# Poverty and inequality estimates of National Income Dynamics Study revisited

DEREK YU

Stellenbosch Economic Working Papers: 05/13
March 2013

KEYWORDS: POVERTY, INEQUALITY, NATIONAL INCOME DYNAMICS STUDY, HOUSEHOLD SURVEYS, MEASUREMENT, SOUTH AFRICA
JEL: I32

DEREK YU
DEPARTMENT OF ECONOMICS
UNIVERSITY OF THE WESTERN CAPE
UNIVERSITY OF STELLENBOSCH
SOUTH AFRICA
E-MAIL: DYU@UWC.AC.ZA

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

# Poverty and inequality estimates of National Income Dynamics Study revisited

DEREK YU[1]

ABSTRACT

The National Income Dynamics Study (NIDS), introduced since 2008, has become an alternative data source for the South African poverty and inequality analyses. In addition to the fact that NIDS is the first national panel study of individuals in South Africa, it is also the only survey that allows the respondents to report income and expenditure as both a single estimate, 'one-shot' amount and an aggregate amount derived from the sum of the amounts for sub-categories. The latter variable, after imputations, was the preferred variable for deriving the poverty and inequality estimates. This paper examines if the poverty and inequality estimates are significantly different, using both the single estimate and the aggregate (before and after imputations) income and expenditure variables.

Keywords: Poverty, Inequality, National Income Dynamics Study, Household surveys, measurement, South Africa
JEL codes: I32

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

## 1. Introduction

In order to evaluate the extent to which a country achieves the objective of poverty and inequality reduction, up-to-date and reliable data are required. In South Africa, the census conducted by Statistics South Africa was the only available data source to examine money-metric poverty and inequality before the transition, using the per capita income and expenditure variables for the analyses. Although the Income and Expenditure Survey (IES) was also a usable dataset, the sample only covered a limited sub-set of households in metropolitan areas, while the 1993 October Household Survey (OHS) excluded people residing in Transkei-Bophuthatswana-Venda-Ciskei (TBVC) from the sample.

A major advance since the advent of democracy has been the availability of more surveys for estimating poverty and inequality. In addition to census, IES and OHS, new surveys were conducted by Statistics South Africa, such as the General Household Survey (GHS) introduced in 2002, the Labour Force Survey (LFS) which replaced the OHS since 2000, and the 2007 Community Survey (CS) which was conducted to fill the information gap between the 2001 and 2011 censuses. In addition, the Project for Statistics on Living Standards and Development (PSLSD) and the National Income Dynamics Study (NIDS) conducted by the Southern African Labour and Development Research Unit (SALDRU) are alternative surveys providing income and expenditure data for the poverty and inequality analyses. Furthermore, although the South African Advertising Research Foundation (SAARF) has been conducting the All Media Products Survey (AMPS) since 1975, the income data of the survey has only been used for these analyses in recent years.

The income and expenditure information was collected by the recall method in all the abovementioned surveys, except for that IES 2005/2006 used the diary method. In the censuses, CS 2007, OHSs, LFSs, GHSs, AMPSs and NIDS, respondents were asked to report income or expenditure as a single estimate, either in actual continuous amount (e.g., OHS 1996-1998 and NIDS) or in bands (e.g., OHS 1999, LFSs, GHSs, AMPSs, censuses and CS 2007). For the latter approach, the number of intervals and the width of each interval differed amongst the censuses and surveys. In contrast, in PSLSD, IESs and NIDS, income and expenditure were derived from the sum of the amounts for sub-categories, e.g., the aggregation approach was used. In other words, NIDS is the only survey that allows the respondents to declare income and expenditure under both single estimate and aggregation approaches.

The majority of recent South African studies used the census and survey data unadjusted to examine poverty and inequality since the transition. Despite using different poverty lines and surveys for the analyses, the results of these studies in general indicated that poverty increased since the transition, before a downward trend took place since 2000 or 2001. In addition, the Gini coefficient, which indicates the extent of inequality, increased continuously (UNDP 2003; Meth and Dias 2004; Leibbrandt, Levinsohn and McCrary 2005; Vermaark 2005; Hoogeveen and Özler 2006; Leibbrandt, Poswell, Naidoo and Welch 2006; Bhorat and Van der Westhuizen 2008; Van der Berg, Louw and Yu 2008; Yu 2008; Bhorat, Van der Berg, Louw and Du Toit 2009; Posel and Rogan 2012).

Some researchers were concerned that the surveys under-captured income or expenditure comparing with the national accounts current income, and that the survey income or expenditure declined rapidly between the two surveys, contradicting the trends observed in the national accounts (for instance, both income and expenditure dropped abruptly between IES 1995 and IES 2000, and such decline was larger than the decrease experienced by the South African economy during the Great Depression of the 1930s). Hence, the survey income or expenditure distribution was adjusted rightwards in line with the national accounts mean income, before poverty estimates were re-examined (Van der Berg and Louw 2004; Pauw and Mncube 2007; Van

der Berg, Burger, Burger, Louw and Yu 2005 & 2009). As expected, the poverty level declined across the surveys after this adjustment took place, but the poverty trends remained the same, as discussed above.

Ardington, Lam, Leibbrandt and Welch (2005) and Yu (2009) were concerned about the high proportion of households with unspecified income in the censuses and CS 2007[2], as excluding these households could lead to unreliable poverty and inequality estimates. Also, the proportion of households reporting zero income was very high in the aforementioned censuses and survey[3], and poverty and inequality estimates would be biased upwards by including these households. Hence, the income of these households (with zero or unspecified income) was imputed by means of sequential regression multiple imputation (SRMI). After imputations, Yu (2009) found lower poverty levels became but poverty trends remained the same. Similarly, the Gini coefficient decreased moderately in both censuses and CS 2007 after imputations, but the inequality trends remained the same between 1996 and 2007, i.e., it increased between 1996 and 2001 but decreased between 2001 and 2007

Finally, the study by Leibbrandt, Woolard, Finn and Argent (2010) derived a revised income variable in PSLSD, IES 2000 and NIDS 2008 by excluding certain income items (e.g., imputed rent, agricultural income, sales of vehicles and fixed property, etc.) so that the per capita income derived across the surveys would be more comparable. The resultant poverty trends were very different from other studies, as a continuous but extremely small decline of poverty was found across the three surveys, while Gini coefficient showed a continuous but negligent increase.

Other factors might also have a played a role to affect the reliability and comparability of poverty and inequality estimates across the surveys, such as whether the income and expenditure information was collected in actual amounts or in bands; whether the single estimate or aggregation method was used to ask respondents to declare the actual amounts; the number of bands and width of each band, as well as the method used (e.g., midpoint method, midpoint-Pareto method, interval regression) to approximate the amount in each band. As NIDS is the only survey that allows respondents to declare the income and expenditure amount from both the single estimate and aggregation approaches, this paper aims to examine if the poverty and inequality estimates derived from these approaches would be significantly different.

The remainder of the paper is structured as follows: Section 2 reviews the merits and drawbacks of the single estimate and aggregation approaches, while Section 3 discusses how the income and expenditure information was captured in NIDS. Section 4 investigates poverty and inequality estimates and trends using the per capita income and expenditure variables derived by these two approaches, focusing on the poverty headcount ratios and Gini coefficients. The NIDS estimates are then briefly compared with the estimates from the other censuses and surveys. Section 5 concludes the paper.

## 2. Single estimate vs. Aggregation approaches to collect income and expenditure information

The single estimate approach captures income and expenditure information by means of one question to ask the respondent to declare total income or expenditure, for instance 'how much money did this household spend on all its expenses in the last 30 days?'; 'What was the total amount of income that this household received from all sources last month?'. This approach is cheaper and less time-consuming, especially for research institutions facing budget and staff

[2] The proportion of households with unspecified income was 11.5% in Census 1996, 16.4% in Census 2001 (before hot deck imputation was conducted by Statistics South Africa) and 11.1% in CS 2007.
[3] 13.0% of households reported zero income in Census 1996, while this proportion was 21.0% and 8.2% in Census 2001 (before hot deck imputation was conducted by Statistics South Africa) and CS 2007 respectively.

constraint. Interviewer and interviewee fatigue could also be avoided to a certain extent. However, this method could confuse the respondents, as they are unsure about what items should be included as part of the total income or expenditure. For instance, some respondents might not regard remittances received, lottery winnings and interest earned from investment as part of total income. This may result in low response rate, and/or under-reporting of total income or expenditure (Deaton 1997: 27; Browning, Crossley and Weber 2002: 7-10). However, Davern, Rodin, Beebe and Thiede Call (2005: 1535) argue that the single estimate approach might work better when capturing income, as asking respondents to declare exact amounts earned from each income source could prove quite burdensome and intrusive for the respondents. People do not like to divulge how much money they earn in too great detail, as a result of the sensitive nature of the questions. In fact, some respondents already find it disturbing to reveal income or even consumption information even if asked to declare the 'one-shot' amount.

A series of questions are asked on all of the sub-items in order to derive the overall income or expenditure amount in the aggregation approach, e.g., questions like 'How much do you earn from income source X?', 'How much do you earn from income source Y?', and so forth are asked, and the total income is derived by adding the amounts from the answers of these questions. If the aggregation approach is used, an issue to consider is the appropriate level of disaggregation. Deaton (2005: 16) claims that the greater the degree of disaggregation of the number of items that are separately distinguished, the more accurate is the measured consumption (expenditure) in total.

This is what happened in the Income and Expenditure Survey (IES)[4], as the degree of disaggregation is greater in the latter survey when compared with NIDS and the 1993 Project for Statistics on Living Standards and Development (PSLSD). For example, the NIDS participants were asked to declare the total spending on the food category 'coffee and tea', but the IES participants were asked to report expenditure on an exhaustive list of coffee and tea items, that is, instant coffee, ground coffee and coffee beans, Ceylon tea, as well as Rooibos and herbal teas. Similarly, while the NIDS participants were asked to report the total expenditure on the non-food category 'cigarettes and tobacco', the IES participants were asked to declare expenditure on a more exhaustive list of items, that is, cigars and cigarillos, pipe and cigarette tobacco, snuff, other items smoked, as well as smokers' requisites (e.g., pipes, cigarette lighters, flints).

However, an expected outcome of the greater level of disaggregation is that the questionnaire is longer and more time-consuming for the participants to complete. For instance, in IES 2000, 30 pages of the questionnaire were devoted to capturing expenditure information, while it was only five pages in the case of NIDS, due to the greater level of disaggregation of the former survey. Deaton (2005: 16) as well as Browning et al. (2002: 12-18) suggest that, if the level of disaggregation is too high, it could be very demanding, expensive, time-consuming and exhausting to both the interviewers and interviewees, and the latter might end up deliberately providing misleading amounts and even not answering some questions (i.e., item non-response). This eventually results in the derivation of an even more inaccurate aggregate consumption amount, compared with the single estimate method. Furthermore, Browning et al. (2002: 19) argue that, for non-durable items, a non-exhaustive list method should be more than enough to obtain reliable information on consumption (expenditure), e.g., the two questions 'expenditure on food at home' and 'expenditure on food outside home' should result in a pretty good predictor of total food expenditure. In contrast, for durable items, they suggest that the exhaustive method works better.

