
Poverty and inequality trends in South Africa using different survey data

DEREK YU

Stellenbosch Economic Working Papers: 04/10

KEYWORDS: SOUTH AFRICA, HOUSEHOLD SURVEY, POVERTY, INEQUALITY, MISSING
DATA, IMPUTATION
JEL: I32

DEREK YU
DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH
PRIVATE BAG X1, 7602
MATIELAND, SOUTH AFRICA
E-MAIL: DEREKY@SUN.AC.ZA



UNIVERSITEIT
STELLENBOSCH
UNIVERSITY



A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE
BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

Poverty and inequality trends in South Africa using different survey data

DEREK YU¹

ABSTRACT

There is an abundance of literature adopting the monetary approach (i.e., using per capita income or expenditure variables) to derive poverty and inequality trends for South Africa since the transition. The most commonly used data sets used for these analyses are the censuses and the Income Expenditure Surveys (IESs) conducted by Statistics South Africa (Stats SA). However, in some recent studies, alternative data sources were used, namely the All Media Products Survey (AMPS) by the South African Advertising Research Foundation (SAARF), as well as the National Dynamic Income Study (NIDS), which is conducted by Southern African Labour and Development Research Unit (SALDRU).

Some of the data sets are problematic in a particular year or in more than one year, which in turn makes the comparison of poverty and inequality results across the years difficult. Examples of these problems are as follows: the serious decline of income and expenditure between the 1995 and 2000 IES; the high proportion of households with zero or unspecified income in the censuses; too few household expenditure bands in the General Household Surveys (GHSs). In addition, in the various studies mentioned above, different poverty lines were used in the poverty analysis, with the most commonly used poverty line values being R250 per month in 1996 Rand, US\$1 a day, US\$2 a day, as well as R211 per month and R322 per month in 2000 Rand (i.e., the two official poverty lines proposed by Woolard and Leibbrandt (2006).

This paper aims to consistently apply the same poverty lines (i.e., the proposed official poverty lines mentioned above) across all the available survey data, in order to explore the poverty and inequality trends over the years, and to find out if these trends are consistent across different surveys during the period under investigation. The data quality problems mentioned above are addressed (if possible), before the poverty and inequality trends are derived.

Keywords: South Africa, Household survey, Poverty, Inequality, Missing data, Imputation
JEL codes: I32

¹ The author gratefully acknowledges the valuable comments by Professor Servaas van der Berg.

Poverty and inequality trends in South Africa using different survey data

1. Introduction

There is an abundance of literature adopting the monetary approach (i.e., using per capita income or expenditure variables) to derive poverty and inequality trends for South Africa since the transition. The most commonly used data sets used for these analyses (e.g., Hoogeveen & Ozler (2006), Leibbrandt *et al.* (2005), Leibbrandt *et al.* (2006), Simkins (2004), van der Berg & Louw (2004), and Yu (2009)) are the censuses and the Income Expenditure Surveys (IESs) conducted by Statistics South Africa (Stats SA). However, in some recent studies, alternative data sources were used. For example, Van der Berg *et al.* (2007) used a data set conducted by an institution other than Stats SA, namely the All Media Products Survey (AMPS) by the South African Advertising Research Foundation (SAARF). In addition, the study by Argent *et al.* (2009) used a newly available data set, namely the National Dynamic Income Study (NIDS), which is conducted by Southern African Labour and Development Research Unit (SALDRU).

Some of the data sets are problematic in a particular year or in more than one year, which in turn makes the comparison of poverty and inequality results across the years difficult. Examples of these problems are as follows: the serious decline of income and expenditure between the 1995 and 2000 IES; the high proportion of households with zero or unspecified income in the censuses; too few household expenditure bands in the General Household Surveys (GHSs). In addition, in the various studies mentioned above, different poverty lines were used in the poverty analysis, with the most commonly used poverty line values being R250 per month in 1996 Rand, US\$1 a day, US\$2 a day, as well as R211 per month and R322 per month in 2000 Rand (i.e., the two official poverty lines proposed by Woolard and Leibbrandt (2006)). This paper aims to consistently apply the same poverty lines (i.e., the proposed official poverty lines mentioned above) across all the available survey data, in order to explore the poverty and inequality trends over the years, and to find out if these trends are consistent across different surveys during the period under investigation. The data quality problems mentioned above are addressed (if possible), before the poverty and inequality trends are derived.

The structure of the paper is as follows: Section 2 will look at the available survey data sets for poverty and inequality analyses. The focus will be whether questions relating to both income and expenditure are asked in survey, as well the way the questions are asked (i.e., if the respondents are asked to declare the exact amount or the relevant income or expenditure category). In addition, the problems (if any) which affect the reliability and comparability of the data across the years as well as the approaches to address the problems will be looked at. The derivation of per capita variables will be discussed in Section 3. This is followed by Section 4, which will compare the national accounts income data with the total income or expenditure derived in each survey. Trends on poverty and inequality will be looked at in Section 5, by using the per capita variables derived in Section 3. Section 6 will conclude the paper.

2. Income and expenditure data sources for poverty and inequality analyses

Various survey data sets are available to provide income and expenditure information for poverty and inequality analyses in South Africa, and almost all of these surveys are conducted by Stats SA. However, some data sources provide both income and expenditure data, while the others only provide one or the other. In addition, in most of these data sources, the respondents are asked to provide their income or expenditure levels in broad bands instead of the actual amounts. Furthermore, a large number of households reported zero or unspecified household income or expenditure in some surveys. Therefore, finding appropriate data to analyse poverty and inequality of the South African economy is a challenge.

In this section, the available surveys for these analyses are looked at. In addition, the data quality problems, if any, of each data set are looked at. The potential problems include the incorrect derivation of the household income or expenditure variable, a high proportion of households reporting zero or unspecified household income or expenditure, too few bands, etc. Approaches (e.g., imputations) used to address the data quality problems will be discussed.

2.1 Population censuses

Since the transition, two population censuses have taken place, in 1996 and 2001. As the cabinet decided that a census would not be conducted in 2006, a gap in information between Census 2001 and the next census (Census 2011) was created. Later, a decision was taken to conduct the 2007 Community Survey (CS 2007)². In the two censuses, the sample is a 10% unit level sample of all households, while in CS 2007, nearly one million households took part.

In all three surveys, household expenditure was not captured, while each member of the household was asked to declare his/her relevant personal income category. Furthermore, in 1996, each household member was asked to declare his/her additional income and remittances received³.

As far as the derivation of the household income is concerned, in Census 1996, it was equal to the sum of personal incomes, additional incomes and remittances received from the members of the household, while it was simply calculated as the sum of all personal incomes in both Census 2001 and CS 2007 (Yu, 2009: 11 – 16). As far as the household income bands in each survey are concerned, these bands were not consistent between 1996 and 2001, while they were exactly the same in 2001 and 2007 nominally.

² Strictly speaking, Census 1996 and Census 2001 are not surveys. However, for the remainder of the paper, all three sources of data will be referred to as surveys.

³ In Census 1996, the three income questions were asked as follows:

- Personal income (Question 20, Section A): “Think of the past year (1 October 1995 to 30 September 1996) and the money each person received. Please indicate this person’s income category before tax. Answer this question by indicating each person’s weekly, monthly or annual income. Include all sources of income, for example housing loan subsidies, bonuses, allowances such as car allowances and investment income. If this person receives a pension or disability grant, please include this amount.”
- Additional income (Question 1.1, Section B): “Think of any additional that this income generates, and that has not been included in the previous section (For example, the sale of home-grown produce or home-brewed beer or cattle or the rental of property. Please indicate this total amount, if anything, during the past year. (1 October 1995 – 30 September 1996). If none enter ‘0.’”
- Remittances received (Question 1.2, Section B): “If this household receives any remittances or payments (for example money sent back home by someone working or living elsewhere or alimony), please indicate the total received during the past year. (1 October 1995 – 30 September 1996). If none enter ‘0.’”

In Census 2001, the personal income question was asked as follows (Question 22, Section A): “What is the income category that best describes the gross income of this person before tax?”

Finally, in CS 2007, the personal income question was asked as (Question P-52, Section G) “What is the income category that best describes the gross monthly or annual income of the person before deductions and including all sources of income?”

In Census 1996, three rules were adopted by Stats SA when the household income was derived from the personal incomes (Yu: 2009: 11). However, when checking the household income variable, it was found out the rules were not applied consistently to some households. Hence, it seems that the household income variable derived originally by Stats SA is not accurate (Yu: 2009: 12 – 13), as shown in Table 1. For the remainder of the paper, the 1996 household income variable derived by the author (by applying the three rules correctly to all households) will be used for the forthcoming analyses. As far as this household income variable is concerned, it can be seen from Table 1 that 13.0% of households had zero income, while 11.5% had unspecified income.

Table 1 Proportion of households in each annual household income category, using the variables derived by Stats SA and the author respectively, Census 1996

	Household income (Derived by Stats SA)	Household income (Derived by the author)
1: None	14.0%	13.0%
2: R1 – R2 400	7.9%	6.4%
3: R2 401 – R6 000	15.5%	16.1%
4: R6 001 – R12 000	13.7%	12.3%
5: R12 001 – R18 000	9.9%	9.7%
6: R18 001 – R30 000	9.5%	8.9%
7: R30 001 – R42 000	5.4%	5.0%
8: R42 001 – R54 000	4.0%	3.8%
9: R54 001 – R72 000	4.2%	4.1%
10: R72 001 – R96 000	2.9%	2.8%
11: R96 001 – R132 000	2.9%	2.9%
12: R132 001 – R192 000	1.8%	1.8%
13: R192 001 – R360 000	1.3%	1.3%
14: R360 001 or more	0.4%	0.4%
99: Unspecified	6.6%	11.5%
	100.0%	100.0%

Looking at Census 2001, 15.6% of respondents had unspecified personal income. Hence, Stats SA applied the so-called hot deck imputation method⁴ to impute the personal income category of these people, before the household income variable was derived. As everyone had specified personal income after the hot deck imputation, this ensured that no household had unspecified household income, as shown in the third column of Table 2. However, if the 1996 household income variable (which contains 11.5% of households with unspecified household income) without any imputations involved and the 2001 household income derived after hot deck imputation was applied on personal incomes are both used for poverty and inequality analyses, the results might not be comparable. Hence, the 2001 household income variable before hot deck imputation would be used for the poverty and inequality analyses. From the second column of Table 2, it can be seen that, before hot deck imputation was applied, 21.0% and 16.4% of households reported zero and unspecified household income respectively.

Finally, in CS 2007, as in Census 1996, no imputations were involved when the household income variable was derived. In addition, some decision rules were applied when the former variable was derived. However, the decision rules differed slightly from those in 1996 (Yu: 2009: 15). The last column of Table 2 shows the proportion of households in each category, and it can be seen that CS 2007 also suffered the problem of a high proportion of households with zero or unspecified household incomes (8.2% and 11.1% respectively).

⁴ Missing values are substituted with observed values drawn from similar responding units, e.g., the observational units are divided into cells and then each missing value within the cell is replaced with a random draw from the observed values.

Table 2 Proportion of households in each annual household income category, Census 2001 and CS 2007

	Census 2001		CS 2007
	Before hot deck imputation	After hot deck imputation	
1: None	21.0%	23.5%	8.2%
2: R1 – R4 800	7.2%	8.1%	5.0%
3: R4 801 – R9 600	15.6%	17.8%	9.0%
4: R9 601 – R19 200	13.3%	16.0%	18.9%
5: R19 201 – R38 400	10.3%	13.0%	19.1%
6: R38 401 – R76 800	7.0%	9.1%	11.4%
7: R76 801 – R153 600	4.9%	6.6%	7.6%
8: R153 601 – R307 200	2.8%	3.8%	5.3%
9: R307 201 – R614 400	1.0%	1.4%	2.8%
10: R614 401 – R1 228 800	0.3%	0.4%	0.9%
11: R1 228 801 – R2 457 600	0.2%	0.3%	0.3%
12: R2 457 601 or more	0.1%	0.1%	0.2%
13: Response not given	16.4%	0.0%	11.1%
	100.0%	100.0%	100.0%

In the two censuses and CS 2007, the annual total household income amount was derived by adding the personal income amounts (plus additional income and remittances received in the case of Census 1996), which were derived using specific rules, as mentioned by Yu (2009: footnotes 8 and 9). Hence, it is possible that households from the same household income category could have different total household income amounts.

To sum up, the major problems of the three surveys discussed above are that the household income variable was derived differently in each survey (hot deck imputation was applied in 2001, and Stats SA did not stick to the decision rules when deriving the 1996 household income variable). In addition, in all three surveys, a high proportion of households reported zero (13.0% in 1996, 21.0% in 2001 and 8.2% in 2007) or unspecified (11.5% in 1996, 16.4% in 2001 and 11.1% in 2007) household incomes. The first problem was already addressed, as discussed above.

As far as the second problem (i.e., a high proportion of households with zero or unspecified income) is concerned, Ardington *et al.* (2005: 5-7) argue that if those with missing data fall disproportionately at the bottom of the income distribution, then levels of poverty will be underestimated. In contrast, if non-response is higher among the wealthy, measures of inequality are likely to be biased downwards. Furthermore, with regard to the higher proportion of households with zero household income, even allowing for South Africa's high unemployment rates, it is highly unlikely that all of these zero income households had no working-age members earning any income. Therefore, when analyzing poverty and inequality, unless the data is missing completely at random, ignoring households with unspecified household income would lead to biased results. Besides, including households that might incorrectly report zero income might lead to over-estimation of poverty and inequality levels.

