Volume 33, Issue 3

Income polarization in Brazil, 2001-2011: A distributional analysis using PNAD data

Fabio Clementi  
*University of Macerata*

Francesco Schettino  
*Second University of Naples*

Abstract

This paper applies a non-parametric tool, the "relative distribution", to identify patterns of changes in Brazil’s household income distribution over the period 2001-2011. Despite the sharp decline in income inequality recently experienced by the country, we are able to document an increased income polarization, which has particularly affected households below the median. The results call directly into question the future sustainability and equity of existing social programs dealing with the unequal distribution of resources.
1. Introduction

Brazil has long been known as one of the countries with the most unequal income distribution in the world. The concentration of incomes in 1960 was already high by international standards, as indicated by a Gini coefficient of 0.504, and continued to increase in the following decades (López-Calva, 2012). Income inequality only declined starting in the mid-1990s: after 1997, the Gini reduced by 0.8% per year; between 2001 and 2007, the average rate of annual decline accelerated to 1.2%, well above the pace of the Latin American region as a whole (Barros et al., 2010). Poverty in the country also declined significantly during the last decade: the absolute number of poor people fell from over 61 million in 2003 to under 40 million in 2009 and the headcount index from 35.8% to 21.4% (Higgins, 2012). Meanwhile, Brazil’s GDP growth managed to overtake the UK as the world’s sixth-largest economy in 2011 (CEBR, 2011).

Although several factors contributed to the recent progress in terms of poverty and inequality reduction—such as economic growth (Barros et al., 2010), expanded access to education during the 1990s (Gasparini & Lustig, 2011), increased demand for unskilled labour (Robinson, 2010) and an increase in the minimum wage (Barros, 2007), it is common opinion that the conditional cash transfer (CCT) programs consolidated and expanded under the administration of the former Brazilian president Luiz Inácio Lula da Silva (2003–2010) have also played an important role. Notwithstanding many critical remarks—focusing principally on the high related costs, CCTs received the appreciation by international institutions and were enthusiastically embraced by many countries as a major social policy instrument (Hall, 2006). “Bolsa Família” is now the largest such scheme in the world: it was budgeted at R$8.3 billion (equivalent to almost 0.4% of GDP) in 2006 and covered around 11 millions families (approximately 46 millions people) over the same year (Lindert et al., 2007). As a result of their excellent targeting, the program’s benefits accounted for something between 21% and 16% of the total fall in Brazilian inequality since 2001 (Soares, 2012). The decline in inequality has been crucial for poverty reduction (accounting for half of the total change between 2001 and 2009) and certainly for making growth friendlier to the poor (López-Calva & Rocha, 2012).

The recent trend in terms of inequality changes is unique in respect to what is being experienced in Brazil’s fellow BRICS countries: Russia, India, China and South Africa (OECD, 2011). However, while there is a substantial literature on inequality and income distribution in Brazil (both in isolation and in a comparative perspective; see e.g. World Bank, 2004, and references therein) relatively little work has been done in terms of analyzing changes in the shape of Brazil’s income distribution in the recent decade. Indeed, the above mentioned evidence heavily relies on summary measures of inequality and not on the whole shape of the income distribution. As noted by Morris et al. (1994), standard measures of inequality may suggest a particular outcome in terms of inequality change—e.g. a fall in the Gini coefficient or Theil index—while implying a radically different pattern of distributional change.

---

1CCTs are direct monetary transfers provided to poor families on condition that they ensure children and adolescents attend school and that they meet basic health care requirements. These conditions attempt both to reduce short-term poverty by direct cash transfers and to fight long-term poverty by investing in the human capital of the poor (see e.g. Fiszbein et al., 2009.)

2For a review see in particular Coggiola (2010).
In particular, they may not capture aspects such as multi-modality and polarization.

In seeking to understand exactly how income inequality fell in Brazil over the last decade, in the present study we look “behind” the usual summary measures and closely examine the patterns of changes that have occurred along the entire Brazilian income distribution. More specifically, it is our aim to investigate whether the favourable combination of economic growth and inequality reduction—from which the country has benefited during the last 15 years or so—has produced significant movements across the income scale, and whether these movements have taken the form of a convergence of the top and bottom percentiles toward the middle income class or of a shrinking of the latter—thereby leading to greater distributional polarization. For this purpose, we apply a non-parametric tool, the “relative distribution”, to survey income data (PNAD) spanning 2001–2011 and covering a large number of households across all federal units of Brazil.