[4] The primary objective 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). Nonetheless, these surveys have also become an important source of information for deriving poverty and inequality estimates and trends.

## 3. Collection of income and expenditure information in NIDS

In NIDS, a stratified, two-stage cluster sample design was adopted (SALDRU 2009). In the first stage, 400 primary sampling units (PSUs) were selected from Statistics South Africa's 2003 Master Sample of 3000 PSUs. The sample of PSUs was a subset of the Master Sample. The explicit strata in the Master Sample were the 53 District Councils (DCs). Next, eight non-overlapping samples of dwelling units were systematically drawn within each PSU. Each of these samples was called a 'cluster'. Fieldworkers were instructed to interview all households living at the selected dwelling unit. If it was found that the dwelling unit was unoccupied or the dwelling no longer existed, the fieldworkers were not permitted to substitute the dwelling unit, but recorded this information on the household control sheet. At the time of writing, two waves of NIDS were conducted, in 2008 and 2010/2011, with 7 301 and 6 809 households reporting income and expenditure information in each survey.

Household expenditure was derived in two ways, with the recall method being adopted in both methods. First, it was derived by adding the respondents' answers on 32 categories of food spending and 53-54 categories of non-food spending[5] (Sections D and E of the household questionnaire), i.e., the aggregation approach. If there was item non-response (e.g., a household did not specify expenditure on one or more food or non-food items) or it was suspected that the food and non-food spending declared by the respondents might not be reliable (e.g., zero food or non-food expenditure, under-reporting or over-reporting of these spending), these expenditure amounts were reset as unspecified before they were imputed (Argent 2009). The imputed food spending and imputed non-food spending were then added to the implied rent expenditure (which took imputed rent into consideration) to derive the aggregate household expenditure.

The household head was also asked to declare 'one-shot' total expenditure in the last 30 days[6], i.e., the single estimate approach. However, the response rates of this question were only 79.4% and 88.7% in the 2008 and 2010 NIDS respectively (see Table 1). SALDRU used the aggregate household expenditure after imputations for poverty and inequality analyses.

Table 1: Response rates of the questions used to derive household income and expenditure in NIDS 2008 and NIDS 2010/2011

| | NIDS 2008 (n = 7 301) | NIDS 2010/2011 (n = 6 809) |
|---|---|---|
| Income | | |
| 1. Single estimate amount | 5 919 (81.1%) | 5 753 (84.5%) |
| 2. Aggregate amount (before imputations) | 6 456 (88.4%) | 6 462 (94.9%) |
| 3. Aggregate amount (after imputations + adding implied rent) | 7 301 (100.0%) | 6 809 (100.0%) |
| Households with specified income under all three approaches | 5 441 (74.5%) | 5 557 (81.6%) |
| Households with specified income under all three approaches and took part in both surveys | 4 212 (57.7%) | 4 212 (61.9%) |
| Expenditure | | |
| 1. Single estimate amount | 5 799 (79.4%) | 6 038 (88.7%) |
| 2. Aggregate amount (before imputations) | 7 211 (98.8%) | 6 362 (93.4%) |
| 3. Aggregate amount (after imputations + adding implied rent) | 7 301 (100.0%) | 6 809 (100.0%) |
| Households with specified expenditure under all three approaches | 5 776 (79.1%) | 5 667 (83.2%) |
| Households with specified income under all three approaches and took part in both surveys | 4 473 (61.3%) | 4 473 (65.7%) |

[5] There were 53 categories of non-food spending in NIDS 2008 but 54 categories in NIDS 2010/2011.

[6] The question was asked (in Questions D31 and D34 of the household questionnaire in NIDS 2008 and 2010 respectively) as "How much money did this household spend on all its expenses in the last 30 days?"

The recall method was also adopted to derive household income. First, with regard to the single estimate approach, the households were asked to declare total household income received in the last month[7]. If respondents did not know the exact income amount, in NIDS 2008, they were given the option to declare the relevant total monthly household income category and could choose from 15 categories[8]. In NIDS 2010/2011, respondents were asked to report whether last month's household income was more than, about equal to or less than a specific amount[9], before the household income in interval terms was converted into continuous amounts using the midpoint-Pareto method. The response rates were 81.1% and 84.5% in NIDS 2008 and NIS 2010/2011 respectively, and only about 8% and 10% of these households reported income in intervals in each survey respectively.

SALDRU (2009) argued that the two 'one-shot' income questions mentioned above could result in lower household income (to be discussed later), and opted to use the respondents' answers on each income component in Section E and Section F of the adult questionnaire (i.e., the aggregation approach). In these two sections, questions on employment income and non-employment were asked, with the respondents being asked to declare the actual amounts. The household aggregate income amount was then derived by adding the respondents' answers on the following six broad components: wage income, government grant income, other government income, investment income, remittances income and agricultural income. Imputation was once again involved if respondents did not specify income earned on certain items or if the income from certain sources as reported by the respondents might not be reliable (e.g., serious under-reporting or over-reporting). The imputed income from these six sources were then added to the implied rent income (which also took imputed rent into consideration, as in what happened to the implied rent expenditure variable discussed previously) to derive the imputed aggregate household income. This explains why all households in the sample had specified aggregate household income after imputations. Finally, this imputed aggregate income variable was used by SALDRU to conduct poverty and inequality analyses.

## 4. Empirical analyses

This section estimates the extent of poverty and inequality using the variables derived from both approaches. Only households with specified 'one-shot' amount, aggregate amount before imputations and aggregate amount after imputations and adding the implied rent, and took part in both surveys are included for the analyses. That is, 4 212 and 4 473 households in the case of income and expenditure respectively (see Table 1).

### 4.1 Total income and expenditure

Figure 1 shows the annual household income and expenditure derived from each approach in 2008 prices for each survey. As expected, household income and expenditure were the highest using the aggregate approach, after imputations and taking implied rent into consideration. When comparing the aggregate amount before imputations and aggregate amount after imputations

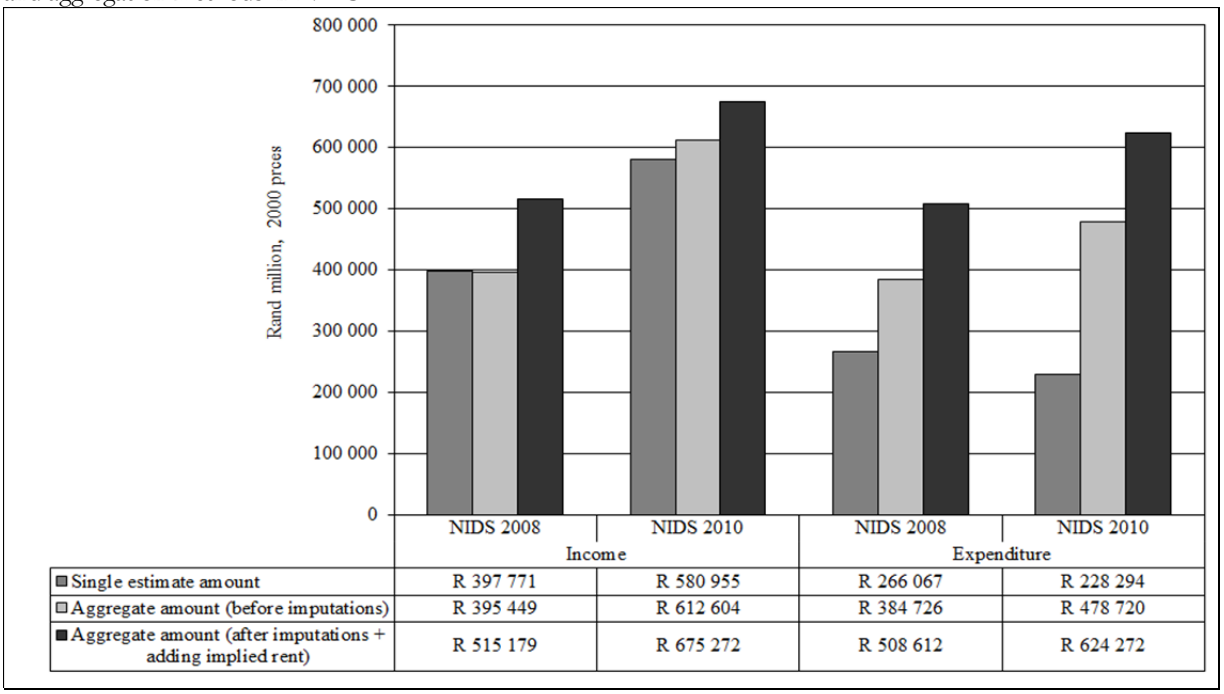
[7] The question was asked (in Questions D38 and D39 of the household questionnaire in NIDS 2008 and 2010 respectively) as "What was the total amount of income (after income tax) that this household received last month? Please note this includes all the household members' salaries and wages, grants, interest, rental income and income from agriculture earned by household members in the last month."

[8] The question was asked (in Questions D39 of the household questionnaire in NIDS 2008) 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-R1 000, 5: R1 001-R1 500, 6: R1 501-R2 500, 7: R2 501-R3 500, 8: R3 501-R4 500, 9: R4 501-R6 000, 10: R6 001-R8 000, 11: R8 001-R11 000, 12: R11 001-R16 000, 13: R16 001-R30 000, 14: R30 001-50 000, 15: R50 001 or more.

[9] The question was asked (in Questions D40.1-40.5 of the household questionnaire in NIDS 2010) as "Would you say the last month's household income was: D40.1 – More than or less than R2 300, D40.2 – more than or less than R1 300, D40.3 – more than or less than R4 700, D40.4 – more than or less than R700, D.50.5 – more than or less than R11 000?"

(and adding the implied rent) in each survey, it was found that in NIDS 2008, only one and five households had the former amount exceeding the latter amount for income and expenditure respectively; this only happened in 2.86% of households in NIDS 2010/2011 for both income and expenditure. It is likely that these households seriously over-reported income or expenditure before imputations, and hence after imputations and even after adding the implied rent, income or expenditure became smaller.