Hence, a method called sequential regression multiple imputation (SRMI) was applied at both person and household levels (they will be referred to as SRMI1 and SRMI2 respectively throughout the paper), before the poverty and inequality analyzes are looked at in Section 5. The SRMI methodology, as well as the decision rules applied on SRMI1 and SRMI2 are discussed in detail by Yu (2009: 27 – 34). After SRMI1 was applied, it can be seen from Table 3 below that no households had unspecified household income, but the proportion of households with zero income remained high (14.8% in 1996, 24.7% in 2001, and 8.6% in 2007). On the other hand, after SRMI2, there was also no household with unspecified household income, and the proportion of households with zero income dropped to below 1% in all three surveys.

Table 3 Proportion of households in each annual household income category in each survey, after SRMI1 and SRMI2 respectively

	Census 1996			Census 2001		Census 2007	
	After SRMI1	After SRMI2		After SRMI1	After SRMI2	After SRMI1	After SRMI2
None	14.8%	0.7%	None	24.7%	0.1%	8.6%	0.0%
R1 – R2 400	6.9%	6.9%	R1 – R4 800	7.9%	7.7%	4.9%	5.1%
R2 401 – R6 000	17.3%	22.5%	R4 801 – R9 600	17.5%	29.7%	8.9%	10.7%
R6 001 – R12 000	13.7%	19.3%	R9 601 – R19 200	15.6%	25.4%	19.6%	24.2%
R12 001 – R18 000	11.2%	13.3%	R19 201 – R38 400	12.8%	14.8%	21.2%	23.4%
R18 001 – R30 000	10.3%	11.0%	R38 401 – R76 800	9.1%	9.6%	14.3%	14.4%
R30 001 – R42 000	5.9%	6.3%	R76 801 – R153 600	6.5%	6.7%	9.8%	9.8%
R42 001 – R54 000	4.5%	4.8%	R153 601 – R307 200	3.8%	3.8%	6.8%	6.9%
R54 001 – R72 000	4.8%	4.8%	R307 201 – R614 400	1.3%	1.4%	3.7%	3.6%
R72 001 – R96 000	3.3%	3.4%	R614 401 – R1 228 800	0.3%	0.4%	1.2%	1.2%
R96 001 – R132 000	3.4%	3.3%	R1 228 801 – R2 457 600	0.2%	0.2%	0.4%	0.4%
R132 001 – R192 000	2.1%	2.0%	R2 457 601 or more	0.1%	0.1%	0.3%	0.3%
R192 001 – R360 000	1.4%	1.4%	Unspecified	0.0%	0.0%	0.0%	0.0%
R360 001 or more	0.4%	0.4%		100.0%	100.0%	100.0%	100.0%
Unspecified	0.0%	0.0%					
	100.0%	100.0%					

2.2 Income and expenditure surveys (IESs)

The IES, also conducted by Stats SA, took place three times since the transition, i.e., the IES 1995 (taking place in September 2005), IES 2000 (taking place in October 2000) and IES 2005-2006 (taking place between September 2005 and August 2006, with sampled households participating for one month and new sub-samples of households starting every month). Although the main purpose of the IES is to collect and provide information on income and expenditure patterns of a representative sample of households so as to update the basket of goods and services required for the compilation of the Consumer Price Index (CPI), these surveys have become an important source of information for poverty and inequality analysis. The number of households taking part in the survey was 29582 in IES 1995, 26263 in IES 2000, and 21144 in IES 2005-2006.

The 1995 and 2000 IESs used the recall method. In this method, a single questionnaire was administered to a household at a selected dwelling unit in the sample, and the responding household was required to recall income from different sources (e.g., remuneration, interest income earned, income from gambling, etc.) as well as expenditure on various food and non-food items, either during the month prior to the survey or for the twelve months prior to the survey, before the total annual household income and household expenditure amounts were derived.

In both surveys, the Standard Trade Classification (STC) approach was used to categorize the income and expenditure items, before the household income and expenditure amounts were calculated. Household income was equal to the sum of regular and irregular incomes, while household expenditure was derived by adding the expenditure amounts on twenty categories⁵, as shown in Table 4 below.

⁵ Strictly speaking, there were twenty-one categories, with the twenty-first category being debt. However, the debt amount was excluded when household expenditure was derived.

Table 4 Derivation of household expenditure, IES 1995 and IES2000

Expenditure category
(1): Housing
+ (2): Domestic workers
+ (3): Food
+ (4): Beverages
+ (5): Cigarettes, cigars, tobacco, etc. and smokers' requisites
+ (6): Personal care
+ (7): Other household consumer goods
+ (8): Household services
+ (9): Household fuel
+ (10): Clothing and footwear
+ (11): Furniture/Equipment
+ (12): Health services
+ (13): Transport
+ (14): Computer and telecommunication equipment
+ (15): Communication for household purposes
+ (16): Education
+ (17): Reading matter and stationery
+ (18): Recreation, entertainment and sports
+ (19): Miscellaneous expenditure (e.g., membership fees, remittances, income tax, insurance)
+ (20): Expenditure on own harvest/livestock
= Total household expenditure

There are two major differences between IES 2005-2006 and the previous IESs. First, the diary method was adopted extensively for the first time in order to record the household's daily acquisitions on a daily basis. However, the recall method was still applied in the income and expenditure items other than non-durable items (e.g., food), as shown in Table 5. Secondly, the Classification of Individual Consumption According to Purpose (COICOP) approach was adopted to categorize various items, before the total consumption and total income were derived. The information is presented in Table 6⁶.

Table 5 Derivation of the annualized income and expenditure figure, IES2005/2006

Type of data item	Reference period		Annualized figure
	[A]: Diary (Survey month)	[B]: Main questionnaire	
Non-durable items	1 month	-	[A] × 12
Semi-durable items	1 month	11 months	[A] + [B]
Durable items	1 month	11 months	[A] + [B]
Services	-	1 or 12 months	[B] (if reference period is 1 month) [B] × 12 (if reference period is 12 months)
Regular income	-	1 and 11 months [#]	Monthly figure + 11-month figure [#]
Irregular income	-	12 months	[B]

[#] In IES2005/2006, respondents were asked to declare income for the previous month and income for the 11 months prior to the survey month for all regular income items. These two figures were then added before the annualized figure was derived.

Thus, it can be clearly seen that the STC approach used in IES1995 and IES2000 is not directly comparable with the COICOP method in IES2005/2006 as far as the derivation of the household expenditure variable is concerned. On the other hand, the derivation of household income is very similar across all three surveys (i.e., adding regular income and irregular income items together), except the inclusion of an imputed rent variable as an income item in IES2005-2006, which will be discussed later.

⁶ For more detailed explanation on the COICOP approach, please refer to Yu (2008: 12-14, 22-32).

Table 6 The main categories of the COICOP, IES2005/2006

Group 1: CPI Consumption (i.e., items included for the compilation of the consumer price index, CPI)
○ Food and non-alcoholic beverages
○ Alcoholic beverages, tobacco and narcotics
○ Clothing and footwear
○ Housing, water, electricity, gas and other fuels
○ Furnishings, household equipment and routine maintenance of the house
○ Health
○ Transport
○ Communication
○ Recreation and culture
○ Education
○ Restaurants and hotels
○ Miscellaneous goods and services
○ Other unclassified expenditure
Group 2: In-kind consumption
Group 3: Income
Group 4: In-kind income
Group 5: Savings
Group 6: Taxes
Group 7: Transfer to others
Group 8: Debts
Group 9: Loss
Group 10: Not CPI consumption (i.e., items not included for the compilation of CPI)
Group 11: Products not in income (i.e., income items that are not included in Group 3)

As explained in Table 4, the items from the first twenty expenditure categories were added to derive the total household annual expenditure in IES1995 and IES2000. In IES2005/2006, the expenditure items fell under group 1 (Consumption), group 2 (In-kind consumption), group 5 (Savings), group 6 (Taxes), group 7 (Transfers to others) and group 9 (Loss). In other words, only the total consumption (but not total expenditure) was derived under COICOP in IES2005/2006. Since the COICOP is very different from the Standard Trade Classification, in order for meaningful comparative analysis to be conducted, there are two options:

- Re-categorize the items in 1995 and 2000, using the COICOP structure, i.e., the 1995 and 2000 total income and consumption variables using the COICOP structure are derived, before they are compared with the 2005/2006 total income and consumption variables.
- Re-categorize the items in 2005/2006 using the STC, i.e., the 2005/2006 total income and expenditure variables using the STC approach are derived, before they are compared with the 1995 and 2000 total income and expenditure variables.

The other problem that could affect the comparability of the surveys was the inclusion of a newly introduced item in 2005/2006 – imputed rent 7% per year of value of dwelling – as both an income item and consumption item. In 2000 prices, this variable amounted to R66 927 million. Thus, this might affect the comparability of the poverty and inequality results across the three IESs when the per capita income and consumption variables are used for the analyses. Hence, when the IES2005/2006 poverty and inequality results using the total income and consumption variables (using the COICOP approach) are looked at, the results with imputed rent and without imputed rent would be shown.

Finally, regardless of whether the STC or COICOP approach was adopted, all households had specified income/consumption/expenditure in all three surveys. No households reported zero income/consumption/expenditure amounts in IES 1995, while only a very negligible proportion of households (less than 1% in each survey) had zero income/consumption/expenditure in the other IESs.

2.3 October Household Surveys (OHSs), Labour Force Surveys (LFSs), and Quarterly Labour Force Surveys

Stats SA has been collecting labour market data since 1993 with the OHS, which was conducted annually between 1993 and 1999, as well as the LFS, which was a biannual survey (conducted in March and September) introduced in 2000 to replace the OHS. The QLFS was introduced in 2008 to replace the LFS, and the former takes place four times a year. Although the main objective of these surveys is to capture the labour market status of the individuals, questions relating to total income and expenditure were asked in some surveys.

In the OHSs, total household income was only captured in 1999⁷. As far as the total expenditure is concerned, the question was asked in the OHS 1996-1999⁸. Between 1996 and 1998, the household was asked to declare total expenditure in actual amounts during the past month, while in 1999, the household was asked to declare the relevant monthly expenditure category⁹. The household expenditure question with the same categorical answers as in OHS 1999 was asked again in LFSs, but only in four surveys (i.e., the 2001 – 2004 September surveys)¹⁰.

Table 7 Proportion of households in each monthly household income or expenditure category, OHSs/LFSs

	OHS					LFS (September)			
	1996#	1997#	1998#	1999##	1999## (Income)	2001	2002	2003	2004
R0 – R399	21.5%	23.3%	26.0%	25.0%	15.9%	31.7%	30.6%	24.6%	21.8%
R400 – R799	24.6%	29.1%	29.2%	26.3%	20.9%	26.1%	26.3%	28.8%	28.6%
R800 – R1199	13.8%	14.8%	12.6%	13.2%	12.4%	11.6%	12.1%	13.9%	14.1%
R1200 – R1799	9.0%	8.8%	7.9%	8.5%	11.4%	7.2%	7.4%	7.6%	8.4%
R1800 – R2499	7.1%	5.5%	6.1%	5.8%	7.0%	5.4%	5.8%	5.9%	6.6%
R2500 – R4999	11.1%	8.7%	9.1%	7.8%	11.0%	8.1%	7.5%	8.4%	8.3%
R5000 – R9999	4.1%	3.8%	3.8%	4.3%	7.2%	4.6%	5.2%	5.7%	6.6%
R10000 or more	1.0%	1.1%	0.9%	1.6%	5.0%	1.7%	2.2%	2.9%	3.1%
Don't know / Refuse / Unspecified	7.6%	5.0%	4.5%	7.6%	9.3%	3.7%	3.0%	2.3%	2.7%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
R0 – R1199	59.9%	67.2%	67.8%	64.5%	49.2%	69.4%	68.9%	67.3%	64.4%

The household expenditure was declared as actual amounts in 1996 – 1998, and their answers were used to categorize them into the expenditure categories in Table 7.

The household weight in OHS1999 was inaccurate (Yu, 2007: 38 – 39), and hence the average of the person weights of the members of the household was used instead as a proxy for household weight.

⁷ The total household income question was asked as follows in OHS 1999 (Question 6.36, Section 6): “What was the total household income in the last month, including wage, salaries, government grants, private pensions and all other sources of income?” The respondent then could choose from ten income categories (See footnote 9 below).

⁸ The question was asked as follows in each OHS:

- OHS 1996 (Section 1.39, Section 1), OHS 1997 (Question 9.40, Section 9) and OHS 1998 (Question 9.40, Section 9): “How much money did this household spend in total, on all items (including food, clothing, housing, transport, medical care, etc), during the past month?”
- OHS 1999 (Question 6.31, Section 6): “What was the total household expenditure in the last month? Include everything that the household and its members spent money on, including food, clothing, transport, rent and rates, alcohol and tobacco, school fees, entertainment and any other expenses.” The respondent could choose from ten expenditure categories (See footnote 9 below).

⁹ There were ten categories for the respondents to choose from: 1: R0 – R399, 2: R400 – R799, 3: R800 – R1199, 4: R1200 – R1799, 5: R1800 – R2499, 6: R2500 – R4999, 7: R5000 – R9999, R10000 or more, 9: Don't know, 10: Refuse.

¹⁰ The household expenditure questions in these four September LFSs (Question 6.25, Section 6 in LFS 2001; Question 7.25, Section 7 in LFS 2002; Question 7.29, Section 7 in LFS 2003; and Question 7.30, Section 7 in LFS 2004) were asked in exactly the same way as in OHS 1999.

In the QLFSs, both household income and expenditure questions were not asked. As the main objective of the OHS/LFS/QLFS is to capture the information on the labour market, the household-level questions (which include total expenditure) were gradually excluded from these surveys, and were asked in the General Household Survey (GHS) instead.

