PART IV

Measurement and Econometric Investigation of Determinants of Inequality in sub-Saharan Africa

Seven measurement errors affecting the assessment of income inequality levels and trends

1. Differences in **statistical assumptions** and **data harmonization** **across countries**
2. Differences over **time** in **survey design** for the **same country**
3. **HBS** increasingly **underestimates capital incomes** and the **total net value added**
4. **Top income earners are undersampled**
5. **Incomes earned** on **assets held abroad** by **SSA nationals are ignored**
6. The distributive impact of **differences in price dynamics** between **food prices** and **overall Consumer Price Index** is ignored
7. The distributive impact of **the provision** of **social benefits across countries** is ignored

15 Building an Integrated Inequality Dataset and the 'Seven Sins' of Inequality Measurement in sub-Saharan Africa[1]

GIOVANNI ANDREA CORNIA AND BRUNO MARTORANO

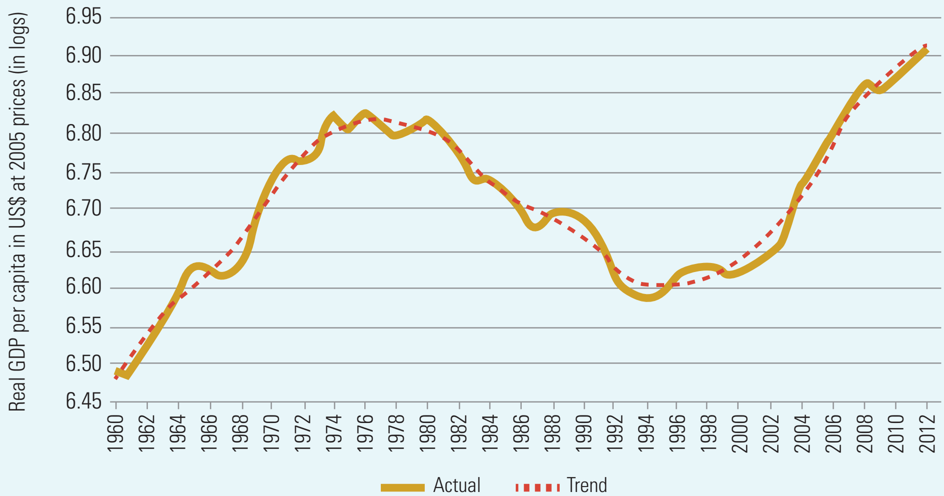
15.1 Introduction

The favourable growth performance of SSA over the last 20 years (figure 15.1) has been accompanied by a modest decline in poverty, from 59 to 48 per cent over 1993-2010, i.e., much less than that recorded in South Asia (Ferreira, 2014). This aggregate trend, however, conceals substantial cross-country variations. The key question, then, is: How can such differences in poverty reduction rates be explained? The standard approach (Bourguignon, 2003) shows that a percentage change in poverty incidence can be decomposed into the percentage changes in the growth rate of GDP per capita and in the Gini coefficient, plus a small residual. 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 per cent in non-resource-rich countries and 2.6 per cent in resource-rich ones. The reason that poverty declined at different rates is, therefore, to be found in the divergence of inequality trends experienced by the SSA countries. Indeed, this chapter and Chapter 2 argue that over 1991-2011, income inequality rose in several countries but declined in a similar number.

Proper documentation of inequality trends in the region is therefore essential to explain cross-countries differences in poverty reduction. However, this task is hindered by the paucity of inequality data, as well as by the lack of a comprehensive database of consistent Gini statistics. This situation becomes more perplexing when considering that over the last two decades, policy formulation has become increasingly 'evidence-based', i.e., based not only on ethical and economic priors, but also on the empirical results generated by a growing number of Household Budget Surveys (HBSs), demographic and health surveys, wealth surveys, multiple indicator cluster surveys, Living Standard Measurement Studies and other surveys. The fields of studies that have benefitted the most from such increase in surveys are those concerning poverty alleviation and the study of income inequality. In most developed and developing regions, academic and policy institutions have built databases tracing the evolution of the Gini coefficient over at least the last 20 years, as in the case of the Luxembourg Income Study (LIS) for the Organisation for Co-operation and Economic Development (OECD) countries, the Socio-Economic Database for Latin America and the Caribbean (SEDLAC) and CEPALSTAT for Latin America, TransMonEE for the European economies in transition, and so on. Finally, during the same period, a number of global inequality databases were created, including the

[1] The authors wish to thank Michael Grimm and an anonymous referee for comments on a previous version of this chapter.

FIGURE 15.1 Trend in the log of aggregate real GDP/capita in SSA, 1960-2012



Source: Ferreira (2014).

World Income Inequality Database (WIID), the Standardized World Income Inequality Database (SWIID), All Gini 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 chapter has two aims. First, Section 15.2 describes the Integrated Inequality Dataset (IID-SSA), which is obtained by comparing the Gini coefficients included in all existing databases or originating from national studies and, on the basis of a standard protocol, the IID-SSA selects the least biased ones. The aim is to collect and assess in a comparative way data on income inequality from all sources, thus making it possible to systematically analyse the region-wide inequality changes recorded during the last two decades. The country trends emerging from the IID-SSA dataset are illustrated in Annex 1 of the UNDP-RBA Working Paper no. 2.[2] The Annex plots the time trend for each of the 29 countries with at least four good-quality and well-spaced Gini points. It also provides information on Gini availability for countries with only one to three Gini data or with zero data. The inequality time series for these countries can be used for a variety of analytical and policy purposes, including the calculation of changes over time in poverty rates or panel regressions of inequality trends. Yet, given the estimation biases discussed in Section 15.3, this information must be used cautiously, i.e., by checking the results generated by the trends 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.

Section 15.3 discusses the biases of the data included in IID-SSA and tries, when possible, to measure their impact on the Gini coefficient in order to alert the researchers of African inequality of the 'seven sins of inequality measurement'[3] in the region. Section 15.3 also presents the methodologies currently

[2] See www.africa.undp.org/content/rba/en/home/library/working-papers/building-the-integrated-inequality-databaseand-the-seven-sins-of.html.

[3] The reader may think that the choice of the term 'seven sins' was inspired by the 'seven cardinal sins' (lust, greed, gluttony, sloth, wrath, envy and pride), which are part of Christian theology, or by T.E. Lawrence's The Seven Pillars of Wisdom, but any reference in this regard is purely coincidental.

adopted to remedy these problems when possible. In addition, Section 15.3 presents a checklist of possible measurement biases that researchers, statisticians and policymakers aiming at assessing the 'real Gini coefficient' of a country should take into account. Indeed, the way the inequality data are usually computed may entail an oversimplification of reality and lead to an underestimation of inequality and policy inaction. The corrections suggested in this chapter require the availability of survey microdata. By making these corrections, it is possible to obtain a better understanding of the real distributive situation of a country. Researchers, policymakers and staff of international institutions are well advised to consider introducing such corrections when working on poverty and inequality at the country level.

15.2 Building a dataset of synthetic inequality statistics

15.2.1 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 standardised database of inequality indexes, such as that produced by SEDLAC or LIS. At present, researchers of SSA inequality rely on inequality statistics originating from one of the following:

a) **WIDER's WIIDv3.0b database**,[4] released in September 2014, which includes fully documented Gini coefficients and decile and quintile distributions for 44 SSA countries, often for long periods of time. For every data point, it includes standard information and documentation on the income concepts used (gross, net, monetary income, earnings and consumption expenditure in cash and kind), basic unit of observation and population coverage (household, family and individual), equivalence scales, sample size and so on. There is also often information about the survey's questionnaire, coverage (national, urban, rural and so on) and availability of reports. Finally, WIIDv3.0b ranks the quality of each Gini or decile distribution with 'scores' 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' and '2' can be used safely in trend and regression analyses. Data points rated '3' or '4' should be only used for ad hoc purposes (e.g. assessing the inequality level of a country). WIIDv3.0b data are computed on HBS produced by National Statistical Offices (NSO), LSMS surveys, POVCAL, and independent field studies.

b) **The World Bank's POVCAL database,**[5] which calculates Gini coefficients on decile distributions derived from survey microdata. POVCAL does not harmonise the microdata according to standardised criteria before calculating the Gini coefficients. Its data partially overlap with WIDER's WIIDv3 data but have a more limited coverage.

c) **The World Bank's International Income Distribution Database (I2D2),**[6] which is a worldwide database drawn from nationally representative household income and consumption surveys, labour force surveys and LSMS surveys comprising a standard set of demographic, education, labour market, household features and income/consumption variables. The I2D2 contains about

[4] See www.wider.unu.edu/research/WIID3-0B/en_GB/database. For more information, see www.wider.unu.edu/research/WIID3-0B/en_GB/WIID-documentation.