Figure 1: Total annual household income and expenditure (Rand million, 2000 prices) derived from single estimate and aggregation methods in NIDS




| | Income | | Expenditure | |
|---|---|---|---|---|
| | NIDS 2008 | NIDS 2010 | NIDS 2008 | NIDS 2010 |
| Single estimate amount | R 397 771 | R 580 955 | R 266 067 | R 228 294 |
| Aggregate amount (before imputations) | R 395 449 | R 612 604 | R 384 726 | R 478 720 |
| Aggregate amount (after imputations + adding implied rent) | R 515 179 | R 675 272 | R 508 612 | R 624 272 |

Interestingly, the single estimate income and aggregate income without imputations were very close in 2008. This finding confirms the argument by Davern et al. (2005) that the single estimate approach could also work quite well when capturing income. In contrast, the single estimate household expenditure was much lower than the aggregate household expenditure before imputations in both surveys. For instance, in 2008, the former amount (R266 067) was only 69.2% as proportion of the latter amount (R384 726), while this proportion was even lower (47.7%) in 2010 (R228 294 vs. R478 720).

Comparing the income and expenditure across the two surveys, regardless of the method used, income increased between NIDS 2008 and NIDS 2010/2011. However, the single estimate expenditure amount dropped rapidly by 14.2% (from R266 067 to R228 294). In contrast, aggregate expenditure before and after imputation increased by 24.4% (from R384 726 to R478 720) and 22.7% (from R508 612 to R624 272) respectively.

The households are divided into quintiles using the single estimate per capita variables, and Table 2 shows that in NIDS 2008, 49.5% reported lower income under the single estimate approach. This proportion increased to 57.0% in NIDS 2010/2011. However, this proportion is the highest for households in quintile 1 (67.3% and 73.1% in each survey), and decreases continuously when moving to the richer quintiles. This proportion is 39.1% and 38.8% in quintile 5 for each survey. Despite the presence of households reporting lower income under the single estimate approach, Table 2 also shows that some households reported higher income under the single estimate approach, and this proportion is higher in the richer quintiles in both surveys (e.g., 60.9% and 61.2% in each survey in quintile 5). Hence, this could explain why the total incomes under single estimate and aggregation approaches are very similar in 2008 (R397 771 vs. R395 449 respectively – see Figure 1).

Table 2: Comparison of single estimate and aggregate income (before imputations) by household quintile, NIDS

| NIDS 2008 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **[1]** | **[2]** | **[3]** | **[4]** | **[5]** | **[6]** | | **Sum of [1]-[5]** |
| Quintile 1 | 23.4% | 15.7% | 11.3% | 7.9% | 9.0% | 32.7% | 100.0% | 67.3% |
| Quintile 2 | 1.7% | 6.0% | 13.8% | 15.2% | 15.0% | 48.2% | 100.0% | 51.8% |
| Quintile 3 | 1.3% | 3.8% | 8.2% | 12.3% | 22.8% | 51.6% | 100.0% | 48.4% |
| Quintile 4 | 1.2% | 1.2% | 6.4% | 6.7% | 25.2% | 59.4% | 100.0% | 40.6% |
| Quintile 5 | 1.0% | 2.1% | 5.7% | 8.1% | 22.3% | 60.9% | 100.0% | 39.1% |
| All households | 5.8% | 5.8% | 9.1% | 10.0% | 18.8% | 50.5% | 100.0% | 49.5% |
| **NIDS 2010/2011** | | | | | | | | |
| | **[1]** | **[2]** | **[3]** | **[4]** | **[5]** | **[6]** | | **Sum of [1]-[5]** |
| Quintile 1 | 17.5% | 13.9% | 15.6% | 14.8% | 11.3% | 26.9% | 100.0% | 73.1% |
| Quintile 2 | 4.2% | 8.9% | 14.6% | 15.6% | 20.3% | 36.5% | 100.0% | 63.5% |
| Quintile 3 | 1.8% | 5.5% | 11.7% | 15.4% | 22.8% | 42.8% | 100.0% | 57.2% |
| Quintile 4 | 0.5% | 3.8% | 9.6% | 12.6% | 25.9% | 47.7% | 100.0% | 52.3% |
| Quintile 5 | 0.0% | 3.1% | 7.2% | 13.7% | 14.7% | 61.2% | 100.0% | 38.8% |
| All households | 4.8% | 7.0% | 11.7% | 14.4% | 19.0% | 43.0% | 100.0% | 57.0% |

Note: household quintiles are derived by using the single estimate per capita income variable.
[1]: 0.0 ≤ (Single estimate / Aggregate amount) < 0.2
[2]: 0.2 ≤ (Single estimate / Aggregate amount) < 0.4
[3]: 0.4 ≤ (Single estimate / Aggregate amount) < 0.6
[4]: 0.6 ≤ (Single estimate / Aggregate amount) < 0.8
[5]: 0.8 ≤ (Single estimate / Aggregate amount) < 1.0
[6]: (Single estimate / Aggregate amount) ≥ 1

Table 3: Comparison of single estimate and aggregate expenditure (before imputations) by household quintile, NIDS

| NIDS 2008 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **[1]** | **[2]** | **[3]** | **[4]** | **[5]** | **[6]** | | **Sum of [1]-[5]** |
| Quintile1 | 17.7% | 29.0% | 19.1% | 14.8% | 8.1% | 11.2% | 100.0% | 88.8% |
| Quintile2 | 8.9% | 21.0% | 23.1% | 16.1% | 12.5% | 18.4% | 100.0% | 81.6% |
| Quintile3 | 6.7% | 14.8% | 19.6% | 13.6% | 14.3% | 31.0% | 100.0% | 69.0% |
| Quintile4 | 7.9% | 16.1% | 17.6% | 13.7% | 12.1% | 32.6% | 100.0% | 67.4% |
| Quintile5 | 2.9% | 10.8% | 11.9% | 20.2% | 13.7% | 40.5% | 100.0% | 59.5% |
| All households | 8.9% | 18.4% | 18.3% | 15.7% | 12.1% | 26.6% | 100.0% | 73.4% |
| **NIDS 2010/2011** | | | | | | | | |
| | **[1]** | **[2]** | **[3]** | **[4]** | **[5]** | **[6]** | | **Sum of [1]-[5]** |
| Quintile1 | 57.0% | 14.8% | 9.2% | 7.4% | 4.2% | 7.3% | 100.0% | 92.7% |
| Quintile2 | 14.5% | 12.3% | 19.3% | 17.1% | 11.3% | 25.4% | 100.0% | 74.6% |
| Quintile3 | 14.2% | 13.8% | 13.4% | 16.3% | 10.7% | 31.7% | 100.0% | 68.3% |
| Quintile4 | 11.7% | 13.0% | 12.9% | 19.4% | 11.4% | 31.7% | 100.0% | 68.3% |
| Quintile5 | 8.2% | 6.6% | 13.8% | 11.1% | 12.7% | 47.7% | 100.0% | 52.3% |
| All households | 21.1% | 12.1% | 13.7% | 14.3% | 10.0% | 28.7% | 100.0% | 71.3% |

Note: household deciles are derived by using the single estimate per capita expenditure variable.
[1]: 0.0 ≤ (Single estimate / Aggregate amount) < 0.2
[2]: 0.2 ≤ (Single estimate / Aggregate amount) < 0.4
[3]: 0.4 ≤ (Single estimate / Aggregate amount) < 0.6
[4]: 0.6 ≤ (Single estimate / Aggregate amount) < 0.8
[5]: 0.8 ≤ (Single estimate / Aggregate amount) < 1.0
[6]: (Single estimate / Aggregate amount) ≥ 1

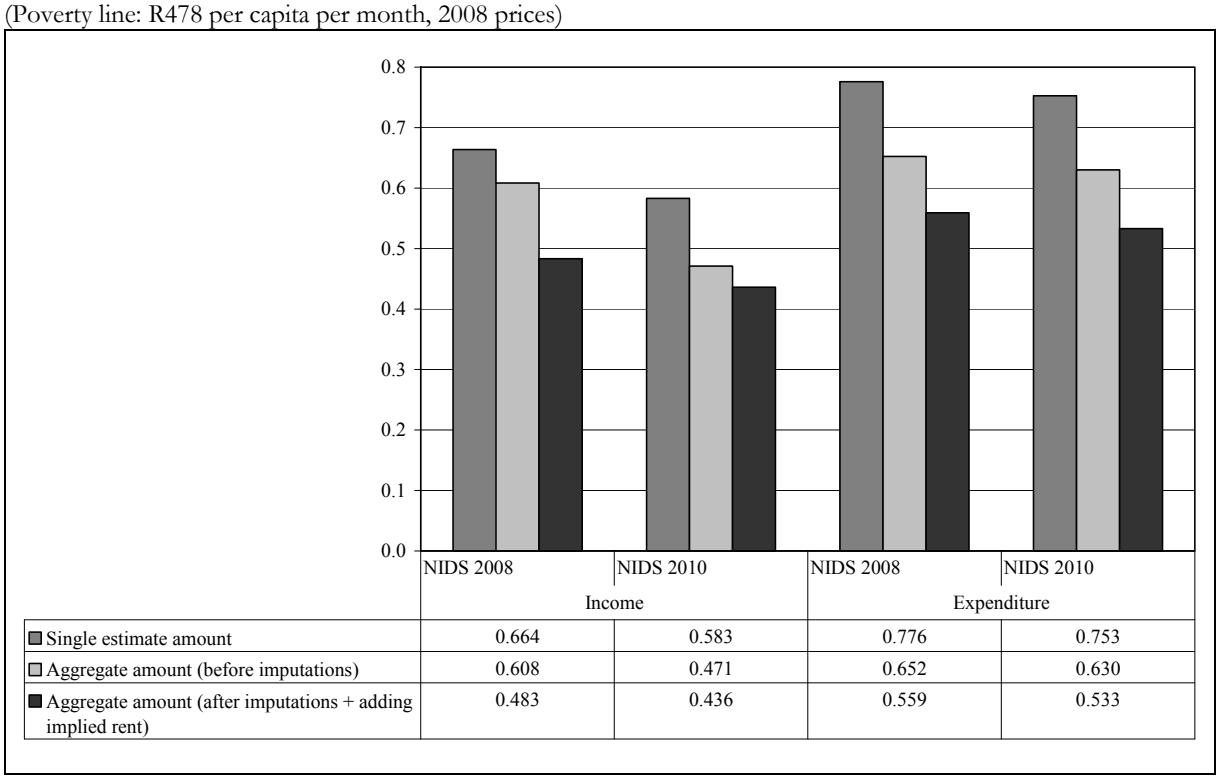
Table 3 above compares the single estimate and aggregate expenditure amounts, once again by dividing the households into quintiles using the single estimate per capita variable. It is found that more than 70% of households reported lower expenditure in the single estimate approach (73.4% in 2008 and 71.3% in 2010/2011). In addition, this proportion is the highest for households in

poorer quintiles (e.g., 88.8% and 92.7% in quintile 1 for each survey), but decreases continuously across the richer quintiles. Furthermore, comparing the results in the last column of Tables 2 and 3, it is obvious that the under-reporting of expenditure under the single estimate approach is relatively more serious than income. This is especially the case in NIDS 2010/2011, thereby explaining the big difference between single estimate income and aggregate income (refer to Figure 1). Finally, as happened with income, Table 3 shows that there are some households reported higher expenditure under the single estimate approach, and this proportion is higher in the richer quintiles in both surveys (e.g., 40.5% and 47.7% in each survey in quintile 5).