The major problem of the OHS/LFS expenditure data is that there are too few expenditure categories, and a very high proportion of households (approximately two-thirds) fell into the first three expenditure categories (See Table 7). Interestingly, this is not the case when looking at household income in OHS 1999, as only 49.2% of households fell into the first three income categories. Comparing the respondents' answers in the household income and expenditure questions in OHS 1999, the results in Table 8 seem to imply that households tended to under-declare their household expenditure, compared with their answers on household income. For example, looking at households in the R800-R1199 income category, only 36.2% fell into the R800-R1199 expenditure category. However, 6.4% (1.8% + 4.6%) reported that the household expenditure was below R800, while the expenditure of 53.2% (22.4% + 12.7% + 13.2% + 3.7% + 1.2%) of the households exceeded R1199.

Table 8 Respondents' answers on household income and household expenditure, OHS 1999

		Household expenditure category								
		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Household income category	[1]	56.0%	4.4%	1.8%	0.8%	0.7%	0.3%	0.3%	0.1%	5.2%
	[2]	25.5%	50.2%	4.6%	2.1%	0.9%	0.6%	0.5%	0.0%	5.2%
	[3]	7.2%	19.2%	36.2%	3.4%	1.7%	0.6%	0.5%	0.4%	4.0%
	[4]	5.3%	12.0%	22.4%	38.4%	5.4%	1.6%	0.2%	0.3%	2.7%
	[5]	1.7%	5.7%	12.7%	16.7%	26.7%	3.2%	0.9%	1.1%	1.7%
	[6]	1.0%	4.2%	13.2%	24.3%	38.6%	39.9%	5.1%	2.9%	2.9%
	[7]	0.2%	0.8%	3.7%	7.6%	16.5%	35.8%	42.6%	3.8%	2.3%
	[8]	0.1%	0.3%	1.2%	2.2%	4.2%	13.3%	43.4%	84.3%	2.3%
	[9]	3.1%	3.2%	4.3%	4.5%	5.4%	4.8%	6.7%	7.1%	73.7%
		100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Note: Income or expenditure categories: [1]: R0 – R399, [2]: R400 – R799, [3]: R800 – R1199, [4]: R1200 – R1799, [5]: R1800 – R2499, [6]: R2500 – R4999, [7]: R5000 – R9999, [8]: R10000 or more, [9]: Don't know / Refuse / Unspecified

Furthermore, in OHS 1996-1998, no households reported zero income or expenditure. In addition, it is impossible to know if any households had zero expenditure in OHS 1999 and the LFSs, since the lowest expenditure category was "R0 – R399". As far as the households with unspecified income or expenditure are concerned (i.e., the third row from the bottom in Table 7), the proportion of households falling into such category was not low (such proportion was nearly 8% in OHS 1996 and OHS 1999), and as mentioned in Section 2.1, excluding these households from poverty and inequality analyses would lead to biased results. Hence, SRMI at household level (i.e., SRMI2) was applied to impute the household expenditure amount (in OHS 1996-1998) and household income or expenditure category (OHS 1999 and LFS 2001-2004) respectively, and the results are presented in Table 9.

Finally, note that the number of households taking part in the OHSs and LFSs mentioned above was about between 26000-28000 in each survey, except in OHS 1996 and OHS 1998, with the sample size just below 20000.

Table 9 Proportion of households in each monthly household income or expenditure category after SRMI2, OHSs/LFSs

	OHS					LFS (September)			
	1996#	1997#	1998#	1999##	1999## (Income)	2001	2002	2003	2004
R0 – R399	21.5%	24.1%	26.6%	25.1%	15.9%	31.8%	30.6%	24.6%	21.8%
R400 – R799	26.4%	30.7%	30.2%	27.9%	22.0%	26.9%	26.7%	29.0%	28.8%
R800 – R1199	15.8%	15.5%	13.2%	15.0%	13.6%	12.2%	12.6%	14.3%	14.4%
R1200 – R1799	11.0%	9.3%	8.4%	9.6%	12.8%	7.7%	7.7%	7.8%	8.8%
R1800 – R2499	8.1%	5.8%	6.5%	6.9%	8.3%	5.8%	6.2%	6.1%	6.9%
R2500 – R4999	12.1%	9.5%	10.0%	8.9%	12.3%	8.7%	8.0%	8.9%	8.7%
R5000 – R9999	4.2%	4.1%	4.1%	5.0%	8.7%	5.1%	6.0%	6.2%	7.2%
R10000 or more	1.0%	1.1%	0.9%	1.7%	6.3%	1.8%	2.3%	3.1%	3.3%
Don't know / Refuse / Unspecified	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
R0 – R1199	63.7%	70.3%	70.0%	68.0%	51.5%	70.4%	69.9%	67.9%	65.0%

The household expenditure was declared as actual amounts in 1996 – 1998, and their answers were used to categorize them into the expenditure categories in Table 7.

The household weight in OHS1999 was inaccurate (Yu, 2007: 38 – 39), and hence the average of the person weights of the members of the household was used instead as a proxy for household weight.

2.4 General Household Surveys (GHSs)

The GHS, which is also conducted by Stats SA, was introduced in 2002. The two main objectives of the survey are to measure the multiple facets of the living conditions of South African households, as well as the quality of service delivery in a number of key service sectors. Six broad areas are covered by the GHS, namely education, health, activities related to work and unemployment, non-remunerated trips undertaken by the household, housing, and household access to services and facilities. The sample size at household level was 24000-28000 in each of the seven GHSs.

Table 10 Proportion of households in each monthly household expenditure category, GHSs

	2002	2003	2004	2005	2006	2007	2008
R0 – R399	31.3%	26.0%	18.5%	19.2%	17.6%	13.5%	9.4%
R400 – R799	27.2%	27.6%	28.6%	28.0%	28.9%	27.2%	23.1%
R800 – R1199	11.9%	13.7%	14.1%	15.0%	17.5%	17.9%	19.1%
R1200 – R1799	7.1%	7.6%	10.4%	10.3%	10.5%	11.9%	12.4%
R1800 – R2499	5.5%	5.9%	6.7%	6.4%	6.5%	7.1%	8.6%
R2500 – R4999	7.2%	7.9%	10.0%	10.3%	9.2%	10.7%	11.5%
R5000 – R9999	4.7%	5.3%	6.4%	6.3%	6.0%	7.3%	8.0%
R10000 or more	1.7%	2.3%	2.2%	2.5%	2.7%	3.0%	5.3%
Don't know / Refuse / Unspecified	3.5%	3.7%	3.1%	2.1%	1.4%	1.5%	2.6%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
R0 – R1199	70.4%	67.2%	61.2%	62.1%	63.9%	58.6%	51.6%

As far as the income and expenditure questions are concerned, the total household income was not captured by the GHS. On the other hand, in all GHSs, the household was asked to declare the relevant monthly household expenditure category¹¹, and the categories were exactly the same as in the LFSs. Table 10 above shows that a similar problem occurs in the GHSs as in the OHSs and LFSs, namely the very high proportion of households falling into the first three expenditure categories. However, it can be seen that such proportion shows a continuous downward trend (from 70.4% in 2002 to 51.6% in 2008).

Furthermore, it is impossible to find out if any households really had zero expenditure, due to the categorization of the expenditure answers. In addition, compared with OHSs/LFSs, the proportion of households with unspecified expenditure in the GHSs (i.e., the third row from the bottom in Table 10) was lower. However, SRMI2 was still applied to impute the household expenditure category of these households, and the results are presented in Table 11.

Table 11 Proportion of households in each monthly household expenditure category after SRMI2, GHSs

	2002	2003	2004	2005	2006	2007	2008
R0 – R399	31.4%	26.0%	18.5%	19.2%	17.6%	13.5%	9.4%
R400 – R799	27.8%	28.2%	29.0%	28.3%	29.1%	27.4%	23.3%
R800 – R1199	12.6%	14.2%	14.6%	15.4%	17.7%	18.2%	19.7%
R1200 – R1799	7.6%	8.3%	10.7%	10.6%	10.7%	12.2%	12.8%
R1800 – R2499	5.9%	6.3%	7.1%	6.7%	6.7%	7.3%	8.9%
R2500 – R4999	7.7%	8.5%	10.5%	10.6%	9.4%	10.9%	11.8%
R5000 – R9999	5.3%	6.0%	7.2%	6.6%	6.2%	7.5%	8.5%
R10000 or more	1.8%	2.5%	2.4%	2.7%	2.7%	3.1%	5.6%
Don't know / Refuse / Unspecified	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
R0 – R1199	71.8%	68.4%	62.1%	62.9%	64.4%	59.1%	52.4%

The survey data discussed so far are all conducted by Stats SA. The next three data sources are conducted by institutions other than Stats SA, and will be the focus of Sections 2.5 – 2.7.

2.5 The Project for Statistics on Living Standards and Development (PSLSD)

The PSLSD survey was conducted in 1993 by SALDRU with assistance from the World Bank. The survey collected a wide range of indicators of standard of living, with the main aim being the collection of statistical information about the conditions under which South Africans lived in order to provide policy makers with the data required for planning strategies to implement the goals outlined in the Government's Reconstruction and Development Programme (RDP). The PSLSD survey only took place once. About 9000 households took part in the survey.

As far as household income is concerned, it was derived by adding the employment income and non-employment income amounts (Section 8 of the questionnaire) as declared by the respondents. The recall method was used. On the other hand, household expenditure was derived by adding the respondents' answers on food spending and non-food spending (Sections 3 and 4 of the questionnaire respectively), and again, the recall method was used. In other words, total household income and expenditure were derived in actual amounts and are not categorical variables.

¹¹ The expenditure question (Question 4.45 in GHS 2002, Question 4.63 in GHS 2003, Question 4.71 in GHS 2004, Question 4.79 in GHS 2005, and Question 4.69 in GHS 2006-2008) was asked as "What was the total household expenditure in the last month? Include everything that the household and its members spent money on, including food, clothing, transport, rent and rates, alcohol and tobacco, school fees, entertainment and any other expenses." The respondent could choose from ten expenditure categories (See footnote 9).

2.6 National Income Dynamic Study (NIDS)

The NIDS, conducted by SALDRU, is the first national panel study of individuals of all ages in South Africa, with the data from the first wave (2008) being released in the last quarter of 2009. The second wave is scheduled to take place in 2010. The main objective of NIDS is to measure and understand who is getting ahead and who is falling behind in South Africa as well as why some people are making progress and the others are not. The five broad areas covered by NIDS are as follows: income and expenditures of the household and the individuals in the household; the assets owned by the household and the services to which the household has access; the level of education and health status; labour market status of the individuals; and membership of community groups. In the first wave, 7305 households took part in the survey.

As far as the total household expenditure is concerned, it was derived in two ways, with the recall method being adopted in both methods. First, it was derived by adding the respondents' answers on food spending, non-food spending and rent expenditure (Sections D and E of the household questionnaire). Secondly, the household head was asked to declare total expenditure in the last 30 days¹². However, since the response rate of this question was only about 79% and it was worried this question would under-estimate total expenditure, it was decided by SALDRU that the total household expenditure would be derived using the first method for all households.

Looking at the derivation of household income, the household was asked (Question D38, Section D of the household questionnaire) to declare total household income amount received in the last month¹³. 5449 households (out of 7305) gave specific answer to this question. For the remaining 1856 households, they were asked (Question D39, Section D of the household questionnaire) to declare the relevant total monthly household income category and could choose from fifteen categories¹⁴. The household income amount was then derived using the mid-point of the interval. Only 474 (out of 1856) households answered this question. In other words, 1382 households (7305 – 5449 – 474) did not give any specific answers on both questions D38 and D39.

SALDRU argued that the two questions mentioned above could under-estimate household income, and opted to use the respondents' answers on each income component in Section E and Section F of the adult questionnaire. In these two sections, questions on employment income and non-employment were asked, with the respondents being asked to declare the actual amounts, and again, the recall method was used. The household income amount was then derived by adding the respondents' answers on the following seven broad components: wage income, government grant income, other government income, investment income, remittances income, implied rent income and agricultural income. This method worked successfully in 7106 households, since the remaining 199 households did not give clear answers in the first five income components mentioned above. Therefore, in these households, the household income amount was derived by adding their one-shot income amount (derived from either question D38 or D39 of the household questionnaire, as discussed above) and the implied rent income together.

To conclude, in NIDS, both the income and expenditure variables were derived as actual amounts and are not categorical variables.

¹² The question was asked as (Question D31, Section D) "How much money did this household spend on all its expenses in the last 30 days?"

¹³ The question was asked as "What was the total amount of income (after income tax) that this household received last month? Please note this includes all the household members' salaries and wages, grants, interest, rental income and income from agriculture earned by household members in the last month."

¹⁴ The question was asked as "Please would you look at the show card and point out the most accurate earnings category for last month's household income?" The fifteen categories are as follows: 1: None, 2: R1-R200, 3: R201-R500, 4: R501-R1000, 5: R1001-R1500, 6: R1501-R2500, 7: R2501-R3500, 8: R3501-R4500, 9: R4501-R6000, 10: R6001-R8000, 11: R8001-R11000, 12: R11001-R16000, 13: R16001-R30000, 14: R30001-50000, 15: R50001 or more.

2.7 All Media Products Survey (AMPS)

AMPS has been conducted either semi-annually or annually by the South African Advertising Research Foundation (SAARF) since 1975. It is mainly used for market research, with data on the usage of a wide range of household goods and services being collected. One respondent aged 16 or above from each household, not necessarily the household head, is asked to fill in the questionnaire. The sample size in each survey ranges between 12000 and 24000.

