UNIVERSITEIT VAN PRETORIA

UNIVERSITY OF PRETORIA

YUNIBESITHI YA PRETORIA

An overview and assessment of different approaches to poverty measurement in South Africa

by

Phineas Masete Letsoalo

A dissertation submitted in partial fulfilment of the requirement for the degree Master of Science in Mathematical Statistics

in the faculty of

Natural and Agricultural Sciences

at the

University of Pretoria

Supervisor: **Dr. Hermi Boraine**
Co-supervisor: **Mr Andre Swanepoel**

November 2015

Declaration

I, Phineas Masete Letsoalo, declare that the dissertation, which I hereby submit for the degree Master of Science at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE: ……..

DATE: ……..

Acknowledgements

I express my gratitude to GOD for the wisdom, strength, courage, guidance and blessings showered upon me throughout studying my Master's Degree.

I am profoundly indebted to my supervisors, Dr. Hermi Boraine and Mr. Andre Swanepoel, for their guidance, patience and inspiration in preparing this dissertation. They always found time in their busy schedules for discussions and guidance. Their patience is highly appreciated.

I thankfully acknowledge the financial assistance provided by the National Department of Social Development.

I express my thanks to Statistics South Africa for availing the data and the associated metadata.

My deepest gratitude goes to my family for their moral support, love, patience, understanding and encouragement. I am more grateful to my wife, Dikefa and my son, Lethabo.

Summary

Measures of poverty are usually estimated using data from national household surveys. The sample design of official household surveys is typically complex, involving multi-stage stratified cluster sampling. Complex sampling affects variance estimation and therefore standard error estimation. Accounting for the complexities of sampling is essential for reliable estimation and analysis. This study presents an overview and assessment of different approaches to poverty measurement in South Africa. The study researches and presents the estimation of different poverty measures and their standard errors in the case of complex multi-stage sampling, using the Income and Expenditure Survey conducted in 2010-2011 by Statistics South Africa.

The study presents background to poverty measurement in South Africa, the importance of measuring poverty, the general definition of poverty and its definition within the South African context, and theoretical concepts and methods for measuring poverty, in general and in the case of complex samples. The background confirms that there has been significant improvement in South Africa since the fall of Apartheid in addressing the inadequate information base for the measurement of poverty and inequality. The review of literature asserts that poverty measurement is essential in providing statistical standards and systematic approach to reporting on levels, contributes to evidence-based decision making in public policy, and also help in assessing the impact of poverty reduction programmes. The review of literature on the definition of poverty found that there is no single universally accepted definition of poverty as there are so many ways to think about what poverty means. Poverty in South Africa is defined and measured in both one-dimensional and multi-dimensional approaches in line with international practices.

Different poverty indices including the Poverty Headcount Index (P_{HC}), Poverty Gap (P_{PG}), Poverty Gap Index (P_{PGI}), Squared Poverty Gap Index (P_{SPGI}), Sen Index (P_S), Sen-Shorrocks-Thon Index (P_{SST}), Watts Index (P_W) and the time taken to exit poverty will be discussed in the general case and in the case of complex samples. The exact method and Jackknife method of estimating variances in the case of complex surveys are presented. An overview of different data sources for poverty measurement in South Africa will also presented.

The distributional properties of a variable used to measure poverty, *in-kind consumption*, suggest that in South Africa, it is critical that poverty is analysed at disaggregated level. The study conducts multiple comparisons using z-test and Bonferroni adjusted confidence intervals to test hypothesis of differences in estimated poverty by gender, population group, settlement type and province of the head of household. The test show significant differences in poverty between these subgroups. The study concludes that it is important to report poverty estimates with standard errors and confidence intervals as these allows inference when analysis of poverty is conducted over time.

Definition of key terms

In order to allow the reader to understand the content of the study, this section provides definitions of key terms used in this study.

Census: A survey in which the entire population is measured.

Classification of Individual Consumption by Purpose (COICOP): International system of classification of goods and services based on individual consumption by purpose.

Consumer Price Index (CPI): An index that measures the price of a fixed basket of consumer goods and services.

Dwelling unit (DU): Structure or part of a structure or group of structures occupied or meant to be occupied by one or households.

Enumeration area (EA): The smallest geographical unit (piece of land) into which the country is divided for census or survey purposes.

Household: A group of persons who live together and provide themselves jointly with food and/or other essentials for living, or a single person who lives alone.

Household head: A person recognised as such by the household, usually the main decision-maker, or the person who owns or rents the dwelling, or the person who is the main breadwinner.

Master Sample (MS): A sample drawn from a population for use on a number of future occasions, so as to avoid ad hoc sampling on each occasion.

Primary Sampling Unit (PSU): Geographical area comprising one or more enumeration areas of the same type (and therefore not necessarily contiguous) that together have at least one hundred dwelling units.

