Income Polarization in Brazil, 2001–2011:
A Distributional Analysis Using PNAD Data

F. Clementi and F. Schettino

1Department of Political Science, Communication and International Relations, University of Macerata, Piazza G. Oberdan 3, 62100 Macerata, Italy
2Department of Law, Second University of Naples, Via Mazzocchi 5, 81055 S. Maria Capua Vetere, Italy

June 7, 2013
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, and continued to increase in the following decades (López-Calva, 2012).

Income inequality only declined starting in the mid-1990s; from 2001 on, inequality levels have fallen steadily (Barros et al., 2010).

Poverty in the country also declined significantly during the last decade (e.g., 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, it is common opinion that social assistance programs have played a crucial role (Hall, 2006).

“Bolsa Família”, now the largest such program in the world, accounted for something between 21% and 16% of the total fall in Brazilian inequality since 2001 (Soares, 2012).
Aim of the Work

- The mentioned evidence heavily relies on summary measures of inequality, but relatively little work has been done in terms of analyzing changes in the shape of Brazil’s income distribution over the recent decade.
- As pointed out by Morris et al. (1994; but see also Voitchovsky, 2005, and Pittau and Zelli, 2006), standard measures of inequality may suggest a particular outcome in terms of inequality change – e.g., a fall in the Gini coefficient – while implying a radically different pattern of distributional change; in particular, they may not capture aspects such as multi-modality and polarization.
- In investigating the recent inequality experience of the Brazilian society, we seek to understand “how” inequality fell by looking behind the usual summary measures and closely examining the actual pattern of distributional changes that have occurred along the entire Brazilian household income distribution.
- For this purpose, we use a non-parametric tool, the relative distribution, which is applied to survey income data (PNAD) spanning 2001–2011 and covering a large number of households across all federal units of Brazil.

The Data
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 (http://stats.oecd.org/).

Furthermore, incomes have been equivalized for differences in household size and weighted by using appropriate sampling weights provided by the IBGE staff.

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>

Source: authors’ calculation on weighted household income data from PNAD

Besides the growth of 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 temporal profile, showing an average yearly decrease that amounts respectively to 1% and 2%.
The Relative Distribution

Background

- Researchers and analysts have developed several summary measures for assessing income inequality (e.g., the Gini coefficient or Theil index).
- However, when used to make relative inequality inference these measures do not always tell the whole story, as comparisons based on a single summary statistic — reflecting an average of the varied effects of income inequality — are likely to mask underlying movements along the income scale that might lead to different economic outcomes in distinct parts of the distribution (e.g., Voitchovsky, 2005; Massari, 2009; Massari et al., 2009).
- The relative distribution is a non-parametric statistical approach introduced by Morris et al. (1994) and Handcock and Morris (1998, 1999) that compares the income (or other) distributions of two populations in a way to consider differences throughout the entire income range.
- It has a simple intuitive meaning and preserves all of the information necessary to compare two distributions.
Definition

- 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* is defined as the ratio of the density of the comparison population to the density of the reference population evaluated at the $r^{th}$ quantile of the reference distribution:

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

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 $Y_0$.
- When no changes occur between the two distributions, $g (r)$ has a uniform distribution; 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 $r^{th}$ quantile of the latter.

Location and Shape Decomposition

- One of the major advantages of this method is the possibility to decompose the relative distribution into changes in *location* and changes in *shape*.
- The decomposition can be written as:

$$\frac{f (y_r)}{f_0 (y_r)} = \frac{f_0 (y_r)}{f_0 (y_r)} \times \frac{f (y_r)}{f_0 (y_r)},$$

Overall
Location
Shape

- $f_0L (y_r)$ is the *median-adjusted* density function:

$$f_0L (y_r) = f_0 (y_r + \rho),$$

where the value $\rho$ is the difference between the medians of the comparison and reference distributions – alternative indices like the mean and/or multiplicative location shift can also be considered.
Distributional Polarization

- A distribution is said to be polarized if there is a tendency to concentrate in the tails rather than the middle (e.g., Wolfson, 1994; Foster and Wolfson, 2010).
- The relative distribution approach also includes a median relative polarization index, re-scaled in order to vary between -1 and 1:

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

- Positive values represent more income polarization and negative values represent less polarization; a value of 0 indicates no differences in distributional shape.
- The MRP index can be additively decomposed into the lower relative polarization index and the upper relative polarization index, which behave similarly as the MRP:

  \[
  MRP = \frac{1}{2} (LRP + URP).
  \]

Covariate Adjustment

- It is possible to adjust the relative distribution for changes in the distribution of a covariate:

  \[
  \frac{f(y)}{f_0(y)} = \frac{f_{0C}(y)}{f_0(y)} \times \frac{f(y)}{f_{0C}(y)}.
  \]

- \( f_{0C}(y) \) is the composition-adjusted density function:

  \[
  f_{0C}(y) = \sum_{k=1}^{K} \pi_k f_{Y_0|Z_0}(y|k),
  \]

  which has the composition of the comparison population but retains the conditional densities of the reference population.
- The composition effect detects the impact of changes in population composition; the residual component reveals changes in the covariate-outcome relationship.
Empirical Results

Changes in Household Income Distribution

- There is a rightward shift of the whole distribution and a change of the shape, especially in the middle income range, from 2001 to 2011.
There is a rightward shift of the whole distribution and a change of the shape, especially in the middle income range, from 2001 to 2011.

The relative distribution is nearly monotonic in its increase, hence implying a decrease of the mass at the lower and middle income ranges and a concomitant spreading out of incomes in the top half of the distribution.
There is a rightward shift of the whole distribution and a change of the shape, especially in the middle income range, from 2001 to 2011.

The relative distribution is nearly monotonic in its increase, hence implying a decrease of the mass at the lower and middle income ranges and a concomitant spreading out of incomes in the top half of the distribution.

Since the median shift is positive, the location effect reduces the share of households in bottom deciles and increases that in the higher ones.
There is a rightward shift of the whole distribution and a change of the shape, especially in the middle income range, from 2001 to 2011.

The relative distribution is nearly monotonic in its increase, hence implying a decrease of the mass at the lower and middle income ranges and a concomitant spreading out of incomes in the top half of the distribution.

Since the median shift is positive, the location effect reduces the share of households in bottom deciles and increases that in the higher ones.

The shape effect indicates a marked change for incomes below the median, with a prominent increase of the fraction of households at the poorest decile of the distribution, and a moderate income growth in the upper part.
There is a rightward shift of the whole distribution and a change of the shape, especially in the middle income range, from 2001 to 2011.

The relative distribution is nearly monotonic in its increase, hence implying a decrease of the mass at the lower and middle income ranges and a concomitant spreading out of incomes in the top half of the distribution.

Since the median shift is positive, the location effect reduces the share of households in bottom deciles and increases that in the higher ones.

The shape effect indicates a marked change for incomes below the median, with a prominent increase of the fraction of households at the poorest decile of the distribution, and a moderate income growth in the upper part.

The fraction of households in the bottom income levels increased consistently by the mid-2000s, while a moderate growth in upper income levels is only apparent toward the end of the decade.
Changes in Household Income Distribution

- There is a rightward shift of the whole distribution and a change of the shape, especially in the middle income range, from 2001 to 2011.
- The relative distribution is nearly monotonic in its increase, hence implying a decrease of the mass at the lower and middle income ranges and a concomitant spreading out of incomes in the top half of the distribution.
- Since the median shift is positive, the location effect reduces the share of households in bottom deciles and increases that in the higher ones.
- The shape effect indicates a marked change for incomes below the median, with a prominent increase of the fraction of households at the poorest decile of the distribution, and a moderate income growth in the upper part.
- The fraction of households in the bottom income levels increased consistently by the mid-2000s, while a moderate growth in upper income levels is only apparent toward the end of the decade.
- The relative polarization indices document a downgrading trend around the mid-2000s and, by 2007, the emergence of a more marked pattern of polarization.
To further interpret the tendency of Brazilian household incomes to polarize, we analyze the changes that occurred in the conditional distributions by region.