## 4.2 Poverty estimates and trends

Figure 2 presents the poverty headcount ratios at the lower bound poverty line (R3 864 in 2000 prices or R5736 in 2008 prices, per capita per annum) proposed by Woolard and Leibbrandt (2006). As expected, both the per capita income and expenditure variables derived after imputations and adding the implied rent are associated with lower poverty rates. In addition, poverty headcount ratios using the per capita expenditure variables without imputations were lower compared with the results using the single estimate per capita variables, due to the fact that the 'one-shot' expenditure was seriously under-estimated (see Figure 1). Furthermore, although the analyses from Figure 1 previously showed that the single estimate and aggregate income (before imputations) levels were very similar in both surveys (especially in 2008), poverty headcount ratios were higher using the per capita single estimate variable. This might be due to the fact that under-reporting of income from the single estimate method is more likely to happen to households in the poorer quintiles, while the likelihood of households declaring single estimate income higher than aggregate income is higher in the richer quintiles, as discussed previously.

Figure 2: Poverty headcount ratios using various income and expenditure variables
(Poverty line: R478 per capita per month, 2008 prices)




| | Income | | Expenditure | |
|---|---|---|---|---|
| | NIDS 2008 | NIDS 2010 | NIDS 2008 | NIDS 2010 |
| Single estimate amount | 0.664 | 0.583 | 0.776 | 0.753 |
| Aggregate amount (before imputations) | 0.608 | 0.471 | 0.652 | 0.630 |
| Aggregate amount (after imputations + adding implied rent) | 0.483 | 0.436 | 0.559 | 0.533 |

The poverty trends across the two surveys were quite similar when using the income and expenditure variables. Regardless of which method was used to derive the income variable, income poverty declined between NIDS 2008 and NIDS 2010/2011. Similar downward trends were observed when using the expenditure variables, despite the fact that the extent of the decline of poverty was less rapid. These results are consistent with the findings from the recent

studies, which found that poverty declined continuously since 2000 or 2001 (see Section 1). However, keep in mind that these recent studies concluded this downward trend in poverty in 2000s by comparing, for instance, IES 2000 with IES 2005/2006, Census 2001 with CS 2007, all 2000-2009 AMPSs, or IES 2005/2006 with NIDS 2008.

As poverty estimates and trends could be subject to the poverty lines chosen, dominance testing should be conducted by means of a cumulative density function (CDF). In a CDF, the vertical axis shows the percentage of total population with an income that is less than or equal to the income value on the horizontal axis. The strength of this approach is that it makes it possible to compare the changes in poverty from one period to the next (or poverty amongst various demographic groups, such as by race, gender, province), independent of any single poverty line. If the CDF for a given period lies above the CDF for the previous period on the horizontal axis, this means that poverty has increased, irrespective of any given poverty line, because the percentage of population with a certain income or less has increased. If the opposite happens, this means poverty has decreased at all poverty lines. The results from Figures A.1 to A.4 in the Appendix clearly show that, regardless of the poverty line chosen, the poverty headcount ratio has always been the greatest, if the single estimate per capita variables are used. Also, the aggregate per capita variables after imputations are associated with the lowest poverty estimates. The gap between the two CDFs using the single estimate method and aggregation method without imputations is clearly wider in the case of expenditure (Figures A.3 and A.4).

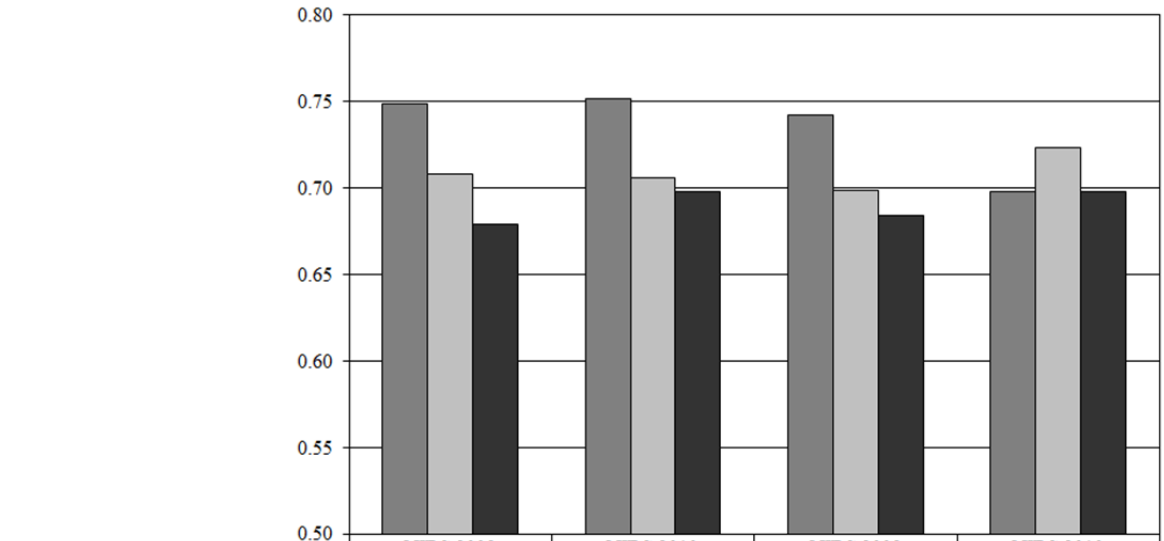
As far as the poverty trends between NIDS 2008 and NIDS 2010/2011 are concerned, Figure A.5-A.10 in the Appendix shows that, regardless of the poverty line chosen, poverty decreased between the two surveys in all three income variables (Figures A.5-A.7). In the case of expenditure, Figure A.8 shows that, comparing the one-shot expenditure variables across the two surveys, the 2010/2011 CDF clearly lies above the 2008 CDF if per capita expenditure is below R150. When expenditure is between R200 and R900, the 2008 CDF lies above the 2010/2011 CDF. At expenditure level beyond R900, the two CDFs are very close to each other. Hence, these findings indicate that the one-shot expenditure poverty headcount trends are sensitive to the poverty line being chosen.

When comparing the aggregate expenditure variables before imputations, Figure A.9 shows that the two CDFs are very close at lower expenditure levels. From about R300 onwards, the 2008 CDF is clearly above the 2010/2011, thereby indicating that poverty decreased between the two surveys. Very similar findings are observed when comparing the aggregate expenditure variables after imputations (see Figure A.10), as the 2008 CDF clearly lies above the 2010/2011 CDF at an expenditure level of beyond R350.

### 4.3 Inequality estimates and trends

This section briefly examines the levels and trends of Gini coefficients in each survey. Looking at the three income variables, in both surveys, the Gini coefficient was the highest using the single estimate per capita income variable. The discussions in Section 4.1 indicate that a higher proportion of households in quintile1 reported a single estimate income amount lower than the aggregate amount (refer to Table 2) and hence it could lead to a wider income distribution and greater inequality. In addition, Figure 3 shows that the imputed income variables resulted in the lowest Gini coefficient in both surveys. Furthermore, inequality decreased very slightly between 2008 and 2010 using aggregate income before imputations (from 0.708 to 0.706). However, using the other two income variables, it was found that inequality increased slightly between the two years.

Figure 3: Gini coefficients using various income and expenditure variables




| | Income | | Expenditure | |
|---|---|---|---|---|
| | NIDS 2008 | NIDS 2010 | NIDS 2008 | NIDS 2010 |
| Single estimate amount | 0.749 | 0.752 | 0.742 | 0.698 |
| Aggregate amount (before imputations) | 0.708 | 0.706 | 0.699 | 0.724 |
| Aggregate amount (after imputations + adding implied rent) | 0.679 | 0.698 | 0.684 | 0.698 |

Looking at the Gini coefficients estimated by using the three expenditure variables, in 2008, the single estimate expenditure variable resulted in the highest Gini coefficient (0.742), while the imputed aggregate variable resulted in the lowest Gini coefficient (0.684). However, the Gini coefficients using these two variables were the same in NIDS 2010/2011 (0.698), while it was the aggregate variable before imputations giving the highest estimate (0.724). Finally, looking at the inequality trends, the Gini coefficient increased between the two surveys using both aggregate expenditure variables. However, a downward trend was observed (decreasing from 0.742 to 0.698) when using the single estimate variable.

## 4.4 Comparison with other surveys

The results of the analyses in Sections 4.2 and 4.3 suggest that poverty and inequality estimates are much higher when using the single estimate income and expenditure variables (except for the expenditure inequality in NIDS 2010/2011). However, when comparing NIDS with other surveys taking place during the same period, Table 4 show that these estimates are higher in the GHSs, but the lowest in AMPSs. The single estimate method was used to capture income or expenditure information in these surveys. Hence, it suggests that the single estimate method does not always lead to higher poverty and inequality estimates. It is possible that factors other than the abovementioned two approaches could play a role to cause the poverty estimates to differ across the surveys. For instance, although the single estimate method was adopted in the aforementioned surveys, all the respondents were only asked to declare income or expenditure in bands. The number of bands is fewer (between eight and ten) in the GHSs but greater in the AMPSs (between 29 and 32). Also, the width of each band is narrower in AMPSs but wider in CS 2007 (especially in the higher-income categories). Thus, fewer and wider brackets are associated with higher poverty and inequality measures. More research needs to be done to examine how these factors could also affect the levels, trends and comparability of poverty and inequality estimates across different surveys.

