

**Estimating the regional distribution
of income in sub-Saharan Africa**

Lucas Chancel
Léo Czajka

December 2017

This version: December 11th, 2017



Estimating the regional distribution of income in sub-Saharan Africa

Lucas Chancel, Léo Czajka

Abstract

This technical note details the methodology followed to construct Sub-Saharan Africa inequality estimates from 1980 to 2016, presented in the World Inequality Report 2018. We combine existing household survey data with available data on top incomes as well as with distributional national accounts. Given the important limitations of the data we use, we stress that much more work will be necessary to estimate the levels and dynamics of income inequality in Sub-Saharan Africa. However, our reconstruction is likely to provide a more realistic representation of the actual level of inequality than standard measures of income inequality based solely on survey data.

Data

Our estimations are based on distributional data extracted from the PovcalNet Database for the Sub-Saharan African Region, as well as on top incomes and distributional national accounts data¹ for African countries available on WID.world for the period considered.

For every household survey conducted in the region, this data source details, by percentiles groups, the coordinates of the Lorenz Curve. We restrict our sample to observations for which the World Bank had direct access to the micro data and to countries where at least 3 surveys were conducted. Our sample eventually covers 25 countries representing more than 75 % of the SSA population, with 4,36 years per country and, inversely, 3 countries per year². The welfare concept measured is household consumption per capita.

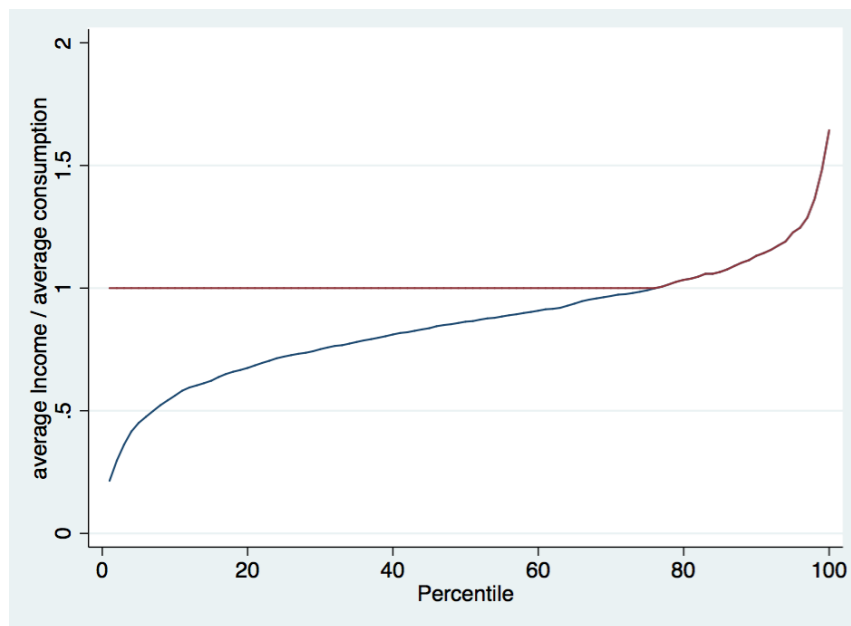
¹ See Alvaredo et al. (2016) for a general discussion of the Distributional National Accounts methodology.

² In alphabetic order: Bénin (2003, 2011, 2015); Burkina Faso (1994, 1998, 2003, 2014); Burundi (1998, 2006, 2013); Cameroon (1996, 2001, 2007, 2014); Côte d'Ivoire (1985-1988, 1993, 1998, 2002, 2008, 2015); Ethiopia (1995, 1999, 2004, 2010); Ghana (1991, 1998, 2005, 2012); Guinea-Bissau (1991, 1993, 2002, 2010); Guinea (1994, 2002, 2007, 2012); Kenya (1994, 1997, 2005); Lesotho (1994, 2002, 2010); Madagascar (1997, 1999, 2001, 2005, 2010, 2010, 2012); Malawi (1997, 2004, 2010); Mali (1994, 2001, 2006, 2009); Mauritania (1995, 2000, 2004, 2008, 2014); Mozambique (1996, 2002, 2008); Niger (1994, 2005, 2007, 2011, 2014); Nigeria (1992, 2003, 2009); Rwanda (2000, 2005, 2010, 2013); Senegal (2001, 2005, 2011); South Africa (1996, 2002, 2006, 2008, 2011); Tanzania (1991, 2000, 2007, 2011); Togo (2006, 2011, 2015); Uganda (1992, 1996, 1999, 2002, 2005, 2009, 2012); Zambia (1991, 1996, 1998, 2002, 2004, 2006, 2015).

From consumption to income distribution

Being smoothed overtime, consumption is generally less unequally distributed than income. Combining estimates from Cogneau et al (2014) and Czajka (2017) we are able to compare consistent series of consumption and income for 5 years. On average, the quantile function of the income distribution is higher than that of the consumption distribution for the top 23%, but lower for the bottom 77% (see Graph 1). The source of such discrepancy can be manifold, especially in the context of sub-Saharan African countries, where self-employment is so important (Deaton, 1997). While we have evidence that top incomes are likely to be underestimated in survey -- and therefore even more so for consumption levels; further research would be needed to establish that residual negatives savings do reflect true indebtedness of the poorest rather than measurement errors. For lack of better options, we keep the bottom 77% of all distributions as such, and multiply percentiles averages of the top 23% by the coefficients plot in graph 1 (red line). Robustness checks show that different assumptions have no impacts on the trends we observe and moderate impacts on the level of overall inequality, and are likely to represent a lower bound of overall inequality levels.