[5] For more information, see http://iresearch.worldbank.org/PovcalNet/index.htm

[6] I2D2 was started in 2005 in the context of the World Development Report on Equity.

50 harmonised variables and covers over 900 surveys from over 160 countries dating back to 1960, although most of the data cover the last two decades. Due to this time-consuming harmonisation process, the I2D2 data facilitate cross-country comparisons in several fields. However, at this time (April 2016), it was possible to access only a few harmonised I2D2 Gini data points. An initial look at these data does not suggest major changes in the trends identified in Chapter 2.

To improve comparability across countries and over time, the data from all surveys are processed according to standard statistical conventions regarding: the definition of household income/consumption expenditure per capita; the definition of the household; the corrections for differences in recall periods; the valuation of the income stream from owner-occupied dwellings; adjustments for non-responses; imputation of missing or unreliable data; the treatment of zero incomes; and the upward adjustments of rural incomes needed to offset differences in rural-urban prices. Thus, by definition, the Gini computed on I2D2 do not coincide with those generated by POVCAL and the NSO, since they rely on different statistical conventions. Furthermore, I2D2 Gini values are computed directly on microdata and should, therefore, be slightly higher than those calculated on decile distributions.

d) **The harmonisation of microdata** to be included in I2D2, and which is ongoing at present. The World Bank has collected around 140 surveys for SSA, although only about 20-30 of them were processed by 2015. The assignment of countries to rising, falling, U-shaped and inverted U-shaped inequality categories presented in Chapter 2 may thus change to some extent if new harmonised data become available.

e) **Milanovic's All the Ginis Dataset**, which compiles data from all sources and adds a few observations drawn from data produced by central statistical offices (CSOs) or surveys launched as part of specific research projects. No adjustments are carried out on the data.

f) **The Luxembourg Income Study (LIS)**, which provides LIS-standardised data for South Africa.

g) **Szolt's SWIID dataset**, which includes Ginis from all sources and years but does not rate the quality or consistency of the data. The majority of that data is obtained through multiple imputation techniques that are not always made explicit. While SWIID offers more complete coverage for countries and years, its content is unclear and depends on opaque and arbitrary assumptions. After a detailed comparison of WIIDv3.0b and SWIID, Jenkins (2014) suggests relying on WIIDv3.0b on 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, it was decided not to use SWIID, even if this entails forgoing a number of (imputed) data that are missing in other databases. Jenkins (2014) notes that in SSA, there is a high prevalence of missing data. Hence a greater proportion of SWIID data for SSA depends on the validity of its imputation model, which, given the high measurement error in basic data, varies considerably.

15.2.2 An Integrated Inequality Dataset for SSA (IID-SSA)

Differences in research results about inequality dynamics thus depend not only on differences in countries/years coverage, but also on the dataset chosen. To overcome this problem and limit the use of low-quality/undocumented data, an Integrated Inequality Dataset for SSA (IID-SSA) was compiled. It selects, for every country/year, the best datum from the five datasets described above or 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 similar (as in WIIDv3.0b and POVCAL), while in others they differ slightly or substantially. As shown in table 15.1, most of the data selected for IID-SSA are from WIIDv3.0b. Fourteen of the 44 countries are from East Africa, nine from Central Africa, five from Southern Africa and 16 from West Africa. There is not a single datum for Equatorial Guinea, Eritrea, Sao Tome and Principe, Somalia and South Sudan. Thus, they are excluded from the dataset.

Some of the data included in IID-SSA may suffer from measurement errors for the reasons discussed in Section 15.3. Yet, a careful selection of data from all available sources reduces some of these errors and increases data consistency and completeness so as to provide the least biased dataset in this field. In this regard, it must be mentioned that most SSA household surveys focus on consumption expenditure per capita. Thus, with the exception of Botswana and Mauritius, which use disposable income per capita, the well-being concept adopted in SSA surveys is 'household consumption expenditure per capita'. This is a concept that reduces the measurement bias but does not make it possible to decompose the changes in total inequality by income sources. For some countries and years, there are surveys providing data on both income and consumption per capita.[7] The distance between the Gini coefficients of the distributions of income and consumption can thus be measured.

To analyse the income dynamics in the region, the authors selected a time series adopting the same income concept and population coverage for each country for the 1993-2011 period,[8] although it is impossible to ensure that the same statistical conventions were adopted in processing the raw data of all surveys. Differences likely remain in statistical conventions across countries and over time that increase the 'noise' in regression analysis.

The authors selected 29 countries of the 44 countries included in the original (all data) IID-SSA with at least four good-quality and well-spaced data derived from surveys adopting time-consistent statistical conventions that depict medium-term inequality trends reasonably well (table 15.1). On average, there are five data points for each of the 29 countries selected that account for 81.8 per cent of SSA's population. Of the countries excluded, only the Democratic Republic of the Congo (which has three data points) has a large population. The other 14 excluded countries are Benin, Chad, Republic of the Congo, Liberia and Sudan (which have one datapoint each), Cabo Verde, Djibouti, Gabon, Namibia and Togo (which have two datapoints each) and Burundi, Comoros, Seychelles and Zimbabwe (which have three data points), for a total of 30 reliable observations, which were not used in the trend analysis.

Altogether, for the period 1991/93-2011, the IIDB-SSA matrix includes 551 (29 x 19) cells, of which 168 (30.5 per cent) are non-zero. To address the missing data problem, the observed data were connected through linear point-to-point interpolations (as in the left panel of figure 15.1). In Annex 1 of the UNDP-RBA Working Paper No. 2, the data points retained are indicated in the last column of each country summary matrix. The interpolated data are coloured in light blue. Finally, to assign each of the 29 countries of table 15.1 to the rising, falling, U-shaped or inverted U-shaped

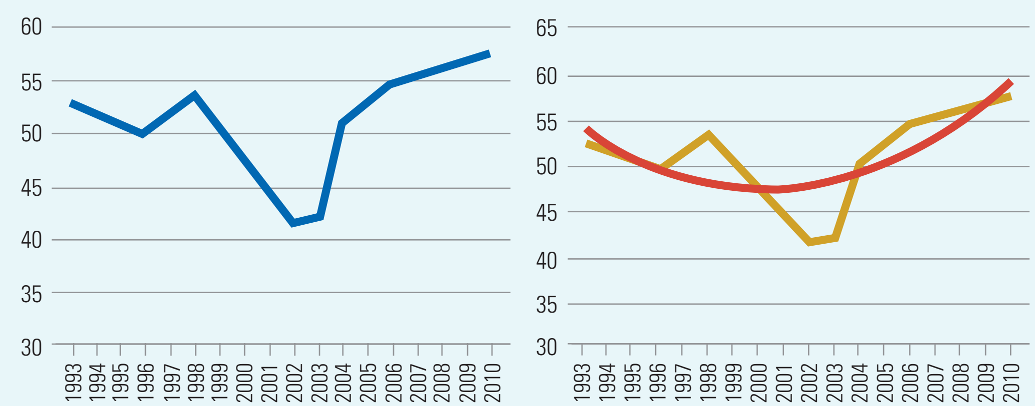
---

[7] For instance, the 2004 and 2011 Integrated Household Surveys for Malawi were standardised by the FAO RIGA project also in terms of household income per capita. They are used for decomposing the Gini changes by income source (Chapter 13).

[8] When aggregating the Gini of these two countries into their respective groups (see below), the authors multiplied them by a correction factor of 0.81. This corresponds to the ratio of the Gini coefficient of the distribution of consumption expenditure to that of disposable income for five countries in the 1980s and early 1990s found by Cogneau et al. (2007).

category, the authors interpolated the Gini time series obtained above with linear and quadratic functions, as shown below in figure 15.2, right panel, with Zambia as an example. In this case, the best fit is clearly a U-shaped pattern. The shape of the trend is decided on the basis of the best R^2 and F-statistics. Finally, each country was assigned to the rising, falling, U-shaped and inverted U-shaped group (see table 15.1).

