growing media concern about the vulnerability of the middle-class, the fall in purchasing power of households, their difficulty in making ends meet (the "fourth week syndrome"), estimates of inequality indices and poverty measures show a substantial stagnation in the last decade. Based on National Statistics of Family Budgets (SFB), Trivellato [20] shows over the period 1990–1997 a substantial stability in the percentage of poor households and a growth in poverty intensity, indicating that poor households have become relatively poorer. Using the same data source, Baldini [1] confirms constant incidence of poverty over the period 1997–2004, despite remarkable differences between regions. Brandolini [5] reports the 1977–1995 time series of Gini coefficients calculated from the Bank of Italy's Survey on Household Income and Wealth (SHIW), showing a cyclical pattern around a relatively stable trend in household income inequality during the 1980s and the early 1990s. The estimates of inequality indices based both on SHIW and SFB, as reported in Boeri and Brandolini [3], remain largely unchanged during the period 1993–2002, after a sharp rise in the early 1990s. The authors note an apparent contradiction between these results and the consumer confidence climate supporting the idea that Italians have perceived a deterioration of their economic conditions and claiming that this inconsistency could be due to several factors, such as disappointed expectations and a horizontal re-distribution across social groups.

These analyses are mainly based on summary measures of location and variation and not on the whole shape of the income distribution, possibly leaving some of the informational content of the above mentioned surveys unexplored. There are several reasons to analyse the entire income shape, permitting to detect the inequality in different parts of the distribution ([21] provides a recent survey). For example, the same degree of inequality can lead to different economic outcomes, depending on whether the inequality is more pronounced in the lower tail of the distribution or in the top deciles. Therefore, approaches that explore the whole content of the income distribution can add insight to the public and academic debate.

The aim of this paper is to determine if short-run changes in the Italian income distribution during the 2000s occurred, despite the evidence of stable indices of poverty, inequality and polarization. We use a relative distribution approach, introduced by Handcock and Morris [10, 11], that provides a non-parametric framework for a straightforward comparison of two income distributions. This method, previously applied to long-run changes in earnings distribution [4, 14], can also detect significant changes in the short-run, signalling underlying structural breaks or the effects of economic and social policies specifically addressed to segments of the population.

The rest of the paper is organized as follows. In the next section the relative distribution method is briefly reviewed. Section 3 is devoted to the illustration of

the main features of income data. The main findings of the changes occurred in the Italian household income distribution are discussed in Section 4. This section also focuses on significant changes of the degree of polarization in the distribution and investigates the degree of polarization within socio-economic groups and over time. Section 5 reports some concluding remarks.

2 The relative distribution approach

The relative distribution is a non-parametric statistical approach introduced by Handcock and Morris [10, 11] that compares the income (or other) distributions of two populations, the "reference" and the "comparison" population, in a way that considers differences throughout the entire income range. Basically, the relative distribution returns the fractions of the "comparison" population that fall in each quantile of the "reference" population. Thus, it is easy to locate and to identify the shifts that have occurred along the income distribution between the two populations.

More formally, let Y_0 be a continuous random variable which represents income for the *reference* population (e.g. households in 2000). Let F_0 be the cumulative distribution function (CDF) of Y_0 and f_0 its probability density function (PDF). The *comparison* population (e.g. households in 2004) generates the continuous random variable Y with F and f its CDF and PDF, respectively. The relative distribution $g(r)$ is defined as the ratio of the density of the comparison population to the density of the reference population evaluated at the rth quantile of the reference distribution:

$$g(r) = \frac{f(F_0^{-1}(r))}{f_0(F_0^{-1}(r))} = \frac{f(y_r)}{f_0(y_r)} \qquad 0 \le r \le 1\ ,\ y_r \ge 0. \tag{1}$$

where $r = F_0(Q_0(r))$, $r \in [0, 1]$, being Q_0 the quantile function of Y_0 and $Q_0(r) = inf\{y_0 \mid F_0(y) = r\} = F_0^{-1}(r) = y_r$. The quantile $Q_0(r)$ can be thought as the value of the income y in the reference distribution below which a proportion r of the ordered income values fall. Thus, $g(r)$ can be interpreted as the ratio of the fraction of households in the comparison population to the fraction of households in the reference population evaluated at the quantile y_r. Intuitively, the *relative data* r are the proportions of households in correspondence to the quantile y_r of the reference population that are realizations of a random variable R defined in [0, 1]. The relative distribution, $g(r)$ is a PDF of the random variable R. The *rescaling* imposed by the quantile function ensures that the density ratio is a proper PDF.

When no changes occur between the two distributions, $g(r)$ is uniform in [0, 1]. A value of $g(r)$ higher (lower) than 1 means that the share of households in the comparison population is higher (lower) than the corresponding share in the reference population, at the rth quantile of the reference population. Put in another way, households of the comparison population have a higher (lower) probability than households in the reference population of having the level of income that corresponds to the rth quantile of the baseline distribution.

In this paper the relative distribution $g(r)$ is calculated as the ratio $\hat{g}(r) = \hat{f}(y_r)/\hat{f}_0(y_r)$, where $\hat{f}$ and $\hat{f}_0$ are obtained as kernel estimates on P quantiles y_r of the reference population. Note that the points on which the two density functions are estimated are the same. A local-polynomial model is finally applied for smoothing the

plug-in estimates $\hat{g}(r)$. The main advantage of the local polynomial smoother is that it is not affected from the boundary bias of the kernel estimator.[1]

This estimation procedure allows us to include sample weights and to use an adaptive bandwidth in the kernel estimator. The weights are attached to each household according to the sample design. The adaptive bandwidth is advised for income data because of sample sparseness [16].

The richness of the relative distribution approach is that it provides a number of tools to isolate several factors which affect the observed outcome. For instance, differences between the reference and the comparison population could be attributed to a change in the average (or the median) income, but they could also be due to differences in shape, that include difference in variation, skewness and other distributional characteristics. It is possible to distinguish between these shifts, namely a *location* effect, due to a change in the first moment, and a *shape* effect, due to a change in higher order moments of the distribution.

The decomposition of the relative distribution in *location* and *shape* effect goes through the definition of an *additive* location-adjusted population $Y_{0L} = Y_0 + \rho$ with the same shape as the reference distribution but with the median of the comparison distribution. The value ρ is the difference between the medians of the two distributions Y and Y_0. The CDF of Y_{0L} is defined as $F_{0L}(y) = F_0(y + \rho)$, and its derivative is the PDF f_{0L}.

