DISEI - Università degli Studi di Firenze

Working Papers - Economics

Building the integrated inequality database and the seven sins of inequality measurement in Sub-Saharan Africa

G. A. Cornia and B. Martorano

Working Paper N. 05/2017

Building the integrated inequality database and the seven sins of inequality measurement in Sub-Saharan Africa

Table of Contents

Authors: Giovanni Andrea Cornia and Bruno Martorano.
All comments can be directed to mailto: giovanniandrea.cornia@unifi.it

I. Introduction and rationale[1]

The favourable growth performance of SSA over the last twenty years (Figure 1) - emphatically referred to by some as 'the SSA Renaissance' or 'Africa Rising' - has been accompanied by a perceptible, but still modest decline in poverty, from 59 to 48 percent over 1993-2010, i.e. much less than that recorded in South Asia (Ferreira 2014). Such aggregate trend however conceals substantial cross country variations. How to explain then such differences in poverty reduction rates? The standard approach (Bourguignon 2003) shows that the percentage change in poverty rates can be decomposed in the percentage change in GDP per capita growth rates and the percentage changes in the Gini coefficient, plus a (generally small) residual.[2] In this regard, it must be noted that in SSA the average GDP growth per capita oscillated in a narrow range, i.e. between 1.7 percent in non-resource rich countries and 2.6 percent in resource-rich ones. The reason why poverty declined at different rates is therefore to be found in the divergence of inequality trends experienced by the countries of the region. This paper as well as Cornia (2014) and the literature quoted therein argue in fact that over the last 20 years the Gini index of inequality rose in several countries but simultaneously fell in a similar number of them.[3]

Figure 1. Trend in the log of aggregate real GDP/capita in sub-Saharan Africa, 1960-2012
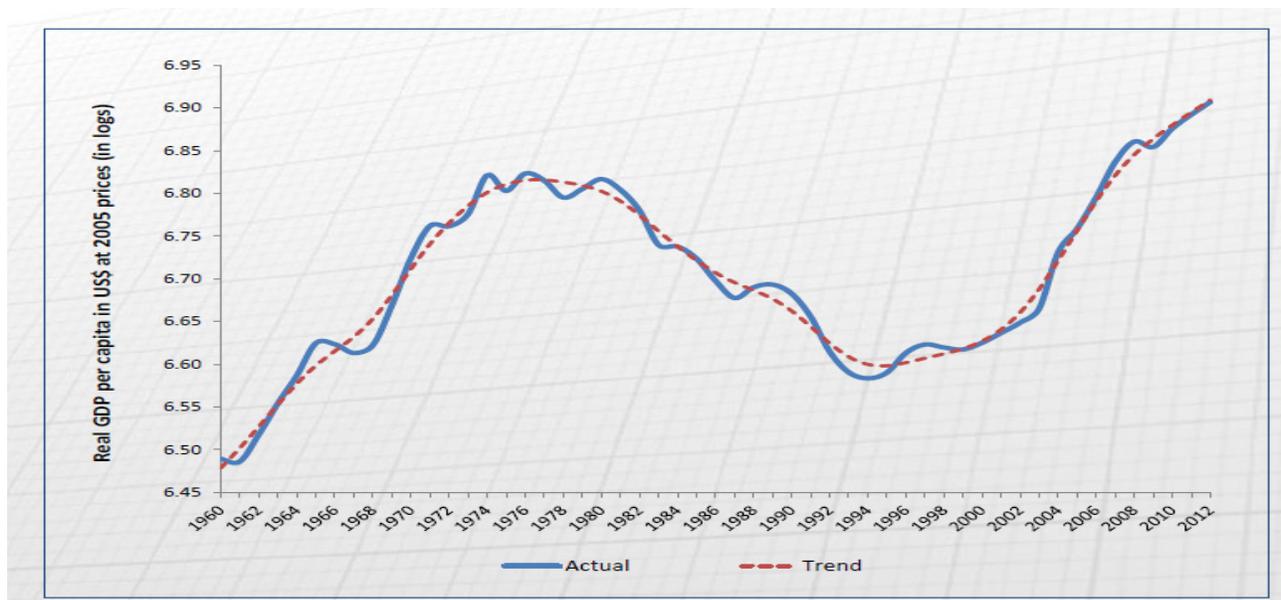
Source: Ferreira (2014).

A proper documentation of inequality trends in the region becomes therefore essential to explain the above mentioned differences in poverty reduction. This task however is hindered by the limited and at time conflicting inequality data in the region and by the lack of a comprehensive database of good-quality and consistent inequality statistics.

---

[2] Over the long term, poverty may also decline due to investments in health, education and social transfers that affect both the GDP/capita growth rate and the distribution of income. Evidence shows that countries that invested heavily in the social sector reduced poverty through an improvement in the human capital of the poor.

[3] UNDP Regional Programme for Africa (RBA-UNDP) Project on SSA Inequality does not deal explicitly with inequality in other dimensions of wellbeing, such as health, education and access to basic services.

This situation is even more penalizing when considering that over the last two decades policy formulation has become increasingly 'evidenced-based', i.e. based not only on ideological and doctrinal priors but also on the empirical evidence provided by a growing number of household budget surveys (HBS), demographic and health surveys, wealth surveys, multiple indicator cluster surveys, multipurpose living standard measurement studies, and other surveys. The field of studies that has benefitted the most from such increase in the number of surveys is that concerning poverty alleviation and the control of inequality. In most developed and developing regions academic and policy institutions have by now built databases tracing the evolution of the Gini coefficient over at least the last 20 years, as in the case of LIS for the OECD countries, SEDLAC and CEPALSTAT for Latin America, TRANSMONEE for the European economies in transition and so on. Finally, during the same period global inequality databases were also created, including WID, WIID, SWIID, Allgini and others which are discussed below.

In view of the problems caused by few and scattered inequality data and the lack of an assessment of their quality and pitfalls, this paper aims at doing two things:[4]

- First, in Section 2 it describes the Integrated Inequality Database (IID-SSA) obtained by comparing the Gini coefficients included in the existing databases, and selecting the least biased Gini's. IID-SSA thus summarizes hopefully in the least distorted and systematic way the existing information on inequality, permitting in this way to analyze the changes recorded in the region in this field during the last two decades, and to draw policy recommendations. The IID-SSA dataset is illustrated in detail in Annex 1. It provides a summary of all Gini coefficients from all international databases and national sources not included already in the former, it selects the best time series of country Gini coefficients for the years 1993-2011 on the basis of a standard protocol, and plots their time trend for the 29 countries with at least four good-quality and well-spaced Gini points. Annex 1 also provides summary information on Gini availability for countries with only 1-3 Gini data. The time series for the 29 above countries can be used for a variety of analytical and policy purposes, be they the calculation of changes in poverty rates over time or panel regressions of Gini trends. Yet, given the data limitations and biases discussed in Section 3, this information has to be used with a pinch of salt, i.e. checking the results they may generate against those predicted by economic theory, economic history and other statistical sources (such as the national accounts) and by introducing whenever feasible the statistical adjustments indicated below.

- Second, in Section 3 it discusses the limitations and biases of the data included in IID-SSA and tries when possible to measure the extent of such biases with the purpose of <u>alerting the researchers of African inequality of the 'seven sins of inequality measurement'</u>[5] most commonly met in the region. Section 3 also presents the approaches currently followed to remedy – when possible - such problems. Such seven problems concern: differences in the design of successive surveys within a country; differences in survey design across countries; under-sampling of top incomes; possible inconsistency between Gini data derived from surveys and the 'labour share' computed on the bais of the national accounts; the neglect of incomes generated by assets held abroad; the distributive impact of different dynamics of food prices and CPI; and the neglect of the public social services in kind provided by the state when calculating the Gini coefficient. In a way, Section 3 represents a 'checklist of possible biases' that researchers, statisticians and policy makers aiming at computing the 'real Gini coefficient' of a country should take into account. Indeed, the usual way the inequality data are computed often constitutes an oversimplification that mostly leads to an underestimate of inequality and lack of policy action. Yet, the corrections suggested in this paper require the availability of survey micro-data (not available to us) and are labor- and assumptions-intensive. But carrying out such corrections allows to compute Gini data that are more precise than those included in IID-SSA data, and get in this way a better grasp of the real distributive situation of a country. Academic economists and staff of UNDP and World Bank are well advised to introduce such corrections when working on poverty and inequality at the country level.

[4] The discussion of the causes of inequality in SSA and of the policy options available to reduce it are discussed in other papers generated as part of the RBA-UNDP project on 'Inequality in SSA' to which the reader is referred to.

[5] The reader may think that the choice of the term 'seven sins' might has been inspired by the 'seven cardinal sins' (lust, greed, gluttony, sloth, wrath, envy and pride) part of Christian theology, or by T. L. Lawrence book's on the 'Seven Pillars of Wisdom'. Yet, any reference to such ideas is purely coincidental.

II. Building a database of synthetic inequality statistics

2.1 Existing inequality databases

One of the problems affecting the analysis of income inequality and its changes in SSA is the lack of a consolidated and sufficiently standardized database of inequality indexes, like that produced by SEDLAC for L. America (http://sedlac.econo.unlp.edu.ar/eng/index.php) or LIS for the OECD countries (http://www.lisdatacenter.org/). At the moment, researchers of SSA inequality rely alternatively on inequality statistics originating either from:

- **WIDER's WIIDv3. 0b database:** (http://www.wider.unu.edu/research/WIID3-0B/en_GB/database/) which was released in September 2014 and which includes fully documented Gini coefficients and decile and quintile distributions for 44 SSA countries, often for long periods of time. For every datapoint the WIIDv3.0b includes standardized information and documentation on the concepts used in each survey concerning income (whether gross, net, monetary, earnings, etc) consumption expenditure (monetary or in kind), basic unit of observation and population coverage (household, family, person), equivalence scales, sample size and so on. There is also information about the survey questionnaire, survey coverage (national, urban, rural and so on) and availability of survey reports. The interest reader may look at the following link for more information http://www.wider.unu.edu/research/WIID3-0B/en_GB/WIID-documentation/). Finally, WIIDv3.0b rates the quality of each Gini or decile distribution with 'scores' going from 1 to 4, mainly on the basis of survey coverage, nature of the questionnaire and data collection methodology. Only good quality data rated '1' or '2' can be used safely in trend and regression analysis. Data-points rated '3' or '4' are not of the same quality and ought to be used only for ad-hoc purposes, under certain conditions and not for panel regressions. WIIDv3.0b data derive from HBS produced by Central Statistical Offices (CSO), LSMS surveys, POVCAL, and independent field studies;

- **The World Bank's Povcal database** (http://iresearch.worldbank.org/PovcalNet/index.htm). It calculates Gini coefficients on the basis of decile distributions derived from surveys microdata. POVCAL does not harmonize the microdata according to standard statistical criteria before computing the deciles distribution and the Gini coefficients. Its data overlap to a good extent with WIDER's WIIDv3 data, but its coverage is thinner;

- **The World Bank's International Income Distribution Database (I2D2)**[6] is a worldwide database drawn from nationally representative HBS, Household Income and Consumption surveys, Labour Force surveys, and LSMS surveys comprising a standardized set of demographic, education, labour market, household socioeconomic features, and income/consumption variables. [7] I2D2 has about 50 'harmonized variables' and covers over 900 surveys from over 160 countries from all over the world[8] and for years at times going back to 1960 though most of the information covers the last two decades. Due to such harmonization process, the I2D2 data facilitate cross-country comparisons in several areas of interest. However, at the time of revising this paper (June 2015), we could access only 9 harmonized Gini data-points. When the entire dataset will be available, we will adjourn the IID-SSA and re-compute the trends. However, an initial look at the nine I2D2 data that have become available over the last two months does not suggest major changes in the trends identified in Cornia (2014).

To improve comparability across countries and over time the survey micro-data are 'harmonized' according to standard statistical conventions concerning: the definition of household income/consumption expenditure per capita; the definition of household members; the corrections for differences in recall periods when transforming income/consumption data into monthly/yearly data; the valuation of the income stream from

---

[6] I2D2 was started in 2005 in the context of the World Development Report on Equity (World Bank 2006). The effort has continued and several World Bank publications and UNDP's HDR utilize this database. I2D2 strives to make the database easy to access and reasonably comparable across time and countries.

[7] Almost all surveys in the I2D2 database are nationally representative. For most of them, the unit of observation is the household member. In a small number of labor force surveys, the survey collects information only on members above a certain age.

[8] The complete list of surveys in the I2D2 is available upon request

owner-occupied dwellings; adjustments for non-responses; imputation of missing or clearly unreliable data; the treatment of zero incomes; and the upward adjustments of rural incomes made to offset differences in rural-urban prices. Thus, by definition, the Gini computed on I2D2 and those of POVCAL and CSO do not coincide since they rely on different statistical conventions. Furthermore, I2D2 Gini's are computed directly on micro-data, and should therefore be somewhat higher that those calculated on decile distributions.

The harmonization of the micro-data to be included in I2D2 is still ongoing at this moment. The World Bank has collected some 140 surveys for SSA, though only about 20-30 of them had been processed by late 2014. The assignment of countries to the of rising, falling, U shaped and inverted U-shape inequality categories, as in Cornia (2014), may thus change somewhat in the future as new harmonized data for past years become available or replace existing IID-SSA Gini's taken from other sources.

- Milanovic's 'All Gini's' dataset which compiles data from all sources and adds a few observations drawn from data produced by CSOs or surveys launched as part of specific research projects. No major adjustments are carried out on the data

- The Luxembourg Income Study (LIS) which provides LIS-standardized data for South Africa.

- Szolt's SWIID dataset which includes Gini's from all sources and years but does not rate the quality or consistency of the data, the majority of which is obtained through multiple imputation techniques which are not always made explicit. While SWIID offers more complete data coverage and for long periods, its content is unclear and depends on opaque and arbitrary multiple assumptions. After a detailed comparison of WIIDv3.0b versus SWIID, Jenkins (2014) suggests to rely on WIIDv3.0b at the condition that 'researchers must take care when selecting observations, to confront the very real data quality issues head on [i.e. by selecting only quality 1 and 2 data] and check whether their conclusions are robust to different treatments of the data'. Because of this conclusion, we decided not to use SWIID data, even if this entails foregoing a number of multiply imputed data which have no equivalent in the other databases. Jenkins (2014) notes that particularly when analyzing inequality changes in SSA, the secondary data on inequality are of poorer quality. In such countries there is also a higher prevalence of missing data, and hence a greater proportion of SWIID data are heavily reliant on the validity of its imputation models, for which, given the high measurement error in basic data, there is greater variability.

Differences in research results about inequality dynamics may thus depend not only on differences between the countries/years considered, but also on the dataset chosen for the analysis. To overcome this problem and limit the use of low-quality/undocumented data, and with the aim of identifying inequality trends in the region, we compiled an 'Integrated Inequality Database for SSA' (IID-SSA) which selects for every country/year the best datum from the first five datasets described above, as well as from a few national sources. IID-SSA contains yearly information for the years 1991/3-2011 for 44 countries with at least one good quality Gini datum. In several cases, the data from the five datasets are identical or very similar (as in the case of WIIDv3.0b and POVCAL), while in others they differ a bit or substantially. As shown in Table 1, most of the data we selected for IID-SSA are from WIIDv3.0b. Of the 44 countries,[9] 14 are from Eastern Africa, 9 from Central-Middle Africa, 5 from Southern Africa and 16 from West Africa. For Equatorial Guinea, Eritrea, Sao Tome and Principe, Somalia and South Sudan there is not even a single datum and are therefore excluded from the dataset. The IID-SSA dataset is enclosed in Annex 1 to this paper.

We are well aware that the data included in IID-SSA may suffer from measurement errors due to the factors discussed in detail in Section 3. Yet, a careful selection of data from all available sources ought to reduce some of these measurement errors by increasing data consistency and completeness so as to provide the 'least biased' dataset in this field. In this regard, it must be mentioned that due to the difficulties to ensure good data

[9] Of these 14 are from Eastern Africa (Comoros, Djibouti, Ethiopia, Kenya, Madagascar, Malawi, Mauritius, Mozambique, Seychelles, Sudan, Tanzania, Uganda, Zambia and Zimbabwe), 9 from Central –Middle Africa (Angola, Burundi,, Cameroon, Central African Republic, Chad, Republic of Congo, Democratic Republic of Congo, Gabon and Rwanda), 5 from Southern Africa (Botswana, Lesotho, Namibia, South Africa and Swaziland) and 16 from Western Africa (Benin, Burkina Faso, Cape Verde, Cote d'Ivoire, Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone and Togo).

on income (due to the high degree of informalization and low monetization of transactions among subsistence farmers and in the urban informal sector) most household surveys focus on consumption expenditure for which measurement and recall errors are smaller. Thus, with the exception of Botswana and Mauritius (which use '*disposable income per capita*'), the wellbeing concept adopted in SSA surveys is '*household consumption expenditure per capita*', a concept that reduces estimation bias but does not allow to decompose the changes in total inequality by income source. However, for several countries (Ghana, Malawi, Burkina Faso, Ethiopia, Angola, Cameroon, CAR, Gabon, Gambia, Kenya, Lesotho, Madagascar, Mali, Mauritania, Nigeria, Senegal, South Africa, Tanzania, Uganda, and Zambia) there are one or two surveys providing data on both income and consumption per capita.[10] These surveys also allow to measure the differences between the Gini coefficient computed on the distribution of consumption per capita and that computed on the distribution of income per capita. Overall, however, and with the exception of Botswana and Mauritius, the wellbeing concept adopted in the 29 countries of Table 1 is '*consumption expenditure per capita*'.

---

[10] This is, for instance the case for the 2004 and 2011 Integrate Household Surveys (IHS 2004 and 2011) whose data have been standardized by the FAO Project called RIGA also in terms of household income per capita through a series of imputations, corrections and standardizations. In fact, these two surveys have been used for the Malawi micro-decomposition of Gini changes over time by income sources and sectors of production (see the Cornia and Martorano June 2015 paper on 'The dynamics of income inequality in a dualistic economy: the case of Malawi, 1990-2011' which is part of the RBA-UNDP project on 'Inequality in SSA').

