Economics Bulletin

Volume 35, Issue 2

# Top incomes in Brazil: preliminary results

Pedro Ferreira de Souza
*Instituto de Pesquisa Econômica Aplicada (Ipea) and University of Brasília (UnB)*

Marcelo Medeiros
*Instituto de Pesquisa Econômica Aplicada (Ipea) and University of Brasília (UnB)*

Fabio A. Castro
*University of Brasília (UnB)*

## Abstract

We present the preliminary results of our analysis of top incomes in Brazil from 2006 to 2012. We describe the evolution of the income shares of the top 1% and the top 5% and estimate a "tax-corrected" Gini coefficient. The data come from personal income tax returns, national accounts and household surveys. The results show that the levels of income inequality in Brazil are higher than those estimated using household surveys, and, contrary to what these surveys show, inequality did not fall over this period: the current trend is one of stability.

**Citation:** Pedro Ferreira de Souza and Marcelo Medeiros and Fabio A. Castro, (2015) "Top incomes in Brazil: preliminary results", *Economics Bulletin*, Volume 35, Issue 2, pages 998-1004
**Contact:** Pedro Ferreira de Souza - pedrosouza@gmail.com, Marcelo Medeiros - marcelo.medeiros.cs@gmail.com, Fabio A. Castro - fabioavcastro@yahoo.com.br.
**Submitted:** October 30, 2014. **Published:** April 22, 2015.

# 1. Introduction

In this research note, we present the preliminary results of our analysis of top incomes in Brazil from 2006 to 2012 combining data from income tax returns, national accounts and household surveys. Following the recent literature on top incomes, we first describe the evolution of the income shares of the top 1% and the top 5%. Next, we merge income tax and survey data to estimate a "corrected" Gini coefficient for Brazil. Our results demonstrate that the levels of income inequality in Brazil are higher than those estimated using household surveys, and, contrary to what these surveys indicate, inequality did not fall over this period: the current trend is one of stability.

This is an ongoing research project and we may still make adjustments to our methodology to maximize comparability both with international data and the historical series for Brazil. Yet, we believe that these adjustments will likely be minor and will not change our general conclusions about the levels and trend of inequality[1].

# 2. Methodology

## 2.1 Tax data

Our estimations rely on personal income tax returns for the period 2006–2012. More specifically, our raw data consisted of tabulations of the number of tax units and total gross income by gross income brackets based on multiples of the minimum wage.

The income definition is as broad as possible, encompassing both taxable and non-taxable monetary incomes, such as wages and salaries, self-employment and unincorporated business income, profits and dividends, pensions and other government transfers, interest, financial returns, and capital gains, among others.

"Gross" refers to incomes that are assessed before personal income and payroll taxes, but after the deduction of corporate and employers' payroll taxes. This income definition is net income after business expenses have been deducted for most business owners, except for self-employed professionals such as physicians and lawyers. Our data are not detailed enough to allow for the latter deductions. Fortunately, they are simply not large enough to change our results in a meaningful way: published statistics for 2012 show that deductions with business expenses corresponded to less than 1% of total income. Thus, even if we assumed – rather implausibly – that only the top 1% incur such expenses, their income share would be only 0.5 percentage points (pp) lower.

According to the Brazilian tax law, tax units can be either individuals or couples. In the latter case, joint filing is optional. In this paper, it is assumed each tax unit is an individual, which is reasonable for the upper reaches of the income distribution, as high-earning couples are unambiguously better-off filing separate tax returns.

[1] Two full-length articles based on this research are forthcoming in the Brazilian journals *Ciência & Saúde Coletiva* (Medeiros, Souza and Castro 2015a, in press) and *Dados* (Medeiros, Souza and Castro 2015b, in press) later this year. The former will also be available in English. Meanwhile, earlier drafts are available as working papers, albeit only in Portuguese (Medeiros, Souza and Castro, 2014a,b). Both working papers discuss our methodology in much greater detail and provide additional robustness checks. In March 2015, Brazil's central statistics office released detailed national accounts data for 2010 and 2011 and revised the Gross Domestic Product (GDP) estimates for both years. Thus, this research note reflects the latest available information (which also includes small corrections to the PNAD data), whereas the forthcoming papers rely on the previous estimates. Fortunately, as expected, such refinements do not change any of the major conclusions of our study.