Therefore, the decomposition can be written as:

$$\underbrace{\frac{f(y_r)}{f_0(y_r)}}_{g(r)} = \underbrace{\frac{f_{0L}(y_r)}{f_0(y_r)}}_{g_L(r)} \times \underbrace{\frac{f(y_r)}{f_{0L}(y_r)}}_{g_S(p)}, \tag{2}$$

where p is the percentile rank in the location-adjusted population Y_{0L} which corresponds to y_r. If the comparison and the reference distributions have the same median,[2] the density ratio for location differences, $g_L(r)$, will be uniform in $[0, 1]$. Conversely, if the two distributions have different median, then $g_L(r)$ is increasing (decreasing) in r if the comparison median is higher (lower) than the reference median. The density ratio for shape differences, $g_S(p)$, represents the relative distribution net of the location effect. The analysis of g_S detects re-distribution that has occurred between the reference and the comparison populations. For instance, $g_S(p)$ would take a (inverse) U-shape, if the comparison population is relatively (less) more spread at around the median than the location-adjusted population. It is thus possible to determine whether there is an increasing income polarization, a

[1] This estimation procedure differs from the one used by Handcock and Morris [10] and developed in the `reldist` R package [9] essentially based on the empirical cumulative distribution function of the relative data. It also differs from the kernel estimation procedure [6] that applies a kernel density estimator directly on the relative data.

[2] Alternative indices like the mean can be considered. The corresponding results do not differ in a significant way, and are not reported here. A *multiplicative* median location shift can also be applied. However, the multiplicative shift has the drawback of affecting the variance and the shape of the distribution. Indeed, the equi-proportionate income changes cause a flattening (or a shrinking) of the shape of the distribution [12].

downgrading – defined as the movement of households into the lower tail of the income distribution–, an upgrading, or a convergence of incomes towards the median.

3 Data

We use two measures of income in our analysis. The first is household annual disposable income, defined as income from all sources, including imputed rents from owner-occupied housing, net of taxes and social security transfers. The second is the same measure exclusive of imputed rents. The inclusion of the imputed rents in the disposable income is still controversial. Its impact on the income distribution is a significant reduction in the inequality levels [8, 19, 22]. A reason to exclude imputed rents is the very low marginal propensity to consume out of real assets. As shown in Paiella [15], the rapid increase in real estate prices in Italy since the end of the 1990s has affected household expenditure only marginally. Accordingly, we perform our analysis using disposable income as well as income net of imputed rents.

Incomes are adjusted for household size by the Italian official equivalence scale[3] and deflated to 2000 prices using the national accounts household expenditure deflator. The choice of the deflator is consistent with the definition of income that we use, inclusive or exclusive of imputed rents.

The data are drawn from the bi-annual survey of the Bank of Italy on household income and wealth (SHIW) from 2000 to 2004 (for a detailed description of the survey, see Banca d'Italia [2]).

Table 1 provides summary measures for household incomes in 2000 and 2004, both gross and net of imputed rents. Between 2000 and 2004, households income shares show a substantially unchanged pattern, apart from a moderate upsurge of income net of imputed rents shares of the poorest 5% and richest 5% of the population. The Gini index displays a negligible increase, while an increment in inequality is detected by the Theil index, at least for incomes net of imputed rents.[4] This could be due to its greater sensitivity to changes at the tails of the distribution. Quite surprisingly, the Wolfson polarization measure shows a modest but significant decline, not detected by the Esteban–Ray (ER) measure[5]. Instead, the polarization index proposed in 2004 by Duclos et al. [7], which is defined on the continuous space and relies on the kernel estimated income distribution, shows a slight increase of polarization between

[3] A scale that assigns 1 to a 2-member household, 0.599, 1.335, 1.632, 1.905, 2.150 and 2.401 to households of one, three, four, five, six and seven or more members, respectively.

[4] Bootstrapped confidence intervals at 95% of the Gini inequality index are, for incomes gross and net of imputed rents, in 2000: [0.312, 0.336], [0.316, 0.339]; while in 2004: [0.322, 0.349], [0.326, 0.354]. Confidence intervals of the Theil index are, for incomes gross and net of imputed rents, in 2000: [0.182, 0.225], [0.184, 0.224]; while in 2004: [0.191, 0.244], [0.196, 0.259]

[5] The ER polarization measure implies a regrouping of the population. In analogy with the Wolfson index, we represent the distribution as a bipolar distribution, using the median as the cut-off value that divides the two presumed groups.

Bootstrapped confidence intervals at 95% of the Wolfson polarization index are, for incomes gross and net of imputed rents, in 2000: [0.277, 0.298], [0.281, 0.304]; while in 2004: [0.246, 0.269], [0.249, 0.273]. Confidence intervals of the Esteban–Ray index are, for incomes gross and net of imputed rents, in 2000: [0.177, 0.186], [0.179, 0.188]; while in 2004: [0.175, 0.186], [0.177, 0.188].

Table 1 Summary measures of Italian household disposable income gross and net of imputed rents (IR): 2000–2004

	Disposable income			
	Gross of IR		Net of IR	
	2000	2004	2000	2004
Mean (2,000 euros)	22,382	24,005	18,302	20,418
Median (2,000 euros)	19,188	20,277	15,540	17,221
Income share (per cent)				
Bottom 5%	0.89	1.08	0.75	0.99
Bottom 10%	2.60	2.85	2.44	2.83
Bottom 20%	7.17	7.38	7.04	7.37
Top 20%	40.37	40.74	40.52	41.01
Top 10%	25.50	26.13	25.65	26.56
Top 5%	16.09	16.88	16.28	17.41
Inequality measures				
Gini	0.32	0.33	0.33	0.34
Theil	0.20	0.21	0.20	0.23
Quintile ratio	5.63	5.52	5.76	5.56
Polarization measures				
Wolfson	0.29	0.26	0.29	0.26
Esteban–Ray ($\alpha = 1.3$)	0.18	0.18	0.18	0.18
Head-count poverty ratio (per cent)				
Line at 50% of mean	18.02	17.21	18.35	17.22
Line at 50% of median	12.76	11.56	12.44	10.92
Line at 60% of mean	26.53	26.30	26.93	27.18
Line at 60% of median	19.19	17.79	18.94	17.66

Authors' calculation on weighted household income data from SHIW. Income data are size-adjusted and expressed in 2000 prices

2000 and 2004.[6] The head count poverty ratio displays a rather modest fall, that is marginally more pronounced when the poverty line is set at the lowest level (50% of the median).

Overall, based on the results reported in Table 1, the income distribution features display only few significant changes, besides the growth of the real mean and median incomes.