Table 1. No. of data points on expenditure consumption per capita** for 29 countries with at least 4 well-spaced Ginis on the distribution of per capita consumption expenditure, 1991/3-2011

Country	Database from which our data were extracted					Data retained for 1993 - 2011			Pop share	Gini trend	D Gini points
	WIID V3	POVCAL	WB-I2D2	Gini All	Nat. Data	Tot Obs	Interpolated	Total			
B. Faso (1994-2009)	…	4		1		5	14	19	2.26	Falling	-10.9
Cameroon* (96-2007)	3					3	16	19	3.05	Falling	-8.8
Ethiopia (95-2011)	…				4	4	15	19	12.82	Falling	0
Gambia (93-2003)	4	…	…	…	…	4	15	19	0.24	Falling	-13.6
Guinea* (94-2007)	3					3	16	19	1.61	Falling	-1
G.Bissau (93-2005)	3			1		4	15	19	0.24	Falling	-9.5
Lesotho (93-2003)	5			1		6	13	19	0.32	Falling	-5.4
Madagascar (94-2010)	4	2		1		5	14	19	3.08	Falling	-7.1
Mali (94-2010)	4					4	15	19	2.01	Falling	-17.5
Niger (94-2008)	4			1		5	14	19	2.22	Falling	-15.2
Senegal (94-2011)	3	1				4	15	19	1.9	Falling	-1
S. Leone* (95-2011)		2		1		3	16	19	0.86	Falling	-18.8
Swaziland (95-2010)	3			1	…	4	15	19	0.18	Falling	-9.2
Tot falling countries	**36**	**7**	**0**	**7**	**4**	**54**	**193**	**247**	**30.79**	**Falling**	**Av. -9.1**
Angola (1995-09)	2	…	1	1	…	4	15	19	2.79	∩ shape	2.8
Mauritania (95-2008)	5	…	…	…	…	5	14	19	0.53	∩ shape	+ 3.5 -3.3
Mozambique (96-2008)	5	…	…	…	…	5	14	19	3.54	∩ shape	+ 2.6 -5.7
Rwanda* (95-2011)	2	1	…	…	…	3	16	19	1.59	∩ shape	+9.1 -4.1
Tot ∩ shaped countries	**14**	**1**	**1**	**1**	**…**	**17**	**59**	**76**	**8.45**	**∩ shape**	**Av +8.4 -7.2**
Botswana* (1994-09)	2	…	1	…	…	3	16	19	0.31	Rising	14.9
Cote Ivoire (95-2008)	4	…	…	…	…	4	15	19	2.93	Rising	8
Ghana (93-2006)	4	1	…	1		6	13	19	3.6	Rising	9
Kenya (94-2006)	3	1	…			4	15	19	6.02	Rising	3.8
Mauritius (91-2011)	3	…	7			10	9	19	0.21	Rising	2.5
South Africa (91-2011)		6		1		7	12	19	8.02	Rising	5.7
Uganda (92-2010)	8	…				8	11	19	4.84	Rising	1.4
Total rising countries	**24**	**8**	**8**	**2**	**0**	**42**	**91**	**133**	**25.93**	**Rising**	**Av. + 6.5**
C. A. R. (1992-2008)	3		…	1	…	4	15	19	0.67	U shape	-5
Malawi (1993-2011)	6	1	…	1	…	8	11	19	2.18	U shape	-23.4 +6.6
Nigeria (1993-2010)	3	1	…	1	…	5	14	19	23.5	U shape	-2.1 +1.8
Tanzania (1993-2010)	4		…		2	6	13	19	6.54	U shape	-2.5
Zambia (1993-2010)		7	…	1		8	11	19	1.93	U shape	4.9
Tot U shaped countries	**16**	**9**	**0**	**4**	**2**	**31**	**64**	**95**	**34.82**	**U shape**	**Av -11.8**
Grand Total	**90**	**25**	**9**	**14**	**6**	**144**	**407**	**551**	**100**	**All**	**…. ….**
% shares	**16.3**	**4.5**	**1.6**	**2.5**	**1.1**	**26.1**	**73.8**	**100**	**100**	**All**	**….. …..**

To analyze the income dynamics in the region, for each country we selected time series using the same income concept and population coverage for the period 1993-2011,[11] though we cannot ensure that the same statistical conventions were adopted in all surveys and in the processing of raw data. Likely, there remain differences across countries and over time in statistical conventions adopted by HBS. This will increase the 'noise' in regression analysis. To select the data needed for the trend analysis, we followed the approach described below.

First, out of the 44 countries included in the original (all data) IID-SSA, we selected 29 countries with at least four good-quality and well-spaced data derived from surveys adopting time-consistent statistical conventions depicting reasonably well medium-term inequality trend (Table 1). On average, there are 5 data-points for each of the 29 countries selected which account for 81.8 percent of the SSA population. The countries excluded account for 18.2 percent. Of the countries excluded, only Congo D.R has a large population. The 15 countries excluded are Benin, Chad, Congo, Republic, and Liberia, Sudan (1 data-point each); Cap Vert, Djibouti, Gabon, Namibia and Togo (2 data each); and Burundi, Comoros, Congo D.R., Seychelles and Zimbabwe (3 data-points), for a total of 30 reliable observations, that we did not use in the trend analysis.

All together, for the period 1993-2011 the IIDB-SSA matrix includes 551 (29 x 19) cells. Of these 144 (26.1 percent) are non-zero. To deal with the problem of missing data we connected the selected data through linear point-to-point interpolation (as in the left panel of Figure 2). In Annex 1, the data-points retained are indicated in the last column of each country data summary matrix and the point to point interpolated ones are colored in yellow. Finally, to assign each of the 29 countries of Table 1 to the rising, falling, U shaped or inverted U shaped category we interpolated the Gini time series obtained with linear and quadratic functions, as shown – as an example - below in Figure 2, right panel, for Zambia (in this case, the best fit is clearly a concave function). We then chose the nature of the trend on the basis of the best statistical fit suggested by the R^2 and F statistics. Finally, we assigned each country to the rising, falling, U shaped and inverted U shaped group. In a subsequent paper of the RBA-UNDP project on 'Inequality in SSA' we will explore by means of a cluster analysis whether the countries belonging to each of the four country groups (falling, rising, U shaped, and inverted U shape trends) share common characteristics such as production structure, population size, initial level of inequality, social variables and others

---

[11] When aggregating the trend of these 2 countries into the their respective groups (see below), we multiplied them by a correction factor of 0.81 which corresponds to the ratio of the Gini coefficient of the distribution of consumption expenditure to that of disposable income found by Cogneau et al (2007) for five countries on the basis of large surveys for the 80s and early1990s. In future panel regressions we will introduce dummy variables to correct for differences between Gini consumption and Gini income.

Figure 2. Example of interpolation of the missing data-points (left panel) and choice of the best interpolated trend (right panel) in the case of Zambia
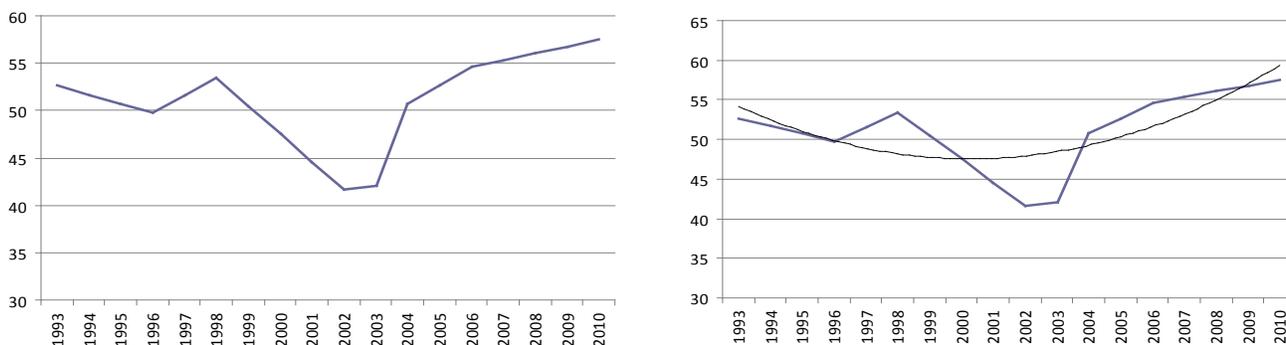
Source: Authors' elaboration.

As shown in Annex 1, we followed the same approach for all 29 countries selected. The figures in Annex 1 show that in most cases, the Gini from different data sources (identified by dots of different colors) point to trends that are similar to that we retained (identified by the blue line). Perceptible differences in levels or trends are evident for only a few years, as in Ghana (early 90s), Lesotho (late 80s), Madagascar (two observations), Mozambique (2008), Nigeria (1992 and 1996) and South Africa (1993-4).

III. Limitations of IID-SSA and the "seven sins of inequality measurement" in sub-Saharan Africa

Hereafter we discuss the statistical problems that may reduce the precision of the estimates of the <u>level</u> of the IID-SSA Gini data. In addition, if the measurement biases discussed below vary in intensity over time, the inequality <u>trend</u> may be affected as well, as would the analyses of the dynamics of income inequality and poverty in the region.

Although substantial progress has been made in recent years, survey data still present several problems that make difficult to identify the real level (and trend) of inequality in SSA. According to Klasen (2014), many factors contribute to this situation, including the weak capacity of the Central Statistical Offices (CSO) of the region as well as the weight of various external actors with different informational needs in deciding the data that has to be collected. These two conditions affect not only the ownership but also the design and comparability of surveys provoking important consequences in terms of data quality and comparability (Sandefur and Glassman, 2013).

Hereafter we discuss in detail the 'seven measurement sins' affecting the assessment of inequality levels and trends in the region. Such sins are not observed only in SSA, and are in fact common to most developing and – to a lesser extent - developed countries. Yet, given the specific characteristics of the region (a highly informal and little monetized economy, large seasonal income/consumption fluctuations, weak statistical institutions, dependence on technical assistance, and weak political checks and balances), such measurement sins are more pronounced in the region. These are discussed in what follows:

3.1 Differences in survey design for different years for the same country.

The region has a comparatively shorter experience with HBS. The methodology of data collection and survey design is evolving so as to reach higher standards, and as a result survey design is often modified in different survey rounds. Sometimes these changes are related to the data availability while other times they respond to the need of improving the quality of the information (Rio Group, 2006). For example, Grimm and Günther (2005) show that the design of the Burkinabé household survey has continuously improved over the years. In particular, they report that the 1994 HBS (EPII) and the 1998 EPII were built on data collected in the pre-harvest period (April-August) while in the previous one (EPI) data were collected in the post-harvest period (October-January). Moreover, "*whereas the EPI has a recall period for food items of 30 days the EPII and the EPIII have a recall*

period for food items of 15 days, and third, the disaggregation of expenditures was continuously increased between 1994 and 2003" (Grimm and Günther, 2005: 10). Likewise, McCulloch et al (2000) report that comparability of different survey rounds represents a serious issue in Mauritania. While the 1987/88 LSMS includes 62 food and 56 non-food items, the 1992 and 1993 Priority Surveys' questionnaire report information for only 12 food items and none on non-food items. More generally, the application of different methodologies for diaries or recall interviews (Gibson, 1999), changes in the reference period (Gibson, Huang and Rozelle, 2003) or in the number of food items included to measure consumption (Lanjouw and Lanjouw, 2001) could jeopardize the comparability of data over time (Jolliffe, 2001) placing serious obstacles to the analysis of inequality in SSA.

As mentioned, the World Bank's I2D2 tries to reduce these problems of comparability over time by harmonizing as far as possible – even ex post – HBS data, by aligning the number of consumption items in different surveys, filling in missing data, and so on. Over the long term this problem should lessen, but it still represent a hindrance in several countries.

3.2 Differences in statistical assumptions and data harmonization across countries.

In recent years, the use of time series and panel econometrics has increased the demand for homogenized questionnaire formats in order to ensure cross–country comparability. Recent examples of this type of projects are the EU-SILC in Europe and MECOVI in Latin America. In SSA, despite a growing number of very different surveys (Figure 3), there are not yet similar initiatives in place (the I2D2 and RIGA projects may fill this gap in the future). As a result, differences across countries in survey design, definitions, degree of disaggregation, income concept, timing, size of surveys, recall period and data processing conventions tend to reduce data comparability. For example, the Malawi Third Integrated Household Survey 2010/2011 provides detailed information for different sources of income while the Burkina Faso Enquete Integrale (2009/10) provides less accurate information especially in terms of private and public transfers. While some of these problems can be handled through the use of dummy variables (as in the case of different income concepts), in others the only solution to ensure comparability is consistent data harmonization.

Figure 3. Types of surveys in African countries, 2000-2011
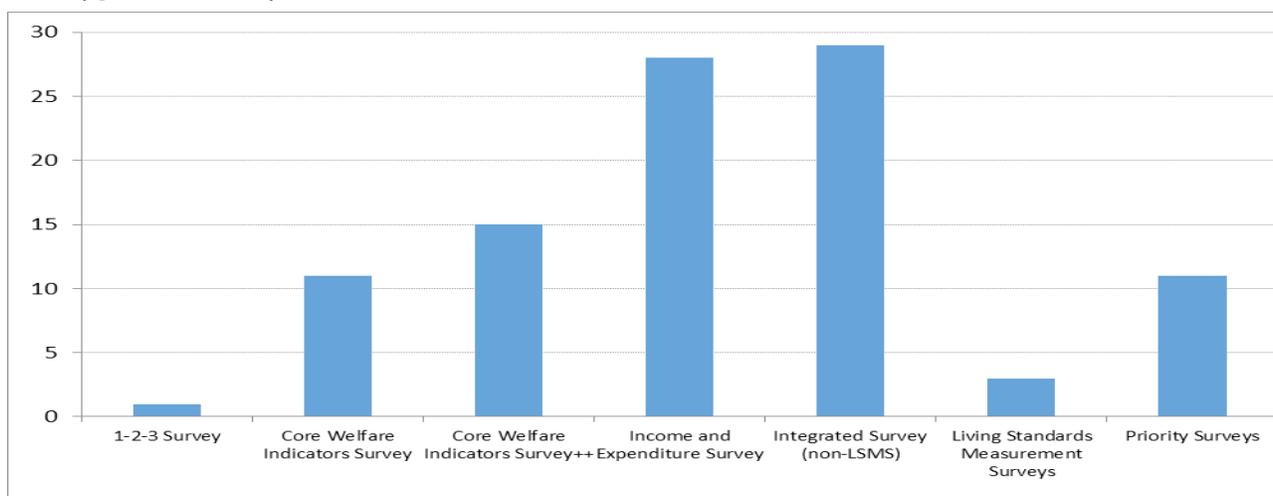
Source: Dabalen, Mungai and Yoshida (2011).

To ensure cross-country data comparability, harmonization ought to start from microdata and adopt for all countries and years the same statistical conventions to define the variables, 'household income/consumption per capita', 'household' (whether it includes external members such as renters, domestic servants and their families); the grouping of capital incomes; the corrections made for differences in recall periods; the imputation of the income/consumption stream from owner-occupied dwellings; the adjustments for non-responses (through matching techniques or the coefficients of a Mincer equation); the imputation of missing incomes and incomes in-kind; the treatment of zero incomes; the grossing-up of income under-reporting; and the upward adjustments of rural incomes to capture differences in rural–urban prices.

This harmonization process improves data comparability but entails that the newly produced Gini coefficients deviate from those generated by CSOs, which may use different statistical assumptions and imputation techniques from those adopted by World Bank. In several Latin American countries, the deviation between standardized SEDLAC and national CSO Ginis is negligible, but in others, it reaches 1. 5-3 points. Yet, it is rare that differences in inequality <u>levels</u> are accompanied by differences in <u>trends</u>. What matters is that the inequality trends coincide, and they generally do.

3.3 Under-sampling of top incomes, top income shares, and integration of HBS-based inequality data with those obtained from tax returns

The conclusions about inequality levels and dynamics reached on the basis of IID-SSA is likely to be biased by the vastly incomplete accounting of top incomes in HBS. This is due to their systematic under-sampling and under-reporting and to the truncation of very-high, low-frequency incomes that are treated as outliers. Such underestimation is more serious with regard to income than consumption data (Deaton and Grosh, 2000) and is more evident in developing countries with a large informal sector, considerable oil-mining resources, and weak institutions. In all these cases, the latent 'true Gini' is higher than the Gini derived from HBS. This situation leads to an underestimation of the <u>level</u> of 'true inequality' at any point in time. In addition, if the underestimation bias changes over time, it may distort the Gini time <u>trend</u> with the possible effect of leading to the identification of fake causal relations.

This measurement bias can be tackled by combining HBS data with data derived from tax returns which allow to estimate the income share of the top 1% or other top percentiles. In this regard, the World Top Incomes Database (WTID) (http://topincomes.parisschoolofeconomics.eu/) has generated a large volume of information for more than twenty countries to date, while other countries are being gradually added. For SSA the WTID already provides information for Mauritius, South Africa and Tanzania (only for the years 1950 – 1970), while similar studies are being conducted for Botswana, Cameroon, Gambia, Ghana, Kenya, Lesotho, Malawi, Nigeria, Seychelles, Sierra Leone, Swaziland, Uganda, Zambia and Zimbabwe.

While studies on the share of top incomes crucially depend on how broad based is taxation in these SSA countries (where often only few corporations and individuals file tax returns) and the extent of tax elusion and evasion, they nevertheless provide additional information on the upper part of the distribution of income which is missed by HBS. For SSA, for instance, there is evidence that the income share of the top 1 percent has risen sharply during the last twenty years in Mauritius and South Africa (Figure 4). The HBS based Gini trend in Table 1 show that inequality has risen during the last decade, but these data underestimate the extent of such increase, as shown below.

Figure 4. Top 1% income share in Mauritius and South Africa, 1990-2011
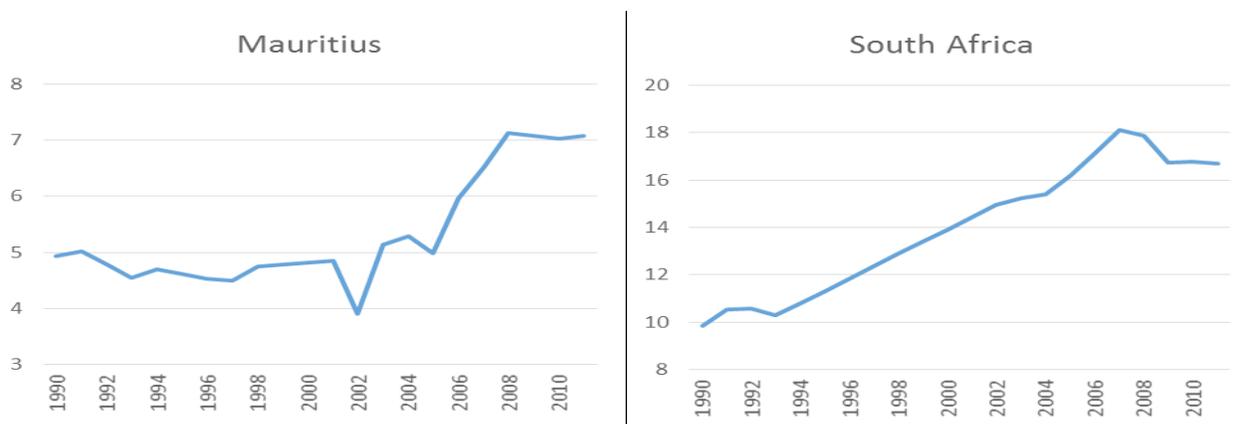
Source: The World Top Incomes Database.