Household expenditure was not captured in all AMPSs. On the other hand, the household income information was collected through showing respondents cue cards divided into 29 or more categories, as shown in Table 12¹⁵. If a respondent refused to answer the question on household income, SAARF imputed household income on the basis of household expenditure implied by the product questionnaire. Thus, it can be seen that AMPS has an advantage that there are more number of income categories, and the income range in each category is narrower, compared with the censuses and the OHS/LFS/GHS.

Table 12 The total annual household income categories in each AMPS

	1993	1994-1996	1997-1999	2000-2001	2002-2006	2007-2008
1	R1-R99	R1-R99	R1-R99	R1-R199	R1-R199	R1-R299
2	R100-R199	R100-R199	R100-R199	R200-R299	R200-R299	R300-R399
3	R200-R299	R200-R299	R200-R299	R300-R399	R300-R399	R400-R499
4	R300-R399	R300-R399	R300-R399	R400-R499	R400-R499	R500-R599
5	R400-R499	R400-R499	R400-R499	R500-R599	R500-R599	R600-R699
6	R500-R599	R500-R599	R500-R599	R600-R699	R600-R699	R700-R799
7	R600-R699	R600-R699	R600-R699	R700-R799	R700-R799	R800-R899
8	R700-R799	R700-R799	R700-R799	R800-R899	R800-R899	R900-R999
9	R800-R899	R800-R899	R800-R899	R900-R999	R900-R999	R1000-R1099
10	R900-R999	R900-R999	R900-R999	R1000-R1099	R1000-R1099	R1100-R1199
11	R1000-R1099	R1000-R1099	R1000-R1099	R1100-R1199	R1100-R1199	R1200-R1399
12	R1100-R1199	R1100-R1199	R1100-R1199	R1200-R1399	R1200-R1399	R1400-R1599
13	R1200-R1399	R1200-R1399	R1200-R1399	R1400-R1599	R1400-R1599	R1600-R1999
14	R1400-R1599	R1400-R1599	R1400-R1599	R1600-R1999	R1600-R1999	R2000-R2499
15	R1600-R1999	R1600-R1999	R1600-R1999	R2000-R2499	R2000-R2499	R2500-R2999
16	R2000-R2499	R2000-R2499	R2000-R2499	R2500-R2999	R2500-R2999	R3000-R3999
17	R2500-R2999	R2500-R2999	R2500-R2999	R3000-R3999	R3000-R3999	R4000-R4999
18	R3000-R3999	R3000-R3999	R3000-R3999	R4000-R4999	R4000-R4999	R5000-R5999
19	R4000-R4999	R4000-R4999	R4000-R4999	R5000-R5999	R5000-R5999	R6000-R6999
20	R5000-R5999	R5000-R5999	R5000-R5999	R6000-R6999	R6000-R6999	R7000-R7999
21	R6000-R6999	R6000-R6999	R6000-R6999	R7000-R7999	R7000-R7999	R8000-R8999
22	R7000-R7999	R7000-R7999	R7000-R7999	R8000-R8999	R8000-R8999	R9000-R9999
23	R8000-R8999	R8000-R8999	R8000-R8999	R9000-R9999	R9000-R9999	R10000-R10999
24	R9000-R9999	R9000-R9999	R9000-R9999	R10000-R10999	R10000-R10999	R11000-R11999
25	R10000-R10999	R10000-R10999	R10000-R10999	R11000-R11999	R11000-R11999	R12000-R13999
26	R11000-R11999	R11000-R11999	R11000-R11999	R12000-R13999	R12000-R13999	R14000-R15999
27	R12000-R12999	R12000-R13999	R12000-R13999	R14000-R15999	R14000-R15999	R16000-R19999
28	R13000-R13999	R14000-R15999	R14000-R15999	R16000-R17999	R16000-R19999	R20000-R24999
29	R14000-	R16000-	R16000-R17999	R18000-R19999	R20000-R24999	R25000-R29999
30			R18000-	R20000-	R25000-R29999	R30000-R39999
31					R30000-R39999	R40000-
32					R40000-	

¹⁵ The question was asked as follows: "Please give me the letter which best describes the total monthly household income of all these people before tax and other deductions. Please include all sources of income, i.e., salaries, pensions, income from investments, etc."

In order to ensure that AMPS is a reliable data source, the data is externally validated using subscriber data from M-Net and new electricity connections made during the past twelve months from Eskom, after the weights are applied. Validation occurs by checking whether the AMPS figures fall within the 95% confidence intervals generated by these data sources. In addition, there is internal validation using historical sales data for consumer durables with long lifespans and low duplication rates, such as microwaves and computers. The consumers indicate whether they have made purchases of these goods during the past twelve months.

2.8 Conclusion

The various available data sources which could be used for poverty and inequality analyses were looked at in this section. Both the income and expenditure data are available in the IES, NIDS and PSLSD, and the data were captured in exact amounts. On the other hand, Census and AMPS did not capture expenditure data, but only captured income data in bands. However, it is possible to derive the household income amount in the censuses, and this amount could differ amongst the households from the same income category. Furthermore, some of the OHSs and LFSs contain information on household expenditure. The respondents were asked to declare the exact amounts in OHS 1996-1998, but were asked to declare the relevant expenditure category in OHS 1999 and in the LFSs. In addition, OHS 1999 captured information on both household income and expenditure. Finally, all GHSs only captured household expenditure in bands, and the bands were exactly the same as those asked in the OHS 1999 and the LFSs in nominal terms. Table 13 below summarizes the derivation of household income and expenditure in each survey.

Table 13 Summary of the derivation of household income and expenditure in each survey

Survey	Year	Income		Expenditure	
		Question asked?	Data captured in bands or exact amounts?	Question asked?	Data captured in bands or exact amounts?
Census	1996 2001 2007 (CS)	Yes	Bands (but it is possible to derive the amount)	No	N/A
IES	1995 2000 2005-2006	Yes	Exact amount was derived by adding the income amounts from various sources	Yes	Exact amount was derived by adding the consumption amounts on various items Note: This amount was consumption, but not expenditure
OHS	1995 – 1999	Yes (Only in 1999)	1999: Bands	Yes (In 4 surveys)	1996, 1997 and 1998: Exact amounts 1999: Bands
LFS	2000 – 2007	No	N/A	Yes (In 4 surveys)	Bands
QLFS	2008 – 2009	No	N/A	No	N/A
GHS	2002 – 2008	No	N/A	Yes	Bands
PSLSD	1993	Yes	Exact amount was derived by adding the income amounts from various sources	Yes	Exact amount was derived by adding the expenditure amounts from various sources
NIDS	2008	Yes	Exact amount was derived by adding the income amounts from various sources	Yes	Exact amount was derived by adding the expenditure amounts from various sources
AMPS	1993 – 2008	Yes	Bands	No	N/A

3. Derivation of the real per capita income, expenditure and consumption variables

In this section, the per capita income, expenditure and consumption variables are derived in 2000 prices, for the poverty and inequality analyzes in Section 5.

As far as the derivation of the per capita variables in the IESs is concerned, if the Standard Trade Classification approach was adopted in all three surveys, both the total household income and expenditure amounts were divided by household size to derive the per capita income and per capita expenditure variables respectively. On the other hand, if the COICOP approach was adopted in all the surveys, the resultant total household income and total consumption amounts were divided by household size, before the per capita income and per capita consumption variables were derived respectively. Note that in IES 2005/2006, per capita income and per capita consumption with and without the imputed rent were derived.

Table 14 Monthly CPIs (Data code in the Quarterly Bulletin of Reserve Bank: 7032) used to convert all nominal values into 2000 prices

Survey	Year	Month	CPI	Survey	Year	Month	CPI	Survey	Year	Month	CPI
Census	1996	Oct	79.78	NIDS	2008	Jan	151.20	AMPS	1993	Jan-Dec	61.16
Census	2001	Oct	106.05	NIDS	2008	Feb	152.06	AMPS	1994	Jan-Dec	66.63
Census	2007	Feb	138.45	NIDS	2008	Mar	153.86	AMPS	1995	Jan-Dec	72.41
IES	1995	Oct	73.20	NIDS	2008	Apr	156.01	AMPS	1996	Jan-Dec	77.73
IES	2000	Oct	101.96	NIDS	2008	May	158.36	AMPS	1997	Jan-Dec	84.42
IES	2005	Sep	129.25	NIDS	2008	Jun	160.32	AMPS	1998	Jan-Dec	90.22
IES	2005	Oct	129.64	NIDS	2008	Jul	163.10	AMPS	1999	Jan-Dec	94.90
IES	2005	Nov	129.88	NIDS	2008	Aug	164.30	AMPS	2000	Jan-Dec	99.97
IES	2005	Dec	130.32	NIDS	2008	Sep	164.80	AMPS	2001	Jul-Dec	106.49
IES	2006	Jan	130.56	NIDS	2008	Oct	165.00	AMPS	2002	Jul-Dec	118.86
IES	2006	Feb	130.91	NIDS	2008	Nov	165.80	AMPS	2003	Jul-Dec	122.04
IES	2006	Mar	131.18	NIDS	2008	Dec	165.80	AMPS	2004	Jan-Dec	123.80
IES	2006	Apr	131.40	NIDS	2008	Missing	160.05	AMPS	2005	Jan-Jun	126.61
IES	2006	May	132.30	GHS	2002	Jul	115.89	AMPS	2006	Jan-Jun	131.56
IES	2006	Jun	133.52	GHS	2003	Jul	121.96	AMPS	2007	Jul-Dec	146.77
IES	2006	Jul	134.44	GHS	2004	Jul	123.89	AMPS	2008	Jan-Jun	155.32
IES	2006	Aug	135.70	GHS	2005	Jul	128.06	LFS	2001	Sep	106.00
OHS	1996	Oct	79.78	GHS	2006	Jul	134.44	LFS	2002	Sep	117.94
OHS	1997	Oct	85.87	GHS	2007	Jul	143.83	LFS	2003	Sep	122.18
OHS	1998	Oct	93.57	GHS	2008	Jul	163.10	LFS	2004	Sep	123.86
OHS	1999	Oct	95.27	PSLSD	1993	Jul	61.73				

Note: New weights (using the results from IES 2005-2006) were adopted for the derivation of the CPI since November 2008 (Data code in the Quarterly Bulletin of Reserve Bank: 7170), and using the new weights, the CPI was derived only from 2002. The monthly CPI values in the last three months of 2008 amounted to 102.6, 103.1 and 103.1 respectively, and thus the inflation between October and November is 0.49% and the inflation between November and December is 0%. Hence, the CPI in November and December 2008 using the old weights (i.e., code: 7032) could be approximated as 165.80 [$165.00 \times (1 + 4.87\%)$] and 165.80 respectively [$165.00 \times (1 + 0\%)$].

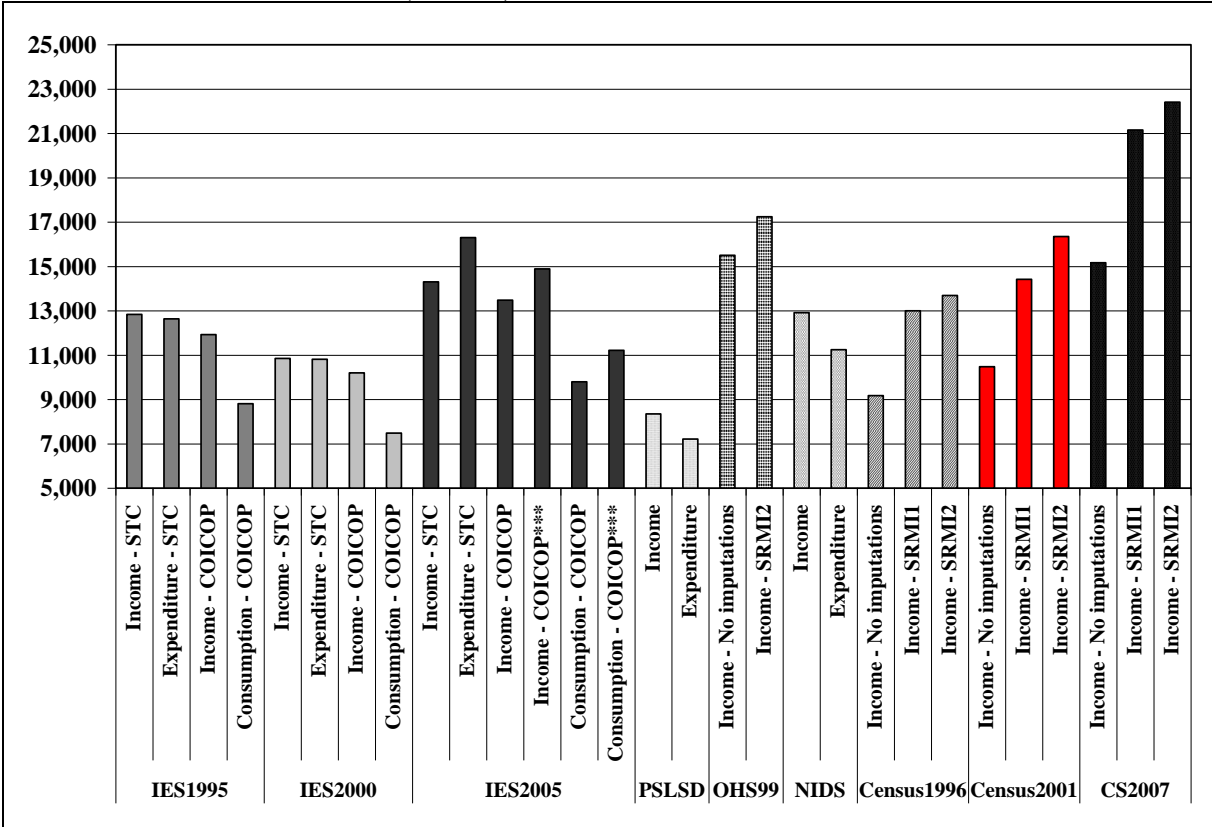
Note: 7 households (out of 7305) in NIDS had missing interview month, and the average of the twelve monthly CPIs in 2008 (i.e., 160.06) was used to convert the nominal amounts into 2000 prices in these households.