Table 4: Poverty headcount ratios (Poverty line: R478 per capita per month, 2008 prices) and Gini coefficients in 2007-2011, selected surveys

| | **Poverty headcount ratio** | **Gini coeffi-cient** |
|---|---|---|
| **Comparing NIDS 2008 with selected 2007-2008 surveys** | | |
| NIDS 2008 income (Single estimate amount) | 0.664 | 0.749 |
| NIDS 2008 income (Aggregate amount, before imputations) | 0.608 | 0.708 |
| NIDS 2008 income (Aggregate amount, after imputations + adding implied rent) | 0.483 | 0.679 |
| NIDS 2008 expenditure (Single estimate amount) | 0.776 | 0.742 |
| NIDS 2008 expenditure (Aggregate amount, before imputations) | 0.652 | 0.699 |
| NIDS 2008 expenditure (Aggregate amount, after imputations + adding implied rent) | 0.559 | 0.684 |
| CS 2007 income (Single estimate amount, before imputations) | 0.529 | 0.759 |
| CS 2007 income (Single estimate amount, after imputations + adding implied rent) | 0.463 | 0.743 |
| AMPS 2008 income (Single estimate amount) | 0.410 | 0.666 |
| GHS 2008 expenditure (Single estimate amount) | 0.706 | 0.787 |
| **Comparing NIDS 2010/2011 with selected 2009-2011 surveys** | | |
| NIDS 2010 income (Single estimate amount) | 0.583 | 0.752 |
| NIDS 2010 income (Aggregate amount, before imputations) | 0.471 | 0.706 |
| NIDS 2010 income (Aggregate amount, after imputations + adding implied rent) | 0.436 | 0.698 |
| NIDS 2010 expenditure (Single estimate amount) | 0.753 | 0.698 |
| NIDS 2010 expenditure (Aggregate amount, before imputations) | 0.630 | 0.724 |
| NIDS 2010 expenditure (Aggregate amount, after imputations + adding implied rent) | 0.533 | 0.698 |
| AMPS 2009 income (Single estimate amount) | 0.414 | 0.644 |
| GHS 2010 expenditure (Single estimate amount) | 0.661 | 0.727 |
| GHS 2011 expenditure (Single estimate amount) | 0.637 | 0.714 |
| NIDS 2010 income (Single estimate amount) | 0.577 | 0.758 |

## 5. Conclusion

NIDS, the first national panel study of individuals in South Africa, is the only survey that allows respondents to report income and expenditure as both a single estimate, 'one-shot' amount and an aggregate amount derived from the sum of the amounts for sub-categories. The latter variable, after imputations and taking the implied rent into consideration, was the preferred variable to derive the poverty and inequality estimates. This paper, using both the single estimate and aggregate variables, found that the poverty and inequality estimates were higher using the single estimate variables. Respondents, especially those at the lower end of the single estimate income and expenditure distributions, reported lower income and expenditure under the single estimation approach.

## 6. References


Ardington, C., Lam, D., Leibbrandt, M. and 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, University of Cape Town.

Argent, J. (2009). *Household income: Report on NIDS Wave 1*. Technical Paper No. 3. Cape Town: Southern African Labour and Development Research Unit, University of Cape Town.

Argent, J., Franklin, S., Keswell, M., Leibbrandt, M. & Levinsohn, J. (2009). *Expenditure: Report on NIDS Wave 1*. Technical Paper No. 4. Cape Town: Southern African Labour and Development Research Unit, University of Cape Town.

Bhorat, H. and Van der Westhuizen, C. (2008). *Economic growth, poverty and inequality in South Africa: The first decade of democracy*. Paper presented at the Development Policy Research Unit Conference, Muldersdrift.

Browning, M., Crossley, T.F. and Weber, G. (2002). *Asking consumption questions in general purpose surveys*. SEDAP Research Paper No. 77. Hamilton: The Program for Research on Social and Economic Dimensions of an Aging Population (SEDAP).

Davern, M., Rodin, H., Beebe, T.J. and Thiede Call, K. (2005). The effect of income question design in health surveys on family income, poverty and eligibility estimates. *Health Services Research*. 40(5): 1534-1552.

Deaton, A. (1997). *The analysis of household surveys: A microeconomic approach to development policy*. Baltimore: The John Hopkins University Press.

Deaton, A. (2005). Measuring poverty in a growing world (or measuring growth in a poor world). *The Review of Economics and Statistics*. 87(1): 1-19.

Hoogeveen, J.G. and Özler, B. (2006). Poverty and inequality in post-apartheid South Africa: 1995-2000. In Bhorat, H. and Kanbur, R. (ed.), *Poverty and policy in post-apartheid South Africa*. Cape Town: Human Sciences Research Council: 59-94.

Leibbrandt, M., Levinsohn, J. and McCrary, J. (2005). *Incomes in South Africa since the fall of apartheid*. NBER working paper 11384. Cambridge, Massachusetts Avenue: National Bureau for Economic Research.

Leibbrandt, M., Poswell, L., Naidoo, P. and Welch, M. (2006). Measuring recent changes in South African inequality and poverty using 1996 and 2001 Census data. In Bhorat, H. and Kanbur, R. (ed.), *Poverty and policy in post-apartheid South Africa*. Cape Town: Human Sciences Research Council: 95-142.

Leibbrandt, M., Woolard, I., Finn, A. and Argent, J. (2010). *Trends in South African income distribution and poverty since the fall of apartheid*. OECD Social, Employment and Migration Working Papers No. 101. Paris: Organisation for Economic Co-operation and Development

Meth, C. and Dias, R. (2004). Increases in poverty in South Africa, 1999-2002. *Development Southern Africa*. 21(1): 59-85.

Pauw, K. and Mncube, L. (2007). *The impact of growth and redistribution on poverty and inequality in South Africa*. DPRU Working Paper 07/126. Cape Town: Development Policy Research Unit, University of Cape Town.

Posel, D. and Rogan, M. (2012). Gendered trends in poverty in post-apartheid period: 1997-2006. *Development Southern Africa*. 29(1): 97-113.

Southern Africa Labour Development Research Unit (2009). *National Income Dynamics Study Wave 1: User document*. Cape Town: Southern Africa Labour and Development Research Unit, University of Cape Town.

United Nations Development Program (UNDP) (2003). *South Africa Human Development Report 2003: The challenge of sustainable development*. Pretoria: United Nations Development Program of South Africa.

Van der Berg, Burger, R., Burger, R.P., Louw, M. and Yu, D. (2005). *Trends in poverty and inequality since political transition*. Stellenbosch Economic Working Papers: 01/2005. Stellenbosch: Stellenbosch University.

Van der Berg, S., Burger, R., Burger, R.P., Louw, M. and Yu, D. (2009). *A series of national account-consistent estimates of poverty and inequality in South Africa*. Stellenbosch Economic Working Papers: 09/2007. Stellenbosch: Stellenbosch University.

Van der Berg, S. and Louw, M. (2004). Changing patterns of South African income distribution: Towards time series estimates of distribution and poverty. *South African Journal of Economics*. 72(3): 546 – 572.

Van der Berg, S., Louw, M. and Du Toit, L. (2009). *Poverty trends since the transition: What we know*. Stellenbosch Economic Working Papers 19/09. Stellenbosch: Stellenbosch University.

Vermaark, C. (2005). *Trends in income distribution, inequality and poverty in South Africa, 1995-2003*. Paper presented at the Economic Society of South Africa Conference, Durban.

Woolard, I. and 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. (2008). *The comparability of Income and Expenditure Surveys 1995, 2000 and 2005/2006*. Stellenbosch Economic Working Papers: 08/2009. 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

Figure A.1: Poverty headcount ratios at different poverty lines: NIDS 2008 income variables

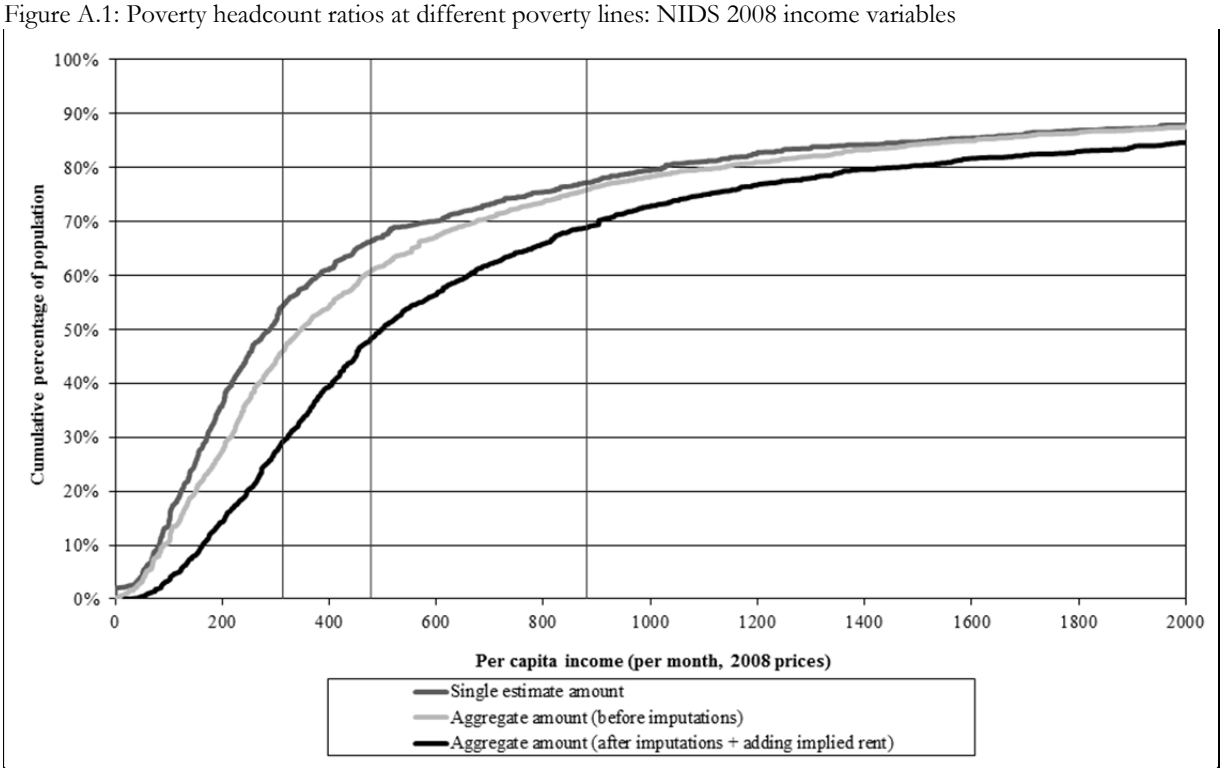



Figure A.2: Poverty headcount ratios at different poverty lines: NIDS 2010/2011 income variables