The approach to 'correct' the HBS Gini consists in using tax returns data (and in particular the share of the top 1% or 0.1%) to compute G*, the 'true Gini coefficient' by means of the following (or similar) formula $G^* = G(1-S) + S$, where S is the income share of the top 1 per cent estimated on the basis of tax returns (Alvaredo 2010). Empirical evidence from developed and developing countries show that G* is higher by several points than the Gini estimated on HBS data. For instance, data for the last decade for Colombia, Argentina and Uruguay show that G* is always higher than G by 3-6 points (Cornia 2015). This gap is, however, fairly constant over time, suggesting that the end of apartheid has not reduced the weight of the elites and may have in fact increased overall inequality as the distance between the corrected and uncorrected trend rises from two to about five points between 1995 and 2010. If this is true, we have a 'level effect' but not a 'trend effect'—which means that the conclusions reached on the basis of the uncorrected Gini G hold – except for a fixed effect. Figure 5 below illustrates well the point in the case of South Africa. It shows a 'level effect' (as the corrected Gini is higher than the HBS one by 2-5 points) but not a major difference in the two trends that are almost parallel, though their distance rises over time by three points.

Figure 5. South Africa, trend in the HBS-based Gini coefficient (bottom line) and the Gini corrected on the basis of tax returns data (upper line),1990-2010.
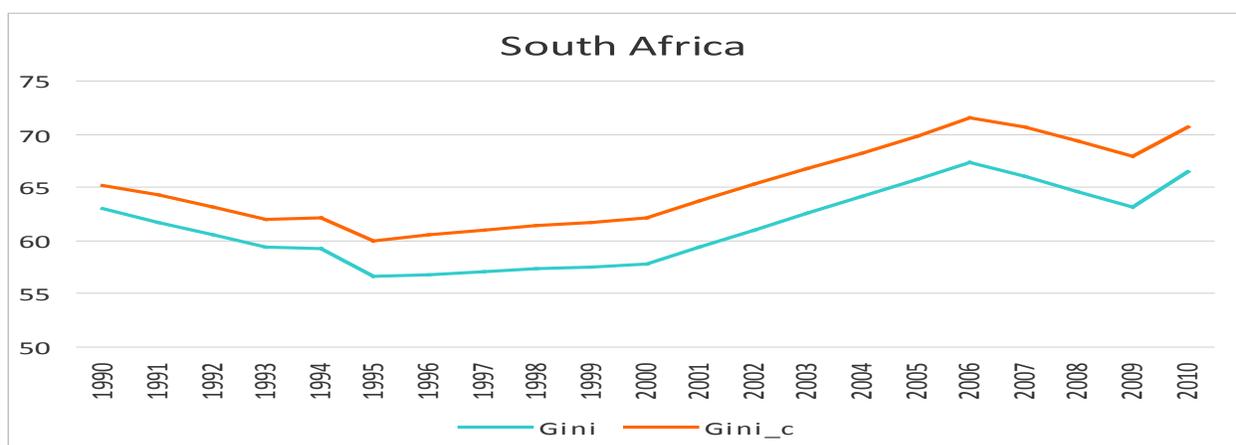
Source: Authors' calculations based on World Top Incomes Database (WTID).

3.4 Cross-checking trends of HBS-based Gini against those in the "labour share"

Another way to check if the trends in HBS-based Gini coefficients are robust is to juxtapose them with those of the 'labor share'(LS) in total net value added. It is possible in fact that under-sampling of top incomes in HBS may prevent a correct representation of the recent increase in capital incomes due to the financialization of the economy and the rise in mining rents. These may be captured by a rise in the 'capital share' or fall in the 'labor share' calculated on the basis of the national accounts. But also this approach has its limitations which depend on the accuracy of national accounts (that are also known to suffer from sizeable estimation errors), on the hypothesis made to compute the LS[12] and on possible offsetting trends in terms of redistribution of gross incomes, for instance through the taxation and redistribution of mining rents.

A second reason to check the HBS-based Gini trends with the LS is that HBS substantially, and often increasingly, underestimate the total net value added. For instance Ravallion (2001) shows on Indian data for the 1990s that the HBS-based mean income per capita was only 60% of the value computed on the basis of the National Accounts, and that such ratio declined over time. In contrast, he found that the difference was not as large in the SSA countries.

[12] There are several definitions of the 'labour share'(LS). The simplest (LS1) is: (compensation of employees) /[total value added – (indirect taxes + consumption of fixed capital)] but this definition poorly fits the reality of developing countries where most people are self-employed. LS2 is more appropriate: (compensation of employees + 2/3 of mixed incomes)/ [total value added – (indirect taxes + consumption of fixed capital)]. There are other theoretical refinements, but in the case of SSA the difficulties in estimating the total value added as well as the factorial distribution of income weaken the empirical strength of more sophisticated estimates of the 'labor share'.

Differences in the level of LS and Gini coefficient are to some extent physiological, as the information on which they are based has been collected in different ways and with different purposes in mind. For example, consumption and income level derived from HBS are based on information self-reported by sampled individuals or households and are subject to large recall errors and other biases. In contrast, income and consumption derived from the national accounts are derived from the accounts on total production and uses of GDP. Next, HBS refer to the income and consumption of households while the total value added measured by the national accounts includes also that of communities (religious, military, rest homes, residential schools and so on). Finally, HBS data generally refer to net incomes (after direct taxes and transfers) while the LS concerns the distribution of gross market income. Thus, it is not surprising to observe different results for income or consumption per capita. The problem arises when the trends in these two distributive indicators move in the opposite direction.

To test whether the trends of HBS-based Gini and LS go in the same direction we rely on Guerriero (2012) who computed labour shares for 25 SSA countries (at times only until the 1990s) using national aggregate data from the United Nations National Accounts Statistics for applying different methodologies to compute alternative LS. She shows that the LS declined over the last few decades in several countries, in particular from the 1980s onwards. These trends (Figure 6) only in part confirm the Gini tendencies identified in Table 1. For example, in Senegal the two trends agree (the LS rose while the Gini coefficient declined). The IS and Gini trends are consistent with each other also for Botswana (the LS fell while the Gini coefficient rose). For Kenya the two trends are consistent (rising Gini coefficient and fall in the LS) since 2003 but not before. In contrast, the fall in the labour share in Lesotho is accompanied by an inconsistent fall in the Gini coefficient (Figure 6). As noted this may be due to accounting problems or to offsetting measures.

Figure 6. Evolution over time in the labour share, selected SSA countries and years
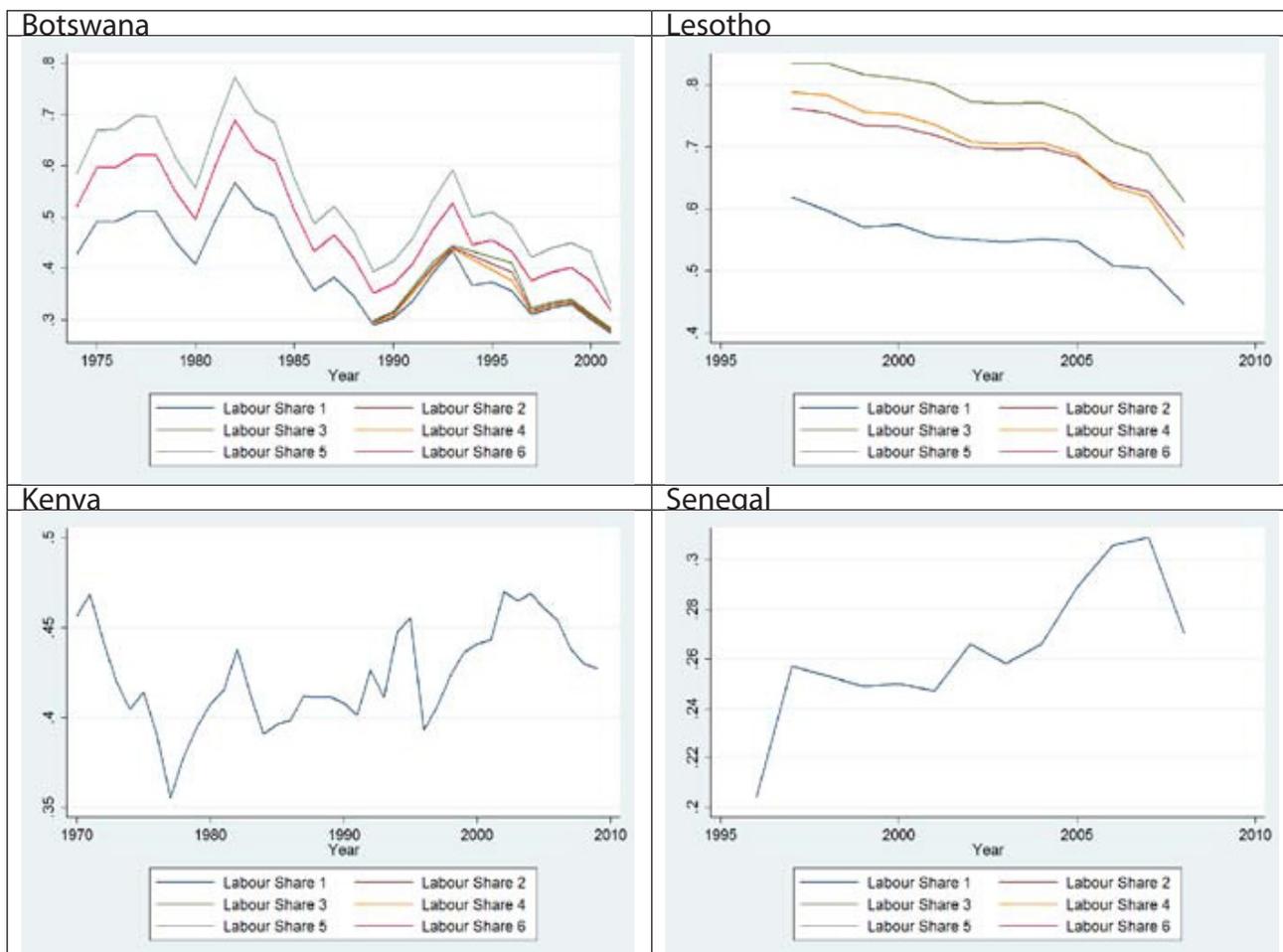
Source: Guerriero (2012).

3.5 Ignoring the incomes accruing on assets held abroad by SSA nationals

Even assuming that the <u>domestic</u> incomes of the rich are fully reported in HBS (or that are added to HBS data on the basis of tax returns, see above), survey data provide a partial picture of the national income distribution whenever SSA national elites hold abroad an important share of national assets, either legally or illegally. Indeed, the incomes received on these assets do not enter the calculation of national income and its distribution.

A rich literature suggests that several SSA countries are a source of substantial capital flights, that substantial assets are held abroad, and that these generate incomes that escape any form of accounting in the home countries. In countries with a liberalized capital account, capital outflows may be the results of a rational portfolio diversification aiming at legally shifting some savings to countries with higher return on assets, lower taxation, or lower risk of default. However, these flows become capital flights if the national norms on taxation and capital controls forbid them. Most importantly, a large part of capital flights consists of the laundering of illicit earnings (from narco-trafficking or theft of national resources) or of shipping abroad resources obtained through the embezzlements of the proceeds of natural resource exploitation. The literature surveyed in Ndikumana (2014) indicates that at least 8 percent of petroleum rents earned by oil rich countries with weak institutions ends up in tax heavens located mainly in advanced countries.

There are two methods for estimating the volume of capital flights, an indirect method and a direct one. The indirect method measures capital flights (KF) as the residual of balance-of-payments components. Following Boyce and Ndikumana (2012), capital flights can be estimated as the difference between the 'inflows of foreign exchange' (debt-creating capital inflows, equal to the change in total debt outstanding owed to foreign residents, adjusted for exchange rate fluctuations, plus foreign direct investments) minus the 'uses of foreign exchange' (the financing of the current account deficit CA, and the change in currency reserves ΔRES). In symbols KF = (Δ DEBTADJ + FDI) − (CA +ΔRES). In principle, the two right hand side terms should equal each other, and their imbalance should be indicative of a capital flight. To such imbalance one must add the value of trade misinvoicing (over-invoicing of imports and under-invoicing of exports) which according to the Global Financial Integrity (a US NGO) constitutes some 2/3 of capital flights. Finally, an additional correction is included for remittance inflow discrepancy (RID) i.e. unrecorded remittances (estimated at 50% of the total in SSA) so that the above equation becomes KF = (Δ DEBTADJ + FDI) − (CA +ΔRES) + MISINV + RID.

Following this method, Ndikumana (2014) estimated that over the years 1970-2010 the accumulated capital flights from 35 main SSA countries totalled US $ 820, and that the estimated capital held abroad in 2010 (capital flights plus accumulated interests and profits) was 1067 bn. in 2010 US$. Capital flights were particularly important in oil rich countries such as Nigeria, Angola, Congo and Sudan. These data would suggest that SSA is a net creditor to the rest of the world since the value of (private) assets held abroad exceeds total (mostly public) liabilities of 283 US$ bn owed to foreign creditors. The volume of capital flights seems to have worsened in recent years in conjunction with the rise in the price of oil and other commodities.

A drawback of the indirect method followed by Ndikumana and others is the assumption that all KF are to be attributed to capital flights, while they could be due to the under-registration of many foreign transactions (including licit transactions) due to weak administrative capacity and economic informality. The assumption that all KF are capital flights is thus questionable, and should be cross-checked using alternative methodologies. To tackle this problem, the 'direct method' of estimation of capital flights focuses on the identification, measurement and analysis of variables that are outcomes of capital exits, such as bank deposits or housing property held abroad by developing countries citizens. This is the approach attempted at the moment at the Paris School of Economics for all countries including SSA (we refer here to the unpublished ongoing work of Cogneau and Rouanet). Data for 1980-2010 on deposits held by foreigners in the 44 countries part of the Bank of International Settlements (BIS) are provided by BIS countries' banks to their central banks. These data are aggregated by central banks and transmitted to BIS. In other words, for all 44 BIS countries Cogneau and co-authors count on information on bank assets of residents of more than 200 countries.

These data show that in 2010 SSA countries held abroad some 5.3 percent of GDP (6.1 if South Africa is excluded) or 48 percent of the value of M2, and that deposits held abroad represent 16.6 percent of domestic money and quasi-money (the same ratio is only 9.7 percent in Latin America and even lower in any other region). This means that an important fraction of monetary savings is found outside the region rather than invested at home. In 2000, the main African oil-producers (Angola, Nigeria, Gabon, Congo, Cameroon), but not Sudan, held abroad deposits of around 7 percent of their GDP. Suggestive evidence cited by Cogneau et al suggests that oil price windfalls are passed to bank deposits abroad, with transmission rates ranging from 2 to 12 percent, with larger countries displaying larger flows and stocks of assets held abroad.

In absolute terms, these results are similar to those identified by Ndikumana. However, all correlations disappear when expressed in proportion of GDP. This discrepancy with Ndikumana's results may suggest a certain inaccuracy of the indirect method. Be that as it may, except for South Africa, SSA appears to be the region with the highest proportion of foreign deposits as a share of domestic money and deposits.

The distributive impact of all this is important but not easy to estimate. Given the massive amount of wealth held in safe havens and that are not incorporated into national income and expenditure accounts, it appears that the standard measures of income inequality and wealth distribution are substantially underestimated. If we accept Ndikumana estimate of 1067 bn 2010 US$ in assets held abroad by SSA citizens in 2010, and if we assume assume an average 5 percent rate of return on assets, then some 53 billions (or 3-4 percent of GDP) of additional income that escape the national accounts, accrue to the top echelon of the SSA society, and cause an average regional underestimate of 2-3 Gini points. Figure 7 illustrates the case of Cote d'Ivoire in 2008. According to our estimations the impact on the Gini coefficient of additional incomes that escape the national accounts is around 1.5 points. Such upward adjustment in the Gini coefficient of the distribution of national income is substantially higher in oil exporting countries.

Figure 7. Estimated impact on the Gini coefficient of additional income that escapes the national accounts, Côte d'Ivoire, 2008
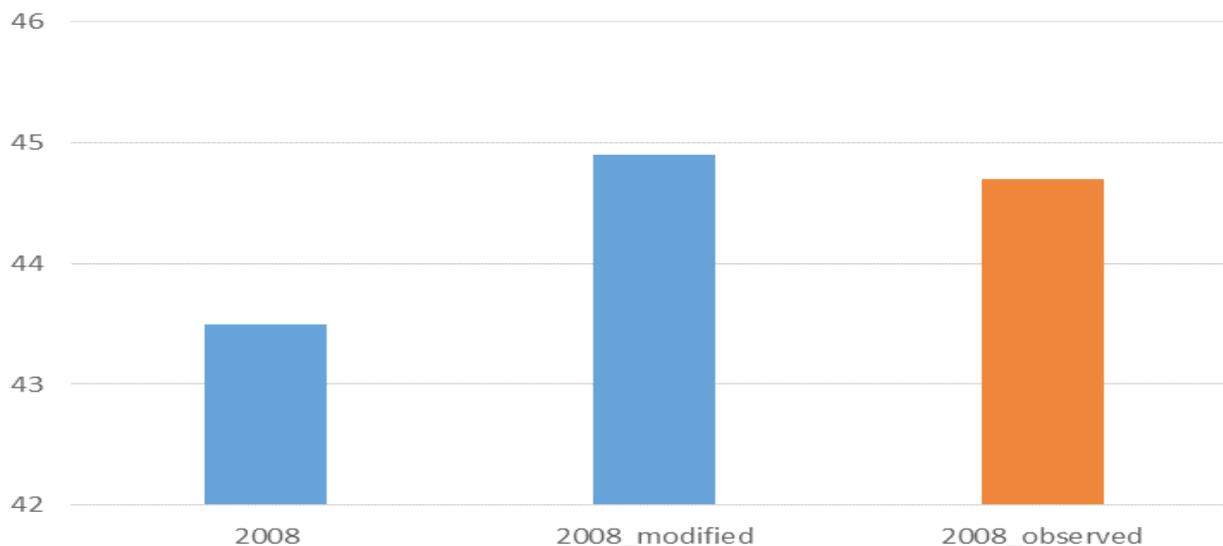
Source: Authors' elaboration based on WIIDv3 data.