Note: As the AMPS metadata did not specify the exact survey month, if the AMPS took place twice a year (i.e., 1993 – 2000 and 2004), the annual CPI was used to deflate the nominal per capita income. However, if only one AMPS took place in the first half of the year (i.e., 2005, 2006 and 2008), the average of the January-June monthly CPI values was used to deflate the nominal per capita income. Similarly, the average of the July-December CPI values was used if only one AMPS took place in the second half of the year (i.e., 2001 – 2003 and 2007).

In PSLSD and NIDS, per capita income was simply derived as total household income amount divided by household size, while per capita expenditure was equal to total expenditure amount divided by household size. Per capita expenditure in OHS 1996-1998 as well as per capita income in the censuses and CS 2007 was also derived in the same way. Looking at OHS 1999, LFS 2001-2004 and GHS 2002-2008 (total expenditure being a categorical variable), the total expenditure amount was approximated as the mid-point of the class interval of each category, while the Pareto method was used to derive the mid-point of the open category (R10000 or more). Total expenditure amount was then divided by the household size to derive the per capita expenditure amount. A similar approach was adopted when the per capita income amount was derived in AMPS and OHS 1999 (household income being a categorical variable).

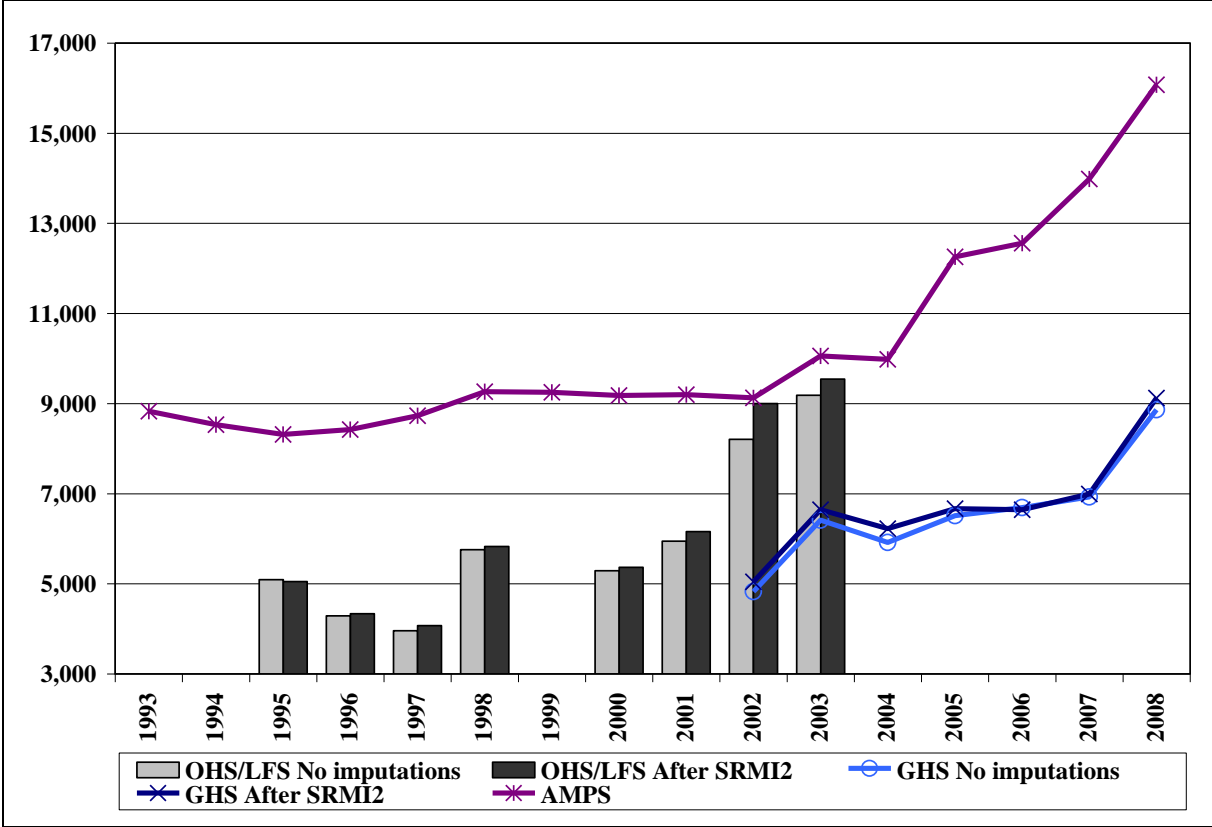
Finally, all nominal amounts were converted into real per capita income 2000 prices, using the South African Reserve Bank’s monthly CPI series (Data code in the Quarterly Bulletin of Reserve Bank: 7032). The CPI values used in each survey are shown in Table 14 above. Furthermore, Figures 1 and 2 show the mean per capita income/expenditure/consumption (2000 prices, per annum) in each survey, and it can be seen that the mean values increased after imputations were applied on the censuses/CS 2007/GHSs/OHSs/LFSs. In addition, the AMPS per capita values are higher than the OHS/LFS/GHS values of the same year.

Figure 1 Mean annual per capita income/expenditure/consumption (2000 prices) in the IESs, PSLSD, OHS 1999 (Income), NIDS, censuses and CS 2007



*** The imputed rent was included as an income or consumption item.

Figure 2 Mean annual per capita expenditure (2000 prices) in the OHSs, LFSs, GHSs as well as mean per capita income (2000 prices) in AMPSS



4. Comparison with national accounts income data

In this section, the total income, expenditure or consumption amounts derived from different surveys in different years are compared with the national accounts income data, so as to see if the surveys under-estimated income/expenditure/consumption seriously. Of course it is possible that the national accounts income figures are incorrect, but this argument will not be discussed further in this paper. The main aim is to see if the under-estimation (if any) of total income/expenditure/consumption in some surveys could affect the poverty results, which will be looked at in Section 5.

Table 15 below shows the derivation of the total income in national accounts, and it can be seen that as a result of the change in the categorization of items in the Quarterly Bulletin of Reserve Bank since 2006, the formula to calculate total income has changed¹⁶. On the other hand, Figure 3 shows that total income (in 2000 prices) almost doubled between 1992 and 2008.

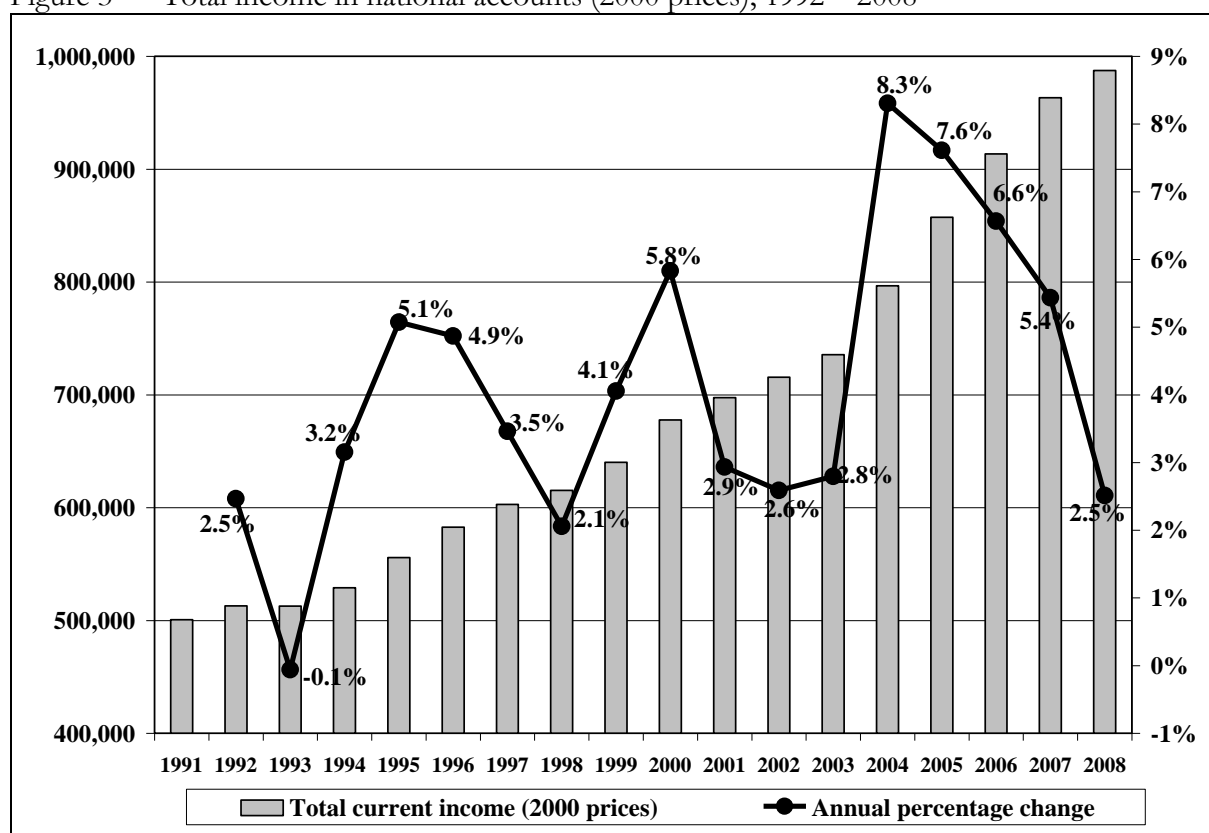
Furthermore, Table 16 as well as Figures 4 and 5 show the total income, expenditure or consumption in each survey as percentage of the national accounts total income in the same year. First, looking at the two censuses and CS 2007, it can be seen that CS 2007 is the survey that captured total income the best. However, all three surveys captured income better after SRMI was applied.

¹⁶ In the national accounts section of the Bulletin’s statistical tables before 2006, the current income of the households was clearly shown as the sum of the five income components as shown in the second to fourth rows of Table 15. However, since 2006, the categorization of the items has changed drastically, and the current income could only be approximated using the items shown in the last four rows of Table 15. These items are under the statistical table “Production, distribution and accumulation accounts of South Africa – Households and non-profit institutions serving households” in the Bulletin.

Table 15 Derivation of the total income in national accounts (the code of each item in the South African Reserve Bank Quarterly Bulletin in brackets)

Old method: Until 2005	
Remuneration	Compensation of employees (6240)
Transfers	Transfers from general government (6257)
Residuals	Property income (6241) + Current transfer from enterprise (6231) + Transfer from rest of the world (6243)
New method: Since 2006	
Remuneration	Compensation of employees (6240)
Transfers	Gross operating surplus/mixed income (6826) + Property income received (6827) - Property income paid (6832) - Consumption of fixed capital (6849)
Residuals	Social benefits received (6836) + Other current transfers received (6837) - Social contributions paid (6840) - Other current transfers paid (6841) + Adjustment for change in net equity of households in pension funds reserves (6845)

Figure 3 Total income in national accounts (2000 prices), 1992 – 2008



Note: The annual percentage change of total income between 1991 and 1992 was 2.5%, while it was -0.1% between 1992 and 1993, and so forth.

Looking at the IESs, IES 1995 was the survey that captured total income and expenditure well. Under the STC approach, these amounts were equal to nearly 95% of the national accounts income amount. One surprising finding from Table 16 is that total income / expenditure / consumption experienced a decline between 1995 and 2000, which contradicted the upward trend in the national accounts total income as seen in Figure 3. This will be discussed in greater detail in Section 5.