The remainder of this paper is structured as follows: Section 2 is devoted to the illustration of the main features of income data; Section 3 reviews the relative distribution method; Section 4 details the results and findings; Section 5 concludes and draws some policy implications.

2. Data and summary statistics

We use data from Brazil’s annual national household survey (Pesquisa Nacional por Amostra de Domicílios, PNAD) for 2001 to 2011. The PNAD is collected every year in September—except in 2010—by the National Census Bureau (Instituto Brasileiro de Geografia e Estatística, IBGE) and is nationally representative at the level of each state. However, until 2003 the PNAD was not representative for the rural areas of the North region (minus the state of Tocantins). Therefore, in order to maintain time-series comparable these areas were excluded from PNAD data for 2004 onward. In this way, our samples have on average about 107,000 observations a year.

All calculations are based on total household income expressed in Brazilian Reais (R$). Current values have been deflated using the consumer price index (yearly series based on 2005) reported by the OECD. Furthermore, incomes have been equivalized for differences in household size and weighted by using appropriate sampling weights provided by the IBGE staff.

Table 1 provides summary measures for annual household income from 2001 to 2011. Besides the growth of the real mean and median incomes, the most notable feature is that income shares of the poorest percentiles of the population increased on average between approximately 2% and 3% per year in the period examined, on the contrary of what observed for the richest percentiles whose shares decreased by around 1% or more. As for inequality, the improvements were also noticeable: the Gini and Theil indices exhibited nearly the same

---


4Available at http://stats.oecd.org/.

5Here we adopt a simple equivalence scale that is most commonly used in international studies (e.g. Atkinson et al., 1995) where total household income is divided by the square root of the number of household members.
temporal profile, showing an average yearly decrease that amounts respectively to 1% and 2%.

All these figures seem to be consistent with the positive (and sometimes enthusiastic) evaluation of Brazil’s economic and social policies under the former President Lula’s administration. But while suggesting important candidate explanations for the distributional change, the statistics reported do not capture the other changes that might have occurred. In particular, the key questions are hinted at but not easily quantified using the standard measures here. How well are the differences captured by simple location shifts? Is there evidence of growing polarization? Are the upper and lower tails of the distributions changing in similar ways? As mentioned in the previous section, in this paper we propose an application based on the “relative distribution”, a non-parametric statistical tool for fully representing differences between distributions that we deem well-suited to deal with the above questions.

3. Assessing income polarization in Brazil: the relative distribution approach

Over the last two decades, the issue of polarization has come to be assigned increasing importance in the analysis of income distribution. Notwithstanding the pains the polarization literature has suffered to distinguish itself from pure inequality measurement—see e.g. Foster & Wolfson (1992), Levy & Murnane (1992), Esteban & Ray (1994) and Wolfson (1994, 1997), it now seems to be fairly widely accepted that polarization is a distinct concept from inequality.

Broadly speaking, the notion of polarization is concerned with the disappearance of the middle class, which occurs when there is a tendency to concentrate in the tails—rather than the middle—of the income distribution. One of the main reasons for looking at income polarization this way, which is usually referred to as “bi-polarization”, is that a well-off middle class is important to every society because it contributes significantly to economic growth, as well as to social and political stability (e.g. Pressman, 2007). In contrast, a society with high degree of income polarization may give rise to social conflicts and tensions. Therefore, in order for such risks to be minimized, it is necessary to monitor the economic evolution of the society using indices that look at the dispersion of the income distribution.