Graph 1: Ratio between income and consumption levels
(Côte d'Ivoire – average over the years: 1993, 1998, 2002, 2008 and 2015)

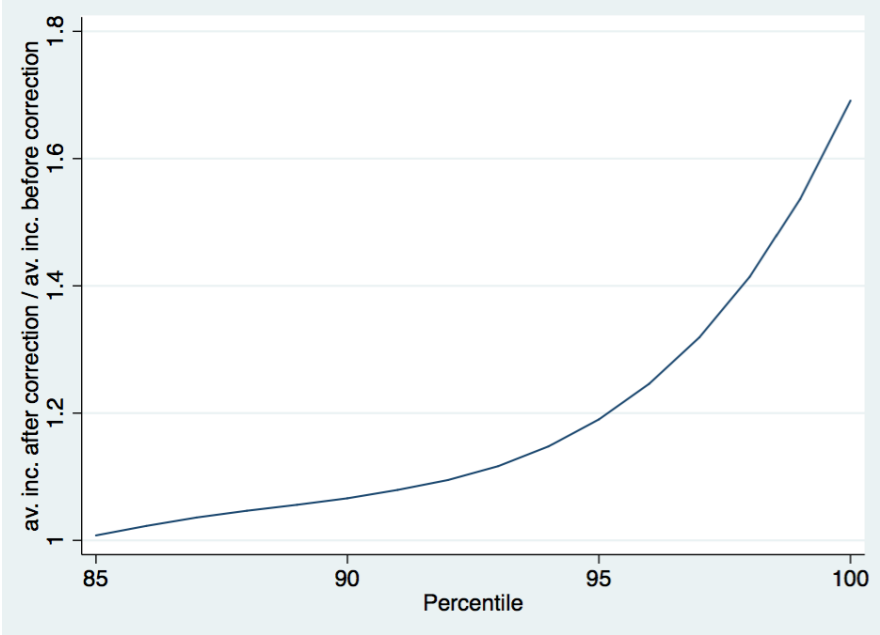


Correction of top incomes

Due to under-reporting and/or non-response bias, surveys tend to significantly underestimate top income levels. For South Africa, access to fiscal data allows us to correct for such bias in a specific way: for each year, we replace the top 1% share by the one estimated by Alvaredo and Atkinson (2010) and uniformly decrease the percentile shares of the bottom 99% so as to compensate for the increase in top 1% share induced by the replacement.

For all other countries, we use the information we have about the underestimation profile of higher percentiles in Côte d'Ivoire. As soon as we are able to develop DINA (Alvaredo et al. 2016) estimates for African countries, we will refine our methodology. We see this correction as however more accurate than reporting inequality levels from survey alone. Combining fiscal and survey data from 2014-2015 in Côte d'Ivoire, we are able to compare top percentile averages before and after adjusting for top income. This analysis shows for instance that the top 1% average is 70% higher according to fiscal data, while the survey seems to have properly captured the bottom 83% (see Graph 2). To adjust top incomes, we multiply percentile averages of the top 17% by the correction coefficients displayed in Graph 2.

Graph 2: Underestimation of top incomes in Côte d'Ivoire, 2015



Interpolation/extrapolation of missing years

For each missing observation between two available survey estimates, we linearly interpolate averages by percentile. Whenever years are missing on either of the two ends of the series (typically in the 1980s or the late 2010s), we keep the first/last available observation constant through all missing years so as to complete the series.

Merging countries

Finally, we weight each percentile shares and averages by the total adult population and national income (in constant international € PPP) as defined according to the WID national accounts series and merge, for each year, all country specific distributions into one single African distribution.

Conclusion

This technical note details the methodology followed to construct Sub-Saharan Africa inequality estimates from 1980 to 2016, presented in the World Inequality Report 2018. We combine existing household survey data with available data on top incomes as well as with distributional national accounts.

Given the important limitations of the data we use, we stress that much more work will be necessary to estimate the levels and dynamics of income inequality in sub-Saharan Africa. However, our reconstruction is likely to provide a more realistic representation of the actual level of inequality that standard measures of income inequality based solely on survey data.

References

Alvaredo, F., Atkinson, A. B., Chancel, L., Piketty, T., Saez, E., Zucman G., "Distributional National Accounts (DINA) Guidelines : Concepts and Methods used in WID.world", WID.world Working Paper, 2016/2

Facundo Alvaredo and Anthony B Atkinson. "Colonial rule, apartheid and natural resources: Top incomes in south africa", 1903-2007. *CEPR*, 2010

Blanchet, T. Chancel, L. "National Accounts Series Methodology", WID.World Working Paper, 2016/1

Cogneau, Denis, Kenneth Hounghbedji & Sandrine Mesplé-Somps, 2014. "The Fall of the Elephant, Two decades of Poverty Increase in Côte d'Ivoire (1988-2008)". WIDER Working Paper 144, pp. 22-29.

Léo Czajka, 2017. "Income Inequality in Côte d'Ivoire 1985-2014", World Wealth and Income Database Working papers series N°8/2017

Angus Deaton. "The analysis of household surveys: a microeconomic approach to development policy". *World Bank Publications*, 1997.