4 The relative distribution analysis

4.1 Changes in income distributions

Figure 1 reports the kernel density estimates of 2000 and 2004 income distributions, gross (Fig. 1a) and net of imputed rents (Fig. 1c). Figure 1a shows a clear shift of

[6] In fact, setting for example $\alpha = 1$, the DER polarization index is equal to 0.32 in 2000 and reaches the value of 0.33 in 2004 for income gross of imputed rents. For income net of imputed rents it raises from 0.32 in 2000 to 0.34 in 2004.

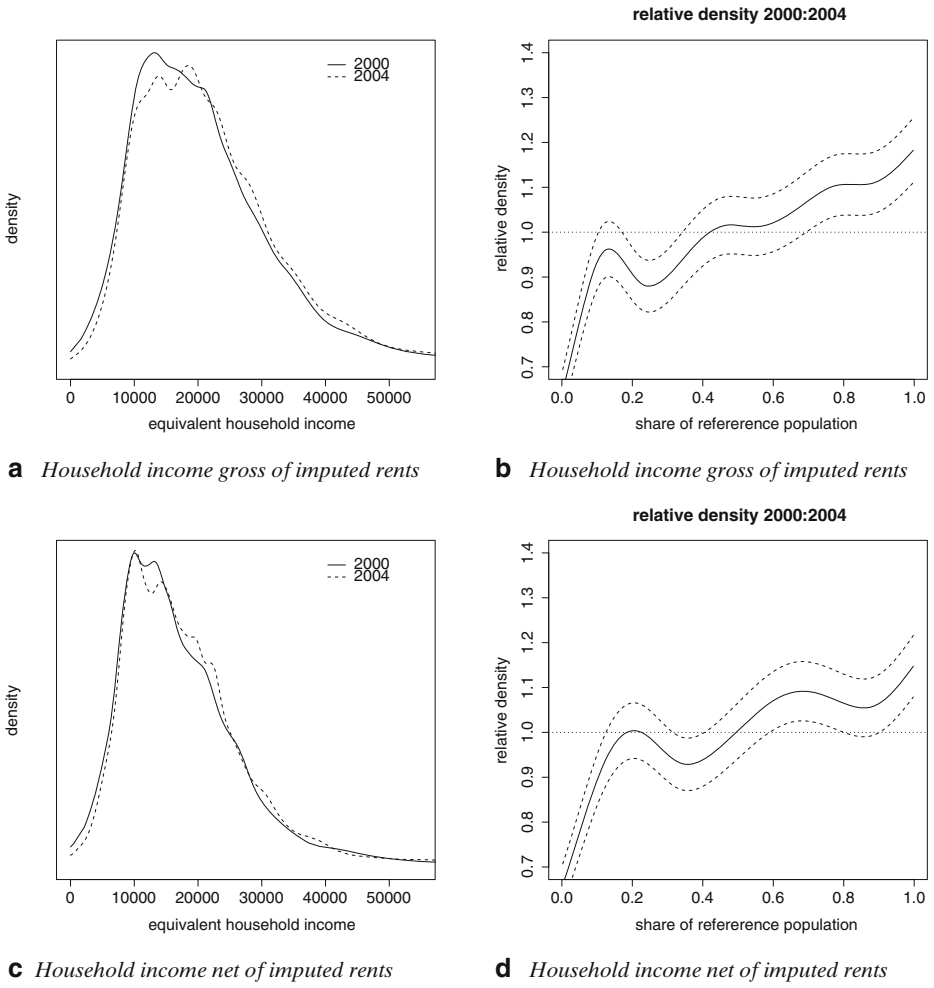


Fig. 1 Comparison between 2002 and 2004 income distributions. Authors' calculation on weighted household income data from SHIW. Income data are size-adjusted and expressed in 2000 prices. **a** and **c** The bandwidths for the estimate of 2000 and 2004 kernel density functions are obtained with the Sheather–Jones criterion. **b** and **d** *Dotted lines* represent 95% confidence intervals

the mass of the distribution rightwards and a change of the shape, especially in the middle income range. The 2004 distribution exhibits two clear modes, while in the 2000 distribution bi-modality is much less evident. There is also a decline in the mass at the middle-income range from 2000 to 2004. The shift rightwards of the 2004 distribution is less pronounced when imputed rents are excluded, the second peak is less marked and a bump in the upper income ranges emerges (Fig. 1c). As shown in Pittau and Zelli [17], the emergence of the modes, and the gap between them, could be interpreted as an increase in polarization, especially when inequality is unchanged.

Further insight is provided by the relative density function that directly compares two densities and indicates whether the upper and the lower tails of the distribution

are growing at the same rate. Figure 1b displays the relative distribution of household income gross of imputed rents along with the 95% confidence interval.[7]

Changes in the distribution are indicated by the generally positive slope of $g_L(r)$: the mass of households below the 2000 median income increased during the early 2000s. More specifically, the relative distribution is less than 1 for $r \leq 0.42$ and more than 1 for values above 0.42 with the peak at $r = 1$. That is, if we choose any percentile between the 1st and the 42nd in the 2000 distribution, the percentage of households in 2004 that earn an amount of income corresponding to the chosen percentile is less than the corresponding percentage of households in 2000. However, income growth between 2000 and 2004 impacts the whole range of the distribution with varying intensity, more positively affecting households in the bottom and in the top deciles.

The relative density for income net of imputed rents, along with the corresponding confidence interval (Fig. 1d), offers instead a somewhat different picture. A decline of the lower middle-income mass is evident, indicating a drop of the share of households between approximately the 24th and the 48th percentiles of the 2000 distribution. The peak of 1.1 is at around the 65th percentile, that is households in 2004 are approximately 10% more likely to fall at the level of 2000 income corresponding to the 65th percentile with respect to households in 2000.

To get a more detailed picture, we can decompose the relative density into location shift and shape effects, reported in Fig. 2 along with their corresponding 95% confidence intervals. Figure 2a and c reports the effect only due to the median shift, that is the pattern that the relative density would have displayed if there had been no change in distributional shape but only a shift of location of the density. Since the median shift is positive, the location effect reduces the share of households in bottom percentiles, increasing instead those in higher percentiles. This effect is more evident for income gross of imputed rents.