Table 1 Summary measures of Brazilian household income, 2001–2011

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>874.7</td>
<td>879.8</td>
<td>837.6</td>
<td>851.1</td>
<td>883.5</td>
<td>940.3</td>
<td>969.4</td>
<td>1,017.3</td>
<td>1,034.4</td>
<td>1,083.9</td>
</tr>
<tr>
<td>Median</td>
<td>462.7</td>
<td>467.2</td>
<td>458.5</td>
<td>480.9</td>
<td>500.0</td>
<td>543.0</td>
<td>570.6</td>
<td>613.4</td>
<td>627.1</td>
<td>672.7</td>
</tr>
<tr>
<td>Income shares</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bottom 5%</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Bottom 10%</td>
<td>1.2</td>
<td>1.2</td>
<td>1.2</td>
<td>1.3</td>
<td>1.4</td>
<td>1.3</td>
<td>1.5</td>
<td>1.4</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>Bottom 20%</td>
<td>3.2</td>
<td>3.3</td>
<td>3.4</td>
<td>3.6</td>
<td>3.8</td>
<td>3.8</td>
<td>3.9</td>
<td>4.0</td>
<td>4.0</td>
<td>4.3</td>
</tr>
<tr>
<td>Top 20%</td>
<td>61.1</td>
<td>60.8</td>
<td>60.0</td>
<td>59.0</td>
<td>58.8</td>
<td>58.3</td>
<td>57.4</td>
<td>56.9</td>
<td>56.3</td>
<td>55.4</td>
</tr>
<tr>
<td>Top 10%</td>
<td>44.8</td>
<td>44.5</td>
<td>43.6</td>
<td>42.7</td>
<td>42.8</td>
<td>42.4</td>
<td>41.4</td>
<td>41.0</td>
<td>40.5</td>
<td>39.8</td>
</tr>
<tr>
<td>Top 5%</td>
<td>31.5</td>
<td>31.1</td>
<td>30.5</td>
<td>29.9</td>
<td>29.8</td>
<td>29.6</td>
<td>28.8</td>
<td>28.5</td>
<td>28.2</td>
<td>27.7</td>
</tr>
<tr>
<td>Inequality metrics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini</td>
<td>0.562</td>
<td>0.557</td>
<td>0.549</td>
<td>0.538</td>
<td>0.535</td>
<td>0.529</td>
<td>0.520</td>
<td>0.514</td>
<td>0.509</td>
<td>0.498</td>
</tr>
<tr>
<td>Theil</td>
<td>0.630</td>
<td>0.626</td>
<td>0.594</td>
<td>0.577</td>
<td>0.572</td>
<td>0.560</td>
<td>0.537</td>
<td>0.525</td>
<td>0.519</td>
<td>0.495</td>
</tr>
</tbody>
</table>

Authors’ calculation on weighted household income data from PNAD.
from the middle toward either or both of the two tails. Measures of income polarization that correspond to this case have been proposed in the literature by Foster & Wolfson (1992), Wolfson (1994, 1997), Wang & Tsui (2000), Chakravarty & Majumder (2001), Rodríguez & Salas (2003), Chakravarty et al. (2007), Silver et al. (2007), Chakravarty (2009), Chakravarty & D'Ambrosio (2010), Lasso de la Vega et al. (2010), and others.

A more general notion of income polarization, which was originally proposed by Esteban & Ray (1994), regards the latter as “clustering” of a population around two or more poles of the distribution, irrespective of where they are located along the income scale. The notion of income polarization in a multi-group context is an attempt at capturing the degree of potential conflict inherent in a given distribution (see Esteban & Ray, 1999, 2011). The idea is to consider society as an amalgamation of groups, where the individuals in a group share similar attributes with the other members (i.e. have a mutual sense of “identification”) but in terms of the same attributes they are different from the members of the other groups (i.e. have a feeling of “alienation”). Political or social conflict is therefore more likely the more homogenous and separate the groups are, that is when the within-group income distribution is more clustered around its local mean and the between-group income distance is longer. In addition to Esteban & Ray (1994), indices regarding the concept of income polarization as conflict among groups have been investigated, among others, by Gradin (2000), Milanovic (2000), D’Ambrosio (2001), Zhang & Kanbur (2001), Reynal-Querol (2002), Duclos et al. (2004), Lasso de la Vega & Urrutia (2006), Esteban et al. (2007), Gigliarano & Mosler (2009) and Poggi & Silber (2010).