Table 16 Comparison of annual total income/expenditure/consumption in various surveys with annual total income in the national accounts in the same year

Survey	Variable	Year	Amount (R million) (2000 prices)	As % of total income in the national accounts
Census/CS	Total income – without any imputations involved	1996	294 475	50.5%
		2001	366 341	52.5%
		2007	629 421	68.9%
	Total income – After SRMI1	1996	339 993	58.3%
		2001	470 360	67.4%
		2007	776 476	85.0%
	Total income – After SRMI2	1996	350 345	60.1%
		2001	506 896	72.7%
		2007	782 283	85.6%
IES	Total income – Standard Trade Classification	1995	527 850	95.0%
		2000	460 572	71.9%
		2005/2006	659 229	72.2%
	Total expenditure – Standard Trade Classification	1995	519 549	93.5%
		2000	458 867	71.7%
		2005/2006	751 153	82.2%
	Total income - COICOP	1995	495 411	89.2%
		2000	441 795	69.0%
		2005/2006	638 786	69.9%
		2005/2006*	705 713	77.3%
	Total consumption - COICOP	1995	365 935	65.9%
		2000	324 026	50.6%
2005/2006		464 459	50.8%	
2005/2006*		531 386	58.2%	
OHS	Total expenditure – No imputations	1996	191 607	32.9%
		1997	172 608	28.6%
		1998	151 399	24.6%
		1999	229 693	35.9%
	Total income – No imputations	1999	607 350	94.9%
	Total expenditure – After SRMI2	1996	197 416	33.9%
		1997	183 153	30.4%
		1998	161 717	26.3%
		1999	252 422	39.4%
	Total income – After SRMI2	1999	746 173	116.5%
LFS	Total expenditure – No imputations	2001	230 514	33.1%
		2002	264 065	36.9%
		2003	370 790	50.4%
		2004	417 062	52.4%
	Total expenditure – After SRMI2	2001	241 690	34.7%
		2002	280 567	39.2%
		2003	414 435	56.3%
		2004	443 144	55.6%
GHS	Total expenditure – No imputations	2002	212 412	29.7%
		2003	287 893	39.1%
		2004	267 470	33.6%
		2005	299 400	34.9%
		2006	312 736	34.2%
		2007	326 385	33.9%
		2008	461 528	46.7%

Table 16 Continued

Survey	Variable	Year	Amount (R million) (2000 prices)	As % of total income in the national accounts
GHS	Total expenditure – After SRMI2	2002	229 177	32.0%
		2003	308 977	42.0%
		2004	289 165	36.3%
		2005	312 468	36.5%
		2006	314 442	34.4%
		2007	334 237	34.7%
		2008	486 045	49.2%
PSLSD	Total income	1993	334 531	65.3%
	Total expenditure	1993	297 679	58.1%
NIDS	Total income	2008	629 044	63.7%
	Total expenditure	2008	547 759	55.5%
AMPS	Total income	1993	336 394	65.6%
		1994	330 381	62.5%
		1995	333 057	59.9%
		1996	349 167	59.9%
		1997	347 982	57.7%
		1998	361 044	58.7%
		1999	360 573	56.3%
		2000	404 993	59.8%
		2001	406 077	58.2%
		2002	403 762	56.4%
		2003	444 193	60.4%
		2004	450 696	56.6%
		2005**	485 001	56.6%
		2006**	516 843	56.6%
		2007**	544 935	56.6%
2008**	558 620	56.6%		

* Including the imputed rent variable

** Originally, the AMPS income showed a rapid 22.4% increase between 2004 and 2005 (the 2005 figure being R551433 million. The growth rate in total income also showed a very rapid increase in 2007 and 2008 (9.9% and 13.9%), and these growth rates are much higher than the growth rates of the national accounts total income during the same period. Thus, it was decided to adjust the 2005 – 2008 AMPS total income in line with the national accounts income growth rates. In other words:

- “Adjusted” 2005 AMPS income = $450696 \times (1 + 7.61\%) = 485001$
- “Adjusted” 2006 AMPS income = $485001 \times (1 + 6.57\%) = 516843$
- “Adjusted” 2007 AMPS income = $516843 \times (1 + 5.44\%) = 544935$
- “Adjusted” 2008 AMPS income = $544935 \times (1 + 2.51\%) = 558620$

Thus, from 2005, the AMPS income was adjusted to remain a constant share of the national accounts income.

It seems the total expenditure was seriously under-captured in the OHSs, LFSs and GHSs, as the total expenditure was only about 30%-50% of the national income, and such proportion only increased slightly even after SRMI2 was applied. This could be due to the fact that the respondents tended to under-estimate their expenditure if asked to declare the “one-shot” amount, and that there were too few expenditure categories for them to choose from (in OHS 1999 as well as LFSs and GHSs). However, an interesting finding is that OHS 1999 captured income very well (See Tables 7 and 9), as the OHS 1999 total household income is equal to 94.9% of the 1999 national income without imputation and 116.5% (i.e., exceeding the national income of 1999) after SRMI2.

As far as the surveys conducted by institutions other than Stats SA are concerned, in both PSLSD and NIDS, total expenditure was under-captured more seriously than total income, when compared with national accounts’ total income. Finally, in almost all AMPSs, total income was equal to approximately 60% of national income.

Figure 4 Total income, consumption or expenditure as percentage of national accounts' total income in the IESs, PSLSD, OHS 1999 (Income), NIDS, censuses and CS 2007

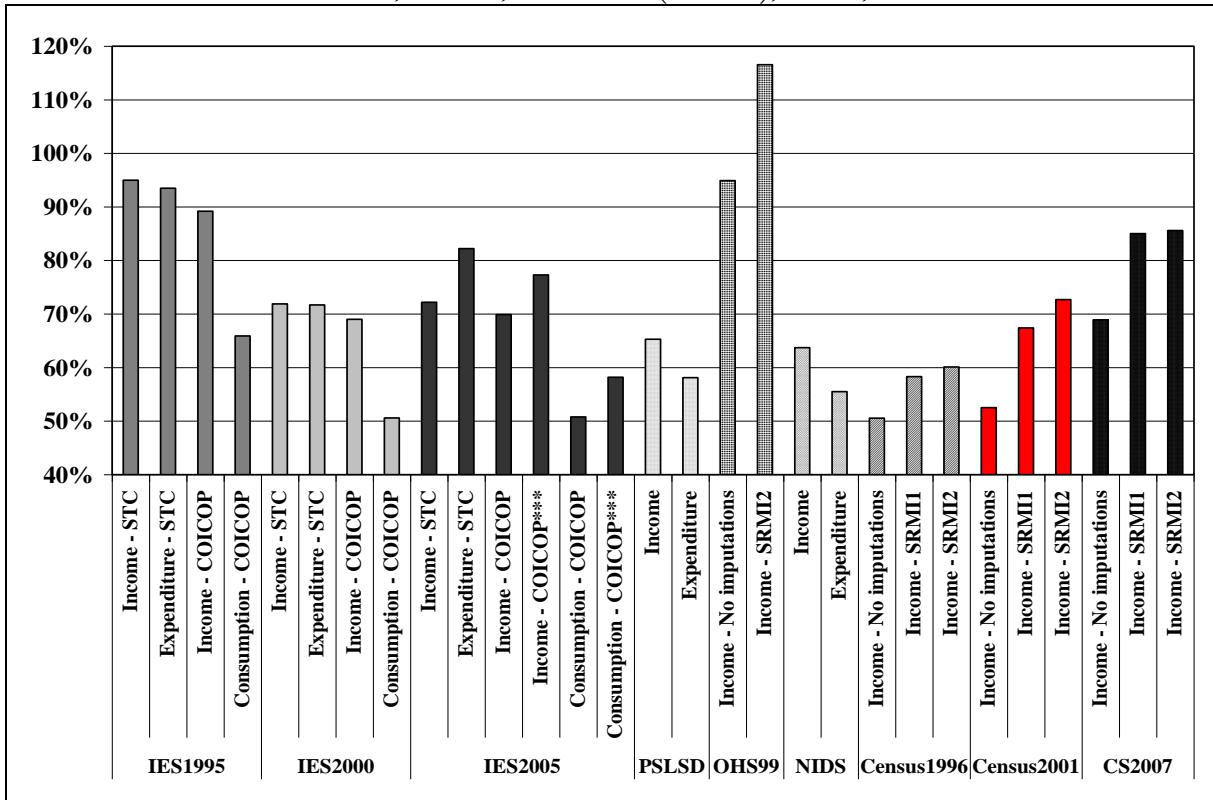
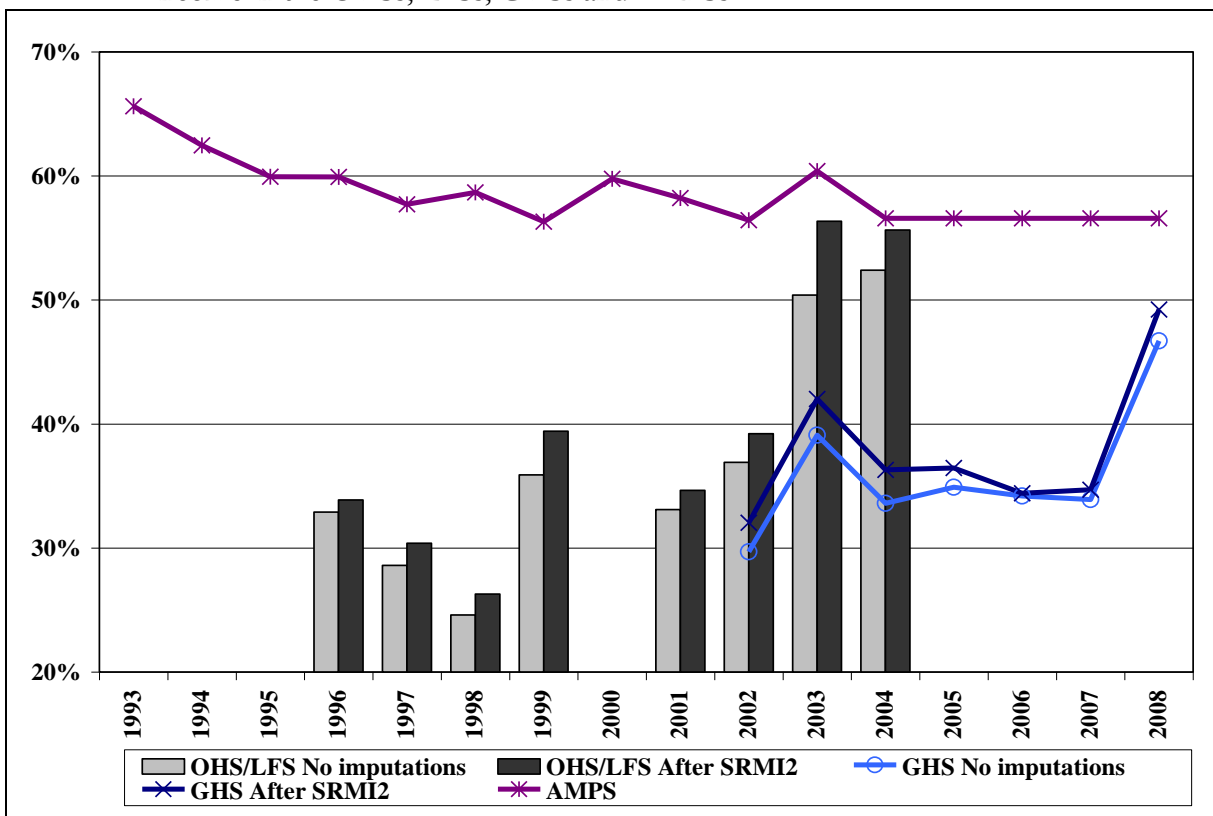


Figure 5 Total income, consumption or expenditure as percentage of national accounts' total income in the OHSs, LFSs, GHSs and AMPSS



To conclude, when compared with the national accounts, IES 1995, CS 2007 and OHS 1999 are the three surveys that captured total income (and expenditure in IES 1995) very well, while serious under-estimation of total expenditure occurred in all the OHSs, LFSs and GHSs.

5. Poverty and inequality trends

In this section, the per capita income, consumption or expenditure variables derived in Section 3 will be used for poverty and inequality analyses. The three official poverty lines in 2000 prices as proposed by Woolard and Leibbrandt (2006) will be used¹⁷. In addition, the analysis takes place at person level, i.e., the product of household weight and household size is the weight variable.

The sampling methodology and the derivation of weights could differ amongst these surveys, which should be taken into account when one compares the poverty and inequality trends in different surveys. In addition, the income or expenditure distribution could also influence the poverty and inequality results (e.g., Kernel density and cumulative density functions might need to be plotted, in order to analyze the poverty and inequality trends of each survey in greater detail). However, these issues require a paper of its own and are beyond the scope of this paper. In this paper, the poverty and inequality results will be presented as they are.

5.1 Poverty trends

Table 17 and Figure 6 below present the results on the poverty headcount ratios. The focus of the discussion here is on the results using the lower bound poverty line of R322 per month (or R3864 per annum)¹⁸.

First, looking at the poverty trends using the two censuses and CS 2007, if no imputations were involved (i.e., nothing was done on households with zero or unspecified household income) it can be seen that poverty increased between 1996 and 2001, before a rapid decline took place between 2001 and 2007. In addition, the 2007 poverty headcount ratio was lower than the 1996 ratio. Furthermore, poverty headcount ratios decreased in all three surveys after SRMI1, and such decrease was greater when SRMI2 was applied. However, the trends discussed above (i.e., upward trend between the two censuses, before a rapid downward trend took place between 2001 and 2007, with the 2007 poverty headcount ratio smaller than the 1996 result) remained the same, after SRMI1 or SRMI2 was conducted. It was argued by Yu (2009: 46) that the extent of poverty increase between the two censuses could be under-estimated because Census 1996 under-captured income more seriously than Census 2001 did, while the extent of the decline of poverty between 2001 and 2007 could be over-estimated because CS 2007 captured income much better (See Section 4).

With regard to the poverty trends using IESs, there was a rapid increase of poverty headcount ratios between IES 1995 and IES 2000, before a downward trend was observed between IES2000 and IES 2005-2006. This trend took place, regardless of whether the STC or COCIOP approach was adopted. However, the IES 2005-2006 poverty headcount ratio was still slightly above the IES 1995 ratio, even after imputed rent was included as an income/consumption/expenditure item. It was argued (Van der Berg et al., 2007: 14) that the extent of increase of poverty could be over-estimated, since there was a large drop of recorded income (or expenditure) between IES 1995 and IES 2000 (while the national accounts income data showed that national income increased between the two years). Such a drop in income between the two surveys was unlikely, as it was larger than the decline experienced by the South African economy during the Great Depression of the 1930s. In addition, such decrease was also larger than the decline experienced by some of the affected countries during the 1998 Asian economic crisis. Thus, it seems certain issues (e.g., differences in sampling methodology) made

¹⁷ The three poverty lines (2000 prices) proposed by these authors are as follows: R211 per month or R2532 per annum (expenditure on food items), R322 per month or R3864 per annum (expenditure on food items and essential non-food items), R593 per month or R7116 per annum (expenditure on food items, essential non-food items and non-essential non-food items). These three amounts were also defined as food poverty line, lower bound poverty line, and upper bound poverty line respectively.

¹⁸ The poverty headcount ratios by race using this poverty line are presented in Table A.1 of the Appendix.

the comparability of the two surveys difficult, and the poverty (and the inequality) results between the two surveys should be interpreted with caution. Finally, since income (or expenditure) was very poorly captured in IES 2000, while IES 2005-2006 was the survey that captured income best, the extent of the decline of poverty between these two surveys could be over-estimated.