3.6 Distributive impact of differences in price dynamics between food prices and overall CPI

The inequality indexes (Gini, Theil, or others) of the distribution of per capita income/consumption are generally computed using current price data. Computing the same indexes at constant prices yields the same results, if the current incomes of all percentiles are divided by the same consumer price index (CPI) or rate of inflation. This common procedure implicitly assumes that all households pay the same price for all goods included in the CPI consumption baskets and that changes over time in these prices affect all households in the same way. In addition, consumption (and often income) are generally recorded on a monthly or weekly

basis, assuming implicitly that such prices are stable throughout the year. These three assumptions introduce a considerable downward bias in the calculation of the Gini index as first, at any point in time, the poor tend to pay more for food (and other items) than better of people, so that their real purchasing power computed on price-unadjusted data is overstated; second, this phenomenon is particularly pronounced during the lean season – as during these months the poor pay even higher prices for food, as lack of credit and storage does not allow them to purchase food when prices are low and store them for future consumption; and third, in periods of rapid food price increases relative to the prices of other goods (as over 2008-11), the CPI of the poor rises faster (as the poor allocate a greater proportion of their expenditure to food) and therefore their real incomes/consumption drop more rapidly than those of the middle-upper class. These three biases are discussed hereafter one by one – together with ways to correctly compute the real distribution of income/consumption, and with the policy measures that could shelter the poor from these adverse changes:

(a) **Differences in food prices at any point in time.** A number of studies (e.g. Gibson and Kim, 2013) have found evidence that the poor pay higher food prices compared to the non-poor. The literature presents a number of reasons for this phenomenon. Mendoza (2011) suggests this is due to the fact that reaching the poor may be more costly, because they live in remote areas characterized by high transport costs and/or lower personal and business security. Poor infrastructure and a risky environments make it costlier for retailers to reach the poor. A price premium is thus charged to recoup these extra costs. Second, even when they are located in urban and peri-urban areas the poor may pay higher prices due to greater liquidity constraints: indeed, the poor may buy food in small quantities, in less competitive markets, during suboptimal periods or on credit, and therefore do not benefit from the discounts granted to quantity/bulk purchases and cash payments. For instance, Mussa (2014) shows on the Malawian 2004 and 2011 Integrated Household Surveys that there is a 'poverty penalty' in inequality measurement. His results show that regardless of location and year, poor households pay more for food compared to non-poor households so that inequality based on a food price-corrected consumption data is much higher than that computed on un-corrected food-price data. According to his estimates, the nominal Gini coefficient underestimates the 'real Gini' by between 3.9 to 7.1 percent, i.e. by between 2 and 3.5 Gini points.

(b) **Food price seasonality.** The strong food price seasonality typical of many developing countries may further worsen the real purchasing power of the poor over the year. For instance, Cornia and Deotti (2015) show that in Niger the prices of millet peak in pre-harvest August during which they are at least 30-40 percent higher than in post-harvest September-October. In years of food crises (as in 2005), the seasonal price increase may be of 100 per cent or more in localized areas (Figure 8). While such price seasonality affects everybody, the poor suffer the most as their lack of liquidity and access to credit, need to repay debts incurred during the prior year by selling millet immediately after the harvest when prices are the lowest, lack of proper postharvest storage facilities, and absence of public interventions to provide affordable credit and build cereal banks increase massively the price they pay for millet and so reduce their real purchasing power, especially during the lean seasons, when food prices continuously escalate. Such problem – which causes a considerable underestimate of consumption/income inequality is extremely common in SSA. For instance, the price of maize - a staple food in Malawi- is sold cheaply immediately after harvest but bought at high cost during the lean season.

Figure 8. Monthly consumer price of millet (CFAF/gg): 2005 vs. 2004 and average 2000-2004
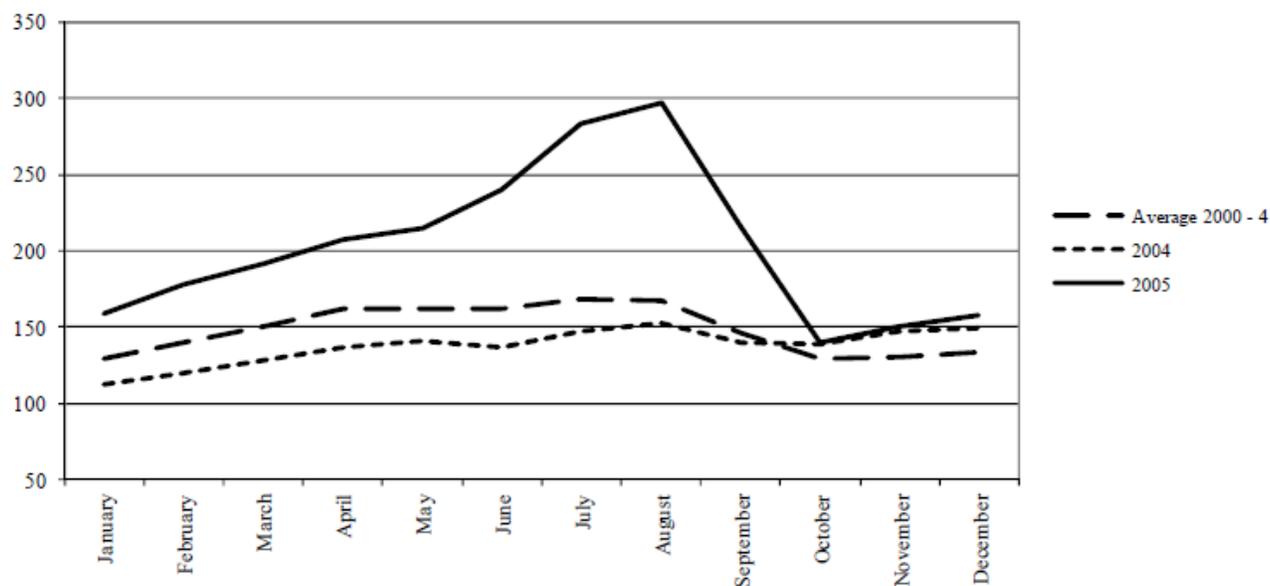
Source: Cornia and Deotti (2015) based on SIMA, National Dataset.

Note: A one-tail t test of the significance of the monthly variations (year on year) confirms at the 10. 9 per cent probability level the hypothesis that the 2005 prices were significantly higher than those recorded over 1994-2004. The significance rises sharply if the test is restricted to May-October.

(c) **Differential price dynamics between food and non-food items.** Faster food price changes over time in relation to other items tend to penalize the poor and worsen the distribution of income or consumption. As noted by Arndt et al (2014: 2) "*Since measures of income inequality are (typically) scale invariant, it follows that there should be no difference between nominal and real measures of income inequality where a single aggregate CPI is used to deflate nominal observations*". Yet, households in the bottom quintile have a different consumption basket than those at the top. In particular, the poor and the poorest assign a much greater proportion (up to 70-80 percent) of their total consumption to food, while those in the top decile assign to food 20-30 percent of their total consumption. This means that whenever the food price index (FPI) and consumer price index (CPI) diverge substantially over time (as observed during the late 2000s), the calculation of the Gini at current prices is substantially biased, as the real purchasing power of the poor is reduced more than proportionally (Grimm and Gunther (2005). These authors show for instance that in Burkina Faso the CPI rose by 23 per cent between 1994 and 1998 while the price of cereals increased more than 50 per cent over the same period (Figure 9). Thus – when taking into consideration the different dynamics of FPI and CPI - the percentage of the population living under the poverty line increased substantially. Likewise, Arndt et al (2014) document on data for Mozambique that income inequality worsened due to the sharp increase in world food prices over 2007–09, as the food consumption of poor households living in urban areas relied heavily on imported food.

Figure 9. Trends in the index number of the official poverty line (Off. PL), Consumer Price Index (CPI) and price of main staples (1994=100), Burkina Faso
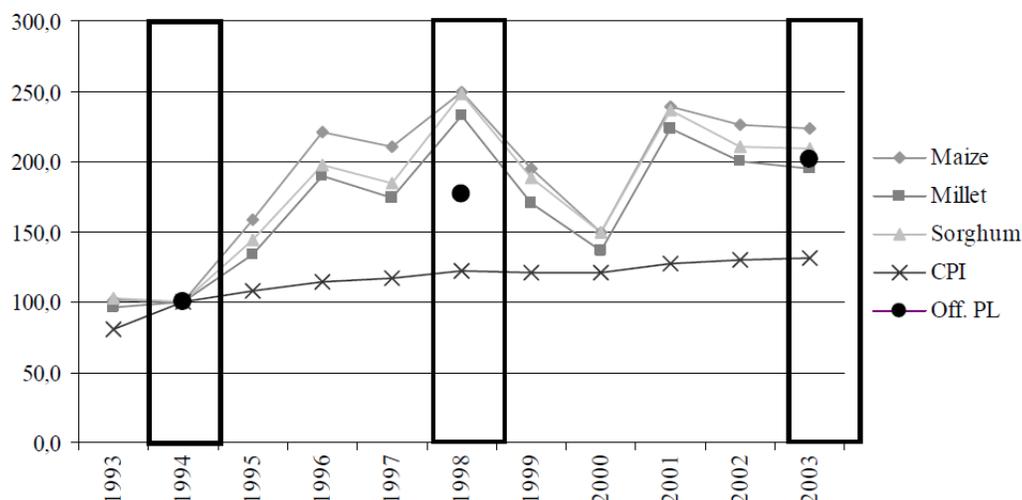
Source: Grimm and Gunther (2005).

Hereafter we further test the distributive impact of the observed changes in the FPI/CPI ratio, by calculating the impact of its changes on the Gini coefficient of four countries for which WIIDv3.0b provides quintile distributions for two years during the 2000s, a period characterized by important food price changes. We selected two countries where inequality rose (i.e. Malawi 2006-11, and South Africa 2000-6). In the first FPI/CPI fell and in the second it rose (Table 2 and Figure 10). We also selected two countries which experienced falling income inequality and where FPI/CPI fell (Mali, 2001-10) or rose (Madagascar, 2001-5).

Table 2. Summary of the impact of changes in the FPI/CPI ratio on the Gini coefficient

Country	Years	Inequality trend	% change in food price index / consumer price index (FPI/CPI)	Δ Gini
Malawi	2006-2011	rising	-9.1	-0.6
South Africa	2000-2006	rising	10.1	0.3
Mali	2001-2010	falling	-20.3	-0.9
Madagascar	2001-2005	falling	17.5	1.5

Source. Authors' elaboration.

To simulate the impact of the FPI versus CPI divergence we used the quintile distributions obtained from WIIDv3.0b and assumed from the literature the following 'plausible food consumption shares' for quintiles in ascending order, i.e. 0.7, 0.6, 0.5, 0.4, and 0.3. We assume that these shares are the same for all four countries considered. To ensure comparability between the values of the Gini coefficients of the first and the second year (as the ratio FPI/CPI had changed significantly), we recalculate at time t+1 the quintile distribution corrected for changes in FPI/CPI by means of the following formula:

$$CQ_{it+1} = [(OQ_{it+1} \cdot sh_{food}) / (FPI/CPI_{t+1}/FPI/CPI_t)] + (1 - sh_{food})$$

where CQ_{it+1}, OQ_{it+1} are the corrected and original quintiles values at t+1 of quintile i, and sh_{food} is its food share in total consumption.

The results presented in Figures 10 and 11 are summarized in Table 2 which shows that the simulated changes in the Gini coefficient are generally moderate, ranging between 0.3 and 1.5, These low values are due in part to the fact that we used the quintiles distribution which lead to a lower Gini than that estimated on micro data and which is reported as the last bar in each figure which is generally 2-3 points higher than that computed on the quintile distribution.

Figure 10. Impact on the Gini coefficient of changes in the FPI/CPI ratio in Malawi (left panel, rising inequality, falling FPI/CPI) and S. Africa (right panel, rising inequality and rising FPI/CPI).
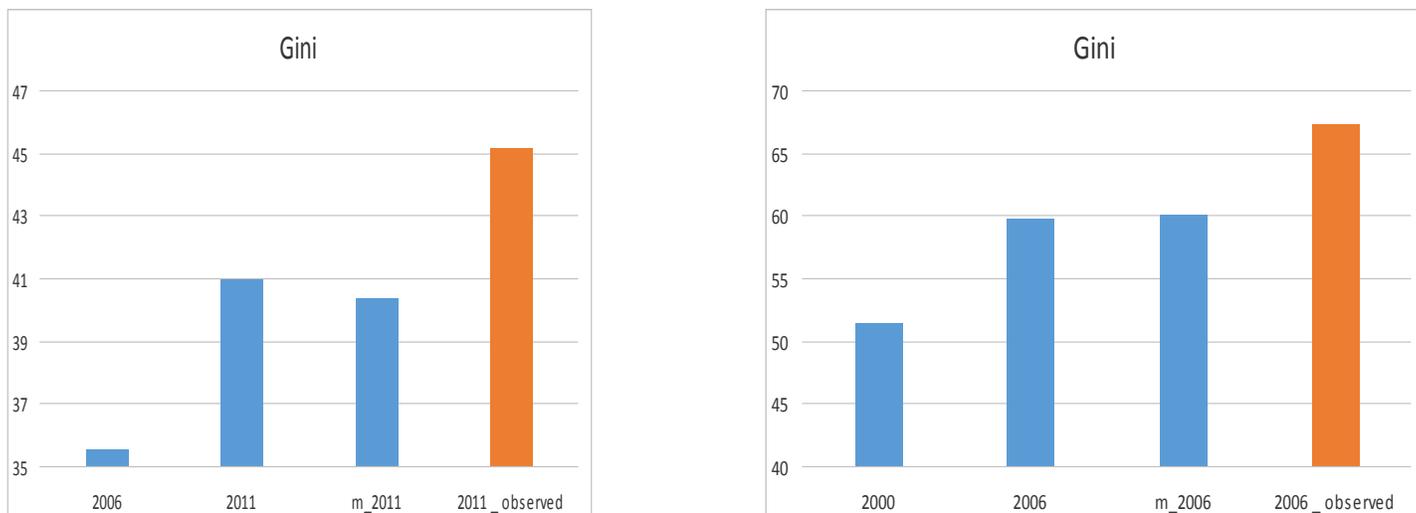
Source: Authors' elaboration.

Notes: The first two bars from the left represents the Gini coefficients computed at current prices on the basis of the quintile distribution reported for the relevant years in WIIDv30b. The bar with an 'm' (modified) in front has been corrected for differences in FPI/CPI. The last bar is the value of the Gini included in the IID-SSA, which is higher because it is computed on micro-data.

Figure 11. Impact on the Gini coefficient of changes in the FPI/CPI ratio in Mali (left panel, falling inequality and falling FPI/CPI) and Madagascar (right panel, falling inequality and rising FPI/CPI)
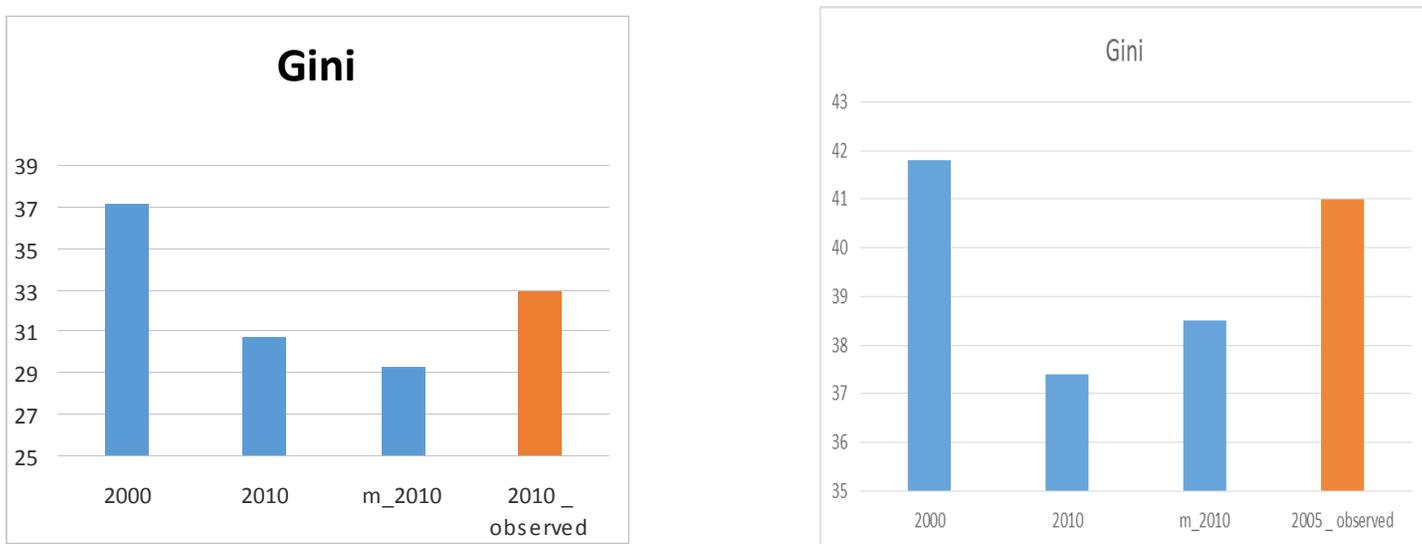
Source: Guerriero (2012).

Note: The first two bars from the left represents the Gini coefficients computed at current prices on the basis of the quintile distribution reported for the relevant years in WIIDv30b. The bar with an 'm' (modified) in front has been corrected for differences in FPI/CPI. The last bar is the value of the Gini included in the IID-SSA, which is higher because it is computed on micro-data.

As one can see, in the four countries selected (where the FPI/CPI price changes were marked) the Gini coefficient changed in a non negligible way (up to 1.5 points).

We now enlarge the test to 18 countries for which we dispose of corrected Gini and FPI/CPI data for the years 2000-2012 (a period during which the FPI/CPI ratio rose in the SSA countries by between 5 to 30 percent, while in only a few it fell) to see whether changes in the latter variable may have affected the values and trends of the Gini coefficients summarized in Table 1 which were used to analyze inequality trends in SSA in Cornia (2014). We test the bivariate relation between the time differences of the FPI/CPI index (x axis) and the first difference between the uncorrected and corrected Gini coefficient (y axis). The test confirms the expected results (Figure 12), i.e. a 0.52 points rise in Gini for an increase when FPI/CPI rises by ten percent. The relation seems stable as suggested by the high value of the R2 (0.62).