Much of the literature so far considered has analyzed summary measures of income polarization. Another strand uses kernel density estimation and mixture models in order to describe changes in polarization patterns over time, not just of personal incomes (as in Jenkins, 1995, 1996, Pittau & Zelli, 2001, 2004, 2006, and Conti et al., 2006) but also of the cross-country distribution of per capita income (see Quah, 1996a, 1996b, 1997, Bianchi, 1997, Jones, 1997, Paap & van Dijk, 1998, Johnson, 2000, Holzmann et al., 2007, Henderson et al., 2008, Pittau et al., 2010, Anderson et al., 2012, and others). The analysis of the income distribution shape provides indeed a picture from which at least three important distributional features can be observed simultaneously (Cowell et al., 1996): income levels and changes in the location of the distribution as a whole; income inequality and changes in the spread of the distribution; clumping and polarization as well as changes in patterns of clustering at

---

6More precisely, there are two characteristics that are considered as being intrinsic to the notion of bi-polarization. The first one, “increased spread”, implies that moving from the central value (median) to the extreme points of the income distribution makes the distribution more polarized than before. In other words, increments (reductions) in incomes above (below) the median will widen the distribution, that is extend the distance between the groups below and above the median and hence increase the degree of bi-polarization. On the other hand, “increased bi-polarity” refers to the case where incomes on the same side of the median get closer to each other. Since the distance between the incomes below or above the median has been reduced, this is assumed to increase bi-polarization. Thus, bi-polarization involves both an inequality-like component, the “increased spread” principle, which increases both inequality and polarization, and an equality-like component, the “increased bi-polarity” criterion, which increases polarization but lowers any inequality measure that fulfills the Pigou-Dalton transfer principle—the requirement under which inequality decreases when a transfer is made from a richer to a poorer individual without reversing their pairwise ranking. This shows that although there is complementarity between polarization and inequality, there are differences as well. See the references cited in the main text for a thorough discussion.
different modes. Finally, a rather recent (yet non-parametric) approach that combines the strengths of summary polarization indices with the details of distributional change offered by the kernel density estimates—the so-called “relative distribution”—has been employed by Alderson et al. (2005), Massari (2009), Massari et al. (2009a, 2009b), Borraz et al. (2011) and Alderson & Doran (2011, 2013) to assess the evolution of the middle class and the degree of household income polarization in a number of middle- and high-income countries in the world.

Recalling the objectives of our study, stated in Section 1, we deem that in the current application the relative distribution approach has some important advantages over the other mentioned methods of investigating income polarization. First, it readily lends itself to simple and informative graphical displays of relative data that reveal precisely where and by how much an income distribution changed over time. Second, by providing the potential for decomposition into location and shape components, it allows one to examine several hypotheses regarding the origins of distributional change—such as whether the change was due to a proportional variation in all incomes that moved the overall distribution either back or forth (while leaving the shape unaltered) or to shape modifications which, by definition, are independent of location shifts. Lastly, it allows to quantify the degree of polarization due to changes in distributional shape only (i.e. net of location shifts), thus enabling one to isolate aspects of inter-distributional inequality that are often hidden when also changes in location are examined.

Basically, the relative distribution method can be applied whenever the distribution of some quantity across two populations is to be compared, either cross-sectionally or over time. To proceed, it is necessary to single out one of the two populations, refer to it as the “comparison” population, and refer to the other as the “reference” population. More formally, let \( Y_0 \) be the income variable for the reference population (e.g. households in 2001) and \( Y \) the income variable for the comparison population (e.g. households in 2011). The relative distribution of \( Y \) to \( Y_0 \) is defined as the distribution of the random variable:

\[
R = F_0(Y),
\]

which is obtained from \( Y \) by transforming it by the cumulative distribution function of \( Y_0 \), \( F_0 \). As a random variable, \( R \) is continuous on the outcome space \([0, 1]\), and its realizations, \( r \), are referred to as “relative data”. Intuitively, the relative data can be interpreted as the set of positions that the income observations of the comparison population would have if they were located in the income distribution of the reference population. The probability density function of \( R \), which is called the “relative density”, can be obtained as the ratio of the density of the comparison population to the density of the reference population evaluated at

---

7Of course, both the location and shape effects—named respectively as “growth” and “inequality” (or “distributional”) effect (Kakwani, 1993; Bourguignon, 2003, 2004)—may also concur together in producing the distributional change.