Table 17 Poverty headcounts at different poverty lines and Gini coefficients, using the per capita variables

Survey	Variable	Year	Poverty headcount			Gini coefficient
			R2532	R3864	R7116	
Census/ CS	Total income – without any imputations involved	1996	0.501	0.606	0.728	0.742
		2001	0.568	0.670	0.789	0.825
		2007	0.397	0.529	0.694	0.774
	Total income – After SRMI1	1996	0.493	0.601	0.726	0.734
		2001	0.547	0.647	0.769	0.817
		2007	0.351	0.478	0.656	0.759
	Total income – After SRMI2	1996	0.441	0.576	0.715	0.694
		2001	0.447	0.592	0.750	0.756
		2007	0.329	0.463	0.650	0.743
IES	Total income – Standard Trade Classification	1995	0.286	0.434	0.622	0.655
		2000	0.429	0.559	0.710	0.711
		2005/2006	0.338	0.488	0.657	0.717
	Total expenditure – Standard Trade Classification	1995	0.300	0.447	0.629	0.660
		2000	0.430	0.564	0.714	0.710
		2005/2006	0.303	0.466	0.654	0.733
	Total income - COICOP	1995	0.319	0.462	0.642	0.660
		2000	0.442	0.572	0.723	0.709
		2005/2006	0.353	0.504	0.667	0.714
		2005/2006*	0.316	0.473	0.652	0.716
	Total consumption - COICOP	1995	0.339	0.502	0.691	0.612
		2000	0.458	0.601	0.753	0.651
2005/2006		0.359	0.531	0.720	0.659	
2005/2006*		0.320	0.500	0.699	0.670	
OHS	Total expenditure – No imputations	1996	0.588	0.702	0.815	0.646
		1997	0.665	0.768	0.875	0.663
		1998	0.667	0.781	0.871	0.662
		1999	0.652	0.742	0.838	0.713
	Total income – No imputations	1999	0.518	0.617	0.745	0.815
	Total expenditure – After SRMI2	1996	0.588	0.702	0.815	0.636
		1997	0.665	0.768	0.875	0.660
		1998	0.667	0.781	0.871	0.659
		1999	0.652	0.742	0.838	0.702
Total income – After SRMI2	1999	0.518	0.617	0.745	0.815	
LFS	Total expenditure – No imputations	2001	0.693	0.773	0.859	0.745
		2002	0.684	0.788	0.853	0.781
		2003	0.678	0.758	0.838	0.813
		2004	0.649	0.738	0.827	0.815
	Total expenditure – After SRMI2	2001	0.693	0.773	0.859	0.739
		2002	0.684	0.788	0.853	0.779
		2003	0.678	0.758	0.838	0.821
		2004	0.649	0.738	0.827	0.815

Table 17 Continued

Survey	Variable	Year	Poverty headcount			Gini coefficient
			R2532	R3864	R7116	
GHS	Total expenditure – No imputations	2002	0.689	0.778	0.861	0.736
		2003	0.681	0.762	0.845	0.772
		2004	0.637	0.733	0.823	0.720
		2005	0.618	0.710	0.840	0.737
		2006	0.619	0.731	0.842	0.753
		2007	0.614	0.695	0.822	0.735
		2008	0.618	0.712	0.829	0.787
	Total expenditure – After SRMI2	2002	0.677	0.768	0.854	0.736
		2003	0.668	0.751	0.837	0.771
		2004	0.627	0.723	0.815	0.723
		2005	0.612	0.705	0.836	0.738
		2006	0.615	0.728	0.839	0.748
		2007	0.611	0.692	0.820	0.735
		2008	0.610	0.706	0.824	0.787
PSLSD	Total income	1993	0.475	0.598	0.745	0.696
	Total expenditure	1993	0.398	0.566	0.750	0.595
NIDS	Total income	2008	0.292	0.459	0.646	0.683
	Total expenditure	2008	0.381	0.528	0.679	0.685
AMPS	Total income	1993	0.438	0.586	0.737	0.672
		1994	0.439	0.593	0.735	0.665
		1995	0.464	0.594	0.741	0.674
		1996	0.473	0.610	0.744	0.678
		1997	0.456	0.589	0.732	0.674
		1998	0.453	0.583	0.725	0.683
		1999	0.469	0.591	0.723	0.685
		2000	0.458	0.582	0.723	0.682
		2001	0.466	0.579	0.717	0.685
		2002	0.434	0.563	0.709	0.670
		2003	0.418	0.554	0.704	0.686
		2004	0.415	0.548	0.703	0.678
		2005	0.391	0.519	0.680	0.683
		2006	0.375	0.504	0.667	0.685
		2007	0.335	0.459	0.617	0.660
2008	0.325	0.443	0.612	0.666		

* Including the imputed rent variable

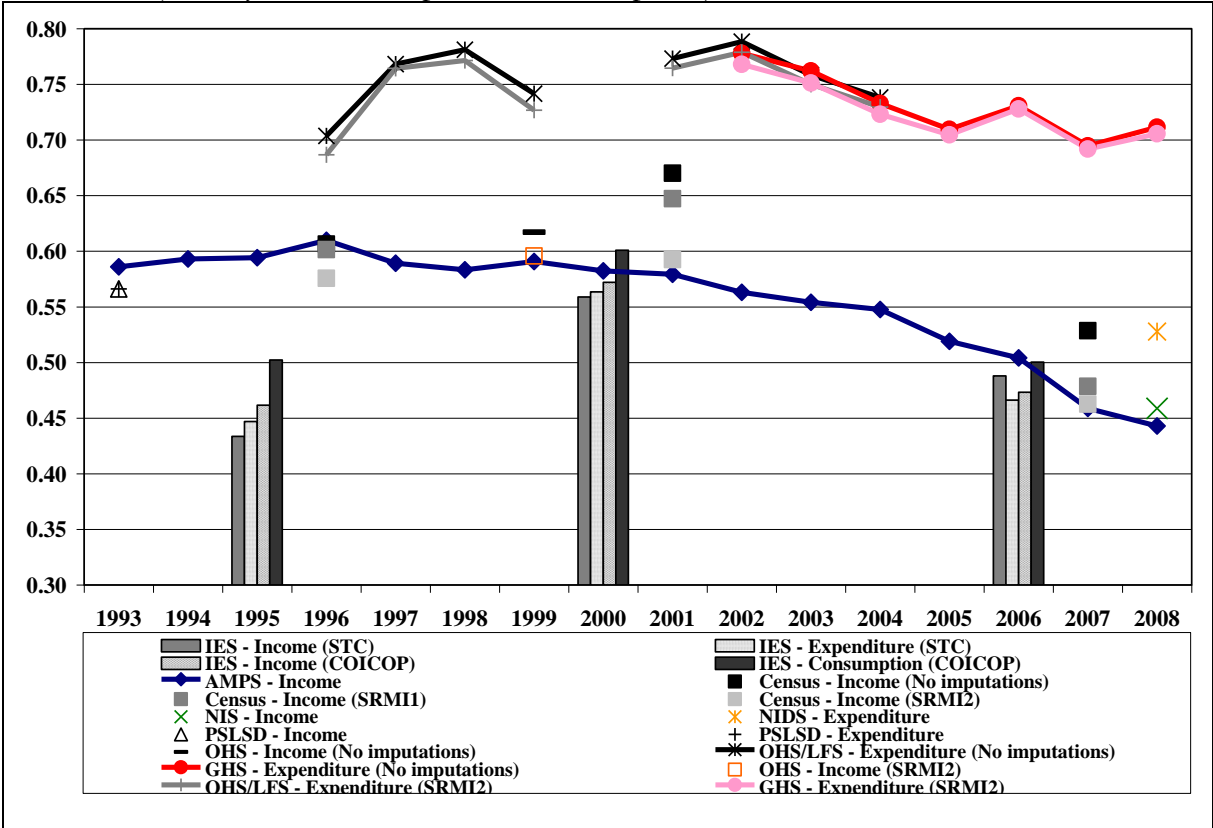
Looking at the OHS and LFS per capita expenditure variable, it can be seen that poverty increased since 1996, until a downward trend was observed from 2002. However, the 2004 poverty headcount ratio was slightly higher than the 1996 ratio. On the other hand, in the GHSs, a continuous downward trend in poverty was observed between 2002 and 2005, before it became stagnant between 2005 and 2008. In addition, the LFS 2002 – 2004 poverty headcount ratios were extremely close to the GHS 2002 – 2004 results (i.e., the overlapping of the black and red lines in Figure 4). Note that due to the serious under-estimation of total expenditure compared with national accounts as discussed in Section 4, the poverty headcount ratios in OHSs, LFSs and GHSs were much higher (always above 0.7) than the results using censuses and IESs (and also AMPS, NIDS and PSLSD, to be discussed below).

Furthermore, after SRMI2 was applied in the OHS/LFS/GHS data, the same poverty headcount trends as discussed above are still observed. In addition, poverty headcount only shows a negligible decrease in the OHSs and LFSs (i.e., the black line vs. the gray line in Figure 6), while the similar negligible decrease took place across all GHSs (i.e., the red line vs. the pink line in Figure 6).

In AMPSs, it can be seen from Figure 6 that there was not too much change in the poverty headcount ratio between 1993 and 1999, but a continuous downward trend took place since 2000. In fact, the poverty headcount ratio dropped to just below 0.45 in 2008. In addition, the AMPS poverty headcount ratios in 1996, 2001 and 2007 were very close to the post-SRMI2 poverty headcount ratios in Census 1996, Census 2001 and CS 2007 respectively.

As only one PSLSD took place, and only one NIDS took place at the time of writing, it is impossible to derive poverty trends from these two surveys. However, the 1993 PSLSD poverty headcount ratios (regardless of whether the income or expenditure variable was used) were slightly below the 1993 AMPS result. On the other hand, due to under-estimation of expenditure in NIDS, the 2008 poverty headcount ratio was higher if per capita expenditure was used (0.528) while the ratio using per capita income (0.459) was quite close to the ratio in AMPS 2008 (0.443).

Figure 6 Poverty headcount, using per capita variables
(Poverty line: R3 864 per annum 2000 prices)



To conclude, in censuses, IESs, AMPSs and OHSs/LFSs, a similar poverty trend was observed – upward trend in the 1990s, before a downward trend took place in the 2000s. The 2002-2005 GHSs showed a downward poverty trend in general, followed by stagnation in 2005-2008.

5.2 Inequality trends

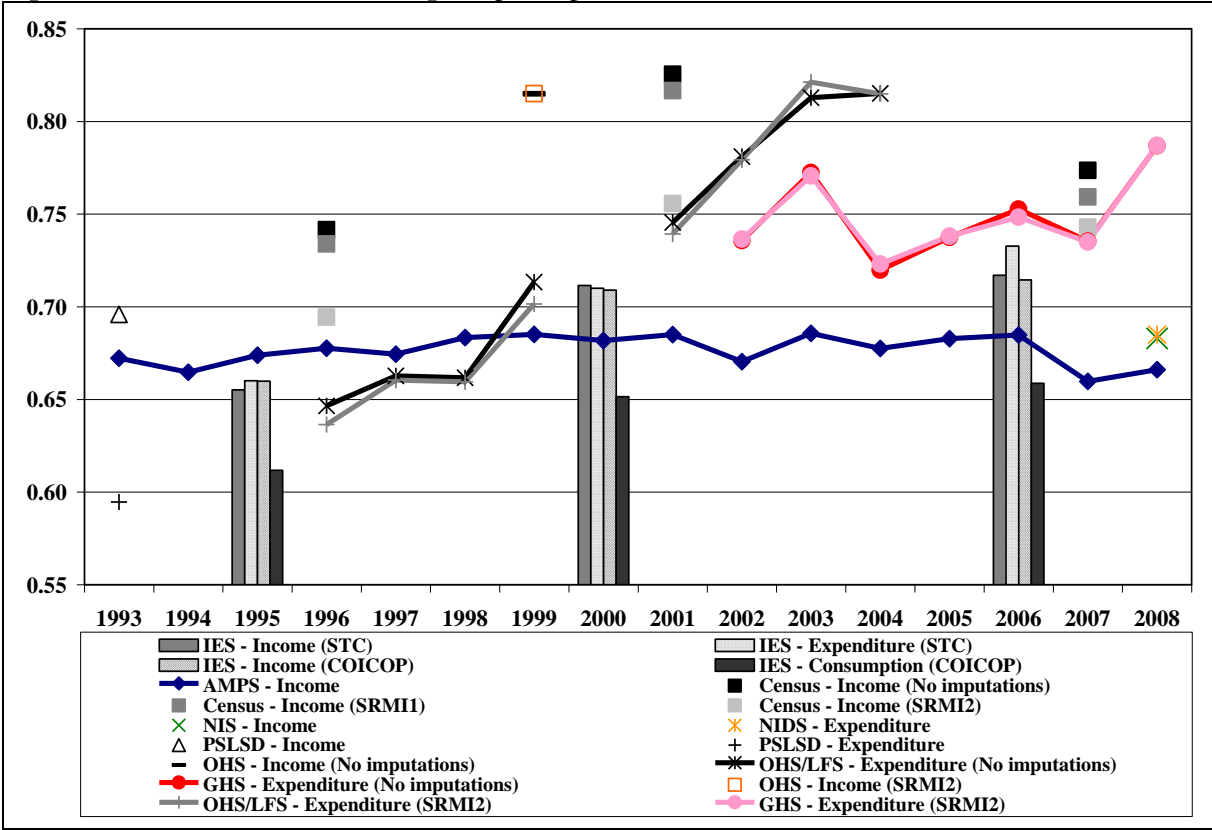
In this section, the Gini coefficients using the per capita variables will be analyzed, and the results are presented in the last column of Table 17 above as well as Figure 7 below¹⁹. In the censuses, the Gini coefficient increased between Census 1996 and Census 2001, before a decline took place in CS 2007. However, the 2007 value was slightly higher than the 1996 value. In addition, this trend remained the same, regardless whether imputation was conducted or not. Furthermore, the Gini coefficients declined in all three surveys after SRMI1, and decreased further after SRMI2.