Figure 12. Relationship between the first difference over time of the FPI/CPI ratio (x axis) and the first difference of the Gini coefficient), 18 sub-Saharan African countries, 2000-2012
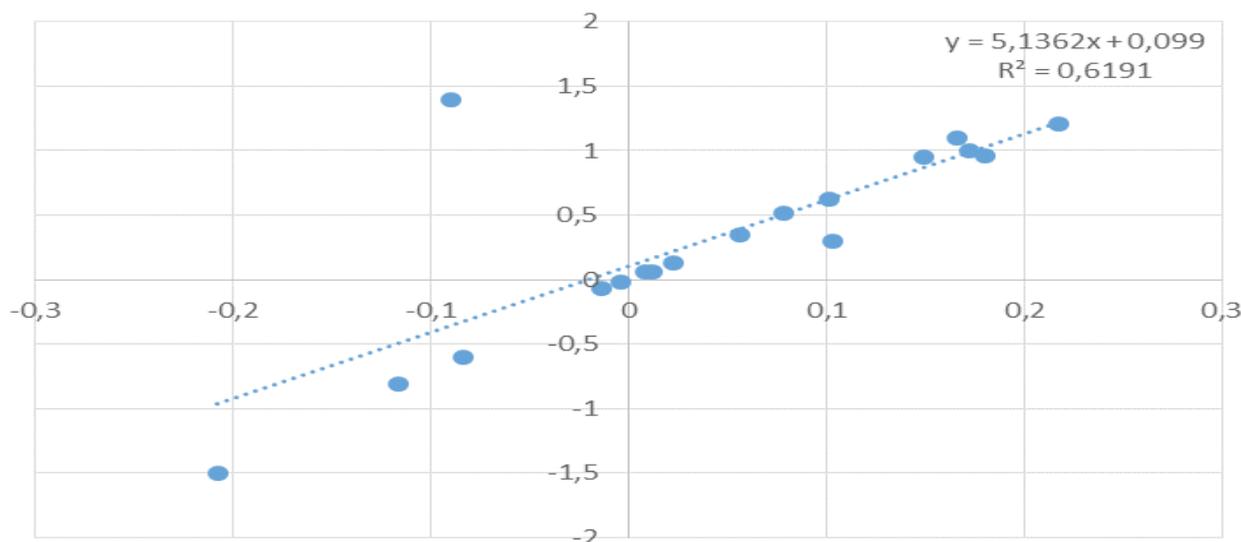
Source: Authors' elaboration.

## 3.7 Distributive impact of differences in the provision of the "social wage" across countries

For sake of completeness, we also briefly mention another aspect that needs to be discussed when looking at the distribution of wellbeing among citizens, though in practice it is difficult to take it into account for a host of data and theoretical reasons. So far, we have discussed the distribution of private income and consumption (which include income transfers from the state, where these exist). Yet the individual and household wellbeing depends also on the mount of services-in-kind provided by the state, with particular reference to health and education. Indeed, any comprehensive welfare comparison (over time and across countries) should take into account the monetary value and incidence of the services supplied in kind by the state to the various quintiles of the population. In the absence of state provision of these social services, households would have to buy them in the market reducing in this way their net income and the consumption of other essential items(e.g. food).

The overall value of public expenditure on health and education in SSA is comparatively low. In particular, the expenditure on health was 2.4 per cent of GDP in 2000 and increased up 2.8 per cent of GDP in 2010. Public expenditure on education was close to 3.5 per cent of GDP in 2000 and 4.3 per cent of GDP in 2010. However, it is necessary to highlight that there is considerable variation across countries. For SSA as a whole, Davoodi et al (2003) show that in the late 1990s their incidence was not progressive, even for primary health care and elementary education (Table 3), but was less regressive than that of private income and consumption, generating in this way a modest redistributive effect. With the emphasis placed during the last decade on

the MDGs, and the spread of democracy it is possible that the incidence of public spending on health and education improved (see below).

Table 3. Benefit incidence of public spending on education and health, sub-Saharan Africa, 1990s, (%, unweighted averages of total sectoral spending)

No. of sample countries	All		Primary education		Secondary education		Tertiary education		Richest
	Poorest	Richest	Poorest	Richest	Poorest	Richest	Poorest		Richest
10	12.8	32.7	17.8	18.4	7.4	38.7	5.2		54.4

No. of sample countries	All		Primary health care		Health centres		Hospitals	
	Poorest	Richest	Poorest	Richest	Poorest	Richest	Poorest	Richest
9	12.9	28.6	15.3	22.7	14.5	23.7	12.2	30.9

Source: Excepts taken from Tables 2 and 3 of Davoodi et al. (2003).

The literature shows that the incidence of education and health spending tends to be more pro-poor in richer than poorer countries. In addition, countries characterized by higher income or consumption inequality (like South Africa) spend a greater amount of resources and have a more pro-poor incidence of public spending, possibly as a result of the policymakers' intention of reducing income disparities. For instance Figure 13 below shows that while the gross income Gini was 0.69, social spending on health and education reduced it by a massive 17 Gini points, while cash transfers reduced it by another 5 points. All this suggests that public expenditure policy (both cash subsidies and services in kind) can be a potent tool to equalize the distribution of overall (private and social) income as confirmed recently by Ostry et al. (2014) on a large country panel.

Figure 13. Impact of cash transfers and social spending on health and education, South Africa, 2009
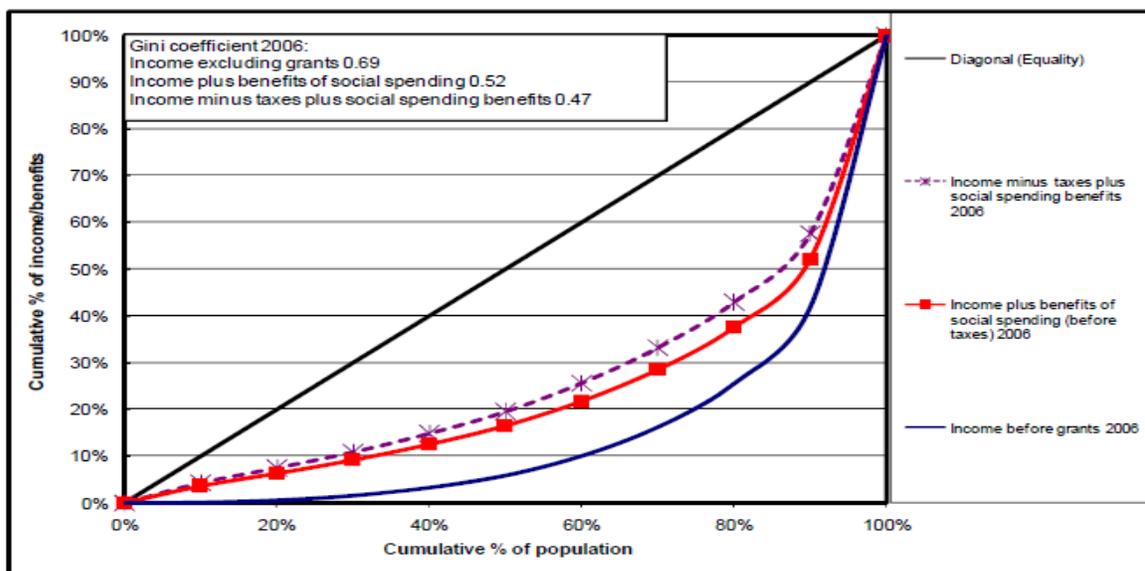
Source: Van der Berg (2009).

In contrast, in poorer low inequality SSA countries (such as those of the Sahel) that are characterized by limited public spending on health and education, the redistributive role of the state via the provision of public social services is more limited.

IV. In conclusion

The paper has illustrated in Section 2 the way IID-SSA has been built and provides an important contribution to the identification of inequality trends in the region that has been analyzed in other studies part of the UNDP project on 'Inequality in SSA'. IID-SSA will be updated at the end of the UNDP project on 'Inequality in SSA'. Hopefully the updating will benefit from the release of the harmonized I2D2 World Bank data. The effect of eventual changes in the level and trends of inequality indexes will be taken into consideration when drafting the final analysis of the causal relationships explaining the inequality dynamics in SSA and the policy recommendations on how to moderate inequality.

In turn, the review carried out in Section 3 has illustrated the main problems encountered in the measurement of income and consumption inequality and the possible corrections needed to compute more realistic inequality figures, especially in the highly informal economies of the region. UNDP, World Bank and academic analysts of country inequality may wish to take them into account when working on inequality and poverty in specific SSA countries. The main recommendations in measuring inequality levels and trends are the following:

a) Any analysis should start from a careful examination of the inequality statistics, so as to make sure that the data utilized refer to the same income concept, geographical coverage, period of the year and so on. The exclusion of inconsistent data – as attempted when building IID-SSA – entails a loss of degrees of freedom but is compensated by greater data cross country comparability and a lower risk of identifying spurious relationships;

b) If possible, survey micro-data should be harmonized ex-ante by using the same questionnaires and statistical conventions, as done in the RIGA project since 2005 and similar initiatives in Europe and Latin America. As for past data, the ex-post harmonization is also useful to improve data comparability but requires making many assumptions. The inequality statistics computed on data harmonized ex-post, as currently done by the World Bank for SSA or by the SEDLAC project for Latin America, differ from those calculated by national CSOs, at time by 1-3 Gini points. However, at least in the Latin American case, this difference seems to concern only the level and not the time trend of such indicators. But there might be exceptions.

c) Even harmonized HBS do not fully and faithfully measure the 'true inequality' existing in a country, as top income earners are undercounted in household surveys and as the returns on assets held in safe havens by national elites are not included in either surveys and national accounts statistics. The discussion presented above shows for instance that in South Africa the inclusion of top incomes raises the Gini coefficient by 3-5 points. Likewise, if included in the distribution of national incomes, the return on assets held abroad would raise the Gini coefficient by another 2 points. Altogether, this means that our IID-SSA data underestimate the 'true Gini' by a massive 5-8 Gini points, possibly more in countries accumulating large rents from the export of valuable primary commodities. The main analytical issue here is whether such underestimation concerns only the level of Gini (a fact which is certain) or also its trend. Figure 5 on South Africa suggests the trend is less affected than the level, but this may not be true in countries like Angola or Equatorial Guinea where the oil discoveries of the last 10-15 years and weak redistributive institutions are likely to have changed not only the Gini level but also its time trend. Overall, the IID-SSA Gini data used in Cornia (2014) represent a lower-bound estimate of the 'true Gini', especially in countries with high asset concentration and exporting valuable primary commodities;

d) Large (15 percent or more) upward deviations of the Food Price Index from the CPI entail an additional increase in the Gini coefficient and poverty rates, and also in this case researchers should therefore taken into consideration such divergence when analyzing trends and designing policy;

e) The trends in the labour share (which has been used with great fanfare recently, as in the case of Piketty's work) can help cross-checking the robustness of the Gini trends described by IID-SSA. Yet, given the accounting problems encountered in the calculation in the labor share in the case of very informalized and rural economies of SSA, they are of more limited use than in formalized economies;

f) The inclusion of 'social services' (health and education) in the calculation of the 'overall (private and public) household income or consumption per capita' likely reduces the Gini coefficient, even in many poor SSA countries, not too speak of middle income countries such as South Africa. More work is however required in this area to examine the overall volume and incidence of these services. This info would be useful for policy-makers aiming at redistributing wellbeing via the provision of such services that have been clearly shown to reduce inequality over the short term and across generations. As in the case of South Africa, the redistributive effect of services in kind appears to be much larger than that of cash transfers.

References

Alvaredo, F. (2010). 'A Note on the Relationship between Top Income Shares and the Gini Coeffi cient' . CEPR Discussion Paper 8071. London : Centre for Economic Policy Research. Available at: <http:// www.cepr.org/pubs/dps/DP8071. asp>

Arndt, Channing, Jones, Sam and Vincenzo Salvucci (2014), "When do relative prices matter for measuring income inequality?," WIDER Working Paper 2014/129.

Bourguignon, F. (2003). The growth elasticity of poverty education: Explaining heterogeneity across countries and time periods. In *Inequality and Growth. Theory and Policy Implications*, T. Eicher and S. Turnovsky, eds. Cambridge: The MIT Press.

Boyce, James K., and Léonce Ndikumana (2012) "Capital Flight from Sub-Saharan African Countries: Updated Estimates", 1970-2010. PERI Working Papers

Cogneau Denis and Léa Rouanet (2015), 'Capital exit from developing countries: Measurement and correlates' January 2015 – PSE, Preliminary document for discussion - Do not quote

Cornia Giovanni Andrea (2014), 'Income Inequality Levels, Trends and Determinants in Sub-Saharan Africa: an overview of the main changes' (first draft, 30 November 2014), UNDP's Project on Inequality in SSA

Cornia Giovanni Andrea (2015), '"Income inequality in Latin America: recent decline and prospects for its further reduction" WIDER WP 20/2015, UNU-WIDER

Cornia Giovanni Andrea and Bruno Martorano (2016), « Inequality and growth in an agricultural-led development model: The case of Ethiopia over 1995 -2011 »

Dabalen, A., Mungai, R. and N. Yoshida (2011), "Frequency and Comparability of Poverty Data in SSA", PREM Knowledge and Learning , April 20, 2011. Presentation available at: http://www.google.co.uk/ url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCYQFjAA&url=http%3A%2F%2Fsiteresources.worldbank.org %2FINTPOVERTY%2FResources%2F335642-1304962627819%2F7920845-1305577604391%2F2_AndrewRoseNobuo. pptx&ei=MZyTVbjSLoWesAH8iqqoDQ&usg=AFQjCNEyaKFFZAiXUyZUy9gXD_Qf7iUB-A&bvm=bv.96952980,d.bGg

Davoodi, Hamid, Erwin Tiongson and Sawitree Asawanuchit (2003), ' How Useful Are Incidence Analyses of Public Education and Health Spending?' IMF Working paper n. 03/227, Washington DC.

Deaton, Angus S., and Margaret Grosh (2000), "Consumption," in Margaret Grosh and Paul Glewwe, eds., Designing household survey questionnaires for developing countries: lessons from 15 years of the Living Standards Measurement Study, Oxford University Press for the World Bank, Vol 1., 91—133

Ferreira, Francisco (2014), "Growth, Inequality and Poverty Reduction in Africa", available at: http://www.studio-cx.co.za/ gtac/wp-content/uploads/2014/11/Francisco-Ferreira-Presentation2.pdf

Gibson J, Kim B. (2013), Do the urban poor face higher food prices? Evidence from Vietnam. Food Policy 41:193-203.

Gibson, John, (1999), "How Robust are Poverty Comparisons to Changes in Household Survey Methods? A Test Using Papua New Guinea Data", Department of Economics, University of Waikato.

Gibson, John, Huang, Jikun and Rozelle Scott (2003), "Improving estimates of inequality and poverty from urban China's Household Income and Expenditure Survey", Review of Income and Wealth, 49(1), 53-68.

Grimm, Michael and Isabel Gunther (2005), "Growth and Poverty in Burkina Faso: A Reassessment of the Paradox," Discussion Papers of DIW Berlin 482, DIW Berlin, German Institute for Economic Research.

Guerriero, Marta (2012), 'The Labour Share of Income around the World. Evidence from a Panel Dataset', Development Economics and Public Policy Working Paper Series WP No. 32/2012, University of Manchester.

Jenkins, Stephen (2014) ,'World Income Inequality Databases: an assessment of WIID and SWIID', No. 2014-31, September 2014. Institute dor Economic and Social Research, https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2014-31.pdf.

Chapter V: References

Jolliffe, Dean (2001), "Measuring absolute and relative poverty: the sensitivity of estimated household consumption to survey design", Journal of Economic and Social Measurement, 27, 1-23.

Klasen, Stephan (2014), Measuring Poverty and Inequality in Sub-Saharan Africa: Knowledge Gaps and Ways to Address them. At: http://blogs.worldbank.org/africacan/measuring-poverty-and-inequality-sub-saharan-africa-knowledge-gaps-and-ways-address-them

Lanjouw, Jean Olson and Peter Lanjouw (2001), "How to Compare Apples and Oranges: Poverty Measurement based on Different Definition of Consumption", Review of Income and Wealth, 47 (1), pp. 25-42.

Mendoza RU. 2011. Why do the poor pay more? Exploring the poverty penalty concept. Journal of International Development 23: 1-28.

McCulloch, N., B. Baulch, B. and M. Cherel-Robson (2000), "Growth, Inequality and Poverty in Mauritania, 1987-1996". Poverty Reduction and Social Development Africa Region, TheWorld Bank, mimeo.

Mussa, R. (2014), Food Price Heterogeneity and Income Inequality in Malawi: Is Inequality Underestimated? MPRA Paper No. 56080, posted 19. May 2014 18:02 UTC. Available at: Online at http://mpra.ub.uni-muenchen.de/56080/

Ndikumana, Leonce (2014). "Capital Flight and Tax Havens: Impact on Investment And Growth in Africa", Révue d'Economie du Developpement, 2014/2.

Ostry, D., A. Berg, C. G. Tsangarides, (2014), 'Redistribution, Inequality, and Growth', an IMF Staff Discussion Note 2014/02

Ravallion, Martin (2001), 'Measuring aggregate welfare in developing countries: How well do National Accounts and Surveys agree? Mimeo. The World Bank, August.

Rio Group (2006) Compendium of Best Practices in Poverty Measurement, Expert Group on Poverty Statistics, Economic Commission for Latin America and the Caribbean, Rio de Janeiro.

Sandefur, Justin, and Amanda Glassman (2013), "The Political Economy of Bad Data: Evidence from African Survey & Administrative Statistics." Center for Global Development, paper presented at UNUWIDER Development Conference, "Inclusive Growth in Africa: Measurement, Causes and Consequences," Helsinki, September 20–21

Van der Berg, Servas (2009), 'Fiscal incidence of social spending in South Africa: A report to the National Treasury'; University of Stellenbosch, 28 February 2009

World Bank (2006), World Development Report, Oxford: Oxford University Press for the World Bank.

World Bank (2013), "International Income Distribution Database (I2D2)". Washington, D.C.: World Bank.

Annex 1
Description of the Integrated Inequality Dataset (IID-SSA)

The database compiles in a comparative way Gini coefficients derived from different sources for the years 1991/5-2011. It has been built in three stages: collection of data from existing sources; selection for every country/year of the best data, interpolating the missing years only for countries with at least four well-spaced data between the late 1980s and 2011, a period which allows to depict the medium-term inequality trends over the last 29 years.

(a) Collection of data from existing sources. In the first stage, we have built a preliminary database which includes 1408 (44x32) cells. In particular, it contains yearly information over the period 1980-2011 referred to 44 countries. As explained in the text above, the Gini data are extracted from:

- WIDER's WIIDv3. OB: www. wider. unu. edu/research/WIID3-0B/en_GB/database/)

- The World Bank's Povcal database: http://iresearch. worldbank. org/PovcalNet/index. htm)

- Branko Milanovic's All the Ginis dataset. Available at: (http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/ EXTRESEARCH/0,,contentMDK:22301380~pagePK:64214825~piPK:64214943~theSitePK:469382,00. html)

- The World Bank's International Income Distribution Database (I2D2). A published reference is still missing as the project is still underway, and only a few data were shared with us,,

- National sources that are not yet included in the international dataset mentioned above.