8Here we limit ourselves to illustrating the basic concepts behind the use of the relative distribution method. For a systematic introduction we refer the reader to Handcock & Morris (1998, 1999)—see also Hao & Naiman (2010, ch. 5). A method very similar in spirit has recently been presented by Silber & Deutsch (2012).
the relative data $r$:

$$g(r) = \frac{f(F_0^{-1}(r))}{f_0(F_0^{-1}(r))} = \frac{f(y_r)}{f_0(y_r)}, \quad 0 \leq r \leq 1, \quad y_r \geq 0,$$

where $f(\cdot)$ and $f_0(\cdot)$ denote the density functions of $Y$ and $Y_0$, respectively, and $y_r = F_0^{-1}(r)$ is the quantile function of $Y_0$.\(^9\) The relative density has a simple interpretation, as it describes where households at various quantiles in the comparison distribution are concentrated in terms of the quantiles of the reference distribution. As for any density function, it integrates to 1 over the unit interval, and the area under the curve between two values $r_1$ and $r_2$ is the proportion of the comparison population whose income values lie between the $r_1^{th}$ and $r_2^{th}$ quantiles of the reference population.

When the relative density function shows values near to 1, it means that the two populations have a similar density at the $r^{th}$ quantile of the reference population, and thus $R$ has a uniform distribution in the interval $[0, 1]$. A relative density greater than 1 means that the comparison population has more density than the reference population at the $r^{th}$ quantile of the latter. Finally, a relative density function less than 1 indicates the opposite. In this way one can distinguish between growth, stability or decline at specific points of the income distribution.

As we have said before, one of the major advantages of this method is the possibility to decompose the relative distribution into changes in location, usually associated with changes in the median (or mean) of the income distribution, and changes in shape (including differences in variance, asymmetry and/or other distributional characteristics) that could be linked with several factors like, for instance, polarization. Formally, the decomposition can be written as:

$$g(r) = \frac{f(y_r)}{f_0(y_r)} = \frac{f_0L(y_r)}{f_0(y_r)} \times \frac{f(y_r)}{f_0L(y_r)}, \quad 0 \leq r \leq 1, \quad y_r \geq 0,$$

where $f_0L(y_r) = f_0(y_r + \rho)$ is a density function adjusted by an additive shift with the same shape as the reference distribution but with the median of the comparison one.\(^10\) The value $\rho$ is the difference between the medians of the comparison and reference distributions. If the latter two distributions have the same median, the density ratio for location differences is uniform in $[0, 1]$. Conversely, if the two distributions have different median, the “location effect” is increasing (decreasing) in $r$ if the comparison median is higher (lower) than the reference one. The second term, which is the “shape effect”, represents the relative density

\(^9\)The income density functions are estimated using a non-parametric kernel method. Once we obtain the relative density functions for different realizations of $R$, we fit a local polynomial to the estimated data points in order to have an accurate description of the relative density. Throughout, we rely on the \texttt{R} statistical package \texttt{reldist} (Handcock, 2011) to implement the relative distribution method.

\(^10\)Median adjustment is preferred here to mean adjustment because of the well-known drawbacks of the mean when distributions are skewed. A \textit{multiplicative} median shift can also be applied. However, the multiplicative shift has the drawback of affecting the shape of the distribution. Indeed, the equi-proportionate income changes increase the variance and the rightward shift of the distribution is accompanied by a flattening (or shrinking) of its shape (see e.g. Jenkins & Van Kerm, 2005).
net of the location effect and is useful to isolate movements (re-distribution) occurred between the reference and comparison populations. For instance, we could observe a shape effect function with some sort of (inverse) U-shaped pattern if the comparison distribution is relatively (less) more spread around the median than the location-adjusted one. Thus, it is possible to determine whether there is polarization of the income distribution (increases in both tails), “downgrading” (increases in lower tail), “upgrading” (increases in the upper tail) or convergence of incomes towards the median (decreases in both tails).

This approach also includes a median relative polarization index (MRP), which is based on changes in the shape of the income distribution to account for polarization. This index is normalized so that it varies between -1 and 1, with 0 representing no change in the income distribution relative to the reference year. Positive values represent more polarization—i.e. increases in the tails of the distribution—and negative values represent less polarization—i.e. convergence towards the center of the distribution. The MRP index for the comparison population can be estimated as (Morris et al., 1994, p. 217):

\[
\text{MRP} = \frac{4}{n} \left( \sum_{i=1}^{n} \left| r_i - \frac{1}{2} \right| \right) - 1, \tag{4}
\]

where \( r_i \) is the proportion of the median-adjusted reference incomes that are less than the \( i^{th} \) income from the comparison sample, for \( i = 1, \ldots, n \), and \( n \) is the sample size of the comparison population.