Figure 2b and d shows the shape effect, the income redistribution across households gross and net of imputed rents, respectively. Not surprisingly, having isolated changes of the shape, a rise of relative density, with respect to the observed relative

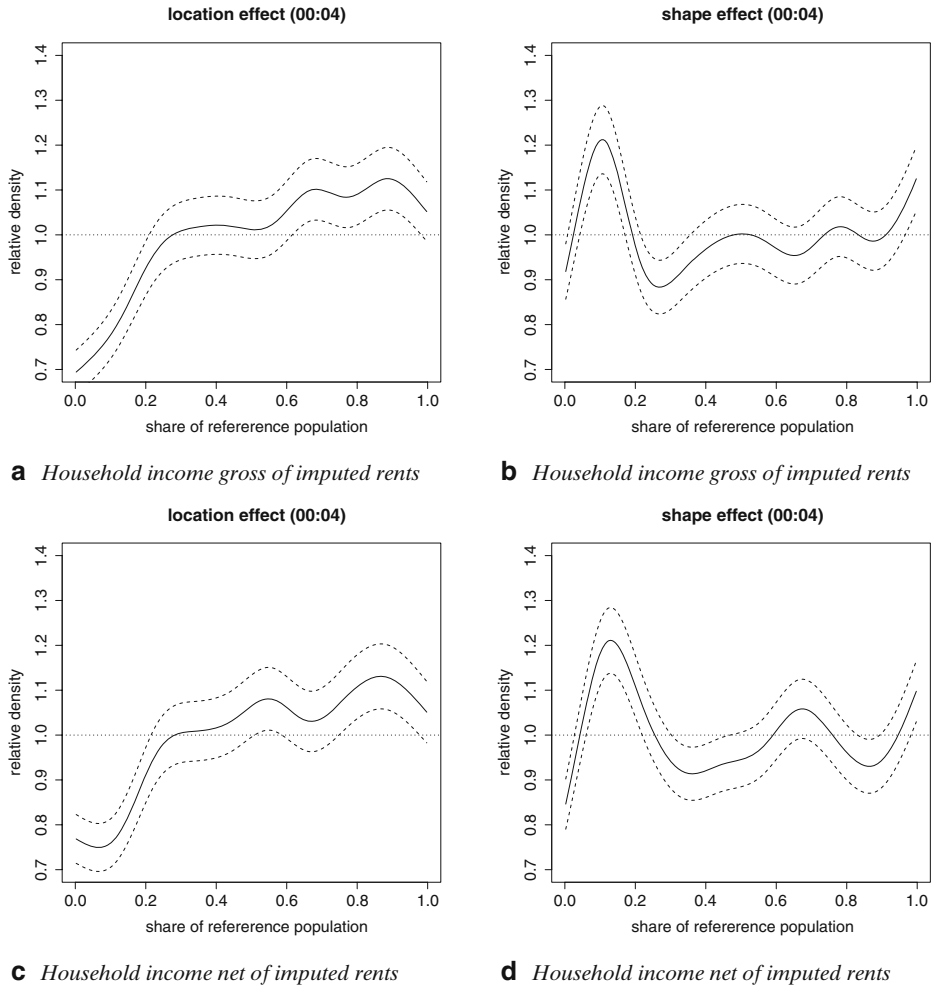
[7] Pointwise confidence intervals for the relative density $g(r)$, $0 \leq r \leq 1$ are based on the asymptotic normal (AN) approximation ([11], p. 144). The normal asymptotic properties of the estimator $g_{n,m}(r)$ of $g(r)$ are derived under regular assumptions: $0 \leq r \leq 1$, $F_0(x)$ and $F(x)$ have continuous and differentiable densities, $f_0(x)$ $f(x)$ respectively. In addition, $K(.)$ has to be a twice continuously differentiable kernel function, satisfying

$$\int_{-1}^{1} K(x)dx = 1; \int_{-1}^{1} xK(x)dx = 0; \int_{-1}^{1} x^2 K(x)dx = \sigma_K^2 > 0 \tag{3}$$

and vanishing outside the bounded interval $[-1; 1]$. Choosing a bandwidth h_m, such that, as $m, n \to \infty$, $h_m \to 0$ with $mh_m^3 \to \infty$, $mh_m^5 \to 0$, $m/n \to k^2 < \infty$, then:

$$g_{n,m}(r) \sim AN\left\{g(r), \frac{g(r)R(K)}{mh_m} + \frac{g(r)^2 R(K)}{nh_m}\right\} \tag{4}$$

where $R(f) = \int f(x)^2 dx$. The second term in the asymptotic variance for $g_{n,m}(r)$ is due to the fact that F_0 is unknown and it must be estimated. In this paper, we use the biweight kernel density function that satisfies the above properties and we estimate h_m using the Sheather–Jones criterion [18].



a *Household income gross of imputed rents*

b *Household income gross of imputed rents*

c *Household income net of imputed rents*

d *Household income net of imputed rents*

Fig. 2 Location shift and shape effects. Authors' calculation on weighted household income data from SHIW. Income data are size-adjusted and expressed in 2000 prices. *Dotted lines* represent 95% confidence intervals

distributions, at bottom percentiles and a relatively small decrease at the top incomes is detected. Overall, both figures display a diminished weight of households having incomes around the 2000 median, and an increase of household shares at the poorest deciles of the distribution. Nonetheless, there are distinctive features of the two median-adjusted relative densities, gross and net of imputed rents.

The median-adjusted relative distribution of incomes, gross of imputed rents, shows a marked change for the incomes below the median, with a decline of the mass between the 19th and the 51st percentile and a prominent increase of the fraction of households between the 5th and the 18th percentile, indicating a clear downgrading of the distribution. On the contrary, the upper part of the relative density does not reveal significant changes apart from an increase at the very top percentiles. The

shape effect of incomes net of imputed rents is more complex, since a double-shrink occurs. The shrink of the lower-middle incomes is still evident, embracing a range between the 25th and the 57th percentile, and a redistribution occurs in the upper part as well. A convergence towards the median emerges, with a decline of the mass in the upper tail between the 74th and the 94th percentile, and an almost equal increase in the upper-median range between the 58th and the 73rd percentile.

4.2 A closer look on polarization

The graphical analysis, as presented in Section 4.1, provides a detailed description of polarization patterns. A link between what we have observed in the graphical analysis and the quantification of the degree of polarization is yielded by the *median relative polarization index*, as introduced by Handcock and Morris [10]. This index keeps track of *changes* in the shape of the distribution and measures the direction and the magnitude of the change.

The median relative polarization index is the mean absolute deviation from the median of the location-matched relative distribution g_S, re-scaled in order to vary between –1 and 1. Positive values represent an increase in income polarization, while negative values imply a convergence of incomes towards the median. A value of zero indicates no differences in distributional shape.