We follow the IBGE’s division of Brazil into five macro-regions: North, Northeast, Central-West, Southeast and South.

The summary statistics (not shown here) document some well-known facts (IBGE, various years): as for the overall population, the increase in mean and median incomes and the relative improvement in the bottom deciles that each region experienced over the last decade were accompanied by a reduction in inequality.

However, the other changes that occurred are not easily captured by these statistics; especially, no evidence supporting the polarization hypothesis emerges.

Therefore, to investigate the degree of polarization over time, we use the median adjustment and obtain the relative polarization indices for each region.

Polarization patterns similar to that observed for the overall income distribution are detected – i.e., a greater polarization in the lower tail and a movement toward the upper income levels by the second half of the 2000s.
We use the covariate adjustment technique to determine whether differences in the rural/urban population composition explain some of the observed changes in the overall income distribution.

The difference in rural/urban population composition had little effect on the 2011 to 2001 relative distribution, whose shape has mainly been influenced by changes in the marginal household income distributions.
We use the covariate adjustment technique to determine whether differences in the rural/urban population composition explain some of the observed changes in the overall income distribution.

The difference in rural/urban population composition had little effect on the 2011 to 2001 relative distribution, whose shape has mainly been influenced by changes in the marginal household income distributions.

Therefore, we analyze the impact of changes in the covariate-response relationship on the overall income distribution by explicitly forming the relative distribution for the two groups defined by the rural/urban categorical covariate.

The losses experienced by rural households between 2001 and 2011 were exclusively due to polarization, while income growth in the upper deciles was produced by both higher median gains and polarization.
We use the covariate adjustment technique to determine whether differences in the rural/urban population composition explain some of the observed changes in the overall income distribution.

The difference in rural/urban population composition had little effect on the 2011 to 2001 relative distribution, whose shape has mainly been influenced by changes in the marginal household income distributions.

Therefore, we analyze the impact of changes in the covariate-response relationship on the overall income distribution by explicitly forming the relative distribution for the two groups defined by the rural/urban categorical covariate.

The losses experienced by rural households between 2001 and 2011 were exclusively due to polarization, while income growth in the upper deciles was produced by both higher median gains and polarization.

For urban households, all of the change in distributional shape was due to a greater polarization in the lower tail, while income growth in the upper deciles appears to have been driven solely by the location shift.
Figure 10 Changes in urban household income distribution between 2001 and 2011

(a) Kernel density

(b) Relative distribution

(c) Location effect

(d) Shape effect

Conclusions
Summary

- 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; we are thus able to examine distributional changes that would not be detected easily from a comparison of standard measures of inequality.
- We document relevant changes in the Brazilian income distribution, despite the substantial falling off in income inequality: the analysis reveals indeed an overall upshift of the distribution, especially from 2005 onward, which masks a tendency to income polarization.
- A within-group analysis shows that all regions experienced greater polarization starting from the mid-2000s; furthermore, the observed spread of income polarization is mainly due to the increase of the relative income gap between wealthier and lower-income households – especially for rural areas – rather than to changes in the composition of the population according to the rural/urban covariate.

Policy Implications

- The recent improvements in Brazil’s income distribution appear to have mainly been propelled by the overall economic growth of the country.
- But as borne out by our results, under a negative growth scenario the shape effect would be brought to prevail, thereby generating a more unequal society.
- Hence, sustaining reductions in both inequality and poverty by making them less growth-dependent represents a key policy challenge for Brazil going forward: tools for a “real” re-distribution of resources – that goes beyond the effects of economic growth – are crucial if the positive trend is to be sustained in the future.
- Among these, making the tax system somewhat more progressive should be a top priority: Brazil’s heavy reliance on indirect taxes burdens the poor and middle-income households disproportionately, whereas the tax burden on the income of the rich is still too low (e.g., Birdsall et al., 2008).
- Furthermore, a large-scale land re-distribution would grant to poorest households the necessary tools to get out of extreme poverty and consequently reduce their actual dependence on social transfers.
References


fabio.clementi@unimc.it
Thank you all!