The MRP index can be additively decomposed into the contributions to overall polarization made by the lower and upper halves of the median-adjusted relative distribution, enabling one to distinguish downgrading from upgrading. In terms of data, the lower relative polarization index (LRP) and the upper relative polarization index (URP) can be calculated as follows:

\[
\text{LRP} = \frac{8}{n} \left[ \sum_{i=1}^{n/2} \left( \frac{1}{2} - r_i \right) \right] - 1, \tag{5}
\]

\[
\text{URP} = \frac{8}{n} \left[ \sum_{i=n/2+1}^{n} \left( r_i - \frac{1}{2} \right) \right] - 1, \tag{6}
\]

with \( \text{MRP} = \frac{1}{2} (\text{LRP} + \text{URP}) \). As the MRP, LRP and URP range from -1 to 1, and equal 0 when there is no change.

4. Empirical results

Figure 1(a) presents kernel density estimates of total household income at the two end points of the 2001–2011 period.\textsuperscript{11} At first glance, we observe a rightward shift of the whole

\textsuperscript{11}To handle data sparseness, the two densities have been obtained by using an adaptive kernel estimator with a Silverman’s plug-in estimate for the pilot bandwidth (see e.g. Van Kerm, 2003). The advantage of this estimator is that it does not over-smooth the distribution in zones of high income concentration, while keeping the variability of the estimates low where data are scarce—as, for example, in the highest income ranges.
Figure 1 Changes in the Brazilian household income distribution between 2001 and 2011. In panel (a), incomes in the upper tiers of the distributions have been truncated for better presentation of the graph, where the vertical lines denote the medians of the two samples. The dotted lines in panels (b)–(d) denote the 95% pointwise confidence limits based on the asymptotic normal approximation (Handcock & Morris, 1999, p. 144)

distribution, which implies an increase of the median income in this period. The increment in the median can be explained by the substantial decline in the mass at the lower and middle income ranges and the concomitant spreading out of incomes in the top half of the distribution. There is also a significant alteration of the shape, especially in the middle income range: the 2011 distribution reveals indeed clear evidence of multi-modality, while the 2001 distribution is nearly unimodal. As shown in Pittau & Zelli (2006), the emergence of more than one mode (and the gap between them) in the distribution of household income could be interpreted as an increase in polarization.

Further insight is provided by the relative density, which directly compares the two distributions. Figure 1(b) shows the fraction of households in 2011 that fall into each percentile
of the 2001 income distribution. Changes in the distribution are indicated by the generally positive slope of the relative density, which implies a decrease of the mass of households below the 2001 median income over the period under consideration. More specifically, the relative distribution is less than 1 for $r \leq 0.44$ and more than 1 for $r > 0.44$. This means that if we choose any percentile between the 1st and the 44th in the 2001 distribution, the fraction of households in 2011 that earn an amount of income corresponding to the chosen percentile is less than the analogous fraction of households in 2001. However, income growth between 2001 and 2011 also positively affected households in the top half of the distribution: the peak of 1.6 is at around the 75th percentile, meaning that households in 2011 are approximately 60% more likely to fall at the level of 2001 income corresponding to the 75th percentile than households in 2001.

To get a more detailed picture, we decompose the relative density into location and shape effects according to Equation (3). Figure 1(c) presents 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 location shift of the density. Since the median shift is positive, the location effect reduces the share of households in bottom percentiles and increases that in the higher ones, hence confirming our prior observation. Figure 1(d) shows the shape effect, which represents the relative density net of the median influence. The visual impression that one gets from the figure above indicates a marked change for incomes below the median, with a decline of the mass between approximately the 9th and the 63rd percentile and a prominent increase of the fraction of households at the poorest decile of the distribution. This means that while the vast majority of households experienced a growth in their real income, the poorest fraction of them failed to catch up with the rest of population. On the contrary, the upper part of the relative density does not reveal significant changes, apart from a slight increase of the mass between the 67th and the 76th percentile and an almost equal increase at the very top income range from the 92nd percentile onward.

The relative distribution method permits us to also analyze how income re-distribution across households took place during 2001 to 2011. For each year within this period, Figure 2(a) shows the shape effect of the household income relative densities using 2001 as the reference year. Following the plot through each successive year, one is offered with the immediate impression that the fraction of households in the bottom income levels increased consistently by the mid-2000s, while the fraction in the middle and the upper declined. However, toward the end of the first decade of 2000s a moderate growth in upper income levels is also apparent, which indicates that the distribution is beginning to polarize.