FIGURE 15.2 Example of interpolation of the missing data points (left panel) and choice of the best interpolated trend (right panel), Zambia



Source: Authors' elaboration.

As shown in Annex 1 of the UNDP-RBA Working Paper No. 2, the same approach was followed for all 29 countries selected. The figures in Annex 1 of the above document show that, in most cases, the Gini from different data sources (identified by dots of different colours) portray trends that are similar to those retained on the basis of the IID-SSA dataset (identified by the orange line). Perceptible differences in levels or trends are evident for only a few countries/years, such as in Ghana (early 1990s), Lesotho (late 1980s), Madagascar, Mozambique, Nigeria and South Africa.

The weighted and unweighted Gini averages of the four groups of countries listed in table 15.1 are plotted in figure 2.3 of Chapter 2, which focuses on an ex ante theoretical discussion of the factors explaining the inequality divergence of the 29 countries analysed. In turn, Chapter 17 presents an econometric investigation of the inequality impact of these factors discussed in Chapter 2.

15.3 Limitations of the IID-SSA and the 'seven sins of inequality measurement' in sub-Saharan Africa

The statistical problems that may reduce the precision of the estimates of the level of the IID-SSA (and other) Gini data are discussed below. In addition, if the measurement biases discussed below vary in intensity over time, the inequality trend may also be affected, as would the analysis 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 it difficult to accurately identify the real level and trend of inequality in SSA. According

TABLE 15.1 Number of data points on the distribution of consumption expenditure per capita for 29 countries with at least four well-spaced Gini data, 1991/3–2011

	Database from which our data were extracted					Data retained for 1993 - 2011					
Country	WIID V3	POVCAL	WB-I2D2	Gini All	Nat'l data	Tot. Obs.	Inter-polated	Total	Pop. share	Gini trend	Difference in Gini points between first and last year
Burkina Faso (1994-2009)		4		1		5	14	19	2.26	Falling	-10.9
Cameroon* (1996-2007)	3					3	16	19	3.05	Falling	-8.8
Ethiopia (1995-2011)					4	4	15	19	**12.82**	Falling	0.0
The Gambia (1993-2003)	4					4	15	19	0.24	Falling	-13.6
Guinea* (1994-2007)	3					3	16	19	1.61	Falling	-1.0
Guinea-Bissau (1993-2005)	3			1		4	15	19	0.24	Falling	-9.5
Lesotho (1993-2003)	5			1		6	13	19	0.32	Falling	-5.4
Madagascar (1994-2010)	4	2		1		5	14	19	3.08	Falling	-7.1
Mali (1994-2010)	4					4	15	19	2.01	Falling	-17.5
Niger (1994-2008)	4			1		5	14	19	2.22	Falling	-15.2
Senegal (1994-2011)	3	1				4	15	19	1.9	Falling	-1.0
Sierra Leone* (1995-2011)		2		1		3	16	19	0.86	Falling	-18.8
Swaziland (1995-2010)	3			1		4	15	19	0.18	Falling	-9.2
Total falling countries	**36**	**7**	**0**	**7**	**4**	**54**	**193**	**247**	**30.79**	**Falling**	**Avg. -9.1**
Angola (1995-2009)	2		1	1		4	15	19	2.79	∩ shape	+18.4 -15.6
Mauritania (1995-2008)	5					5	14	19	0.53	∩ shape	+ 3.5 -3.3
Mozambique (1996-2008)	5					5	14	19	3.54	∩ shape	+ 2.6 -5.7
Rwanda* (1995-2011)	2	1				3	16	19	1.59	∩ shape	+9.1 -4.1
Total ∩-shaped countries	**14**	**1**	**1**	**1**		**17**	**59**	**76**	**8.45**	**∩ shape**	**Avg. +8.4 -7.2**
Botswana* (1994-2009)	2		1			3	16	19	0.31	Rising	14.9
Côte d'Ivoire (1995-2008)	4					4	15	19	2.93	Rising	8.0
Ghana (1993-2006)	4	1		1		6	13	19	3.6	Rising	9.0
Kenya (1994-2006)	3	1				4	15	19	**6.02**	Rising	3.8
Mauritius (1991-2011)	3		7			10	9	19	0.21	Rising	2.5
South Africa (1991-2011)		6		1		7	12	19	**8.02**	Rising	5.7
Uganda (1992-2010)	8					8	11	19	**4.84**	Rising	1.4
Total rising countries	**24**	**8**	**8**	**2**	**0**	**42**	**91**	**133**	**25.93**	**Rising**	**Avg. + 6.5**
Central African Republic (1992-2008)	3			1		4	15	19	0.67	U shape	-17.7 +12.7
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	-4.9 +2.4
Zambia (1993-2010)		7		1		8	11	19	1.93	U shape	-11.0 + 15.9
Total U-shaped countries	**16**	**9**	**0**	**4**	**2**	**31**	**64**	**95**	**34.82**	**U shape**	**Avg. -11.8 +7.9**
Overall 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**	

Source: Author's compilation of the databases listed above and population data provided by UN DESA, Population Division (2015).

to Klasen (2014), many factors contribute to this situation. They include the weak capacity of the NSO of the region and the weight of various external actors with different informational needs in determining which data have to be collected. These two conditions affect both the ownership and the design and comparability of surveys. This generates important consequences in terms of data quality and comparability (Sandefur and Glassman, 2013).

The following is a detailed discussion on the seven measurement sins affecting the assessment of inequality levels and trends. Such sins are common to several low-income countries and, to a lesser degree, developed countries. Yet, given the characteristics of the region (a highly informal and little-monetised economy, large seasonal fluctuations in income and consumption, weak statistical institutions, dependence on technical assistance and weak political checks and balances), these measurement sins are more pronounced in SSA and are thus discussed below.

15.3.1 Differences over time in survey design for the same country

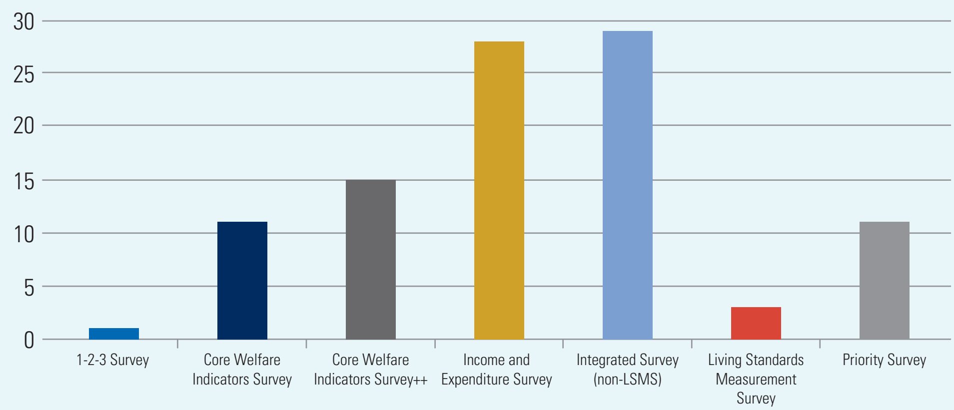
The region has less experience with the HBS than other developing regions. The data collection methodology is evolving to meet higher standards and, as a result, survey design is often modified in subsequent rounds. These changes are sometimes related to data availability, while on other occasions, they respond to the need to improve the quality of the information (Rio Group, 2006). For example, Grimm and Günther (2005) show that the design of the Burkinabé Household Budget 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 for 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).

Similarly, McCulloch, Baulch and Cherel-Robson (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 reported 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) and changes in the reference period (Gibson, Jikun and Scott, 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), causing serious problems to the analysis of inequality in SSA. As mentioned, the World Bank's I2D2 aims to reduce these problems of comparability over time, including by: harmonizing HBS data as far as possible, even ex post; equalising the number of consumption items in different surveys; and filling in missing data. Although some progress has been recorded in terms of data comparability, some of the problems mentioned earlier still hinder the analysis of inequality changes in many countries.

