Top Incomes Impacts on Inequality, Growth and Social Welfare: Combining Surveys and Income Tax Data in Brazil

See **paper**

Marcelo Neri¹

Marcos Hecksher²

¹ - FGV Social and FGV EPGE; marcelo.neri@fgv.br
² - IPEA and ENCE/IBGE

This paper is part of the Brazilian chapter of the “Inequality in the Giants” project supported by UNU-Wider

Objective

This paper evaluates the impacts of combining household surveys with income tax return files, in terms of growth, inequality and social welfare in Brazil from 2007 to 2015. This empirical exercise holds the promise to add higher and more realistic top incomes values into traditional surveys. (Piketty 2014; Medeiros, Souza and Castro 2015a, 2015b, Morgan 2018)

While the previous literature focused on the impacts of these data combination exercises on income inequality, there are new sources of understanding about the economic causes and social consequences behind these changes. In particular, by looking jointly at mean and inequality estimates obtained we get additional evidence on the nature of measurement error issues involved. It also allow an assessment cumulative welfare implications.
Individual Monthly Income by Population Quantile in 2007 (R$)

Source: PNAD (National Household Survey/IBGE) and IRPF (PIT – Personal Income Tax/RFB) – unit individuals 18 years of age or +

Lorenz Curves for PNAD, PIT AND mixed PNAD-PIT databases in 2007 and 2015

Source: PNAD/IBGE; PIT/SRF and Combined databases
## PNAD and IRPF overlap points: Inequality (Gini)

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2015</th>
<th>total var.</th>
<th>annual var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNAD</td>
<td>0.625</td>
<td>0.582</td>
<td>-0.043</td>
<td>-0.005</td>
</tr>
<tr>
<td>Fit 0.911</td>
<td>0.698</td>
<td>0.690</td>
<td>-0.008</td>
<td>-0.001</td>
</tr>
<tr>
<td>Fit 0.900</td>
<td>0.690</td>
<td>0.690</td>
<td>-0.008</td>
<td>-0.001</td>
</tr>
<tr>
<td>Fit 0.866</td>
<td>0.690</td>
<td>0.690</td>
<td>-0.008</td>
<td>-0.001</td>
</tr>
</tbody>
</table>

Source: PNAD/IBGE; IRPF/SRF and Combined databases.

## PNAD and IRPF overlap points: Mean Income (constant R$ at 2015 prices)

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2015</th>
<th>total % var.</th>
<th>annual % var.</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNAD</td>
<td>1,333</td>
<td>1,521</td>
<td>14.2%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Fit 0.911</td>
<td>1,675</td>
<td>2,100</td>
<td>25.4%</td>
<td>2.9%</td>
</tr>
<tr>
<td>Fit 0.900</td>
<td>2,107</td>
<td>2,107</td>
<td>25.8%</td>
<td>2.9%</td>
</tr>
<tr>
<td>Fit 0.866</td>
<td>2,108</td>
<td>2,108</td>
<td>25.9%</td>
<td>2.9%</td>
</tr>
</tbody>
</table>

Source: PNAD/IBGE; IRPF/SRF and Combined databases.
PNAD and IRPF overlap points:
Social Welfare (Sen 1976)

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2015</th>
<th>total var. %</th>
<th>annual var.%</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNAD</td>
<td>500</td>
<td>636</td>
<td>27.2%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Fit 0.911</td>
<td>505</td>
<td>651</td>
<td>28.9%</td>
<td>3.2%</td>
</tr>
<tr>
<td>Across Bases</td>
<td>1.02%</td>
<td>2.41%</td>
<td>1.7%</td>
<td>0.2%</td>
</tr>
</tbody>
</table>

Source: IBGE/PNAD and combined PNAD-IRPF databases

Real Income growth rate per income deciles of the population (2007-2015)

Source: IBGE/PNAD; IRPF/SRF and Combined databases
Partial Results of Combining Surveys and PIT Data

Static - The level of inequality measure rises when higher top incomes replace previous lower estimates based on surveys, this same exercise also increases by construction, the mean and the social welfare levels associated with it.

Dynamic - The movement of these combined estimates present a slower inequality trend fall than pure household surveys, at the same time income mean growth trends rose at a faster pace which implies possibly higher social welfare growth rates than suggested by previous surveys estimates.

Income Inequality Just within PIT has fallen
Income Distribution Within PIT – By Municipality

Real Growth Rate of Per Capita GDP in Brazil – 1901-2016 (% annual)

Source: IBGE and Ipeadata
Let's check the different pieces of this puzzle

Real Growth Rates of Average Income per income tax file (% annual)

![Graph showing real growth rates of average income per income tax file from 2008 to 2015.]

Source: RFB. Deflation CPI (IPCA/IBGE)

Macro Perspective on the PIT X GDP Income Growth Puzzle

In the 2007-2015 period PIT mean income grows 4.72% per year against 1.23% of GDP. Gap of 3.49 percentage points per year (ppy). Overestimation of financial gains 0.35 ppy.

**Deflators** – If we applied to nominal GDP the CPI (IPCA) instead the implicit GDP deflator growth gap would falls from 3.48 ppy to 1.75 ppy. Leaving both nominal National Accounts and PIT records untouched. Neri (2009; 2014) also allows reconciling almost all differences between GDP and standard PNAD income growth differences.

**Formalization** - share of occupied population contributing to social security rose 2.56 ppy. Accounts for the whole gap plus 0.8 ppy. As time passed the IRS observed better incomes.

**PIT encompasses both growth and formalization effects.**

- GDP x PNAD x PIT income growth gaps are due mostly to deflators and formalization. From 2007 to 2011 the annual growth of PIT taxpayers’ mean income (10.1%) was much higher than GDP’s one (3%). Once again, deflators and formalization can explain the bulk of the gap.

(Macro story)

- **Overall conclusion of the data combination exercise** - Slowed the pace of inequality fall but accelerated mean income growth. According to Sen’s index, higher gains in social welfare after the data sets combination.
Main Meso Findings - Main Puzzle: While the population got older, PIT taxpayers became younger and declared + elderly dependents.

✓ Microentrepreneurs formalization and transfiguration
✓ Illusory real financial income growth
✓ Profile of exempt incomes reinforces these three effects and breaks the idea that if you declare you pay tax so we can use PIT for higher incomes
✓ It’s risky to conclude on the trend of Brazilian inequality using PIT data. The comparison of PIT tables through time can produce biased trends while formalization advances, with profiles and shares of taxpayers in total population changing considerably in disagreement with demographic trends.

✓ In sum: PIT suggest that the level of inequality is much higher than that observed in household surveys. Combining both data sets also increases inequality with respect to surveys but with Pareto improvement by construction.
✓ Moving from level to changes: Inequality within just PIT fell. Combining surveys with PIT data: inequality may have fallen less but growth was much higher. Do we trust such high growth rates? Was there a dynamic Pareto improvement after all?