(b) **Selection of the best data to build the IID-SSA.** For every country/year, we selected the best Gini datum from the five data sources indicated above so as to increase data consistency and completeness through a careful comparative analysis of the collected data.

The selection process, has followed the following protocol

- Whenever available, we selected data from WIDER's WIIDv3.0b database that provides detailed information on data sources and their quality. These data are consistent in terms of income definition, coverage and measurement units, and all have a quality rating '1' or '2',

- Whenever WIDER's WIIDv33.0b data were not available, or were of low quality, or were non –comparable with the rest of the time series (e.g. because of the use of a different income concept), we selected Gini's from the World Bank's POVCAL database. For example, we chose POVCAL data in the case of Burkina Faso in the 1990s because WIDER's WIIDv3.0b data were classified as low quality. The same criterion was followed for S.Africa and Zambia in the 2000s so as to maintain consistency in the income definitions,

- Whenever the first two databases did not provide any Gini data for a given year/country, we selected data (when available) from the "All the Ginis Dataset" (2 per cent of all data selected). It is necessary to underscore that – on average - data from this dataset generally overlap to a large extent with WIDER's WIIDv3.0b data,

- As a fourth choice, we used some data from the World Bank's 'International Income Distribution Database" (I2D2) (1 per cent of all data selected),

- Last, for Tanzania, we integrated the previous information with national data (two data points) while for Ethiopia we use data from Cornia and Martorano (2015) that calculated Gini on the basis of national survey data refereed to 1995/6, 1999/2000, 2004/5 and 2010/11.

(c) **Construction of complete time series, 1991/5-2011**. Out the 44 countries included in the first stage, we retained only the 29 with at least four well-spaced and consistent data. To build time series for these countries we interpolated the observed Gini point-to point by means of linear trends as shown in the figures reported in the following pages. The interpolated data are colored in yellow.

As new data come out all the time an update of IID-SSA is planned just before the end of the project.

Angola (consumption data)

year	IID-SSA	Milanovic	POVCAL	nat. sources	I2D2	WIID3
1995	40.2	40.2				
1996	43.9					
1997	47.6					
1998	51.2					
1999	54.9					
2000	58.6	58.1	58.6			58.6
2001	56.9					
2002	55.1					
2003	53.4					
2004	51.7					
2005	50					55
2006	48.2					43
2007	46.5					
2008	44.8				44.8	
2009	43		42.7			43
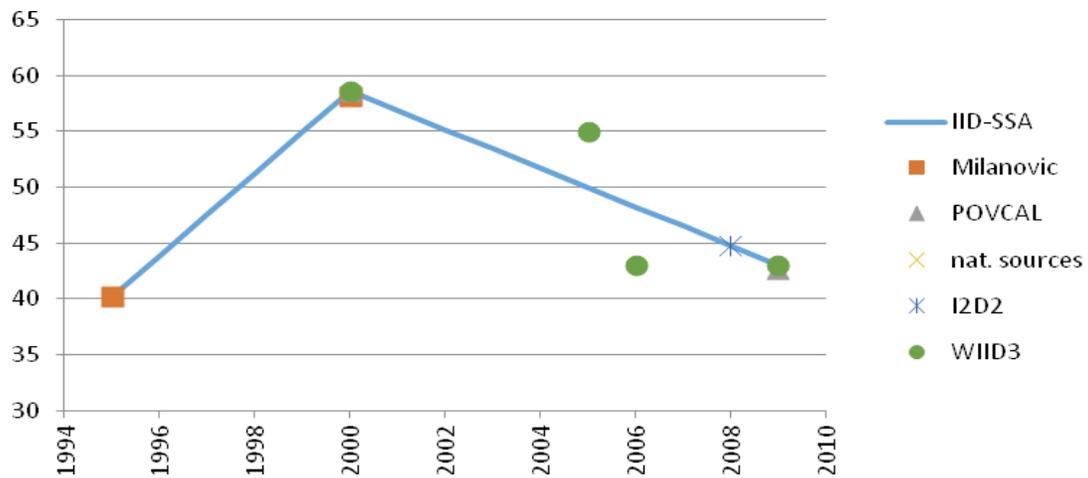
Botswana (income data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1994	53.7	61.0	61.0			**53.7**
1995	54.1					
1996	54.5					
1997	54.9					
1998	55.3					
1999	55.7					
2000	56.1					
2001	56.5					
2002	56.9					
2003	57.3					**57.3**
2004	59.2					
2005	61.1					
2006	62.9					
2007	64.8					
2008	66.7					
2009	68.6				**68.6**	
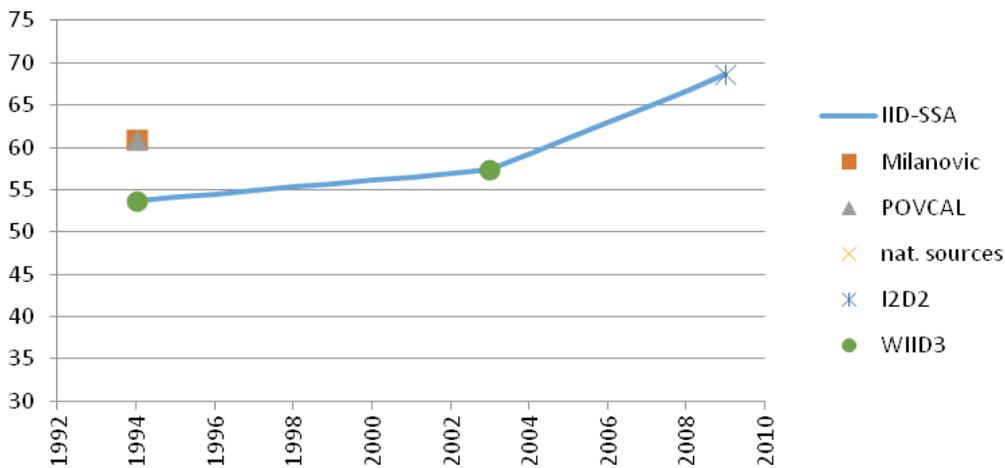
Burkina Faso (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1994	50.7	46.5	**50.7**		54.3	48.2
1995	49.8	39.0				39.0
1996	48.8					
1997	47.9					
1998	46.9	46.9	**46.9**			46.7
1999	52.4	**52.4**				
2000	48.7					
2001	45.0					
2002	41.3					
2003	39.6	37.6	**39.6**			39.5
2004	38.0					
2005	38.3					
2006	38.7					
2007	39.1					
2008	39.4					
2009	39.8	39.8	**39.8**			39.8
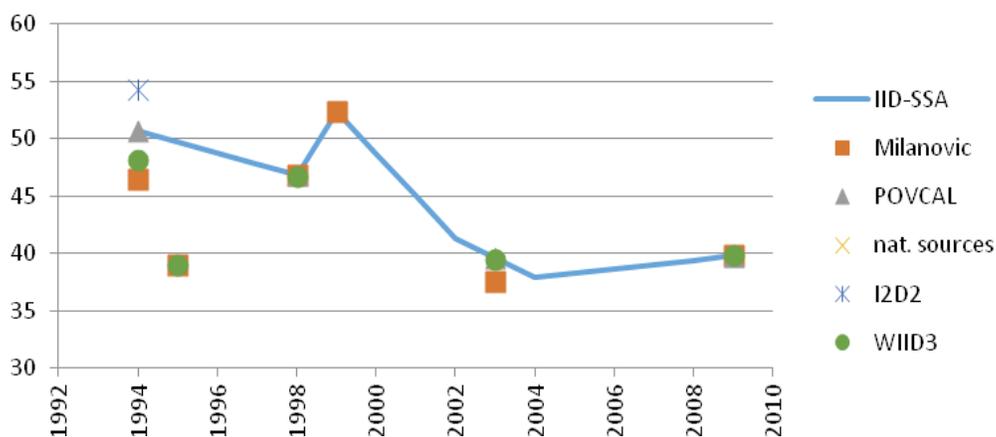
Cameroon (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1996	47.8	43.2	40.7		47.1	**47.8**
1997	47.2					
1998	46.5					
1999	45.9					
2000	45.2					
2001	44.6	40.4	40.4		45.7	**44.6**
2002	43.7	44.1				
2003	42.7					
2004	41.8					
2005	40.9					
2006	39.9					
2007	39.0	39.0	38.9		37.1	**39.0**
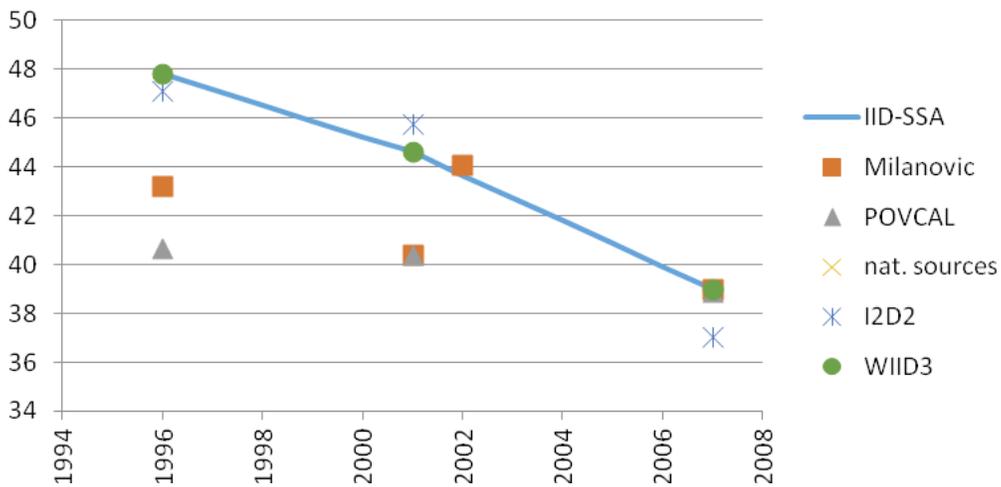
Central African Republic (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1992	61.3	61.3	61.3			**61.3**
1993	59.5	**59.5**				
1994	57.9					
1995	56.3					
1996	54.7					
1997	53.1					
1998	51.5					
1999	50.0					
2000	48.4					
2001	46.8					
2002	45.2					
2003	43.6	43.3	43.6			**43.6**
2004	46.1					
2005	48.7					
2006	51.2					
2007	53.8					
2008	56.3	56.2	56.3			**56.3**

Côte d'Ivoire (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1995	36.7	36.5	36.7			**36.7**
1996	37.7					
1997	38.8					
1998	39.8	45.2	43.8			
1999	40.8					**40.8**
2000	42.6					
2001	44.4					
2002	46.2	44.2	48.4			**46.2**
2003	46.0					
2004	45.7					
2005	45.5					
2006	45.2					
2007	45.0					
2008	44.7	41.5	41.5		50.4	**44.7**
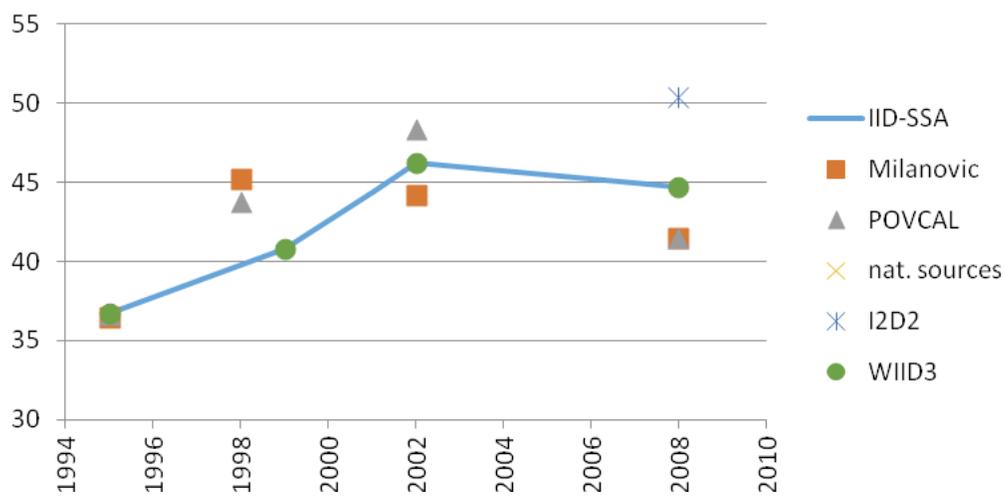
Ethiopia (consumption data)

year	IID-SSA	Milanovic	POVCAL	nat. sources	I2D2	WIID3
1995	30	40	40	**30**	32.4	40
1996	29.8	38.5				
1997	29.6					44.1
1998	29.4	50.1				
1999	29.2					
2000	29	27.9	30	**29**	38.5	30
2001	29.4					
2002	29.8					
2003	30.2					
2004	30.6	28.9				
2005	31	29.8	29.8	**31**		29.8
2006	30.8					
2007	30.7					
2008	30.5					
2009	30.3					
2010	30.2					29.8
2011	30		33.6	**30**		
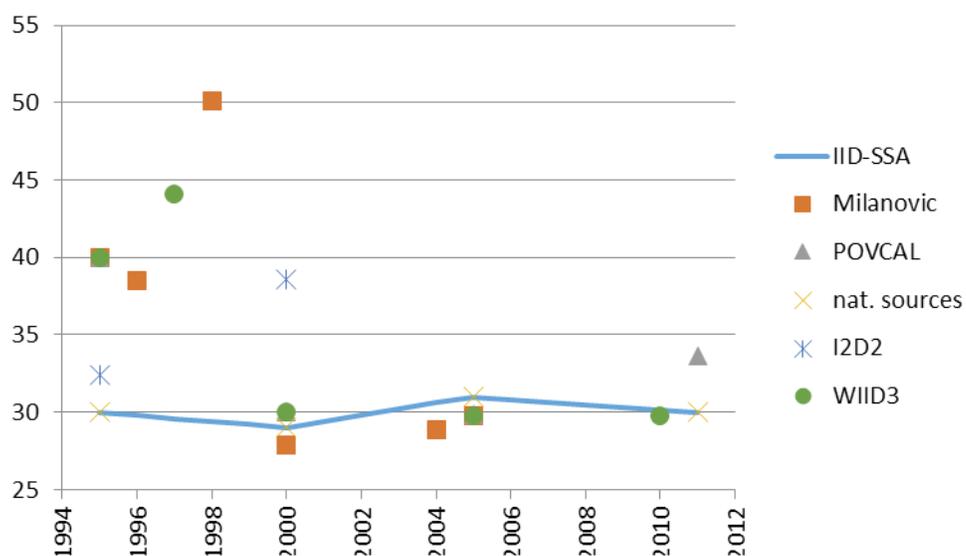
The Gambia (consumption data)

year	IID-SSA	Milanovic	POVCAL	nat. sources	I2D2	WIID3
1993	60.9					**60.9**
1994	59.4					**59.4**
1995	57.1					
1996	54.8					
1997	52.5					
1998	50.2	49.9	50.2		53.7	**50.2**
1999	49.6					
2000	49					
2001	48.5					
2002	47.9					
2003	47.3	46.9	47.3			**47.3**
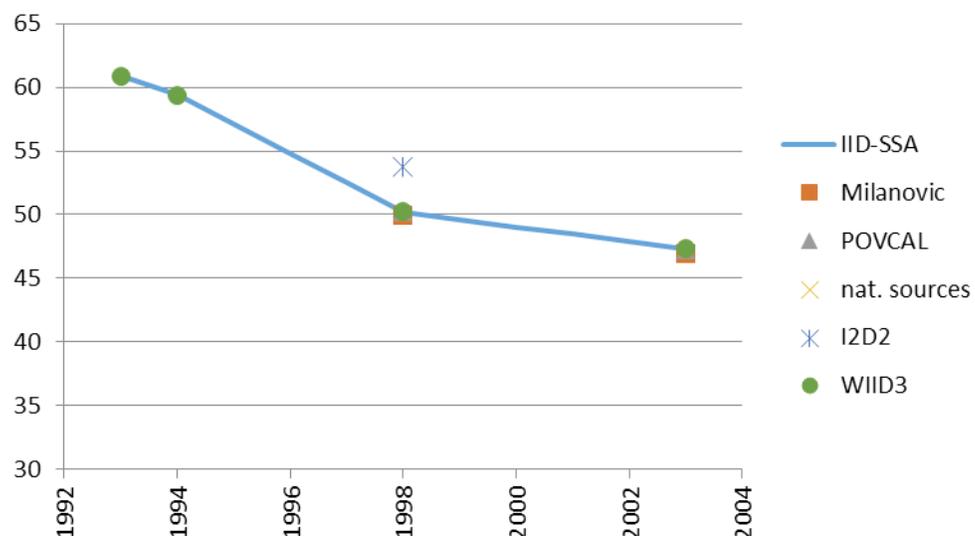
Ghana (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1993	33.8	33.0				**33.8**
1994	33.5					
1995	33.3					
1996	33.0					
1997	32.7	32.7				**32.7**
1998	43.4	40.5	40.8		43.2	**43.4**
1999	40.8					**40.8**
2000	41.1					
2001	41.3					
2002	41.6					
2003	41.9					
2004	42.1					
2005	42.4	**42.4**			44.2	
2006	42.8	42.8	**42.8**			
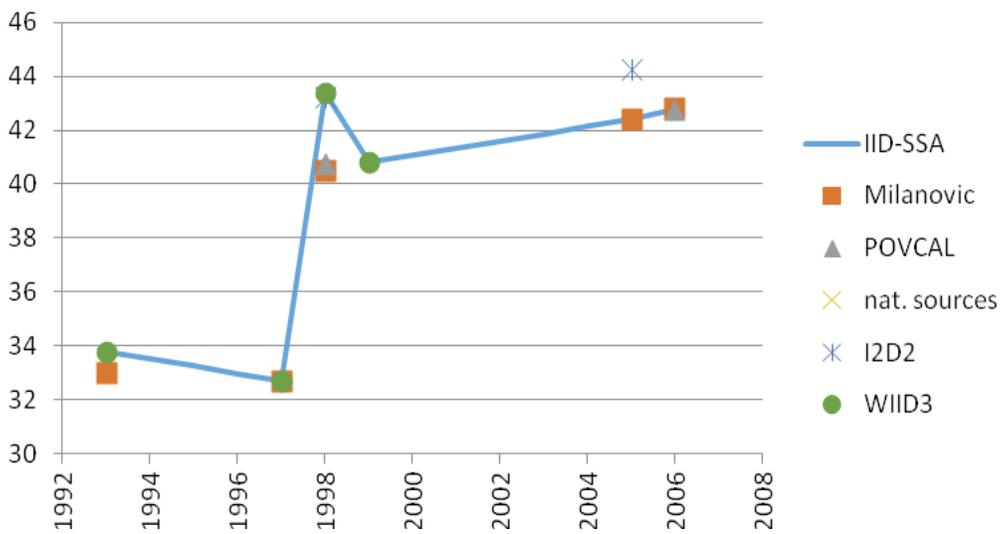
Guinea (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1994	40.4	39.5	44.9		49.5	**40.4**
1995	40.2	40.4				
1996	40.0					
1997	39.8					
1998	39.6					
1999	39.4					
2000	39.2					
2001	39.0					
2002	38.8					
2003	38.6	42.9	40.3			**38.6**
2004	38.8					
2005	39.0					
2006	39.2					
2007	39.4	39.6	39.4			**39.4**