Formally, the median relative polarization index of Y with respect to Y_0 is defined as:

$$\mathrm{MRP}(F, F_0) = 4 \int_0^1 \left| r - \frac{1}{2} \right| g_S(r) dr - 1 \tag{5}$$

and can be estimated as:

$$\widehat{\mathrm{MRP}}(F, F_0) = \frac{4}{m} \sum_{j=1}^{m} \left| \hat{R}_j - \frac{1}{2} \right| - 1 \tag{6}$$

where $\hat{R}_j = F_{0n}(Y_j - \rho)$ are the estimates of the location-matched relative data and m is the sample size of the comparison population.[8]

The median polarization index can be decomposed into the contributions from the lower and upper tails of the distributions, emphasizing the change of the overall polarization due to incomes above and below the median of the relative distribution.

[8]The R_js are here estimated with a kernel-type estimator of F_0 as in Molanes-López and Cao [13]:

$$F_{0n} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{M}\left(\frac{y - Y_{0i}}{h_0} \right)$$

where $\mathbf{M}$ is the cumulative distribution function of the kernel M, h_0 is the bandwidth and n the sample size of the reference population.

Table 2 Relative polarization indices (2000–2004)

Household disposable income									
Gross of IR					Net of IR				
Index	Value	CI (95%)		*p* val.	Index	Value	CI (95%)		*p* val.
MRP	0.040	0.022	0.058	0.00	MRP	0.036	0.018	0.054	0.00
LRP	0.119	0.093	0.146	0.00	LRP	0.131	0.105	0.158	0.00
URP	−0.039	−0.064	−0.014	0.00	URP	−0.060	−0.085	−0.035	0.00

Authors' calculation on weighted household income data from SHIW. Income data are size-adjusted and expressed in 2000 prices. The *p* values refer to the null hypothesis that polarization does not change. MRP is the median relative polarization index; LRP is the lower relative polarization index; URP is the upper relative polarization index

The lower and upper polarization indices are defined, respectively, by:

$$\mathrm{LRP}(F, F_0) = 8 \int_0^{1/2} \left| r - \frac{1}{2} \right| g_S(r) dr - 1 \tag{7}$$

$$\mathrm{URP}(F, F_0) = 8 \int_{1/2}^{1} \left| r - \frac{1}{2} \right| g_S(r) dr - 1. \tag{8}$$

They vary between −1 and 1 and can be estimated in a similar way.

Table 2 reports the relative polarization indices computed for household income gross and net of imputed rents, along with their 95% confidence intervals. In both cases, the median relative polarization indices, MRP, are significantly positive, implying an increase of the overall polarization from 2000 to 2004. This total outcome results from two contrasting effects that can be detected by the lower and upper indices, LRP and URP: more polarization in the lower tail and convergence towards the median in the upper tail of the distribution, with the former effect prevailing. In other words, the shift towards the lower tail of the distribution is not offset by the shift of the upper tail towards the median, revealing an overall downgrading of household incomes over the period.

To give an idea of the magnitude of the changes of polarization, the estimated value of MRP for incomes gross of imputed rents equal to 0.04 is equivalent to a 4% population shift from the median of the distribution to the upper and the lower quartiles.[9] Note that the value 0.04 of the MRP index is the average of the lower and upper indices that have contrasting effects: LRP = 0.12 and URP = −0.04. Therefore, the value of 0.04 can be interpreted as a 6% population shift from the median of the distribution to the lower quartile and a simultaneous 2% population shift from the upper quartile of the distribution towards the median.

[9]This is because of the interpretation of MRP in terms of a proportional shift of mass in the distribution from more central to less central values: MRP = $4d\delta_p$, being d the distance between the median and the tails of the distribution, measured on the unit interval, and δ_p a net change in the mass.

4.3 Decomposition by employment status

The classification we used to study the horizontal re-distribution between households in different socio-economic groups is according to the labor market status of the household head: employee,[10] self-employed, retired from work, other non-employed. There are, of course, many other possible decompositions. However, Boeri and Brandolini [3] document the importance of classifying the population according to the employment status of household head to explain the cyclical evolution of income inequality in Italy in the last decade.

In a fashion similar to the decomposition into location and shape effects, the distributional impact of employment status can assume two forms, the *compositional* and the *residual* effects. The former is due to a change in the population shares according to the employment status, the latter effect is instead due to the change in the relationship between the covariate (employment status of household head) and the response variable (household income).

In our analysis the *compositional* effect is likely to be negligible because of the marginal changes of sample structure observed in a short time period. Therefore, the shape of the relative density is mainly due to changes in the conditional income distributions by household head's employment. Consequently, over relatively short periods of time, changes in the patterns of the relative conditional distributions by employment status give us a deeper view of the polarization process.

Table 3 reports summary statistics according to the employment status of the household head. In the first row are reported the shares of households in each sub-group. This group partition is consistent with the main income source of households. For example, households whose head is an employee get most of their income from wages. Therefore, changes in the income distribution of these households are manly affected by shifts in the wage distribution. Similar considerations apply for self-employment incomes and transfers.

The kernel estimates for income, gross and net of imputed rents, along with densities of three household typologies are reported in Fig. 3. The shapes of the overall income distribution seem to be essentially influenced by the distribution of households with the head employed. The households distribution whose head is employed(dashed lines), moves from a bi-modal shape in 2000 to an approximate tri-modal shape in 2004. The three modes in 2004 distribution are more evident for income net of imputed rents and are farther apart in 2004, suggesting an upsurge of polarization not captured by summary measures (see Table 3).

An increase in the number of households with low income is evident in 2004 from the income distribution of households whose head is retired (dashed and dotted lines). At the same time, the growth in the upper tail of the overall distribution appears to be due to households whose head is self-employed (dotted lines).