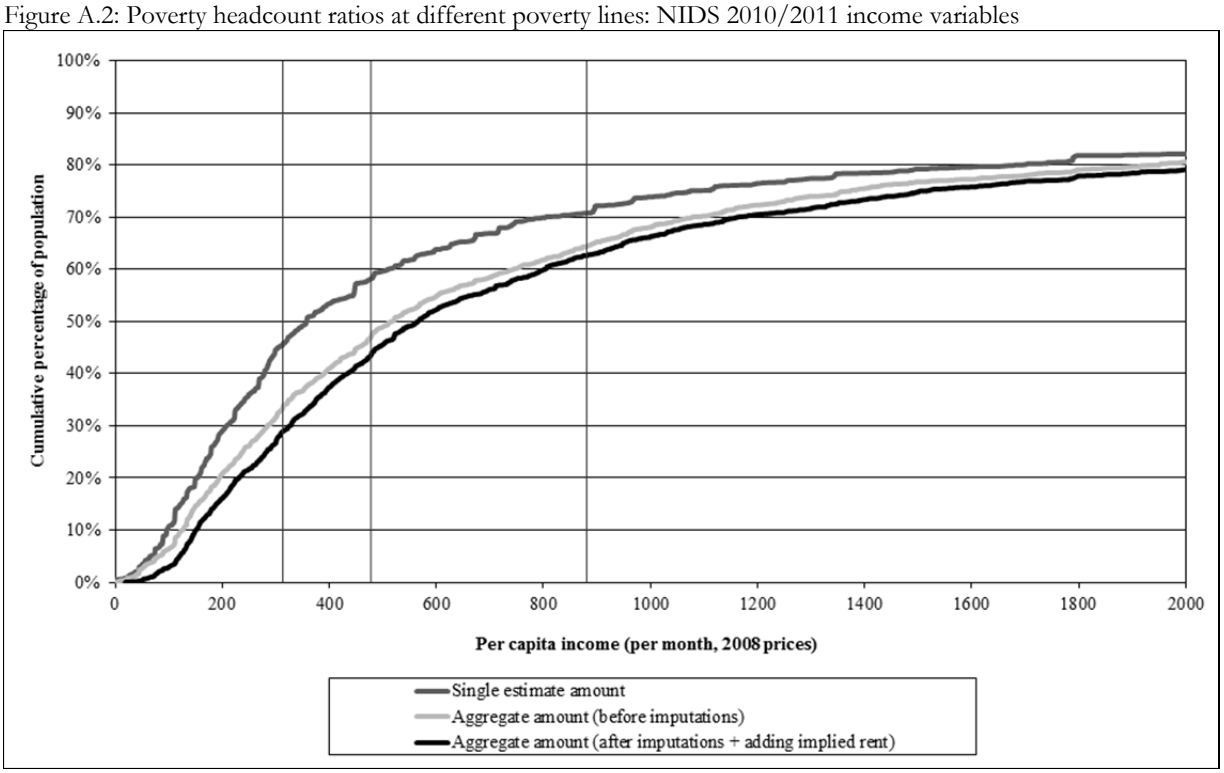


Figure A.3: Poverty headcount ratios at different poverty lines: NIDS 2008 expenditure variables

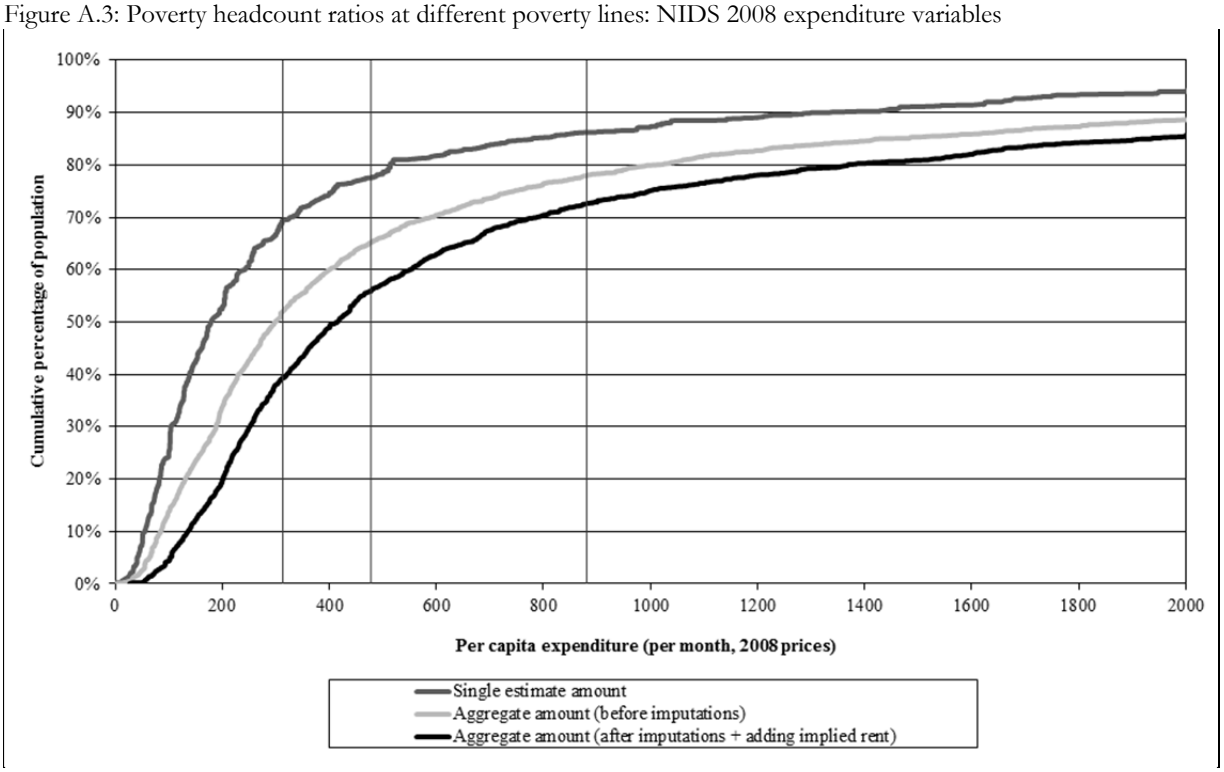



Figure A.4: Poverty headcount ratios at different poverty lines: NIDS 2010/2011 expenditure variables

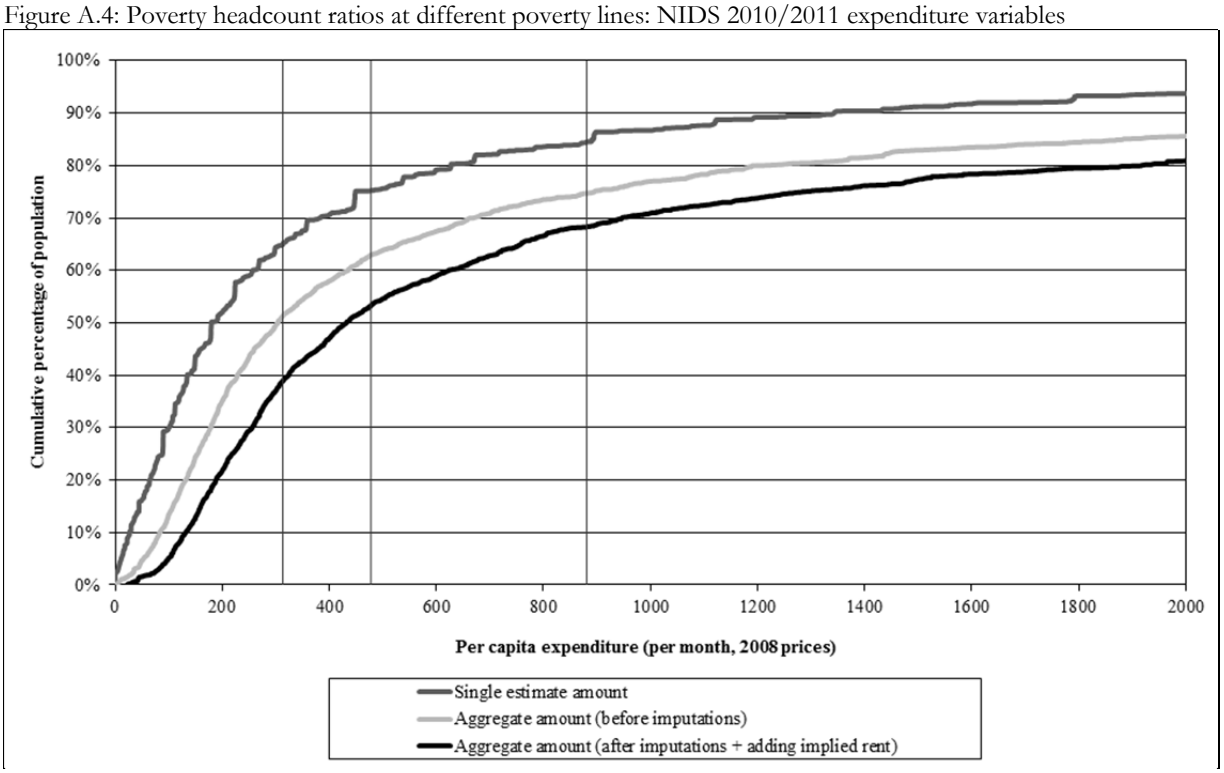


Figure A.5: Poverty headcount ratios at different poverty lines: NIDS single estimate income variables

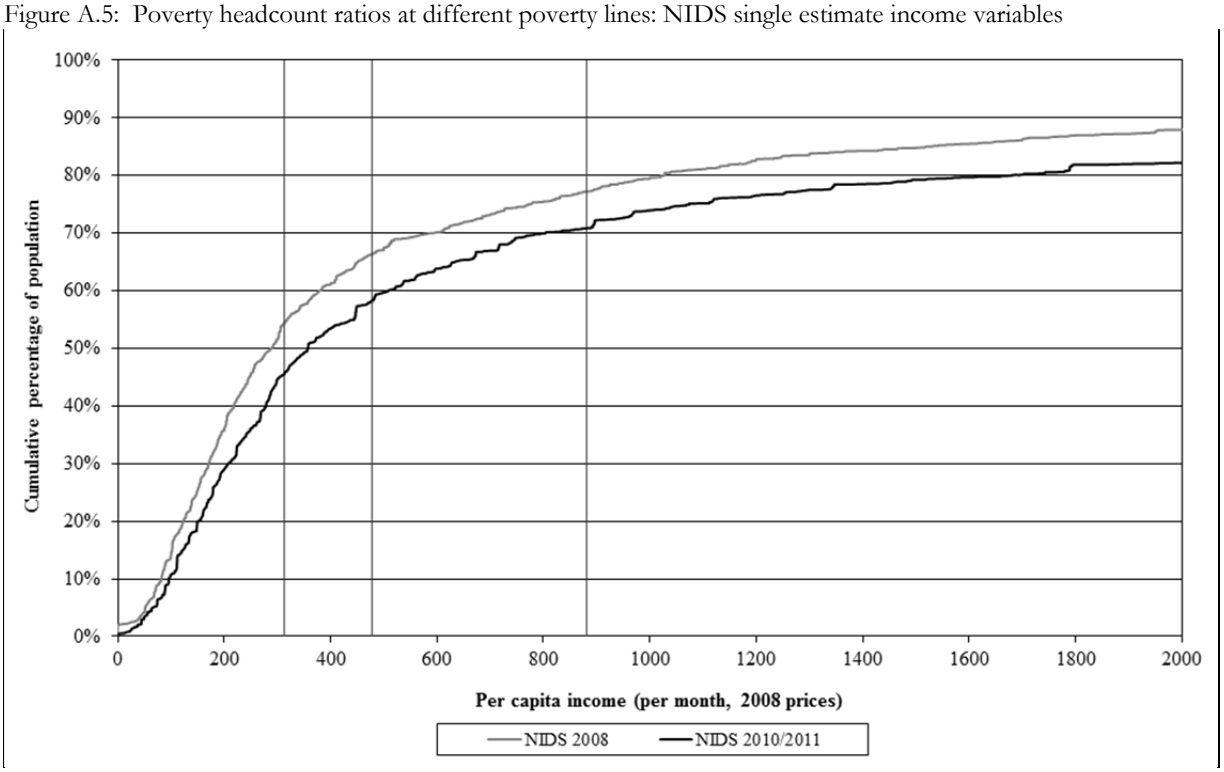



Figure A.6: Poverty headcount ratios at different poverty lines: NIDS aggregate income variables without imputations