Guinea-Bissau (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1993	47.8	47.8	47.8		48.7	**47.8**
1994	40.0					**40.0**
1995	39.4					
1996	38.9					
1997	38.3					
1998	37.8					
1999	37.2					
2000	36.6					
2001	36.1					
2002	35.5	35.3	35.5			**35.5**
2003	36.4					
2004	37.4					
2005	38.3	**38.3**				
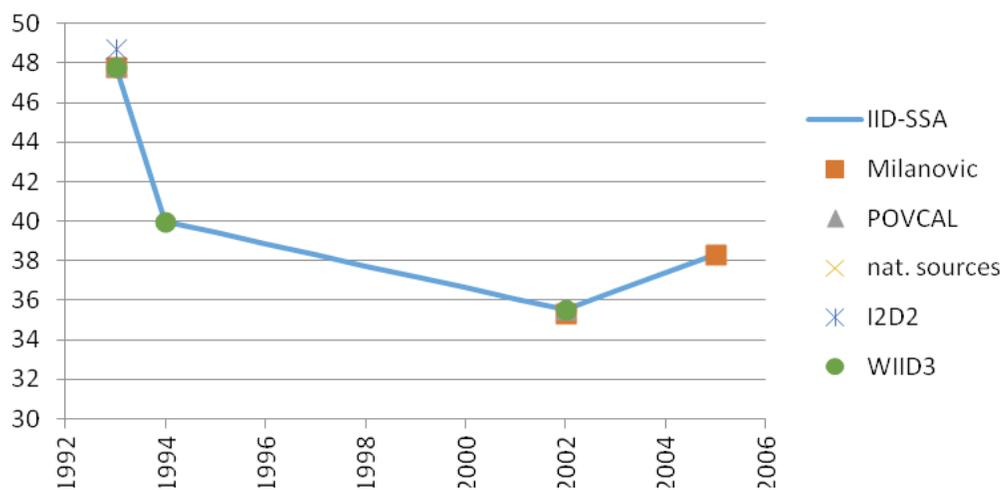
Kenya (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1994	42.1	42.1	**42.1**			44.5
1995	42.2					
1996	42.4					
1997	42.5	44.4	42.5			**42.5**
1998	43.2					
1999	43.8					
2000	44.5					
2001	45.1					
2002	45.8					
2003	46.4					
2004	47.1					
2005	47.7	47.1	47.7		46.3	**47.7**
2006	45.9					**45.9**
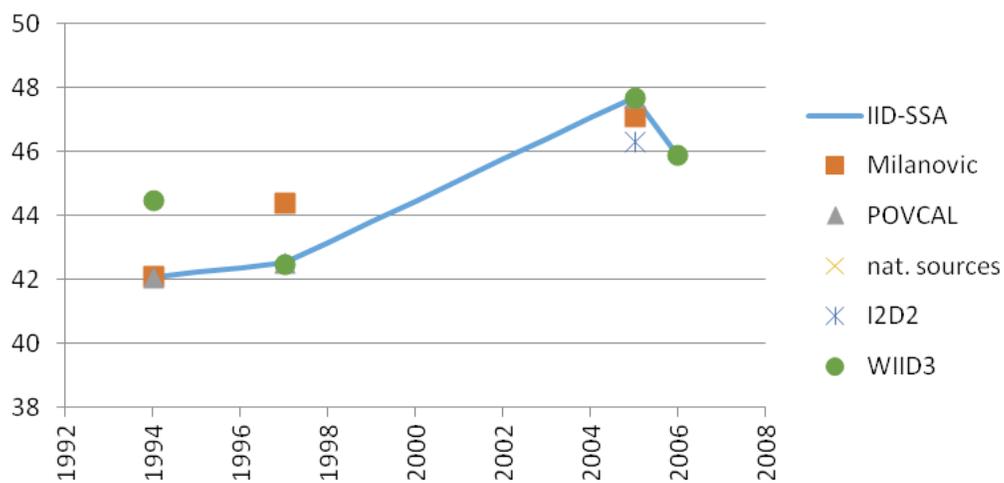
Lesotho (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1993	57.9	56.5	57.9			**57.9**
1994	63.2	63.2	63.2			**63.2**
1995	67.0	66.3				**67.0**
1996	65.3					
1997	63.5					
1998	61.8					
1999	60.0					**60.0**
2000	57.3					
2001	54.7					
2002	52.0	**52.0**				
2003	52.5	52.5	52.5			**52.5**
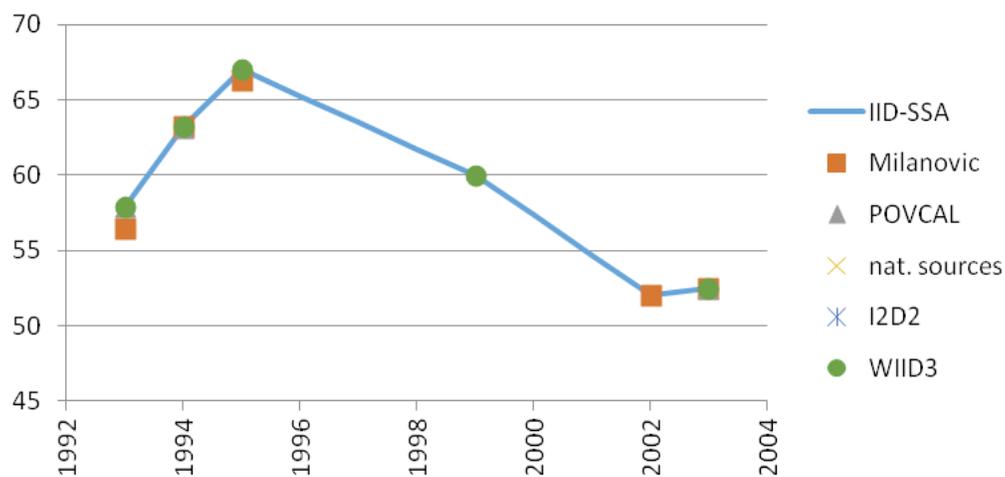
Madagascar (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1993	46.4	44.5	46.1		59.5	**46.4**
1994	44.6					
1995	42.8					
1996	41.0					
1997	39.2	39.2	**39.2**			40.6
1998	37.9	**37.9**				
1999	41.8	41.8	**41.8**			40.2
2000	43.5					
2001	45.3	45.9	47.5		47.4	**45.3**
2002	44.2					
2003	43.2					
2004	42.1				43.0	
2005	41.0	69.7	47.2			**41.0**
2006	40.7					
2007	40.3					
2008	40.0					
2009	39.6					
2010	39.3	44.1	44.1		44.4	**39.3**

42
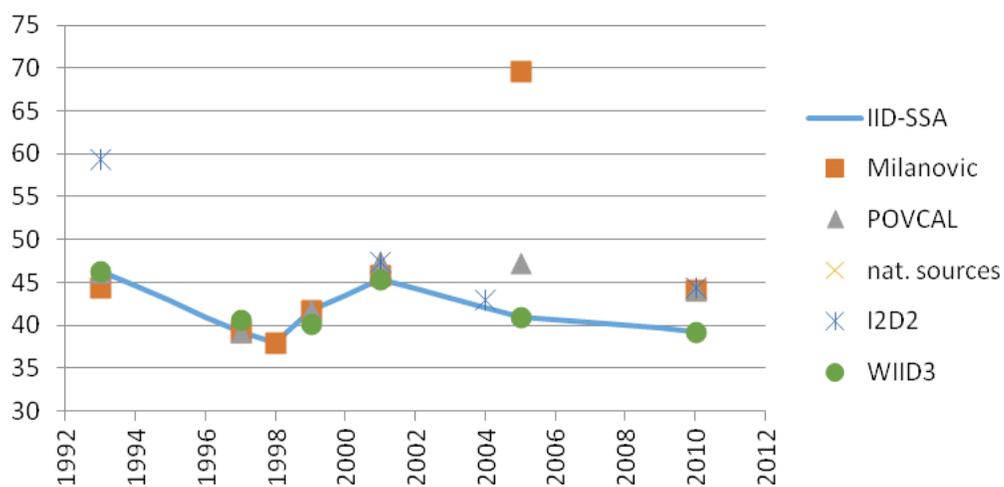
Malawi (consumption data)

Year	IID-SSA	Mila	Povcal	N a t .	I2D2	WIID3
1993	62.0	62.0				**62.0**
1994	59.1					
1995	56.2					
1996	53.2					
1997	50.3	49.5			72.1	**50.3**
1998	50.3	50.3	50.3			**50.3**
1999	48.4					
2000	46.5					
2001	44.7					
2002	42.8					
2003	40.9					
2004	39.0	39.0	39.0		44.1	**39.0**
2005	38.6	**38.6**			41.0	
2006	39.0					**39.0**
2007	40.2					
2008	41.5					
09	42.7					
2010	43.9	45.2	**43.9**		45.5	
2011	45.2					**45.2**
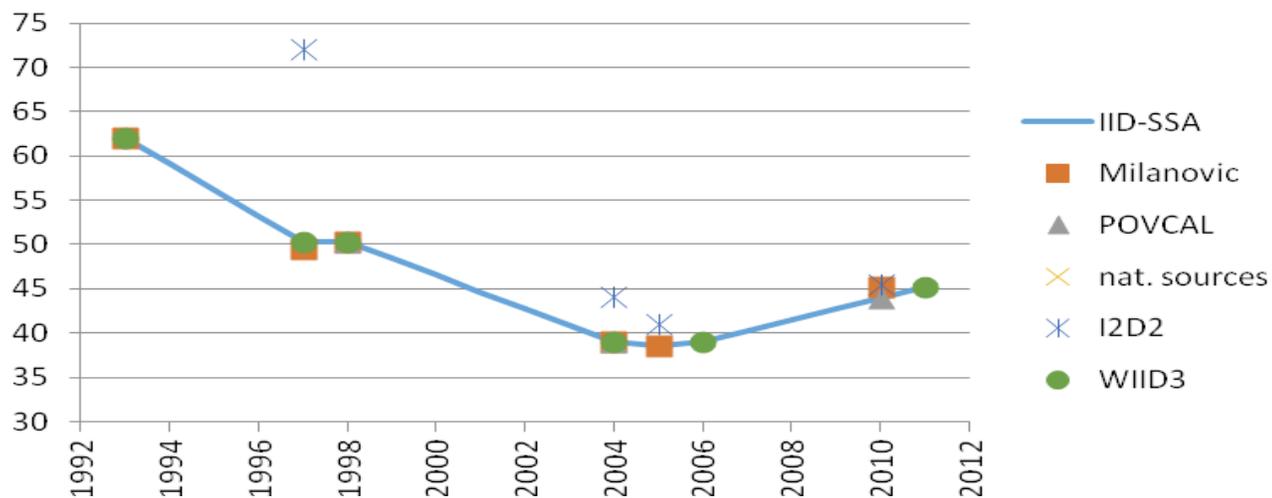
Mali (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1994	50.5	48.9	50.6		48.5	**50.5**
1995	49.0					
1996	47.5					
1997	46.0					
1998	44.6					
1999	43.1					
2000	41.6					
2001	40.1	38.9	40.0			**40.1**
2002	39.9					
2003	39.7					
2004	39.4					
2005	39.2					
2006	39.0	38.6	39.0			**39.0**
2007	37.5					
2008	36.0					
2009	34.5					
2010	33.0	33.0	33.0			**33.0**

44
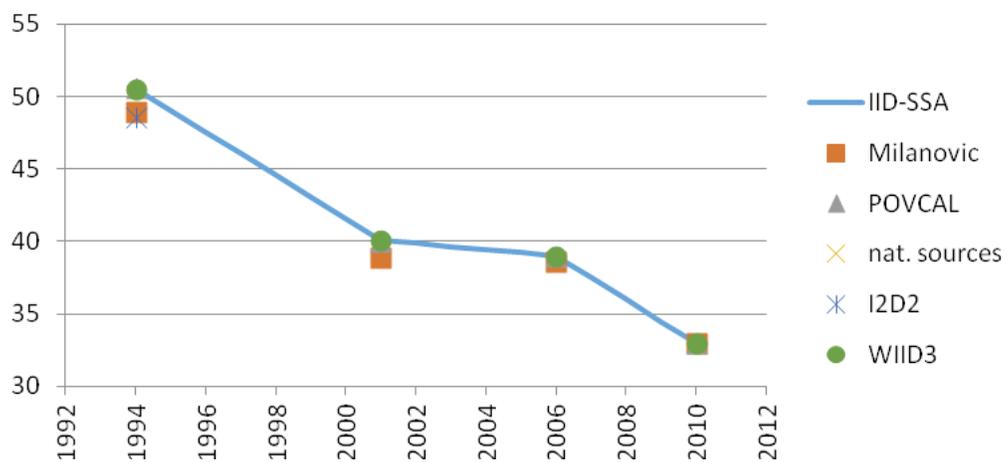
Mauritania (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1995	37.8	38.0				**37.8**
1996	37.3	37.3	37.3			**37.3**
1997	37.7					
1998	38.2					
1999	38.6					
2000	39.0	38.9	39.0		41.2	**39.0**
2001	39.6					
2002	40.2					
2003	40.7					
2004	41.3	41.3	41.3			**41.3**
2005	40.5					
2006	39.7					
2007	38.8					
2008	38.0	40.5	40.5			**38.0**
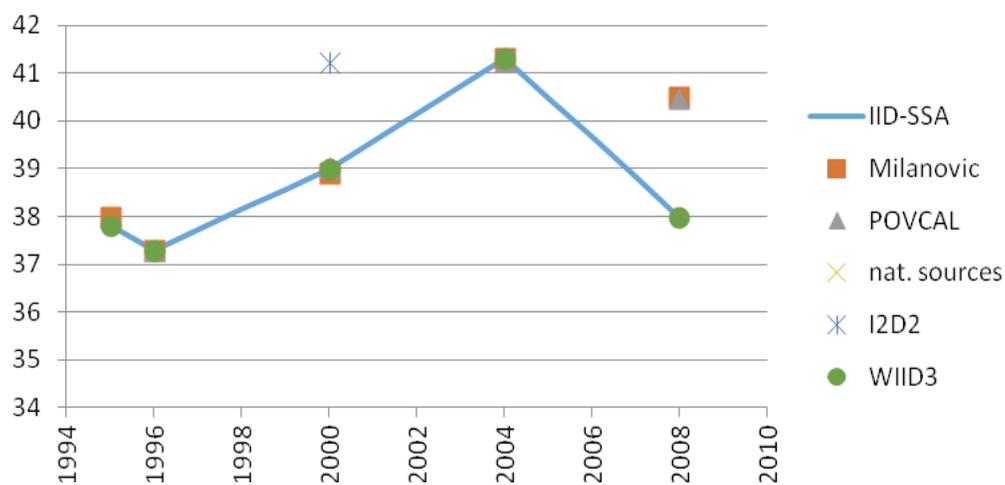
Mauritius (income data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1991	37.0	36.7				**37.0**
1992	37.3					
1993	37.7					
1994	38.0					
1995	38.4					
1996	38.7					**38.7**
1997	38.4					
1998	38.1					
1999	37.7				30.4	
2000	37.4					
2001	37.1					**37.1**
2002	37.4					
2003	35.7				**35.7**	
2004	36.1				**36.1**	
2005	36.5				**36.5**	
2006	36.4					
2007	36.3				**36.3**	38.8
2008	36.4				**36.4**	
2009	38.8				**38.8**	
2010	39.1					
2011	39.5				**39.5**	