The definition of the total population is somewhat arbitrary. Recent papers about top incomes usually calculate total tax units as all individuals aged 15, 18, or 20 and above (Atkinson, Piketty and Saez, 2011, p. 20-23). Our figures rely on population estimates calculated by Brazil's central statistics office, Instituto Brasileiro de Geografia e Estatística (IBGE), and define an intermediate age cut-off, computing total tax units as all individuals aged 18 and above. Our results seem to be robust compared to other plausible thresholds: on average, the top 1% income share rises by 0.7 pp when the population control includes all individuals aged 15 and above and declines by 0.5 pp when all individuals aged 20 and above are counted.

The results presented in section 3 of this paper rely on national accounts data to construct the control for total income. For the years 2006–2012, this income denominator is given by the sum of employees' compensation *plus* capital incomes *plus* mixed incomes *plus* property incomes *plus* transfer incomes *minus* imputed rentals for owner-occupied housing. Alternatively, one could express this sum as the households' gross primary incomes *minus* the imputed rentals for owner-occupied housing. For 2012, due to the unavailability of detailed national accounts data, we set the income denominator at 68.3% of GDP, that is, the average ratio observed for the years 2006–2011. Again, our results seem to be robust if compared to different plausible income denominators (Medeiros, Souza and Castro, 2014b)[2].

As usual, our income brackets do not correspond to the percentage groups of interest. For instance, 1.6% of the population was in our top bracket in 2012. Therefore, we follow the standard approach of the top incomes literature and estimated the exact population quantiles via the Pareto interpolation. In particular, we follow Feenberg and Poterba (1993) and calculate the Pareto parameters using the lower limit of the income bracket closest to the desired quantile. Estimates of the relevant parameters of the Pareto distribution are presented in the Appendix.

## 2.2 Survey data

In order to provide a benchmark for the tax data results, we also present estimates from the major household survey in Brazil, the *Pesquisa Nacional por Amostra de Domicílios* (PNAD). The PNAD is a large, nationally representative, multi-purpose survey carried out annually (except in Census years, such as 2010) by IBGE, with sample sizes usually in excess of 100,000 households. It is the most used data source for the study of income inequality in Brazil, although it is widely recognized that it severely underestimates capital incomes.

The raw PNAD data was slightly adjusted to enhance compatibility with the income definition pertaining to the tax data. Formal public and private sector workers, including military personnel, had their labor market earnings multiplied by 13.3 in order to account for their annual wages plus the $13^{th}$ salary and annual leave (equivalent to 1/3 of their monthly wage). Pensions were multiplied by 13 due to the $13^{th}$ salary. All other incomes were multipled by 12. Cases with missing data in any income variable were dropped.

[2]Both the working paper and the forthcoming full-length versions of this research set the control for total income at 66.7% of GDP for the reasons mentioned in footnote 1. This refinement does not change in any way the major conclusion of section 3, that is, that income tax data reveals top income shares to be at best stable between 2006 and 2012. It does, however, mitigate the apparent rise in top shares during this period: for instance, we originally reported a rise 3.7 pp rise in the top 1% income share, whereas our revised figures suggest a 1.6 pp increase. This reinforces the conclusion that inequality has at best remained stable between 2006 and 2012.

Such adjustments have very small effects on inequality levels. In 2012, for instance, the income shares of the top 1% and top 5% are 0.2 pp higher in the original data, while the Gini coefficient remains virtually unchanged.

## 3. Top income shares in Brazil, 2006–2012

Table 1 shows the evolution of the income shares of the top 1% and 5% over time. The top panel presents the results from the tax data, showing a slight upward trend, particularly at the very top. Whereas the top percentile increased its income share by 1.6 pp, the rest of the top vingtile gained only 0.8 pp. It is not entirely clear whether this rise of the top 1% can be explained by the business cycle: the correlation is stronger up until the international financial crisis, as real GDP growth ranged between 4% and 6% between 2006 and 2008 and then dipped to –0.3% in 2009, and top income shares mimicked this trend. Afterwards, the picture gets more complicated: the share of the top 1% hits its nadir in 2010, a year of exceptionally high growth for the Brazilian economy (7.5%) and then rose slightly in the low-growth years of 2011 and 2012.

Unfortunately, our data are restricted to a very short time span, so it is hard to draw general lessons: these results are more suggestive than conclusive. Likewise, considering how imperfect our data might be, it is more reasonable at this point to conclude that, at best, there are no signs of declining income concentration at the top, as opposed to what household surveys indicate.