15.3.2 Differences in statistical assumptions and data harmonisation across countries

In recent years, the use of time series and panel econometrics has increased the demand for homogenised questionnaire formats to ensure cross-country comparability. Recent examples of this type of projects are the EU Statistics on Income and Living Conditions and the Programme for Improvement of Surveys and Measurement of Living Conditions in Latin America and the Caribbean. In SSA, despite a growing number of different surveys (figure 15.3), similar initiatives are not yet in place, although the I2D2 and RIGA projects have started filling this gap. As a result, cross-country differences in survey design, definitions, degree of disaggregation, income concept used, timing and size of the survey, 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 survey, *Enquête Intégrale sur les Conditions de Vie des Ménages* (2009/10) provides less accurate information especially in terms of private and public transfers. While some of these problems can be addressed through the use of dummy variables (as in the case of different income concepts), in others, the only solution that can ensure comparability is consistent data harmonisation.

FIGURE 15.3 Types of surveys in African countries, 2000-2011



Source: Dabalen et al. (2011).

To ensure cross-country data comparability, harmonisation should start from microdata and adopt, for all countries and years, the same statistical conventions to define variables such as: 'household income/consumption per capita'; 'household' (whether or not 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; treatment of zero incomes; the grossing-up of income underreporting; and the upward adjustments of rural incomes to capture differences in rural and urban prices.

This harmonisation process improves data comparability but means that the newly produced Gini coefficients differ from those generated by NSO, which may use different assumptions and imputation techniques from those adopted by international agencies or databases. In several Latin American countries, the deviation between standardised SEDLAC and NSO Ginis is negligible, but in others, it reaches 1.5–3 points. Yet, it is rare that differences in inequality levels are accompanied by differences in trends. What matters is that the inequality trends coincide – and they generally do.

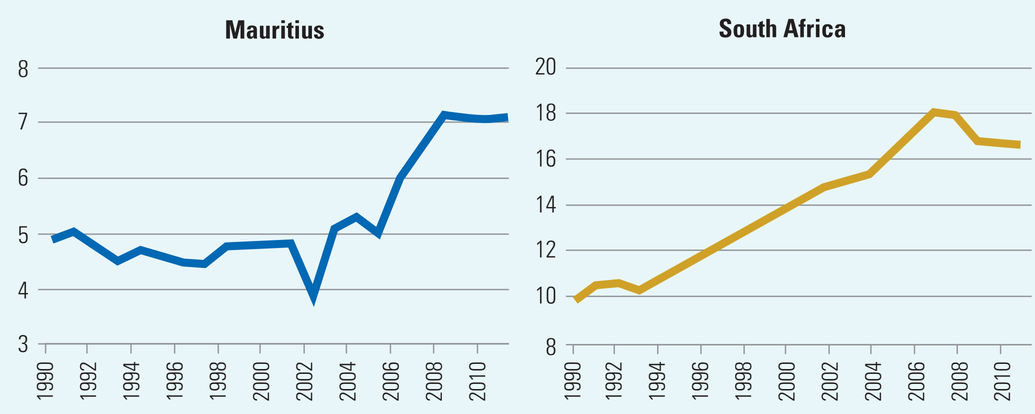
15.3.3 Undersampling of top incomes

The conclusions about inequality levels and dynamics reached on the basis of IID-SSA is likely to be biased by the incomplete accounting of top incomes in all HBS. This is due to their systematic under-sampling and underreporting and to the truncation of very high incomes that are treated as outliers. Such underestimation is more serious with regard to income than to 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 the HBS. This situation leads to an underestimation of the level of 'true inequality' at any point in time. In addition, if the underestimation bias changes over time, it may distort the Gini trend, which could result in identifying spurious causal relations.

The undercounting of top incomes can be tackled by combining HBS data with data derived from tax returns, which make it possible to estimate the income share of the top 1 per cent or other top percentiles. In this regard, the World Top Incomes Database (WTID)[9] has generated a large volume of information for more than 20 countries to date, while other countries are being added gradually. For SSA, the WTID already provides information for Mauritius, South Africa and United Republic of Tanzania (only for 1950-1970), while similar studies have been launched for Botswana, Cameroon, The Gambia, Ghana, Kenya, Lesotho, Malawi, Nigeria, Seychelles, Sierra Leone, Swaziland, Uganda, Zambia and Zimbabwe.

Studies on the share of top incomes crucially depend on how broad based taxation is in these countries (where often only few corporations and individuals file tax returns) and the extent of tax avoidance and evasion. Nevertheless, they 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 per cent has risen sharply during the last 20 years in Mauritius and South Africa (figure 15.4). The HBS-based Gini trend in table 15.1 shows that inequality has risen during the last decade. However, these data underestimate the extent of such increase, as shown below.

FIGURE 15.4 Top 1% income share in Mauritius and South Africa, 1990-2011



Source: The World Top Incomes Database (WTID).

[9] See http://topincomes.parisschoolofeconomics.eu

To 'correct' the HBS Gini estimates, one can rely on the income share of the top 1 per cent or 0.1 per cent declared to the tax office to compute G*, the 'true Gini coefficient,' by using the 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 shows 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). In South Africa (figure 15.5), this gap rises sharply between 1990 and 1995, but then almost stabilises, suggesting that the end of apartheid moderated the rising power of the elites. In fact, since 1995-1996, there is, therefore, a 'level effect' and a modest 'trend effect', which indicates that the conclusions reached on the basis of the uncorrected Gini (G) hold, to some extent.

15.3.4 Cross-checking trends in HBS-based Gini against the trends in the labour share

Another way to check if the trends in HBS-based Gini coefficients are realistic is to juxtapose them with those of the labour share (LS) in total net value added. It is possible, in fact, that the under-sampling of top incomes in HBSs may prevent a correct representation of the recent increase in capital incomes due, for instance, to the rise in mining rents. These effects may be captured by a rise in the 'capital share' computed on the basis of the national accounts. However, the estimation of the latter also has problems related to the accuracy of national accounts, the hypothesis made to compute the LS[10] and 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 trend in the LS is that HBSs substantially and, often, increasingly underestimate the total net value added. For instance, based on Indian data for the 1990s, Ravallion (2001) shows that the HBS-based mean income per capita was only 60 per cent of the value computed on the basis of the national accounts and that this ratio declined over time. In contrast, he found that the difference was not as large in the SSA countries.

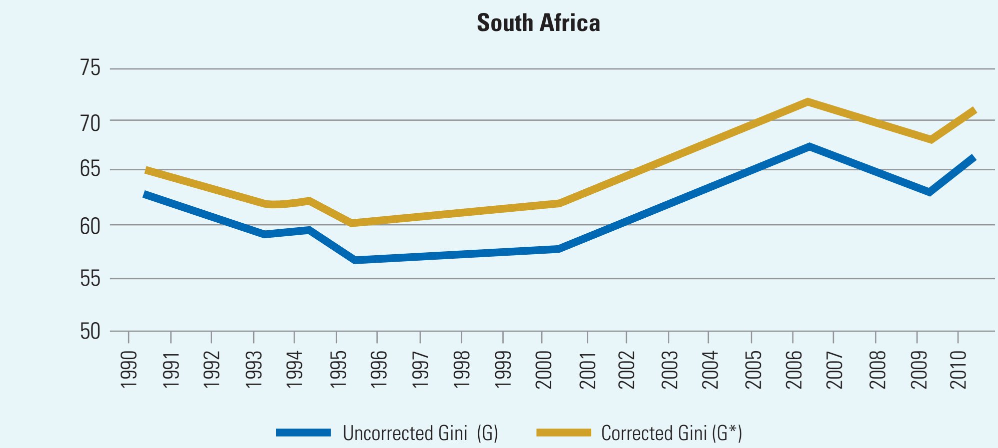
Differences in the level of LS and Gini coefficient are to some extent physiological, because the information on which they are based is collected in different ways and for different purposes. For example, consumption and income levels derived from HBSs are based on information that is self-reported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

To test whether the trends in HBS-based Gini and LS move in the same direction, an analysis of the Guerriero (2012) results was performed. This author computed labour shares for 25 SSA countries (at times, only until the 1990s) using national data from the United Nations National

[10] There are several definitions of the LS. The simplest, LS1, is: (compensation of employees) /[total value added – (indirect taxes + consumption of fixed capital)]. However, this definition poorly fits the reality of countries where most people are self-employed. In those cases, 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 value added weakens the strength of more elaborate estimates of the LS.

Accounts Statistics and applying different methodologies to compute alternative LS. Guerriero's results show that the LS declined over the last few decades in several countries, in particular from the 1980s onwards. These trends (figure 15.6) only partially confirm the Gini tendencies identified in table 15.1. For example, in Senegal, the two trends coincide (the LS rose, while the Gini coefficient declined). The LS and Gini trends are also consistent with each other in Botswana (the LS fell, while the Gini coefficient rose). In Kenya, the two trends are consistent (rising Gini coefficient and falling LS) since 2003, but not before. In contrast, the fall in the labour share in Lesotho is inconsistent with the fall in the Gini coefficient.