A link between what we have observed in the graphical analysis and the quantification of the degree of polarization is yielded by the relative polarization indices. These indices can be used to keep track of changes in the shape of the income distribution across the whole 2001–2011 period by measuring the magnitude and direction of differences between

---

12We have chosen 2001 as the reference year throughout the analysis. Choosing an alternative reference year would change the view provided by the relative distribution graphs and the value of the relative polarization indices in each year, but it would not affect the year-to-year comparisons that are of interest here (Morris et al., 1994, p. 210). Furthermore, using 2001 as the reference year allows us to examine the longest span available in the PNAD series for Brazil.

13The relative distribution, and therefore its shape effect, is by definition flat in the reference year (Morris et al., 1994, p. 211).
the distribution in each successive year and that of the 2001 reference year. Figure 2(b) plots the set of three indices computed from the PNAD data using Equations (4)–(6). The MRP index falls initially, indicating a small, though significant convergence in incomes during the early 2000s. After this, however, the index rises consistently, becoming significantly positive after 2005. The LRP and URP indices track the MRP pretty closely over the entire time period. The LRP index falls until 2003, indicating significant convergence from the lower part of the income distribution toward the middle. By 2004 the index climbs almost steeply (and consistently) until the end of the time span, although it does not reach significance immediately—because it is rising from lower levels—but only from 2005 onward. A similar story can be told for the URP index. Here too, the first part of the period brings some movement from the top half of the income distribution toward the middle: except for 2002 and 2006, the value of the URP remains significantly negative during the first half of the 2000s. Instead, by 2007 it begins to rise and becomes significantly positive, indicating that an upgrading in the distribution also took place in the second half of the 2000s.

The obtained evidence holds vis-à-vis a further check involving the polarization indices proposed by Foster & Wolfson (1992) and Duclos, Esteban, & Ray (2004). To avoid biased comparisons among different measures of polarization, in line with the approach used above—which requires that differences in location be removed for the relative polarization indices to be computed, we proceed by first estimating “growth-neutral” counter-factual distributions, i.e. what the distribution of income among households in any one year would have been if the median had remained a the level of the 2001 reference year; then, for each year in the period subject to analysis, we compute the values of the aforementioned indices using the estimated counter-factual. In this way, we are sure that the yearly comparisons of polarization levels are only affected by changes in the shape of the distribution and not

---

14By definition, the value of the three indices always equals 0 in the reference year (Morris et al., 1994, p. 209).
also by the growth process occurred over the same study period. Figure 3 shows visually the polarization estimates based on the Foster-Wolfson and the Duclos-Esteban-Ray indices, alongside with the 95% confidence intervals represented by vertical bars. The two indices have nearly the same temporal profile, even if the Duclos-Esteban-Ray index is less stable. As shown by the vertical bars, both measures decline significantly between 2002 and 2003. Thereafter, the Foster-Wolfson index rises continuously, and the rise is almost always statistically significant. The Duclos-Esteban-Ray index tracks the Foster-Wolfson pretty closely

Both the Foster-Wolfson and the Duclos-Esteban-Ray measures are relative indices of polarization, meaning that they remain invariant under equiproportionate changes in all incomes. Therefore, unless the growth process is distribution-neutral—i.e. all levels of income grow at the same rate, when using such measures the contribution of growth to the estimated level of income polarization might more than offset the contribution that is attributable to changing distribution, which is of major concern here. That is why we have decided to proceed the way we did. See e.g. Ravallion (2010) for a similar approach to the measurement of income polarization.

The Foster-Wolfson and the Duclos-Esteban-Ray measures, as well as their confidence intervals, have been estimated using the latest version of DASP, the Distributive Analysis Stata Package (Arnar & Duclos, 2013), which is freely available at http://dasp.ecn.ulaval.ca/.

Figure 3 Polarization measures for the income distribution of Brazilian households, 2001–2011. Vertical bars represent 95% confidence intervals. The Duclos-Esteban-Ray index has been computed with the polarization sensitivity parameter $\alpha$ set at 0.5.
over the period of rising polarization, even though the pairwise comparisons indicate that the change is now statistically significant only between 2007 and 2008 and from 2009 to 2011. The results seem therefore to portray similar tendencies as that depicted by polarization evaluated using measures based on the relative distribution.\(^{17}\)

In sum, rather being *solely* a story of declining inequality, the recent changes in Brazil’s income distribution bring about a story of polarization. In fact, we are able to document a downgrading trend around the mid-2000s and, by 2007, the emergence of a more marked pattern of polarization. The latter, however, is not symmetric, as the LRP index is always more positive than the URP, indicating more polarization in the lower than in the upper tail.