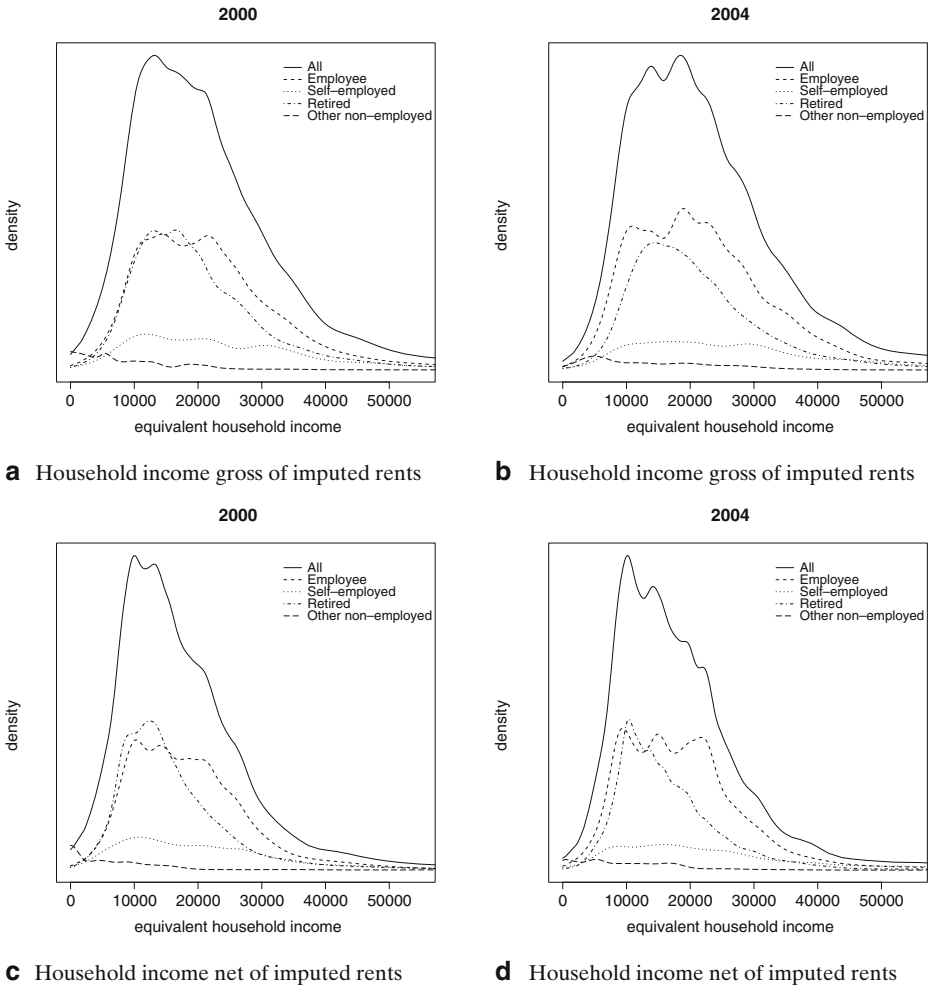
The median income of households whose head is self-employed is higher than that of all the other groups, as expected. The growth in median income is positive but modest for employed households, positive and large for self-employed and negative for retired households. Income shares for households with employed head do not

[10] Households whose head is a contingent worker are included in the employee sub-group.

Table 3 Summary measures for household incomes, by employment status of household head

	Disposable income															
	Employee				Self-employed				Retired				Other non-employed			
	Gross of IR		Net of IR		Gross of IR		Net of IR		Gross of IR		Net of IR		Gross of IR		Net of IR	
	2000	2004	2000	2004	2000	2004	2000	2004	2000	2004	2000	2004	2000	2004	2000	2004
% of the sample	44.43	46.89	44.43	46.89	14.03	12.67	14.03	12.67	38.03	37.42	38.03	37.42	3.50	3.01	3.50	3.01
Mean	22,266	23,207	18,839	19,459	27,909	34,815	23,129	29,501	21,476	22,066	16,787	16,834	11,573	15,071	8,616	10,212
Median	20,459	20,818	17,232	17,761	22,116	26,285	18,787	20,588	18,069	19,275	14,073	14,741	6,676	12,724	5,087	8,727
Income shares (per cent)																
Bottom 5%	1.20	1.19	1.18	1.18	0.79	0.80	0.70	0.65	1.32	1.48	1.29	1.52	0.00	0.34	0.00	0.00
Bottom 10%	3.14	3.02	3.15	3.15	2.17	2.03	1.94	1.73	3.29	3.62	3.34	3.81	0.01	1.19	0.00	0.04
Bottom 20%	8.02	7.64	8.14	7.84	5.91	5.38	5.60	4.74	8.20	8.86	8.35	9.34	0.92	3.99	0.00	1.33
Top 20%	36.97	37.96	36.56	37.31	44.56	49.01	45.00	50.86	40.07	37.85	39.97	36.80	56.98	43.84	62.95	44.20
Top 10%	22.10	23.05	21.93	22.75	29.01	34.61	29.57	36.74	25.74	23.30	25.82	22.29	39.74	26.72	45.50	27.54
Top 5%	12.98	13.96	13.14	13.85	18.87	24.65	19.20	26.63	16.51	14.23	16.73	13.43	27.30	16.94	33.18	14.69
Inequality measures																
Gini	0.28	0.31	0.28	0.30	0.37	0.43	0.38	0.46[a]	0.31	0.29	0.31	0.27	0.59	0.44	0.67	0.47
Theil	0.14	0.15	0.14	0.15	0.28	0.39	0.29	0.46[a]	0.20	0.15	0.20	0.13	0.54	0.31	0.58	0.32
Quintile ratio	4.61	4.97	4.49	4.76	7.54	9.11	8.04	10.74	4.89	4.27	4.79	3.94	62.24	11.00	–	33.34
Polarization measures																
Wolfson	0.26	0.24	0.26	0.24	0.37	0.34	0.37	0.38	0.27	0.24	0.26	0.24	0.55	0.39	0.61	0.43
Esteban–Ray (α =1.3)	0.16	0.17	0.16	0.17	0.21	0.22	0.22	0.24	0.17	0.16	0.17	0.15	0.30	0.25	0.34	0.27

Authors' calculation on weighted household income data from SHIW.
Income data are size-adjusted and are expressed in 2000 prices
[a] Indicates that 95% confidence intervals in two periods do not overlap



a Household income gross of imputed rents

b Household income gross of imputed rents

c Household income net of imputed rents

d Household income net of imputed rents

Fig. 3 Overall density and sub-groups densities. Authors' calculation on weighted household income data from SHIW. Income data are household size-adjusted and are expressed in 2000 prices

remarkably change for both bottom and top incomes and inequality shows a slight increase. Considering households whose head is self-employed, richer households, especially those in the top decile and in the top 5%, receive a bigger fraction of income in 2004 than in 2000. This may explain the considerable rise in the inequality indices for this group. It is worthwhile noting that these distributions are the most unequal and polarized among those we consider. Conversely, households with head unemployed in the bottom tail of the income distribution increase their income shares, while those in the top tail receive a smaller amount in 2004, especially when income is net of imputed rent. Consequently, inequality and polarization indices show a small reduction for this group.