In any case, there is no doubt that the *level* of income inequality is very high. On average, about 24% of all income accrued to the top 1%, and the top 5% as a whole received over 42% of all income. Such figures are considerably higher than in other countries. Recent studies have shown the income share of the top percentile to lie between 18–22% in the United States and Colombia and between 10–15% in many other countries, such as Germany, the United Kingdom, and Canada. The income share accruing to the top percentile in the most egalitarian countries is normally between 7% and 10% (Atkinson and Piketty, 2007, 2010; Atkinson, Piketty and Saez, 2011, Alvaredo, Atkinson, Piketty *et al.*, 2013; Vélez, 2012).

Finally, it is worth noting that, in 2012, the income threshold of the top percentile was about PPP$ 107,500 (R$ 203,000), whereas the cut-off for the top vingtile was about PPP$ 30,500 (about R$ 57,500). Average incomes were, approximately, PPP$ 292,500 (about R$ 553,000) for the top 1% and PPP$ 104,500 (about R$ 198,000) for the top 5%. In the same year, the average income for individuals aged 18 and above was PPP$ 11,000 (R$ 21,000)[3].

Comparing the top and bottom panels, it becomes clear the PNAD substantially underestimates the concentration of income at the top. In particular, the top 1% is the root of the problem, as the survey data is relatively accurate regarding the top 1%–5% income share, that is, the top vingtile excluding the top percentile. However, it is also the case that the trend over time of the top 1%–5% differs according to the data source: it rises by 0.8 pp in the tax data, but falls by 2.2 pp in the surveys.

[3]Brazilian *reais* were converted to international dollars using the PPP conversion rate of 1.89, following the United Nations' Millennium Development Goals indicators (see `http://mdgs.un.org/unsd/mdg/SeriesDetail.aspx?srid=699`). The World Bank reports a lower PPP conversion factor for Brazil in 2012 (1.52), which would make all figures about 24% higher (see `http://data.worldbank.org/indicator/PA.NUS.PPP?order=wbapi_data_value_2012+wbapi_data_value`).

Table 1: Top income shares in Brazil – 2006–2012 (%)

| Top income shares | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | Δ (pp) | Average |
|---|---|---|---|---|---|---|---|---|---|
| Income tax data | | | | | | | | | |
| Top 1% | 22.8 | 23.6 | 26.0 | 24.0 | 23.3 | 24.0 | 24.4 | 1.6 | 24.0 |
| Top 5% | 41.3 | 40.5 | 45.2 | 42.9 | 41.3 | 41.9 | 43.7 | 2.4 | 42.4 |
| *Top 1%–5%* | *18.5* | *16.9* | *19.2* | *18.9* | *18.0* | *17.9* | *19.2* | *0.8* | *18.4* |
| Survey data | | | | | | | | | |
| Top 1% | 14.8 | 14.4 | 14.0 | 14.1 | n/a | 13.6 | 14.0 | –0.8 | 14.2 |
| Top 5% | 35.7 | 35.0 | 34.2 | 33.9 | n/a | 33.0 | 32.8 | –2.9 | 34.1 |
| *Top 1%–5%* | *21.0* | *20.6* | *20.1* | *19.8* | *n/a* | *19.4* | *18.8* | *–2.2* | *20.0* |

Source: Medeiros, Souza and Castro (2014b), updated with new GDP estimates (see footnote 1 above).

Notes: The PNAD is not carried out in Census years such as 2010. The Top 1% and the Top 1%–5% shares might not always sum up exactly to the Top 5% due to rounding.

The absolute income thresholds reveal a similar pattern. In 2012, the cut-off for the top 5% was 20% higher in the income tax data than in the PNAD, while the thresholds for the top 1% was almost 70% higher. These figures were lower at the beginning of the period, which in turn explains the discrepancies revealed by both data sources: in contrast with the income tax data, the PNAD suggests a slight but continued decrease of top income shares. Since we have good reason to expect the tax data to be more reliable, we can conclude the survey results are misleading. The bulk of the evidence so far points towards stability or even increasing concentration at the top.