46
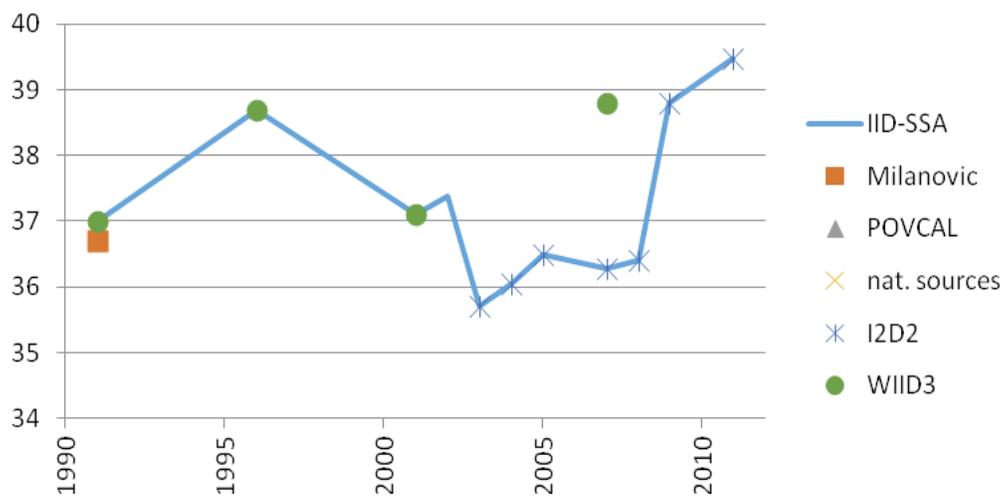
Mozambique (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1996	44.5	44.5	44.5			**44.5**
1997	44.8					
1998	45.1					
1999	45.5					
2000	45.8					
2001	46.1					
2002	46.4	46.4				**46.4**
2003	47.1	47.1	47.1			**47.1**
2004	44.2					
2005	41.3	41.3				**41.3**
2006	41.3					
2007	41.4					
2008	41.4	45.6	45.7		58.9	**41.4**
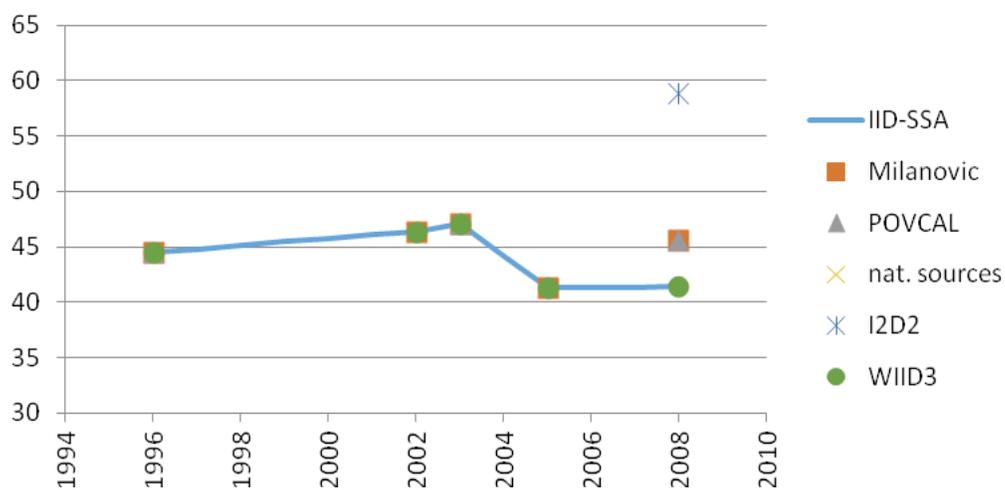
Niger (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1994	49.8	41.5	41.5			**49.8**
1995	50.6	50.2				**50.6**
1996	49.9					
1997	49.3					
1998	48.6					
1999	47.9					
2000	47.2					
2001	46.6					
2002	45.9					
2003	45.2					
2004	44.6					
2005	43.9	43.4	43.9		46.0	**43.9**
2006	40.6					
2007	37.3	**37.3**				
2008	34.5	34.6	34.6			**34.5**
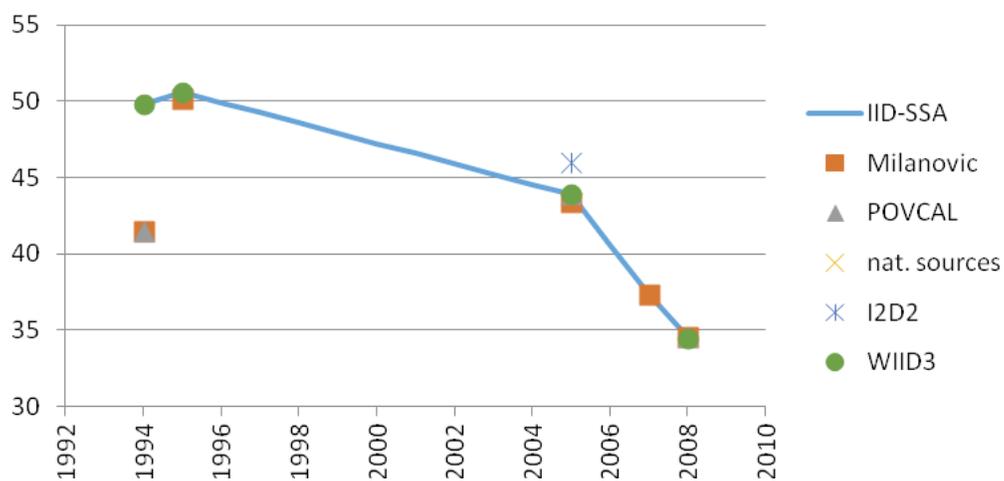
Nigeria (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1993	45.0	**45.0**				
1994	45.5					
1995	46.0					
1996	46.5	51.1	**46.5**			52.9
1997	46.1					
1998	45.7					
1999	45.3					
2000	44.9					
2001	44.5					
2002	44.1					
2003	43.7	41.8			45.2	**43.7**
2004	42.9	40.9	42.9			**42.9**
2005	43.2					
2006	43.5					
2007	43.8					
2008	44.1					
2009	44.4					
2010	44.7	46.8	48.8			**44.7**
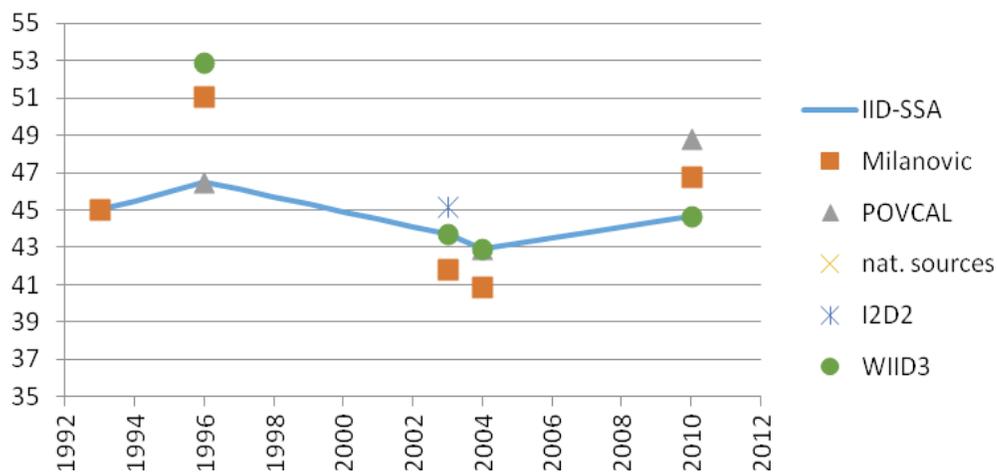
Rwanda (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1995	44.0					
1996	45.5					
1997	47.0				50.7	
1998	48.5					
1999	50.0					
2000	51.5	51.5	**51.5**			45.1
2001	51.8					
2002	52.0					
2003	52.3					
2004	52.6					
2005	52.8				58.7	
2006	53.1	53.1	53.1			**53.1**
2007	52.3					
2008	51.5					
2009	50.6					
2010	49.8				57.5	
2011	49.0	50.8	50.8			**49.0**
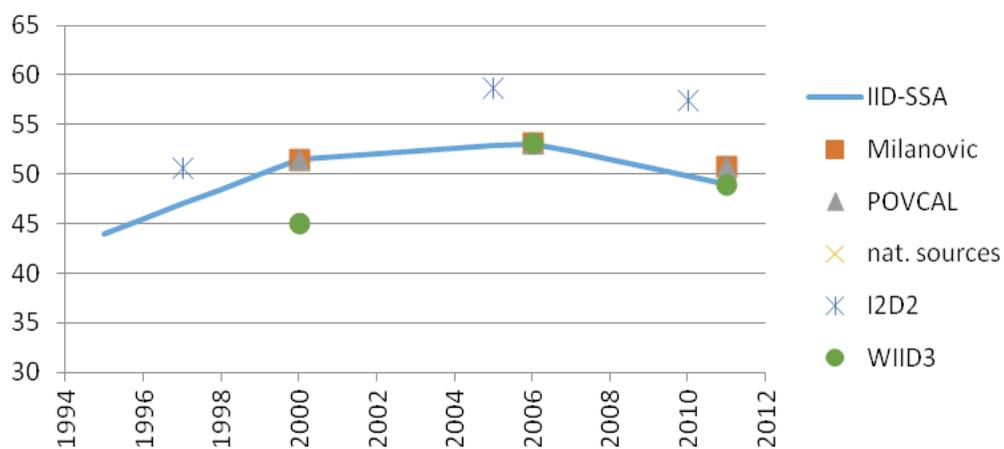
Senegal (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1994	41.3	41.4	41.4			**41.3**
1995	41.3	51.9			54.4	
1996	41.3					
1997	41.3					
1998	41.3					
1999	41.3					
2000	41.3					
2001	41.3	46.3	41.3		46.8	**41.3**
2002	40.8					
2003	40.3					
2004	39.7					
2005	39.2	38.9	39.2		43.9	**39.2**
2006	39.4					
2007	39.6					
2008	39.8					
2009	39.9					
2010	40.1					
2011	40.3		**40.3**			
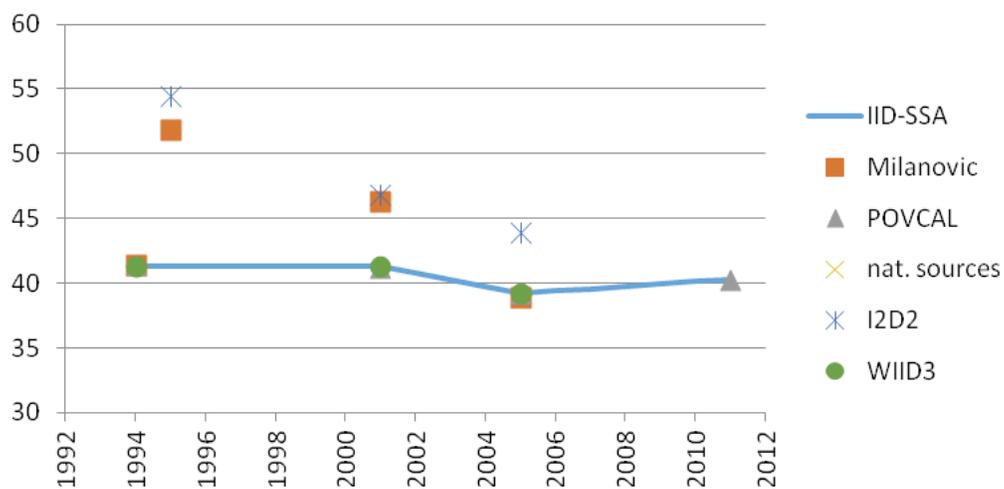
Sierra Leone (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1995	54.2					
1996	52.7					
1997	51.3					
1998	49.8					
1999	48.3					
2000	46.9					
2001	45.4					
2002	44.0					
2003	42.5	38.1	**42.5**		43.1	39.0
2004	42.4					
2005	42.2	**42.2**				
2006	41.1					
2007	39.9					
2008	38.8					
2009	37.6					
2010	36.5					
2011	35.3		**35.4**			
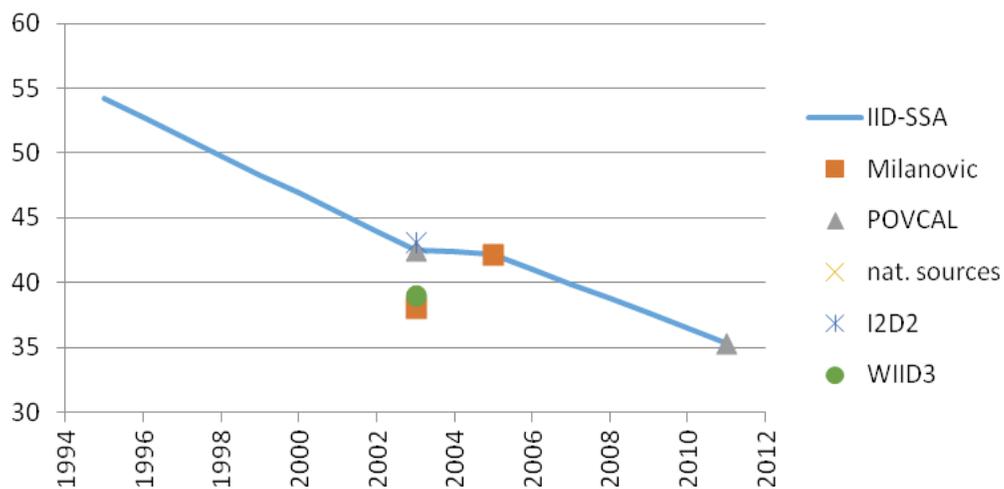
South Africa (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1993	59.3	57.7	**59.3**			
1994	59.3	**59.3**				
1995	56.6	62.0	**56.6**			59.0
1996	56.8					
1997	57.1					54.5
1998	57.3					
1999	57.5					
2000	57.8	57.3	**57.8**			
2001	59.4					
2002	61.0					
2003	62.6					
2004	64.2					
2005	65.8					
2006	67.4	67.4	**67.4**			
2007	66.0					
2008	64.6	69.8				59.4
2009	63.1	63.1	**63.1**			
2010	64.1	66.5				59.4
2011	65.0		**65.0**			
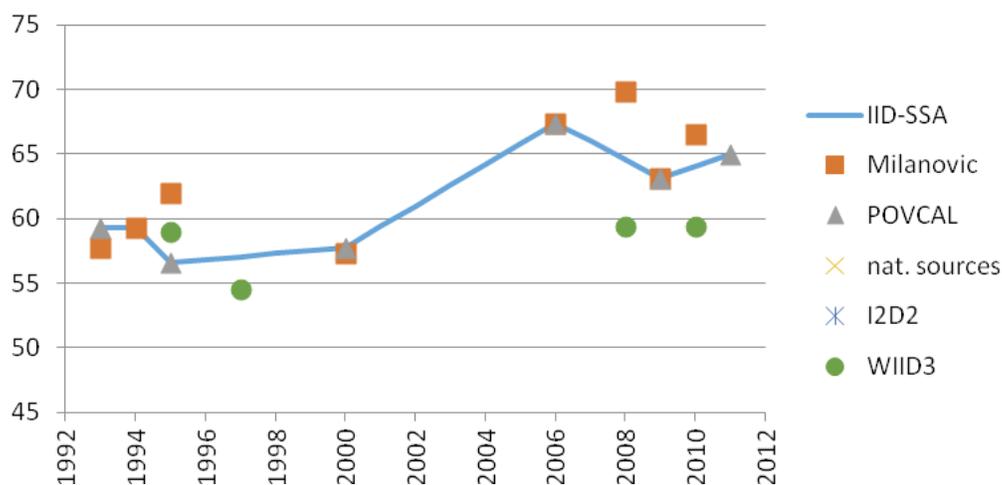
Swaziland (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1995	60.7	60.7	60.7		60.8	**60.7**
1996	59.0					
1997	57.3					
1998	55.5					
1999	53.8					
2000	52.1				54.6	
2001	50.4	50.7	50.7			**50.4**
2002	49.9					
2003	49.4					
2004	48.9					
2005	48.3					
2006	47.8					
2007	47.3					
2008	46.8					
2009	46.3	**46.3**				
2010	51.5	51.5	51.5			**51.5**
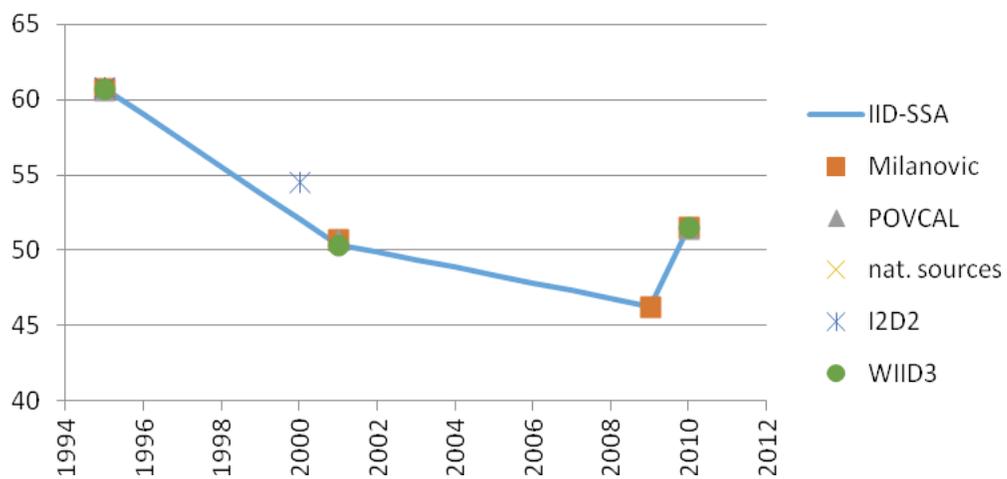
Tanzania (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1993	39.5	36.3				**39.5**
1994	38.8					
1995	38.1					
1996	37.4					
1997	36.7					
1998	36.0					
1999	35.3					
2000	34.6	34.6	34.6			**34.6**
2001	35.0	34.4				**35.0**
2002	35.0					
2003	35.0					
2004	35.0					
2005	35.0					
2006	35.0					
2007	35.0	37.6	37.6			**35.0**
2008	36.0			**36.0**		
2009	36.5					
2010	37.0			**37.0**		
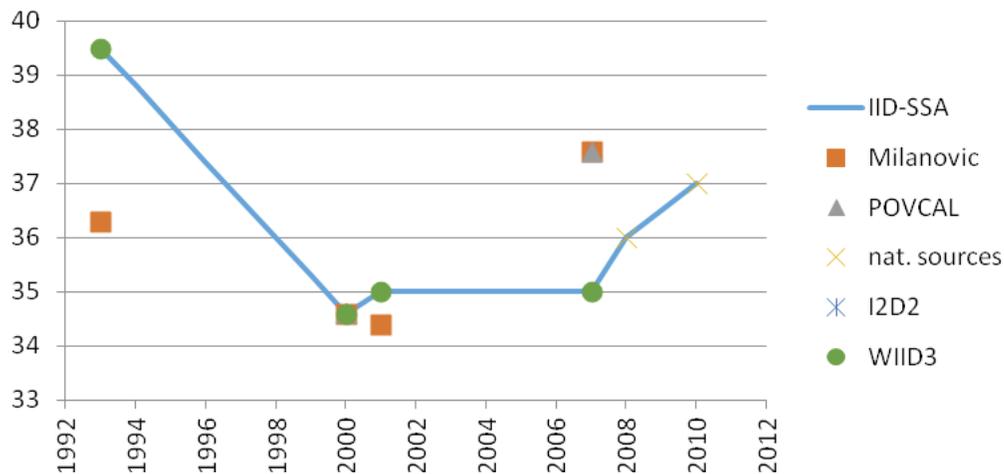
Uganda (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1992	40.9	38.0	42.6			**40.9**
1993	40.0	39.2				
1994	39.0					
1995	38.0					
1996	37.1	37.1	37.1			**37.1**
1997	39.1					
1998	41.1					
1999	43.1	43.1	43.1			**43.1**
2000	44.9					**44.9**
2001	43.8					
2002	42.8	43.6	45.8			**42.8**
2003	42.1					
2004	41.5					
2005	40.8	42.2				**40.8**
2006	41.3	42.6	42.6			
2007	41.7					
2008	42.1					
2009	42.6	42.6	44.3			**42.6**
2010	42.3	42.9				**42.3**

Zambia (consumption data)

Year	IID-SSA	Milanovic	Povcal	Nat. sources	I2D2	WIID3
1993	52.6	51.3	**52.6**			51.2
1994	51.7					
1995	50.7					
1996	49.8	49.8	**49.8**			51.9
1997	51.6					
1998	53.4	53.1	**53.4**			54.6
1999	50.5					
2000	47.5					
2001	44.6					
2002	41.6	**41.6**				
2003	42.1	42.1	**42.1**			42.1
2004	50.7	50.3	**50.7**			50.8
2005	52.7					
2006	54.6	54.6	**54.6**			
2007	55.3					
2008	56.1					
2009	56.8					
2010	57.5	57.4	**57.5**			