Table 4 Polarization indices per household head employment status (2000–2004)

	Disposable income								
	Gross of IR					Net of IR			
	Index	CI (95%)		*p* val.		Index	CI (95%)		*p* val.
Employee									
MRP	0.051	0.023	0.079	0.00	MRP	0.046	0.019	0.074	0.00
LRP	0.126	0.085	0.166	0.00	LRP	0.159	0.119	0.201	0.00
URP	−0.024	−0.061	0.014	0.11	URP	−0.067	−0.105	−0.029	0.00
Self-employed									
MRP	0.140	0.088	0.191	0.00	MRP	0.116	0.064	0.167	0.00
LRP	0.328	0.252	0.404	0.00	LRP	0.193	0.117	0.269	0.00
URP	−0.049	−0.118	0.020	0.08	URP	0.039	−0.031	0.108	0.14
Retired									
MRP	0.042	0.013	0.070	0.00	MRP	0.028	−0.001	0.056	0.03
LRP	0.170	0.128	0.213	0.00	LRP	0.135	0.094	0.177	0.00
URP	−0.087	−0.125	−0.049	0.00	URP	−0.080	−0.119	−0.041	0.00
Other non-employed									
MRP	0.303	0.204	0.402	0.00	MRP	0.162	0.061	0.263	0.00
LRP	0.619	0.484	0.755	0.00	LRP	0.359	0.205	0.512	0.00
URP	−0.013	−0.143	0.117	0.42	URP	−0.035	−0.161	0.091	0.30

Authors' calculation on weighted household income data from SHIW. Income data are size-adjusted and expressed in 2000 prices. The *p* values refer to the null hypothesis that polarization does not change. MRP is the median relative polarization index; LRP is the lower relative polarization index; URP is the upper relative polarization index

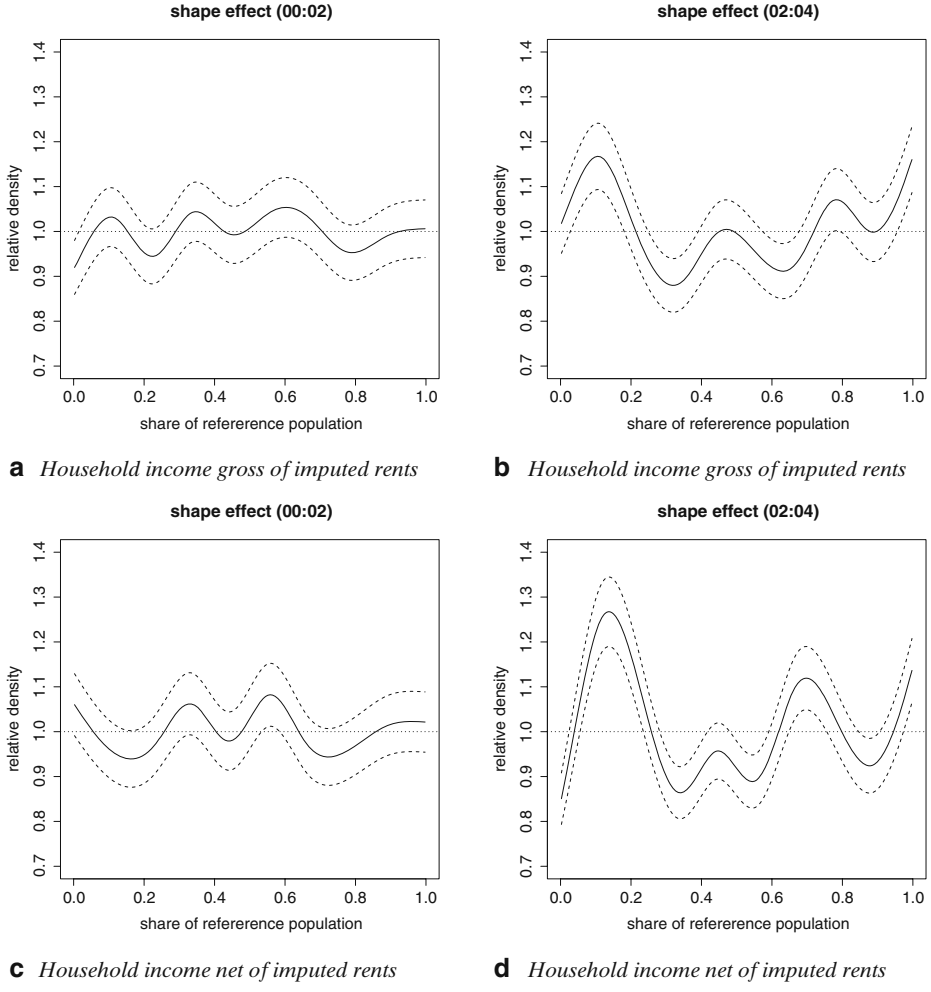
The set of relative polarization indices, based on the median-adjusted relative distribution computed for each group separately, is reported in Table 4. The MRP indices are all significantly positive, indicating a marked increase of polarization in the sub-group distributions. The magnitude of MRP takes the highest value for the self-employed income distribution. In all cases, the growth of polarization stems from a significant increase of polarization in the lower tail of each distribution, not compensated by a significant convergence in the upper tail.

The pattern of polarization seems to be slightly different for households whose head is self-employed, especially for incomes net of imputed rents. This different behavior is essentially due to a shift from the self-employed median income towards the top of the distribution. The polarization is less marked for households with a retired head. The reduction in their incomes in the upper tail of the distribution is larger than that observed in any other group.

4.4 Temporal decomposition

Here we document that the more significant change in the shape of the Italian income distribution occurs over the period 2002–2004 rather than over 2000–2002. To highlight changes that took place within each sub-period, we decompose the shape

effect g_S into two period effects in a similar way as described in Section 2: the 2000–2002 and the 2002–2004 effect. Figure 4 shows the shape effects in the two periods of the household income relative densities, gross and net of imputed rents, along with the corresponding confidence intervals. The overall shape effect (2000–2004) resembles the shape effect estimated for the 2002–2004. In fact, for both incomes, gross and net of imputed rents, the shape effect that occurs in the period 2000–2002 is barely insignificant (see Fig. 4a and c). The median polarization indices, MRP, shown in Table 5 confirm this pattern. No significant changes in the polarization measure are detected in the first two years. The values of the indices in the 2002–2004 are instead positive and their magnitude is almost equivalent to the corresponding magnitude for indices estimated on the whole period. However, a shift away from the median



a *Household income gross of imputed rents*

b *Household income gross of imputed rents*

c *Household income net of imputed rents*

d *Household income net of imputed rents*

Fig. 4 Temporal decomposition of the shape effects. Authors' calculation on weighted household income data from SHIW. Income data are household size-adjusted and expressed in 2000 prices. *Dotted lines* represent 95% confidence intervals

Table 5 Temporal decomposition of polarization indices (2000–2004)

	Disposable income								
	Gross of IR					Net of IR			
	Index	CI (95%)		*p* val.		Index	CI (95%)		*p* val.
2000–2002									
MRP	0.011	−0.007	0.029	0.12	MRP	0.012	−0.006	0.030	0.10
LRP	0.095	0.068	0.121	0.00	LRP	0.080	0.054	0.106	0.00
URP	−0.073	−0.098	−0.048	0.00	URP	−0.056	−0.081	−0.031	0.00
2002–2004									
MRP	0.047	0.029	0.066	0.00	MRP	0.042	0.024	0.060	0.00
LRP	0.120	0.094	0.147	0.00	LRP	0.135	0.109	0.161	0.00
URP	−0.025	−0.050	−0.001	0.02	URP	−0.051	−0.076	−0.026	0.00

Authors' calculation on weighted household income data from SHIW. Income data are size-adjusted and expressed in 2000 prices. The *p* values refer to the null hypothesis that polarization does not change. MRP is the median relative polarization index; LRP is the lower relative polarization index; URP is the upper relative polarization index

for the lower incomes as well as a squeeze of the upper tail towards the median is also perceptible in the first period.