FIGURE 15.5 Trend in the HBS-based Gini coefficient (bottom line) and the Gini corrected on the basis of tax returns data (upper line), South Africa, 1990-2010



Source: Authors' calculations based on WTID.

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

Even assuming that the domestic incomes of the rich are fully reported in HBSs (or are added to HBS data on the basis of tax returns), survey data provide a partial picture of the national income distribution whenever SSA nationals hold an important share of their assets abroad. Indeed, the incomes received on these assets do not enter the calculation of national income and its distribution.

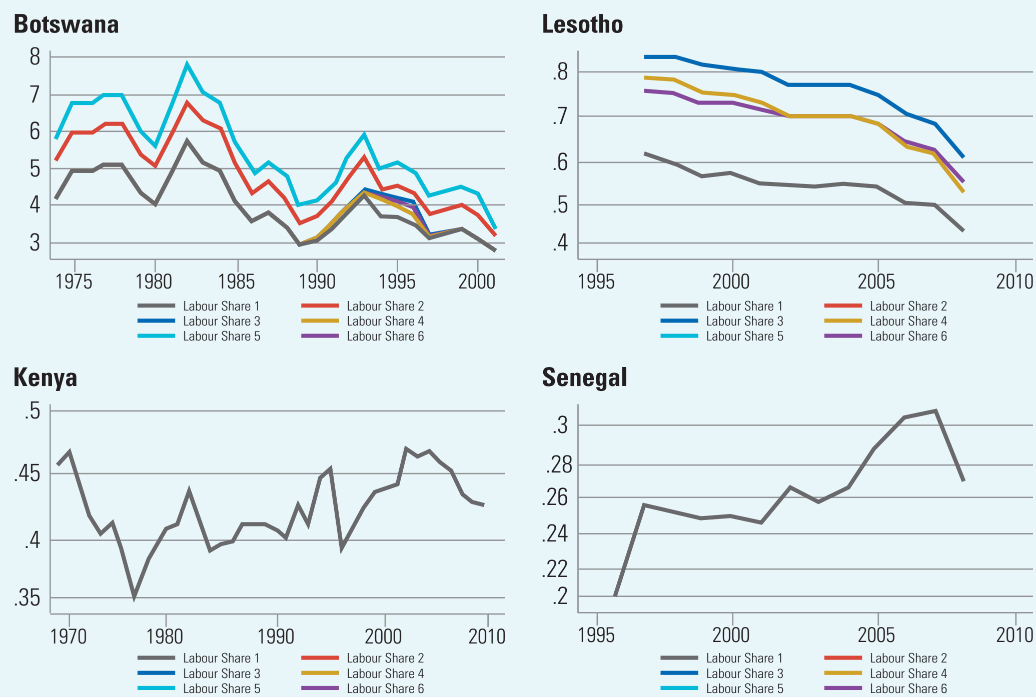
A rich literature suggests that several SSA countries are a major source of capital flight, that substantial assets are held abroad and that these assets generate income that escapes any form of accounting and taxation in the home countries. The release of the Panama Papers in spring 2016 illustrated the extent of this problem in SSA countries.

In countries with a liberalised capital account, capital outflows may be the results of rational portfolio diversification aiming at legally shifting some savings to countries with higher returns on assets, low taxation or low risk of default. However, these flows constitute capital flight if the national norms on taxation and capital controls forbid them. Most importantly, a large share of capital flight involves laundering of illicit earnings (from narcotrafficking or theft of national resources) or shipping resources abroad obtained by embezzling the proceeds of natural resource exploitation. The literature surveyed in Ndikumana (2014) indicates that at least 8 per cent of petroleum rents earned by oil-rich countries with weak institutions ends up in tax havens.

There are two methods to estimate the volume of capital flight: indirect and direct. Following Boyce and Ndikumana (2012), using the indirect method, capital flight (KF) may be estimated as the difference between 'inflows of foreign exchange' (debt-creating capital inflows adjusted for exchange rate fluctuations, plus foreign direct investments) minus 'uses of foreign exchange' (the financing of the current account deficit - CA - and 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. Thus, any difference indicates capital flight. To this imbalance, one can add the value of trade misinvoicing (overinvoicing of imports and underinvoicing of exports - MISINV)which, according to the NGO Global Financial Integrity, constitutes around two-thirds of capital flight. Finally, an additional correction can be included for remittance inflow discrepancy (RID), i.e., unrecorded remittances (estimated at 50 per cent in SSA) such that the above equation becomes KF = (Δ DEBTADJ + FDI) – (CA +ΔRES) + MISINV + RID.

Following this method, Ndikumana (2014) estimated that capital flight from 35 SSA countries over the period 1970-2010 totalled US$820 billion and that the estimated capital held abroad in 2010 (capital flight plus accumulated interests and profits) was US$1,067 billion. Capital flight was

FIGURE 15.6 Evolution over time in the labour share in selected SSA countries and years



Source: Guerriero (2012).

particularly important in oil-rich countries such as Nigeria, Angola, Republic of the Congo and Sudan. These data suggest that SSA is a net creditor to the rest of the world since the value of private assets held abroad exceeds the total (mostly public) liabilities of US$283 billion owed to foreign creditors. Capital flight seems to have risen in recent years, in conjunction with a surge in commodity prices. As revealed by the Panama leaks, in many cases capital flight refers to multinational companies operating in the oil sector. As a result, the increasing price of oil in the late 2000s may be associated with an increasing flight of oil rents to tax havens.

A drawback of the indirect method is the assumption that all capital flight is illicit, while it could be due to underregistration of licit foreign transactions due to weak administrative capacity. To tackle this problem, the direct method focuses on measuring the outcome of capital flight, i.e., the volume of bank deposits or housing property held abroad by citizens of developing countries. In their still-unpublished work, Cogneau and Rouanet follow this approach (personal communication with the authors) for deposits held abroad by foreigners, including from 44 SSA countries for 1980-2010. The basic information source for this estimate is data provided by the Bank of International Settlements.

These data show that in 2010, SSA citizens held deposits abroad equal to approximately 5.3 per cent of GDP (6.1 per cent, if South Africa is excluded), 48 per cent of the value of broad money supply (M2), or 16.6 per cent of domestic money and quasi-money. (This ratio is 9.7 per cent for Latin America and is lower in other regions.) This indicates that an important fraction of savings is placed abroad, rather than invested at home. In 2000, the main African oil-producing countries (Angola, Nigeria, Gabon, Republic of Congo and Cameroon) transferred abroad deposits of around 7 per cent of their GDP. Suggestive evidence cited by Cogneau and Rouanet shows that between 2 per cent and 12 per cent of oil price windfalls are transferred to foreign bank deposits, with larger countries displaying larger flows and stocks of assets held abroad. In absolute terms, these results are similar to those of Ndikumana, but differ if expressed as a share of GDP. Be that as it may, except for South Africa, SSA is the region with the highest foreign deposits relative to domestic money.

The distributive impact of all this is important but not easy to estimate. Given the large amount of wealth held in safe havens, the standard measures of income inequality and wealth distribution are substantially underestimated. If one accepts the Ndikumana estimate of US$1,067 billion in assets held abroad by SSA citizens in 2010 and assumes an average 5 per cent rate of return on assets, then some US$53 billion escape the national accounts. Assuming this income accrues to the top echelon of society, the average regional Gini would rise by two to three points. The authors' calculations for Côte d'Ivoire in 2008 show that if capital flight was taken into account, the Gini coefficient would rise by around 1.5 points. This upward adjustment would be substantial for oil exporters.