## 4. The "tax-corrected" Gini coefficient

As in other Latin American countries, many survey-based studies have highlighted a prolonged fall of income inequality in Brazil in the past decade (Barros, Foguel and Ulyssea, 2006). The Gini coefficient for labor market earnings has been declining since the late 1990's, but it was only after the beginning of the $21^{st}$ century that the Gini for household *per capita* income began its prolonged fall and reached its lowest levels in decades (Barros, Carvalho, Franco *et al.*, 2006; Soares, 2006). For instance, according to the PNAD data, the Gini coefficient for total gross income among individuals aged 18 and over dropped almost 7% between 2006 and 2012, falling monotonically from 0.638 to 0.596. In this section, we assess whether this picture changes when the survey distribution is complemented by the tax data distribution, which covers between 18% and 20% of the adult population. Thus, we calculate a "tax-corrected" Gini coefficient for 2006, 2009, and 2012.

Merging both datasets is not a trivial undertaking. After exploring a few different techniques, we settled for the simplest one: for the bottom 90% of the population, we considered the incomes collected by the household survey, and, for the richest 10%, we opted for the tax data information. Indeed, a visual inspection of both distributions shows absolute gross incomes are very similar between the $85^{th}$ and the $90^{th}$ percentiles, and then diverge quite abruptly (see Figure 4 in Medeiros, Souza and Castro, 2014a).

Table 2 displays the estimated Gini coefficients. Once we correct the underestimation of top incomes, the apparent decline in inequality almost vanishes, as the "corrected" Ginis are not only much higher but also much more stable. In fact, half of all income growth between 2006 and 2012 in the "corrected" PNAD data accrued to the top 5%, while only 12% accrued to the poorest half of the population.

In other words, the Brazilian income distribution is so skewed that the slight increase in top income shares discussed in the previous section almost completely offset the redistribution among the poorest 90% of the population, as captured by survey data. The overarching story regarding the Gini coefficient seems to be more about stability than change.

Table 2: Original and "corrected" Gini coefficients for Brazil – 2006, 2009 and 2012

| Source | 2006 | 2009 | 2012 | Δ (%) |
|---|---|---|---|---|
| A) Original PNAD[a] | 0.638 | 0.616 | 0.596 | –6.6 |
| B) "Tax-corrected" PNAD | 0.696 | 0.698 | 0.688 | –1.1 |
| B / A (%) | 109 | 113 | 115 | – |

Source: Medeiros, Souza and Castro (2014a).

[a] "Original" refers to the survey data after adjustments to ensure comparability with the income tax data

These results are quite robust to different thresholds, as discussed more fully in Medeiros, Souza and Castro (2014a). Merging the PNAD and the income tax data at the $85^{th}$ or the $95^{th}$ percentiles, instead of the $95^{th}$, has no consequential effects on either inequality levels or trends. There is a perceptible decline in the Gini coefficient only when we set the threshold at the $99^{th}$ percentile, that is, in the extreme case when we extract absolute incomes for the bottom 99% of the population from the survey data and rely on the tax data only for the top 1%. Even in this case, the drop in inequality is far more muted (about 3%) than the original PNAD suggests.

An alternative procedure would be to assume that the income share of the top $p$% should come from tax data and then follow the formulas presented by Alvaredo (2011), using the survey data only to estimate the Gini for the bottom $(1-p)$%. This approach yields lower inequality levels, but generally the same trend over time as ours: for instance, using the top 5% income share from tax data and relying on surveys solely to assess inequality among the bottom 95% suggests the Gini coefficient dropped from 0.671 in 2006 to 0.660 in 2012, a 1.7% decline. This figure is very similar to the one reported in Table 2 and much lower than the 6.6% drop reported by the PNAD.

## 5. Next steps

Provided we are granted access to the relevant data, in the near future we intend to:

a. Revise our estimates using more refined tabulations (ideally, up to 0.1% of the population);

b. Update the series with 2013 data and extend it back in time to 1927;

c. Decompose total income by factor components;

d. Increase the comparability of our series to data currently available at the *World Top Incomes Database* for several countries;

e. Improve the matching with household surveys other than the PNAD.

# 6. References


Alvaredo, F. (2011) "A Note on the Relationship Between Top Income Shares and the Gini Coefficient" *Economics Letters* **110**, 274–277.

Alvaredo, F., A. B. Atkinson, T. Piketty and E. Saez (2013) "The Top 1 Percent in International and Historical Perspective" *Journal of Economic Perspectives* **27**(3), 3–20.

Atkinson, A. B. and T. Piketty (Eds.) (2007) *Top Incomes over the* $20^{th}$ *Century: a Contrast Between Continental European and English-speaking Countries*, Oxford University Press: Oxford.