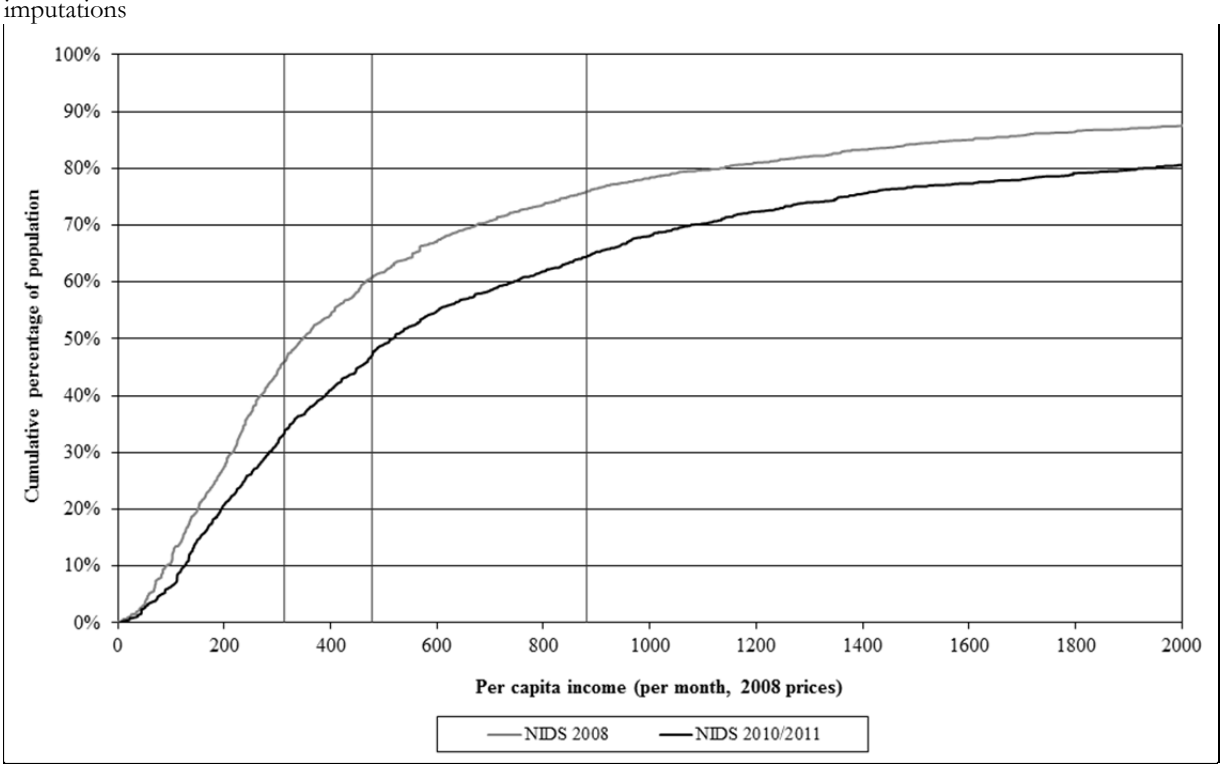


Figure A.7: Poverty headcount ratios at different poverty lines: NIDS aggregate income variables after imputations and adding implied rent

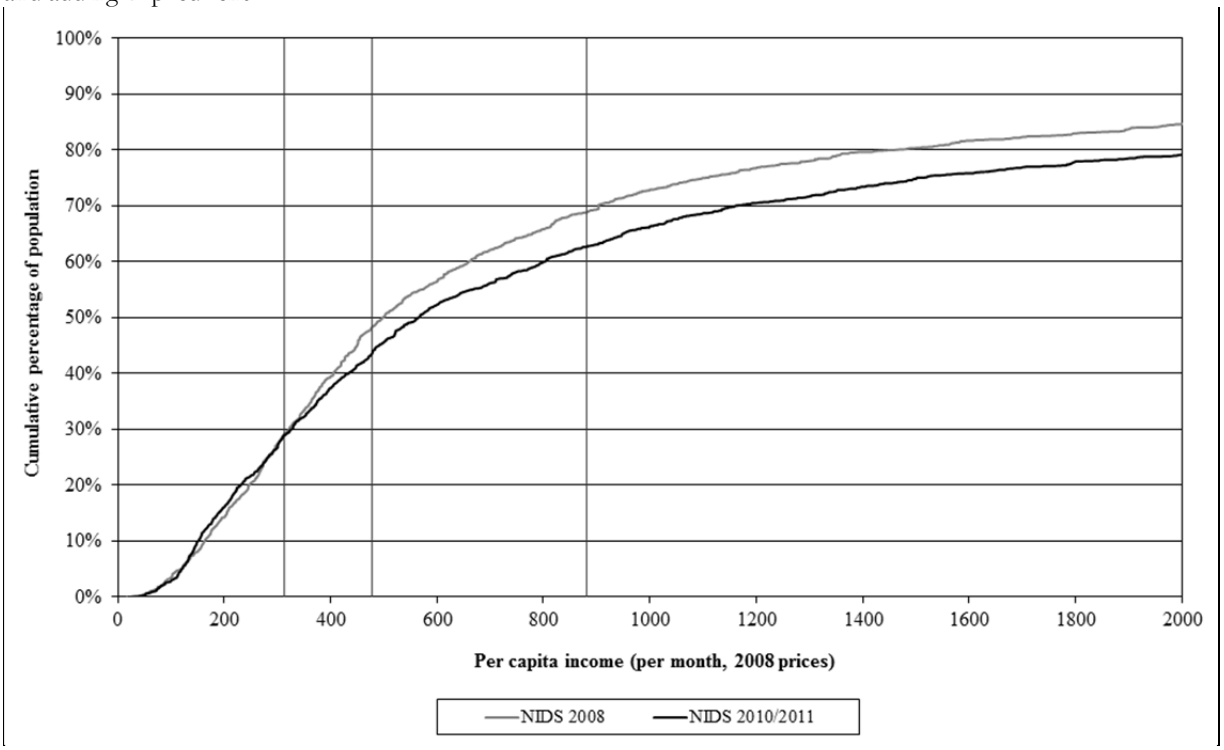



Figure A.8: Poverty headcount ratios at different poverty lines: NIDS single estimate expenditure variables

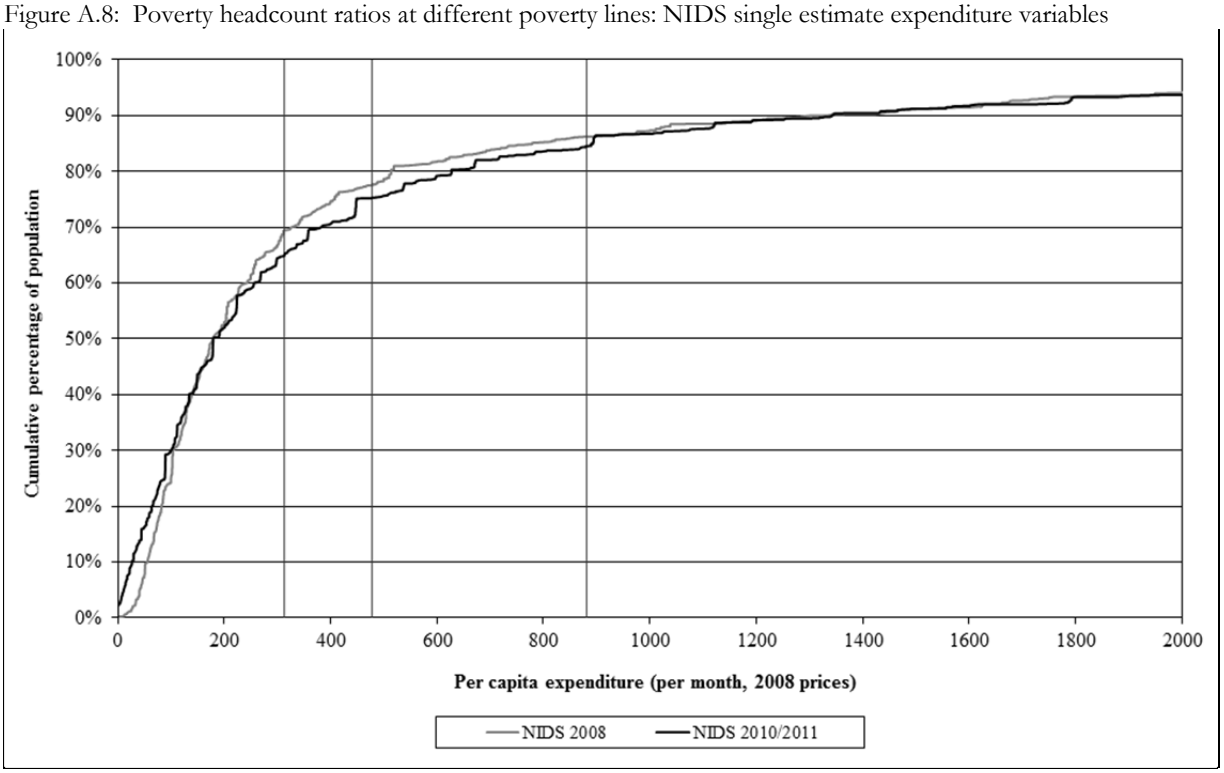


Figure A.9: Poverty headcount ratios at different poverty lines: NIDS aggregate expenditure variables without imputations