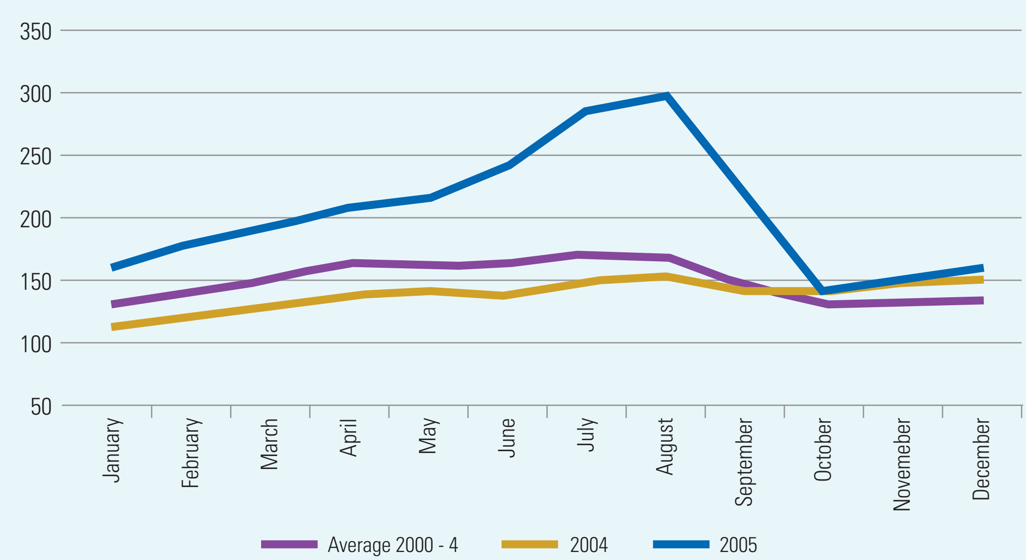
15.3.6. Distributive impact of differences in price dynamics between food prices and overall Consumer Price Index (CPI)

The inequality indices of the distribution of per capita income/consumption are generally computed at current prices. This assumes, implicitly, that all households pay the same price for the goods they consume, that changes over time in these prices affect all households in the same way and that consumption prices recorded over a month or week are stable throughout the year. These assumptions do bias the Gini index downward and are, therefore, discussed below.

a) **Differences in food prices at any point in time.** A number of studies (e.g. Gibson and Kim, 2013) found evidence that the poor pay higher food prices than the non-poor. Mendoza (2011) suggests that this is because reaching the poor may be more costly as they live in remote areas characterised by high transport costs and lower security. 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. Therefore, they do not benefit from discounts granted for bulk purchases and cash payments. For instance, Mussa (2014) shows, based on Malawian 2004 and 2011 surveys, that regardless of location and year, poor households pay more for food than non-poor households. Inequality based on food price-corrected consumption data is thus higher than that computed on uncorrected data. According to his estimates, the nominal Gini coefficient underestimates the 'real Gini' by between two and 3.5 Gini points.

b) **Food price seasonality.** The marked food price seasonality typical of many developing countries further worsens the real purchasing power of the poor. For instance, Cornia and Deotti (2014) show that in Niger, the prices of millet in August peak at levels that are 30-40 per cent higher than in post-harvest September. In years of food crises (as in 2005), the seasonal price increase may be 100 per cent or more (figure 15.7). While such seasonality affects everyone, the poor suffer the most due to their lack of liquidity and access to credit, need to repay debts incurred by selling millet immediately after the harvest at low prices and lack of proper post-harvest storage facilities. In addition, lack of cereal banks increases massively the price they pay for millet, especially during the lean season, when food prices escalate sharply. This problem causes a considerable underestimation of consumption/income inequality and is extremely common in SSA.

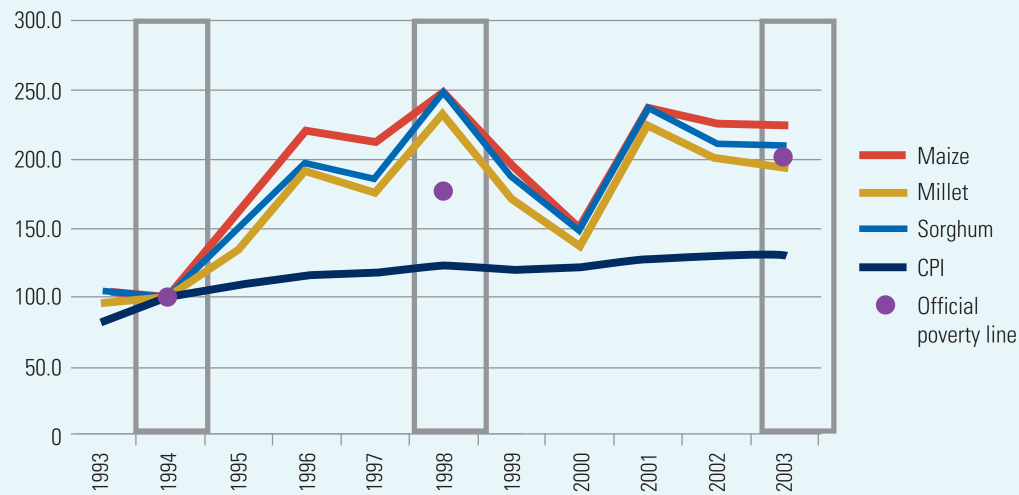
FIGURE 15.7 Monthly consumer price of millet (CFAF/kg): 2005 vs. 2004 and average 2000-2004



Source: Cornia and Deotti (2014).

c) **Differential price dynamics between food and non-food items.** As noted by Arndt, Jones and Salvucci (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 poorest assign up to 70-80 per cent of their total consumption to food, while those in the top decile assign 20-30 per cent. Thus, whenever the food price index (FPI) and the consumer price index (CPI) diverge substantially (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 Günther, 2005). These authors show, for instance, that the CPI rose by 23 per cent in Burkina Faso between 1994 and 1998, while the price of cereals rose more than 50 per cent (figure 15.8). Similarly, Arndt, Jones and Salvucci (2014) documented that inequality in Mozambique worsened due to a sharp rise in world food prices over 2007–2009, as the urban poor relied heavily on imported food.

FIGURE 15.8 Trends in the index number of the official poverty line, CPI and price of main staples (1994=100), Burkina Faso



Source: Grimm and Günther (2005).

The distributive impact of the observed changes in the FPI/CPI ratio is then tested below, 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 characterised by steep food price rises. Two countries where inequality rose – Malawi and South Africa – were selected. FPI/CPI fell in the first and rose in the second (table 15.2 and figure 15.8). Two countries that both experienced a reduction of inequality were also selected; one experienced falling FPI/CPI (Mali) and the other experienced rising FPI/CPI (Madagascar).

TABLE 15.2 Summary of the impact of changes in the FPI/CPI ratio on the Gini coefficient

Country	Years	Inequality trend	% change in 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-CPI divergence, the authors used the quintile distributions provided by WIIDv3.0b and assumed, for all four countries, the following quintiles 'plausible food consumption shares', i.e., 0.7, 0.6, 0.5, 0.4 and 0.3. To ensure comparability between the values of the Gini of the first and second year (given that the FPI/CPI ratio had changed significantly), the authors recalculate at time t+1 the quintile distribution corrected for changes in FPI/CPI through the 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} are the food shares in total consumption. The results presented in figures 15.9 and 15.10 are summarised in table 15.2, which shows that the simulated changes in the Gini coefficient are generally moderate, ranging between 0.3 and 1.5. This is due, in part, to the use of quintiles distribution that generate lower Gini increases than the Gini increases estimated on micro-data, which are reported in the last bar in each figure. This is generally two to three points higher than that computed on quintile distributions. As can be observed, the Gini coefficient changed by up to 1.5 points in the four countries selected (where the FPI/CPI price changes were marked).