¹⁹ The Gini coefficients by race of household head are presented in Table A.2 of the Appendix.

A different trend was observed when looking at the IESs, as there was a continuous worsening of the income inequality problem, regardless of whether the STC or COICOP approach was adopted to derive the per capita variables. In addition, the increase of Gini coefficient between IES 1995 and IES 2000 was more rapid, while there was only a slight increase in the coefficient between IES 2000 and IES 2005-2006. A similar trend was observed in the OHSs/LFSs (i.e., a continuous upward trend). In the GHSs, an upward trend was also observed in general (although the Gini coefficient declined between 2003 and 2004, and again between 2006 and 2007). In addition, there was no obvious change in the Gini coefficients as well as inequality trends, after SRMI2 was applied on the OHS/LFS/GHS data.

In AMPSs, the Gini coefficients stayed within the 0.66-0.69 range throughout the years under study. In the 1993 PSLSD, the Gini coefficient was much lower (0.595) if per capita expenditure was used but much higher (0.696) if per capita income was used. Finally, the Gini coefficient was similar (approximately 0.68) in the 2008 NIDS, regardless of which per capita variable was used. In addition, the 2008 NIDS Gini coefficient was quite close to the 2008 coefficient using the AMPS per capita income variable.

Figure 7 Gini coefficients, using the per capita variables



To summarize, a continuous worsening of the income inequality problem was observed in IESs, OHSs, LFSs and GHSs, while AMPS showed that the Gini coefficient stagnated between 1993 and 2008. Finally, using the two censuses and the CS 2007, Gini coefficient increased between 1996 and 2001, before it declined in 2007. Thus, the income inequality trends differed amongst the surveys.

6. Conclusion

This paper has applied the proposed official poverty lines to explore the poverty trends in various surveys since 1993. In general, an upward trend was observed in the 1990s, before a downward trend took place in the 2000s. In addition, the Gini coefficients were looked at to derive the inequality trends. While AMPS showed that the Gini coefficient stagnated between 1993 and 2008, a continuous worsening of the income inequality problem was observed in IESs, OHSs, LFSs and GHSs. Furthermore, income inequality worsened between the two censuses, before it declined in CS 2007.

As mentioned at the beginning of Section 5, the poverty and inequality trends were presented as they are (with imputations applied in some surveys). However, it must be kept in mind that the datasets used in this paper are, strictly speaking, cross sectional data and are not designed to be used as a time series, due to changes in the sampling design, sampling frame, questionnaire structure, the way the income and expenditure questions are phrased, and the experience of the interviewers. All these factors could influence the comparability of the poverty and inequality results across different surveys during the period under investigation.

References

Ardington, C., Lam, D., Leibbrandt, M. & Welch, M., 2005. *The sensitivity of estimates of post-apartheid changes in South African poverty and inequality to key data imputations*. CSSR working paper no. 106. Cape Town: Centre for Social Science Research.

Argent, J., Finn, A., Leibbrandt, M. & Woolard, I., 2009. *Poverty: Analysis of the NIDS wave 1 dataset*. Discussion Paper No. 13. Cape Town: Southern Africa Labour and Development Research Unit, University of Cape Town.

Van der Berg, S., Louw, M. & Du Toit, L., 2007. *Poverty trends since the transition: What we know*. Stellenbosch: Stellenbosch University.

Woolard, I. & Leibbrandt, M., 2006. *Towards a poverty Line for South Africa: Background note*. Cape Town: Southern Africa Labour and Development Research Unit, University of Cape Town.

Yu, D., 2007. *The comparability of the Statistics South Africa October Household Surveys and Labour Force Surveys*. Stellenbosch Economic Working Papers: 17/07. Stellenbosch: Stellenbosch University.

Yu, D., 2008. *The comparability of Income and Expenditure Surveys 1995, 2000 and 2005/2006*. Stellenbosch Economic Working Papers: 11/08. Stellenbosch: Stellenbosch University.

Yu, D., 2009. *The comparability of Census 1996, Census 2001 and Community Survey 2007*. Stellenbosch Economic Working Papers: 21/09. Stellenbosch: Stellenbosch University.

Appendix

Table A.1 Poverty headcounts by race, using the per capita variables
(Poverty line: R3864, 2000 prices)

Survey	Variable	Year	Black	Coloured	Indian	White
Census/ CS	Total income – without any imputations involved	1996	0.720	0.393	0.158	0.055
		2001	0.765	0.453	0.189	0.068
		2007	0.609	0.320	0.145	0.050
	Total income – After SRMI1	1996	0.718	0.385	0.151	0.062
		2001	0.755	0.436	0.177	0.072
		2007	0.562	0.282	0.126	0.059
	Total income – After SRMI2	1996	0.693	0.365	0.125	0.033
		2001	0.701	0.369	0.119	0.024
		2007	0.551	0.256	0.098	0.016
IES	Total income – Standard Trade Classification	1995	0.533	0.289	0.034	0.007
		2000	0.660	0.347	0.128	0.072
		2005/2006	0.576	0.289	0.141	0.019
	Total expenditure – Standard Trade Classification	1995	0.548	0.313	0.028	0.006
		2000	0.673	0.368	0.101	0.013
		2005/2006	0.545	0.346	0.094	0.005
	Total income - COICOP	1995	0.566	0.322	0.034	0.009
		2000	0.674	0.370	0.130	0.079
		2005/2006	0.594	0.306	0.148	0.019
		2005/2006*	0.559	0.275	0.136	0.015
	Total consumption - COICOP	1995	0.611	0.382	0.067	0.010
		2000	0.713	0.412	0.120	0.025
2005/2006		0.622	0.390	0.100	0.006	
2005/2006*		0.587	0.358	0.090	0.004	
OHS	Total expenditure – No imputations	1996	0.823	0.596	0.311	0.102
		1997	0.885	0.672	0.301	0.085
		1998	0.895	0.677	0.384	0.110
		1999	0.845	0.621	0.304	0.105
	Total income – No imputations	1999	0.720	0.408	0.166	0.033
	Total expenditure – After SRMI2	1996	0.741	0.556	0.524	0.406
		1997	0.886	0.670	0.290	0.079
		1998	0.894	0.666	0.367	0.099
		1999	0.843	0.603	0.274	0.088
Total income – After SRMI2	1999	0.712	0.382	0.146	0.026	
LFS	Total expenditure – No imputations	2001	0.879	0.695	0.253	0.096
		2002	0.888	0.720	0.306	0.096
		2003	0.865	0.649	0.217	0.055
		2004	0.843	0.609	0.210	0.061
	Total expenditure – After SRMI2	2001	0.877	0.685	0.240	0.088
		2002	0.887	0.714	0.300	0.086
		2003	0.864	0.648	0.214	0.050
		2004	0.842	0.600	0.222	0.056
GHS	Total expenditure – No Imputations	2002	0.875	0.689	0.324	0.097
		2003	0.857	0.655	0.319	0.079
		2004	0.837	0.595	0.225	0.060
		2005	0.817	0.542	0.224	0.055
		2006	0.833	0.567	0.371	0.081
		2007	0.796	0.541	0.284	0.060
	Total expenditure – After SRMI2	2002	0.874	0.682	0.305	0.088
		2003	0.855	0.645	0.316	0.070
		2004	0.835	0.587	0.212	0.054

Table A.1 Continued

Survey	Variable	Year	Black	Coloured	Indian	White
GHS	Total expenditure – After SRMI2	2005	0.816	0.537	0.222	0.052
		2006	0.833	0.564	0.368	0.078
		2007	0.795	0.538	0.281	0.057
		2008	0.815	0.536	0.192	0.081
PSLSD	Total income	1993	0.735	0.298	0.094	0.025
	Total expenditure	1993	0.699	0.343	0.026	0.005
NIDS	Total income	2008	0.548	0.273	0.107	0.012
	Total expenditure	2008	0.624	0.336	0.170	0.027
AMPS	Total income	1993	0.728	0.391	0.107	0.008
		1994	0.736	0.385	0.101	0.012
		1995	0.736	0.351	0.109	0.008
		1996	0.747	0.385	0.126	0.011
		1997	0.725	0.354	0.111	0.012
		1998	0.726	0.346	0.116	0.010
		1999	0.731	0.367	0.122	0.011
		2000	0.707	0.364	0.098	0.014
		2001	0.707	0.347	0.104	0.015
		2002	0.678	0.381	0.123	0.019
		2003	0.672	0.350	0.104	0.020
		2004	0.661	0.346	0.098	0.015
		2005	0.622	0.368	0.079	0.024
		2006	0.607	0.334	0.067	0.019
		2007	0.554	0.307	0.058	0.016
2008	0.537	0.270	0.073	0.010		

* Including the imputed rent variable

Table A.2 Gini coefficients by race, using the per capita variables

Survey	Variable	Year	Black	Coloured	Indian	White
Census/ CS	Total income – without any imputations involved	1996	0.698	0.557	0.510	0.480
		2001	0.781	0.659	0.628	0.620
		2007	0.700	0.653	0.657	0.603
	Total income – After SRMI1	1996	0.693	0.550	0.501	0.477
		2001	0.778	0.644	0.616	0.605
		2007	0.690	0.636	0.620	0.583
	Total income – After SRMI2	1996	0.620	0.528	0.481	0.459
		2001	0.654	0.601	0.582	0.566
		2007	0.663	0.615	0.608	0.559
IES	Total income – Standard Trade Classification	1995	0.564	0.488	0.472	0.438
		2000	0.630	0.555	0.519	0.510
		2005/2006	0.615	0.593	0.559	0.512
	Total expenditure – Standard Trade Classification	1995	0.569	0.499	0.463	0.434
		2000	0.610	0.548	0.486	0.484
		2005/2006	0.620	0.638	0.687	0.520
	Total income - COICOP	1995	0.571	0.493	0.469	0.442
		2000	0.628	0.553	0.513	0.511
		2005/2006	0.609	0.588	0.560	0.513
		2005/2006*	0.538	0.565	0.519	0.438
	Total consumption - COICOP	1995	0.527	0.452	0.453	0.436
		2000	0.550	0.477	0.434	0.433
		2005/2006	0.543	0.553	0.517	0.449
		2005/2006*	0.599	0.591	0.555	0.496

* Including the imputed rent variable

Table A.2 Continued

Survey	Variable	Year	Black	Coloured	Indian	White
OHS	Total expenditure – No imputations	1996	0.590	0.512	0.452	0.465
		1997	0.525	0.507	0.409	0.510
		1998	0.561	0.488	0.421	0.476
		1999	0.596	0.558	0.564	0.562
	Total income – No imputations	1999	0.702	0.694	0.702	0.590
	Total expenditure – After SRMI2	1996	0.628	0.518	0.627	0.611
		1997	0.521	0.503	0.406	0.496
		1998	0.557	0.479	0.411	0.461
		1999	0.584	0.537	0.527	0.534
Total income – After SRMI2	1999	0.692	0.688	0.683	0.571	
LFS	Total expenditure – No imputations	2001	0.621	0.609	0.531	0.575
		2002	0.673	0.652	0.647	0.610
		2003	0.667	0.641	0.650	0.641
		2004	0.686	0.651	0.622	0.649
	Total expenditure – After SRMI2	2001	0.616	0.598	0.517	0.555
		2002	0.669	0.638	0.644	0.597
		2003	0.670	0.645	0.659	0.643
		2004	0.683	0.646	0.617	0.648
GHS	Total expenditure – No imputations	2002	0.621	0.605	0.564	0.556
		2003	0.679	0.628	0.649	0.598
		2004	0.613	0.559	0.602	0.524
		2005	0.631	0.562	0.597	0.578
		2006	0.638	0.592	0.633	0.609
		2007	0.641	0.640	0.627	0.567
		2008	0.683	0.724	0.664	0.604
	Total expenditure – After SRMI2	2002	0.619	0.594	0.555	0.550
		2003	0.674	0.615	0.641	0.591
		2004	0.611	0.556	0.589	0.521
		2005	0.630	0.558	0.594	0.575
		2006	0.634	0.586	0.626	0.601
		2007	0.640	0.637	0.624	0.564
		2008	0.682	0.720	0.666	0.597
PSLSD	Total income	1993	0.559	0.441	0.479	0.451
	Total expenditure	1993	0.449	0.405	0.371	0.322
NIDS	Total income	2008	0.594	0.528	0.539	0.483
	Total expenditure	2008	0.593	0.564	0.521	0.457
AMPS	Total income	1993	0.513	0.493	0.415	0.390
		1994	0.510	0.468	0.393	0.393
		1995	0.538	0.469	0.426	0.387
		1996	0.532	0.479	0.413	0.392
		1997	0.543	0.479	0.433	0.391
		1998	0.547	0.471	0.436	0.406
		1999	0.556	0.486	0.465	0.405
		2000	0.582	0.502	0.458	0.421
		2001	0.585	0.517	0.472	0.420
		2002	0.581	0.523	0.468	0.435
		2003	0.577	0.517	0.506	0.477
		2004	0.570	0.516	0.506	0.457
		2005	0.592	0.537	0.474	0.510
		2006	0.594	0.531	0.481	0.520
		2007	0.592	0.498	0.486	0.497
2008	0.601	0.506	0.533	0.521		