——— (2010) *Top Incomes: a Global Perspective*, Oxford University Press: Oxford.

Atkinson, A. B., T. Piketty and E. Saez (2011) "Top Incomes in the Long Run of History" *Journal of Economic Literature* **49**(1), 3–71.

Barros, R. P., M. Carvalho, S. Franco and R. Mendonça (2006) "A Queda Recente da Desigualdade de Renda no Brasil" in *Desigualdade de Renda no Brasil: uma Análise da Queda Recente* by R. P. Barros, M. N. Foguel and G. Ulyssea, eds., Ipea: Brasília.

Barros, R. P., M. N. Foguel and G. Ulyssea (Eds.) (2006) *Desigualdade de Renda no Brasil: uma Análise da Queda Recente – vol. 1*, Ipea: Brasília.

Feenberg, D. and J. Poterba (1993) "Income Inequality and the Incomes of Very High-Income Taxpayers: Evidence from Tax Returns" in *Tax policy and the economy* by J. Poterba, ed., MIT Press: Cambridge, Mass., vol. 7, 145–177.

Medeiros, M., P. H. G. F. Souza and F. A. Castro (2014a) "A Estabilidade da Desigualdade de Renda no Brasil, 2006 a 2012: Estimativa com Dados do Imposto de Renda e Pesquisas Domiciliares" Available at `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2493877`.

——— (2014b) "O Topo da Distribuição de Renda no Brasil: Primeiras Estimativas com Dados Tributários e Comparação com Pesquisas Domiciliares, 2006/2012" Available at `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2479685`.

——— (2015a) "A Estabilidade da Desigualdade de Renda no Brasil, 2006 a 2012: Estimativa com Dados do Imposto de Renda e Pesquisas Domiciliares" *Ciência & Saúde Coletiva* **20**(4), in press.

——— (2015b) "O Topo da Distribuição de Renda no Brasil: Primeiras Estimativas com Dados Tributários e Comparação com Pesquisas Domiciliares, 2006/2012" *Dados* **58**(1), in press.

Piketty, T. (2001) *Les Hauts Revenus en France au $XX^e$ Siècle: Inègalités et Redistributions, 1901–1998*, Éditions Grasset & Fasquelle: Paris, France.

Soares, S. S. D. (2006) *Distribuição de Renda no Brasil de 1976 a 2004 com Ênfase no Período entre 2001 e 2004*, Texto para Discussão n. 1166 Ipea: Brasília.

Vélez, J. L. (2012) *Income and Wealth at the Top in Colombia: an Exploration of Tax Records, 1993–2010*, Public Policy and Development Master Dissertation, Paris School of Economics, Paris, France.

## Appendix: estimating the parameters of the Pareto distribution

The relevant parameters of the Pareto distribution were estimating following Piketty (2001), using frequency and income information. Let $y^*$ be the average income above the income threshold $y$, then $\beta = y^*/y$ and Pareto's alpha is $\alpha = \beta/(\beta - 1)$. The parameter $k$ is given by $y(1 - F(y))^{(1/\alpha)}$, where $F(y)$ is the cumulative distribution function of $y$.

The income cut-off the top $p\%$ is estimated by $y_p = k/(p^{(1/\alpha)})$, and the average income for the top $p\%$ is estimated by $y_p^* = y_p\beta = y_p\alpha/(\alpha - 1))$.

Table A1 presents our estimates for the more interesting parameter $\alpha$ both according to both the income tax and survey data. Note, however, that the parameter was estimated from grouped data in the case of the income tax returns, whereas the survey results were calculated directly from individual-level microdata.

Table A1: Estimated $\alpha$ for the top 1% and top 5% – Brazil, 2006-2012

| Income group | Source | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|---|---|---|---|---|---|---|---|---|
| Top 1% | Income tax data | 1.618 | 1.536 | 1.538 | 1.591 | 1.570 | 1.541 | 1.580 |
| Top 1% | Survey data | 2.512 | 2.439 | 2.554 | 2.481 | *n/a* | 2.474 | 2.303 |
| Top 5% | Income tax data | 1.417 | 1.393 | 1.379 | 1.402 | 1.396 | 1.382 | 1.411 |
| Top 5% | Survey data | 1.989 | 1.985 | 1.974 | 2.021 | *n/a* | 1.954 | 1.938 |

Source: Medeiros, Souza and Castro (2014b).

Note: The PNAD is not carried out in Census years such as 2010.