5. Concluding remarks

We have used the relative distribution approach to analyze changes in the Brazilian household income distribution between 2001 and 2011. This method provides a non-parametric framework for taking into account all of the distributional differences that could arise in the comparison of distributions over time. In this way, we are able to summarize multiple features of the income distribution that would not be detected easily from a comparison of standard measures of inequality.

The paper documents relevant changes in the Brazilian income distribution, despite the substantial falling off in income inequality. The analysis of size-adjusted household incomes indicates an overall upshift of the distribution, especially from 2005 onward, which partly masks a tendency to income polarization. In fact, having controlled for the median increase, a more clear rise in polarization is detected, mainly due to a downgrading of lower incomes that overcompensated the convergence of higher incomes toward the median. By contrast, starting from 2007 the process of polarization of household incomes is more pronounced, with both the lower and upper tails shifting away from the median of the distribution.

Overall, these findings suggest that the recent improvements in Brazil’s income distribution have been propelled mainly by the overall economic growth of the country, while social policy programs would have played a key role in affecting the shape of the distribution—leading to a greater polarization at both the top and bottom tails of the income distribution. The observed movements of households toward low and high incomes (and away from the middle) could be justified, on the one side, by deductions and exemptions on taxes that are granted as political privileges to landowners (rents) and financial capitalists (profits), and, on the other side, by the heavy reliance on indirect taxation that disproportionately burden the income of poor and middle-income households, who consequently bear a significant share of the total cost for social programs (e.g. Birdsall et al., 2008, ch. 4.).

\(^{17}\)While not shown here due to space constraints, the above-mentioned indices have also been calculated using the relative *median-adjusted* data, defined in terms of the comparison distribution in each successive year relative to the 2001 reference distribution, where the latter has been location-adjusted so as to make the medians of the comparison and reference distributions coincide—in short, the same data used to obtain the relative polarization indices. Once more, after initially experiencing a small (though significant) convergence in household incomes during the initial years, a marked polarizing trend emerges by the mid-2000s, and this trend climbs consistently and significantly to the end of the period under study.
Hence, sustaining reductions in both inequality and poverty by making them less growth-dependent represents a key challenge for Brazil going forward: as borne out by our results, under a scenario of poor performance growth the shape effect would be brought to prevail, thereby generating a more unequal society. Considering the recent halt in Brazil’s economic growth that followed the global economic crisis, this paper suggests adopting policies well targeted to a “real” re-distribution of resources, i.e. aimed at allowing structural improvements in the income distribution that go beyond the effects of economic growth. Among these, making the tax system somewhat more progressive by increasing the tax burden on the income of rich households (including business profits as well as financial and agricultural rents) would improve the overall distribution of income and, at the same time, free up precious resources for domestic demand (especially by the middle class). Furthermore, reform programs to alleviate the unequal distribution of land would grant to poorest households—in particular those living in the North and Northeast regions of Brazil—the necessary tools to get out of extreme poverty and consequently reduce their actual dependence on social transfers.\footnote{Brazil has one of the most unequal distribution of land in the world. The concentration of property in Brazil is so skewed that the largest 3.5% of landholdings represent 56% of total agricultural land (Hidalgo et al., 2010). The Gini coefficient of land inequality remained stable between 1967 and 1998, measuring around 0.84 in both the beginning and end of the period (Hoffmann, 1998). Since then, it increased to 0.856 in 1995 and 0.872 in 2006 (IBGE, 1997, 2009). Some regional differences exist, but land inequality in all regions is high when compared internationally (Hoffmann & Ney, 2010).}

The paper can be extended in several directions. Perhaps the most obvious extension is to examine how different sources of household income might have impacted the observed increase of income polarization. Also, the decomposition of the relative distribution according to covariates measured on households would allow one to detect the contribution of different household characteristics such as geographic location, gender, age and ethnicity to the observed changes. Due to the richness of data available from the PNAD and the many opportunities offered by the relative distribution approach, we are in a good position to readily expand our analysis in the near future.

References