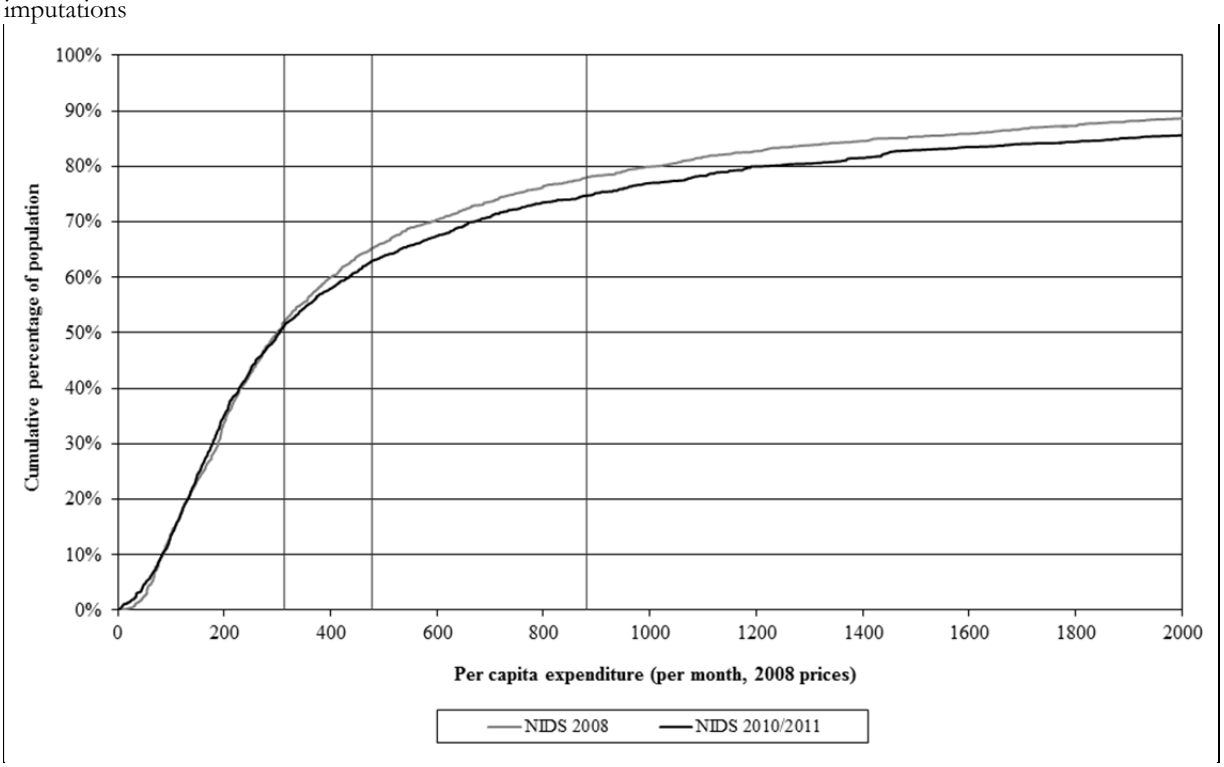



Figure A.10: Poverty headcount ratios at different poverty lines: NIDS aggregate expenditure variables after imputations and adding implied rent

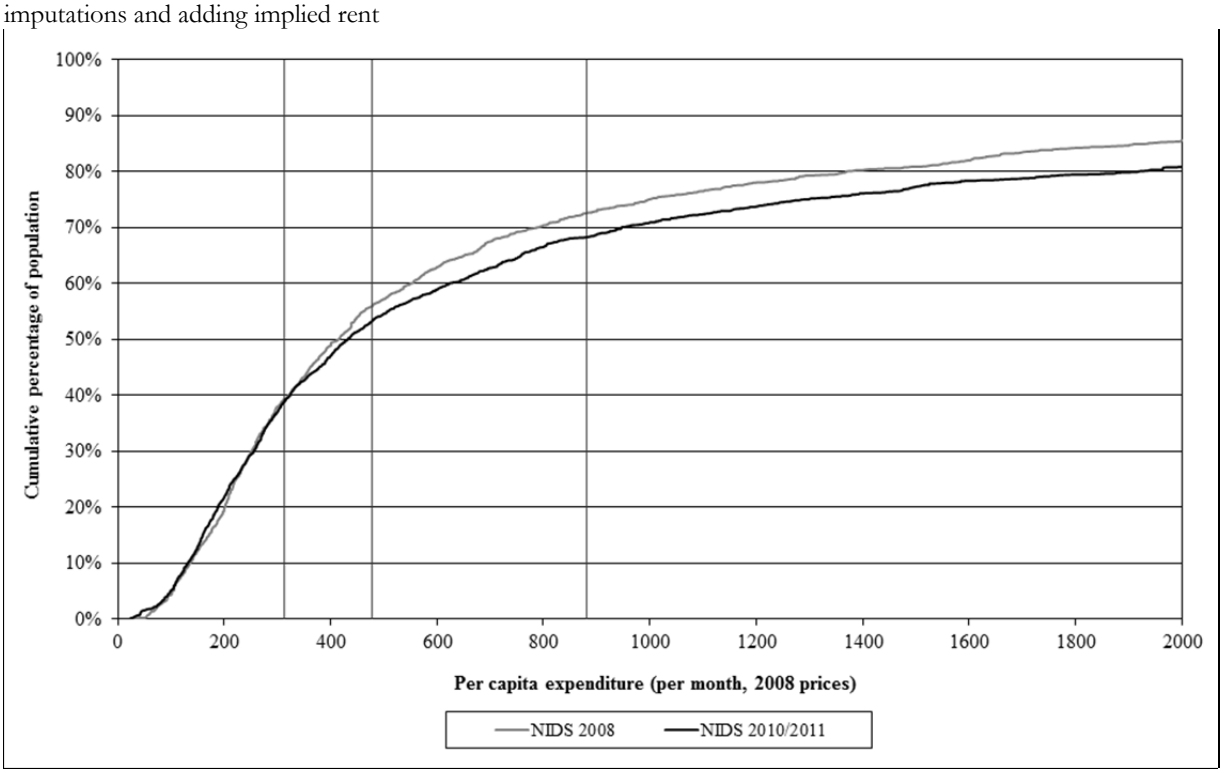


Figure A.11: Lorenz curves: NIDS 2008 income variables

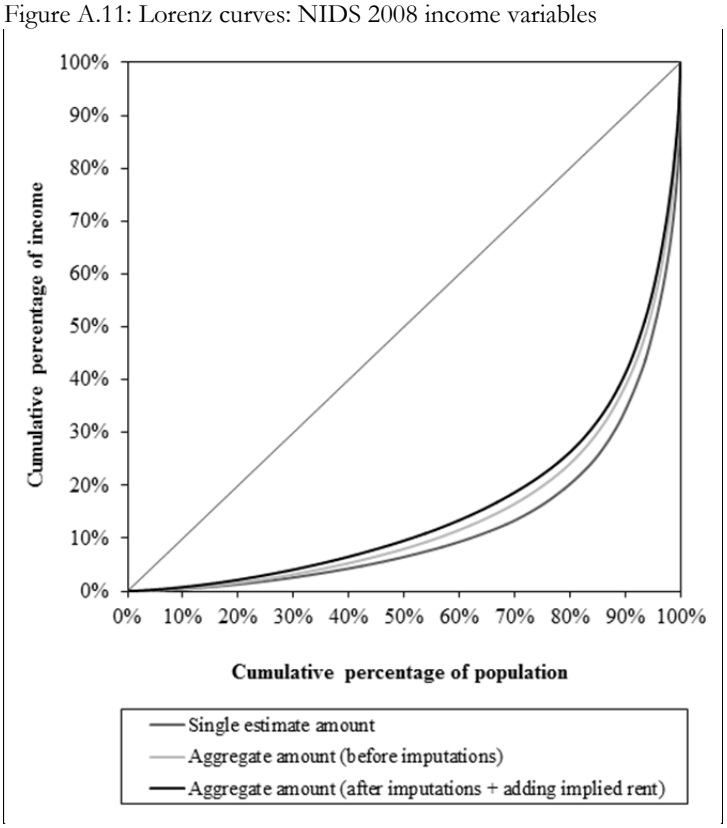



Figure A.12: Lorenz curves: NIDS 2010/2011 income variables

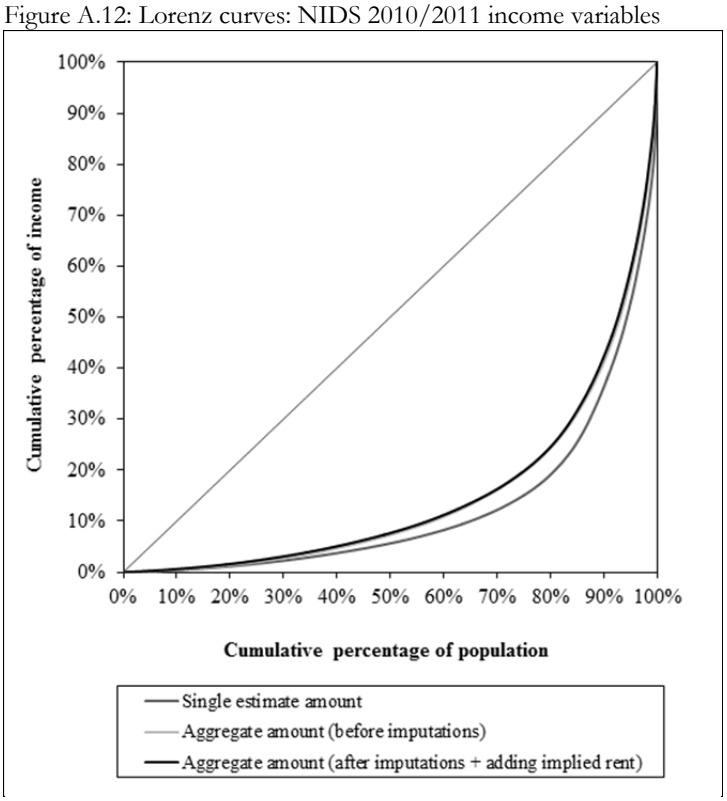


Figure A.13: Lorenz curves: NIDS 2010 expenditure variables

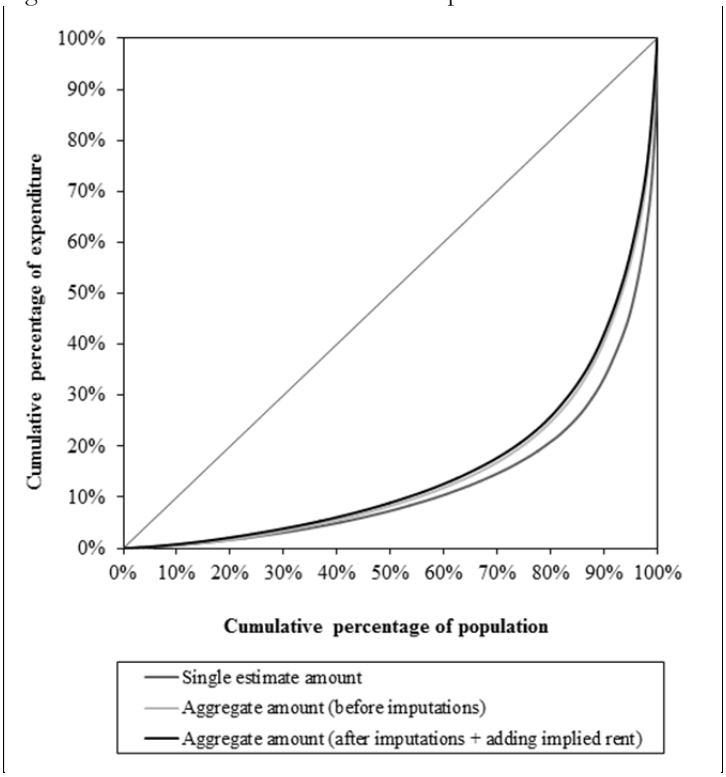



Figure A.14: Lorenz curves: NIDS 2010 expenditure variables