FIGURE 15.9 Impact on the Gini coefficient of changes in the FPI/CPI ratio, Malawi (left panel, rising inequality and falling FPI/CPI) and South Africa (right panel, rising inequality and rising FPI/CPI)

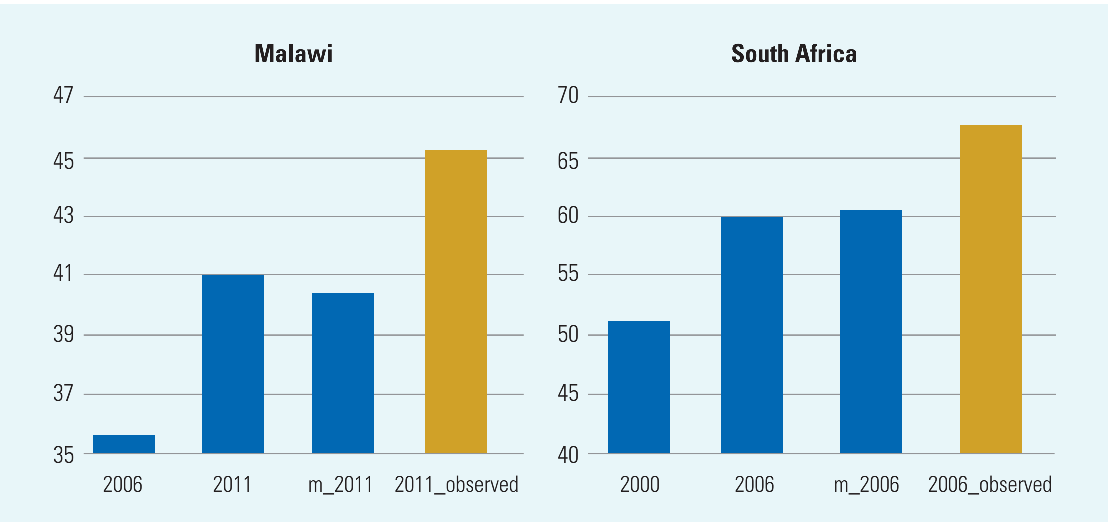
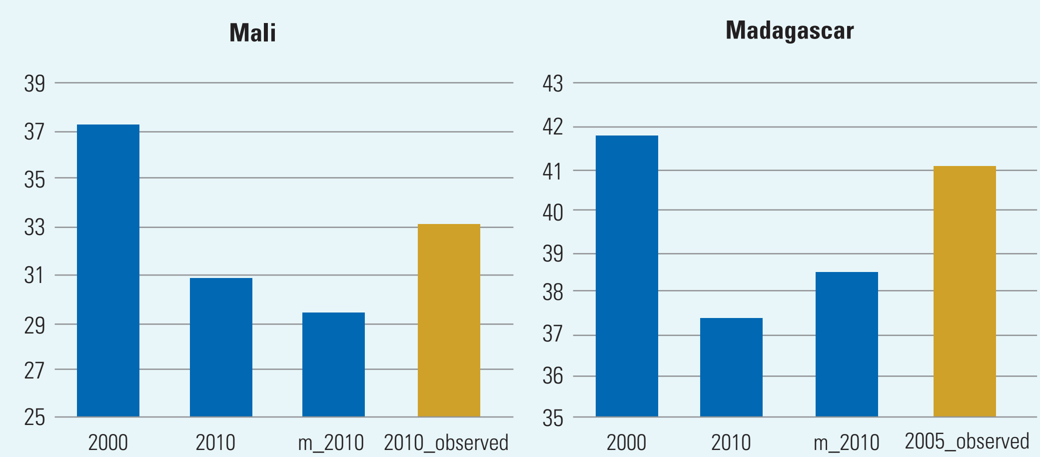


FIGURE 15.10 Impact on Gini coefficient of changes in the FPI/CPI ratio in Mali (left panel, falling inequality and falling FPI/CPI) and Madagascar (right panel, falling inequality and rising FPI/CPI)



Source: For both tables 15.9 and 15.10, authors' elaboration.
Notes: The first two bars on the left represent 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) preceding 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 microdata.

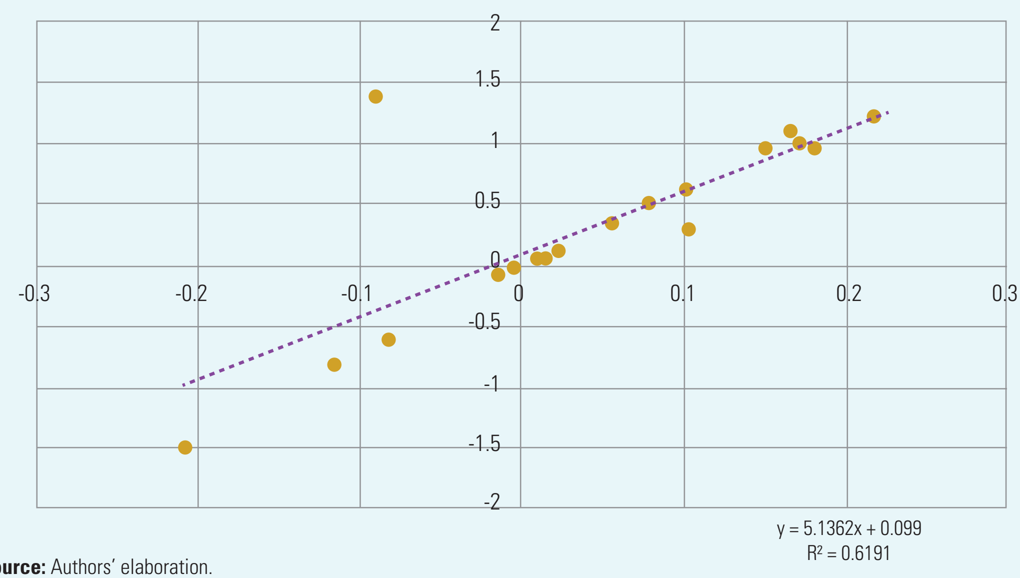
The test was then expanded to 18 countries for which the authors have corrected Gini and FPI/CPI data for 2000-2012 (a period during which the FPI/CPI ratio rose in SSA by between 5 and 30 per cent, while it fell in only a few). This expansion aimed to determine whether changes in the latter variable may have affected the values and trends of the Gini coefficients summarised in table 15.1 and used to analyse inequality trends in SSA in Chapter 2. The bivariate relationship was tested 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 shows (figure 15.11) that the Gini coefficient rises by 0.52 points for every ten-point increase in FPI/CPI. The relation is stable, as suggested by a R^2 of 0.62.

15.3.7 Distributive impact of differences in the provision of social benefits across countries

The discussion has focused thus far on the distribution of private income and consumption (which include cash transfers from the state, where they exist). Yet, individual and household well-being also depends on the amount of in-kind services provided by the state, with particular reference to health and education. Indeed, any comprehensive welfare comparison should take into account the value and incidence of the services supplied in-kind by the state to the various quintiles of the population. In the absence of these state services, households have to buy them in the market, which reduces their ability to consume other essential items, such as food.

The overall value of public expenditure on health and education in SSA is comparatively low. In particular, the expenditure on health rose from 2.4 per cent of GDP in 2000 to 2.8 in 2010. In turn, expenditure on education rose from 3.5 per cent of GDP in 2000 to 4.3 in 2010. However, there is

FIGURE 15.11 Relationship between the first difference over time of the FPI/CPI ratio (x axis) and the first difference of the Gini coefficient, 18 SSA countries, 2000-2012

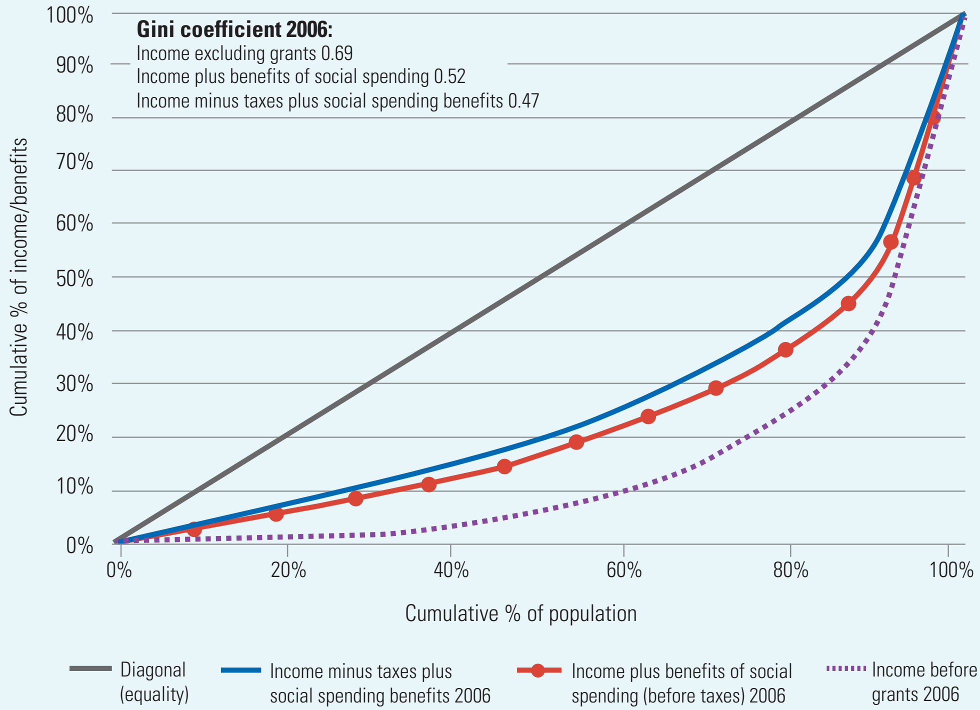


Source: Authors' elaboration.

considerable variation across countries. As shown in Chapter 2, table 2.4, covering a sample of ten countries and nine SSA countries in the late 1990s, the incidence of public health and education expenditure was not progressive, even for primary health care and elementary education. Yet, it was less regressive than the distribution of private income and consumption, thus generating a modest redistributive effect. With the emphasis on the MDGs during the last decade, the incidence of social spending likely improved, thereby generating a favourable effect on the distribution of private income/consumption, including the social wage.