5 Concluding remarks

We have used the relative density method to analyse changes in the Italian household income distribution between 2000 and 2004. In contrast to methods that rely on summary statistics, this nonparametric method uses all information about the shape of the distributions. We are able to document important changes in the income distribution, despite substantial stability in income inequality and polarization measures. The analysis of the size-adjusted household incomes, gross of imputed rents, shows a dominant location effect, an overall upshift of the distribution, that masks a tendency to polarization in household incomes. In fact, having controlled for the median increase, a rise in polarization is detected, mainly due to a downgrading of lower incomes. The analysis of incomes net of imputed rents shows a somewhat more complex picture: the overall rightward shift of the distribution does not offset a fall in the lower-middle mass. In this sense, while it is not appropriate to talk about a general "household impoverishment", this evidence indicates a clear dwindling of the middle-income class. The distribution of imputed rents obscures significant distributional changes occurred in the 2000s. This is due to the fact that in Italy home ownership is quite homogeneous across household income deciles. The process of polarization of incomes net of imputed rent is more pronounced, with the lower tail contributing positively and the upper negatively, but the net overall polarization is approximately equivalent in size to the one estimated for incomes gross of imputed rents. Changes that we find are concentrated over the period 2002–2004. A within group analysis, according to the employment status of the household head, shows that all groups experienced greater polarization during this period. Similar patterns in both the tails of the distributions are detected for households whose head is

employed or retired - a greater polarization in the lower tail and a convergence in the upper tail. By contrast, the more polarized distribution within households whose head is self-employed is due to a shift away from the median of both tails.

Acknowledgements This paper has benefited from helpful comments of two anonymous referees, Paul Johnson, Andrea Brandolini, participants of 2nd meeting of the Society for the Study of Economic Inequality, CEIS seminar and DSE-ISFOL workshop. Responsibility for remaining errors is ours.

References

1. Baldini M.: Le molte conferme sulla povertà in Italia. www.lavoce.info, 13 October (2005)
2. Banca d'Italia: Italian Household Budgets in 2004, Supplements to the Statistical Bulletin, Banca d'Italia, XVI, No. 17 (2006)
3. Boeri, T., Brandolini, A.: The age of discontent: Italian households at the beginning of the decade. G. Econ. Ann. Econ. **63**(3/4), 449–487 (2004)
4. Bernhardt, A., Morris M., Handcock, M.S.: Women's gains or men's losses? A closer look at the shrinking gender gap in earnings. Am. J. Sociol. **101**(2), 302–328 (1995)
5. Brandolini, A.: The distribution of personal income in post-war Italy: source description, data quality and the time pattern of income inequality. G. Econ. Ann. Econ. **58**(2), 183–239 (1999)
6. Ćwik, J., Mielniczuck, J.: Data-dependent bandwidth choice for a grade density kernel estimate. Stat. Probab. Lett. **16**, 397–405 (1993)
7. Duclos, J.Y., Esteban, J., Ray, D.: Polarization: concepts, measurement, estimation. Econometrica **72**, 1737–1772 (2004)
8. Frick, J.R., Grabka, M.M.: Imputed rent and income inequality: a decomposition analysis for Great Britain, West Germany and the U.S. Rev. Income Wealth **49**(4), 513–537 (2003)
9. Handcock, M.S.: Reldist: Relative Distribution Methods. R package version 1.5-3, http://www.stat.washington.edu/~handcock/RelDist (2005)
10. Handcock, M.S., Morris, M.: Relative distribution methods. Sociol. Method. **28**, 53–97 (1998)
11. Handcock, M.S., Morris, M.: Relative Distribution Methods in the Social Sciences. Springer-Verlag, New York (1999)
12. Jenkins, S.P., Van Kerm, P.: Accounting for income distribution trends: a density function decomposition approach. J. Econ. Inequal. **3**, 43–61 (2005)
13. Molanes-López, E.M., Cao, R.: Plug-in bandwidth selectors for the kernel relative density estimator. Ann. Inst. Stat. Math. (2008) (forthcoming)
14. Morris, M., Bernhardt, A.D., Handcock, M.S.: Economic inequality: new methods for new trends. Am. Sociol. Rev. **59**, 205–219 (1994)
15. Paiella, M.: Does Wealth Affect Consumption? Evidence for Italy. Banca d'Italia, Temi di discussione No. 510, Roma (2004)
16. Pittau, M.G., Zelli, R.: Testing for changing shapes of income distribution: Italian evidence in the 1990s from kernel density estimates. Empir. Econ. **29**, 415–330 (2004)
17. Pittau, M.G., Zelli, R.: Trends in income distribution in Italy: a non-parametric and a semi-parametric analysis. J. Income Distrib. **15**, 90–118 (2006)
18. Sheather, S.J., Jones, M.C.: A Reliable data-based bandwidth selection method for kernel density estimation. J. R. Stat. Soc. Ser. B **53**, 683–690 (1991)
19. Smeeding, T.M., Saunders, P., Coder, J., Jenkins, S., Fritzell, J., Hagenaars, A.J.M., Hauser, R., Wolfson, M.: Poverty, inequality, and family living standards impacts across seven nations: the effects of non-cash subsidies for health, education and housing. Rev. Income Wealth **39**(3), 229–256 (1993)
20. Trivellato, U.: Il monitoraggio della povertà e della sua dinamica: questioni di misura ed evidenze empiriche. Statistica **58**(4), 549–574 (1998)
21. Voitchovsky, S.: Does the profile of income inequality matter for economic growth? J. Econ. Growth **10**, 273–296 (2005)
22. Yates, J.: Imputed rent and income distribution. Rev. Income Wealth **40**(1), 43–66 (1994)