The literature shows that the incidence of education and health spending tends to be more pro-poor in wealthier than in poorer countries. In addition, democratic countries characterised by high inequality (such as South Africa) spend a greater amount of well-targeted public resources, possibly as a result of policymakers' desire to reduce income disparities. Figure 15.12 on South Africa 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 five points. All this suggests that public expenditures in cash and in-kind can be a potent tool to equalise the sum of the distribution of private and social income, as shown recently by Ostry, Berg and Tsangarides (2014) on a large country panel. In contrast, in poorer SSA countries, such as in the Sahelian region, that are characterised by limited public social spending, the redistributive role of the state via the provision of public services is more limited.

FIGURE 15.12 Impact of cash transfers and social spending on health and education, South Africa, 2006



Source: Van der Berg (2009).

15.4 Conclusions

This chapter first illustrated the procedure followed in building the IID-SSA dataset, which provides an important contribution to identifying inequality trends in the region, as analysed in Chapter 2. The second part of the chapter illustrated the main problems encountered in measuring income and consumption inequality in SSA and the possible corrections needed to produce more realistic figures. Policymakers, international agencies and country analysts in the region 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 may be summarised below.

Any analysis should start from a careful examination of the inequality statistics so as to ensure that the data utilised refer to the same income concept, geographical coverage, period of the year and so on. The exclusion of inconsistent data when building the IID-SSA dataset entails a loss of information, but is compensated by greater cross-country comparability and a lower risk of identifying spurious relations. If possible, survey microdata should be harmonised ex-ante by using the same questionnaires and statistical conventions, as done in the RIGA project since 2005 and in similar initiatives in Latin America. The ex-post harmonisation of past data is also useful, but requires relying on many assumptions that may be questionable at times. The inequality statistics computed on data harmonised ex-post, as currently performed by the World Bank for SSA, differ from those calculated by NSOs, at times by one to three Gini points. However, in the Latin American case, this difference only concerns the level and not the trend of such indicators. Nonetheless, there may be exceptions.

Even the harmonisation of HBSs does not necessarily approximate the 'true inequality' of a country because top incomes are undercounted in the HBS and the returns on assets held in safe havens by the elites are not included in either the surveys or the national accounts. As shown above for South Africa, the inclusion of top incomes raises the Gini coefficient by three to five points. Similarly, the inclusion of return on assets held abroad in the distribution of national incomes raises it by another two points. This means that data used here likely underestimate the 'true Gini' by five to eight Gini points and, possibly, more in the case of exporters of valuable commodities. The key analytical issue here is whether such underestimation only concerns the level of Gini (a fact that is certain) or, also, its trend. Figure 15.4 on South Africa suggests the trend is less affected than the level, but this may not be true in countries such as Angola or Equatorial Guinea, where recent oil discoveries and weak redistributive institutions are unlikely to have improved the Gini level. Overall, the IID-SSA Gini presented in table 15.1 are a lower-bound estimate of the 'true Gini'. This is especially true in countries with a high asset concentration and that export valuable primary commodities.

A 15 per cent or greater increase in the Food Price Index relative to the CPI entails an additional increase of the Gini coefficient. In this case, researchers should also take into account this divergence when analysing trends and designing policy. The trends in the labour share can help cross-check the robustness of the Gini trends identified by IID-SSA. Yet, given the problems encountered in measuring it in highly informal economies, such comparison may be of more limited use than in industrialised economies. Finally, the inclusion of social services in the calculation of overall (private and public) household income/consumption per capita likely reduces the Gini coefficient even in poor SSA countries, as well as in middle-income countries like South Africa. More work is needed, however, to estimate the volume and incidence of these public services. This information is useful for policymakers seeking to improve the distribution of well-being by providing those social services that have been shown to reduce inequality over the short term and across generations. In the case of South Africa, the redistributive effect of in-kind services 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 coefficient. CEPR Discussion Paper 8071. London, Centre for Economic Policy Research. www.cepr.org/pubs/dps/DP8071.asp

Arndt, C., S. Jones, and V. Salvucci. 2014. When do relative prices matter for measuring income inequality? UNU-WIDER Working Paper 2014/129.

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

Boyce, J. K. and L. Ndikumana. 2012. Capital flight from sub-Saharan African countries: updated estimates, 1970-2010. PERI Working Papers.

Cogneau, D., T. Bossuroy, Ph. De Vreyer, C. Guénard, V. Hiller, P. Leite, S. Masplé-Somps, L. Pasquier-Doumer and C. Torelli. 2007. Inequalities and Equity in Africa. Paris, Agence Française de Développement.

Cornia G.A. 2015. Income inequality in Latin America: recent decline and prospects for its further reduction. WIDER WP 20/2015, UNU-WIDER.

Cornia G.A. and L. Deotti. 2014. Prix du mil, prix du mil, politique publique et malnutrition des enfants: le cas du Niger en 2005. *Revue d'économie du développement.* De Boeck Université, 22(1), 5-36.

Dabalen, A., R. Mungai, and N. Yoshida. 2011. Frequency and Comparability of Poverty Data in SSA. PREM Knowledge and Learning. http://goo.gl/c0tSaZ

Deaton, A. S. and M. Grosh. 2000. Consumption. *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Study.* Margaret Grosh and Paul Glewwe, eds. Oxford University Press for the World Bank 1, 91-133.

Ferreira, F. 2014. Growth, inequality and poverty reduction in Africa. www.studio-cx.co.za/gtac/wp-content/uploads/2014/11/Francisco-Ferreira-Presentation2.pdf

Gibson, J. 1999. How Robust Are Poverty Comparisons to Changes in Household Survey Methods? A Test Using Papua New Guinea Data. Department of Economics. Hamilton, New Zealand, University of Waikato.

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

Gibson, J., H. Jikun and R. 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, M. and I. Günther. 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, M. 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, S. 2014. World Income Inequality Databases: An Assessment of WIID and SWIID. No. 2014-31, September 2014. Institute of Economic and Social Research, www.iser.essex.ac.uk/research/publications/working-papers/iser/2014-31.pdf

Jolliffe, D. 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, S. 2014. Measuring Poverty and Inequality in Sub-Saharan Africa: Knowledge Gaps and Ways to Address Them. http://blogs.worldbank.org/africacan/measuring-poverty-and-inequality-sub-saharan-africa-knowledge-gaps-and-ways-address-them

Lanjouw, J. O. and P. Lanjouw. 2001. How to compare apples and oranges: poverty measurement based on different definition of consumption. *Review of Income and Wealth*, 47(1): 25-42.

Mendoza R.U. 2011. Why do the poor pay more? Exploring the poverty penalty concept. *Journal of International Development* 23, 1-28.

McCulloch, N., B. Baulch. and M. Cherel-Robson. 2000. Growth, Inequality and Poverty in Mauritania, 1987-1996. Poverty Reduction and Social Development Africa Region. The World Bank. Mimeo.

Mussa, R. 2014. Food price heterogeneity and income inequality in Malawi: Is inequality underestimated? MPRA Paper No. 56080. http://mpra.ub.uni-muenchen.de/56080

Ndikumana, L. 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, M. 2001. Measuring aggregate welfare in developing countries: how well do national accounts and surveys agree? Mimeo. World Bank.

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

Sandefur, J. and A. 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, 20-21 September.

UNDESA (United Nations Department of Economic and Social Affairs) Population Division. 2015. World Population Prospects 2015. New York.

Van der Berg, S. 2009. Fiscal Incidence of Social Spending in South Africa: A Report to the National Treasury. University of Stellenbosch.