Sampling frame: A list, map, or other specification of units in the population from which a sample may be selected.

Sampling unit: A unit that can be selected for a sample. We may want to study individuals, but do not have a list of all individuals in the target population. Instead, households serve as the sampling units, and the observation units are the individuals living in the households.

Sampling weight: The sampling weight of unit i is interpreted as the number of population units represented by unit i.

Standard error (SE): The square root of the estimated variance of a statistic.

Table of Contents

Chapter 1: Introduction

1.1. Introduction

The year 2015 marks twenty-one years since the end of Apartheid in South Africa. The first democratically elected government inherited a country described by the World Bank as one of the world's most unequal economies, with a Gini co-efficient of 0.58 (Hunter et al 2003). The country was also characterised by huge inequalities in the quality of education, healthcare and basic infrastructure, such as access to safe drinking water, sanitation and housing. For instance, while only a quarter of all Blacks had access to piped water in their houses, Indians and Whites had universal access in 1995 (Hoogeveen and Özler 2004). While there has been significant progress in provision of quality education, healthcare, housing and provision of basic services (Leibbrandt et al 2006), South Africa still faces the triple challenges of unemployment, poverty and inequality.

The government's commitment to reducing poverty and inequality was articulated well in the Reconstruction and Development Programme (RDP) commitment to "meeting basic needs" that informed the democratic government's policy framework from 1994. This commitment was reiterated in 2004 by former President Thabo Mbeki during his Inauguration Address, when he said:

> "*Endemic and widespread poverty continues to disfigure the face of our country. It will always be impossible for us to say that we have fully restored the dignity of all our people as long as this situation persists. For this reason, the struggle to eradicate poverty has been, and will continue to be, a central part of the national effort to build the new South Africa*". (Mbeki 2004)

The commitment was reiterated again in the National Development Plan (NDP) and it reads, "No political democracy can survive and flourish if the mass of our people remain in poverty, without land, without tangible prospects for a better life. Attacking poverty and deprivation must therefore be the first priority of a democratic government" (NDP 2030, 2011, p.14). The two fundamental objectives anchored in South Africa's current guiding framework for development, the NDP and Vision for 2030, are elimination of poverty and reduction of inequality. Therefore the elimination of poverty and reduction of inequality have always remained at the forefront of South Africa's political agenda. If poverty reduction is to be effective, the nature of poverty and income inequality needs to be better understood. The NDP acknowledges that attacking poverty and deprivation remains the priority of government. However, the positive impact of provision of 'social wage' in helping reduce

poverty in the country is expressed. The 'social wage' is a combination of several poverty reduction and social development programmes being implemented by government in order to improve the lives of the poor and reduce their cost of living (StatsSA 2014a). This social wage includes free basic services such as primary health care; no-fee paying schools, RDP housing, social grants and provision of piped water, sanitation and electricity to households.

Depending on the approach used in measuring poverty, the poverty line, the source of data, whether poverty is measured at the household or individual level, it can be observed that the estimated level of poverty in South Africa differs over a range between 8% and 45.5% (StatsSA 2014a; StatsSA 2014b; Leibbrandt M et al 2010). The wide gap in these poverty estimates is an area of concern, given that globally, there are analytical tools to identify and locate the poor, to describe their characteristics and to measure the extent of poverty at different levels of aggregation. Another concern is that these estimates are sometimes reported without mentioning possible sampling and non-sampling errors. As argued by Houghton and Khandker (2009), it is more useful and honest to report poverty estimates with sampling errors than reporting only the point estimate. Another good practice for reporting poverty estimate from survey data is to report confidence intervals, which provide a range within which the population poverty estimate is likely to fall with a specified level of confidence, something that is not constantly practiced in South Africa. It is also difficult to measure progress in poverty alleviation over time.

This study focuses on an overview and assessment of the different approaches to poverty measurement in South Africa. Among others, the overview includes importance of measuring poverty, the definition of poverty, concepts related to poverty, different types of poverty, different approaches to poverty measurement and a brief overview of different data sources used for poverty measurement in South Africa. As part of the assessment of different approaches to poverty measurement in South Africa, Income and Expenditure Survey 2011 data will be used to calculate different poverty indices, their standard errors and confidence intervals. It is however critical that background to poverty measurement in the country and the importance of measuring poverty is discussed before any of the above can be done. The justification is that having an understanding of where the country comes from in measuring poverty as well as the importance thereof will assist in assessing the approaches currently employed in measuring the levels of poverty in South Africa.

1.2. Background to poverty measurement in South Africa

The analysis of poverty and inequality in South Africa dates a way back with the First Carnegie Inquiry on poverty undertaken in 1922 focusing on the 'poor white' problem. Much later, in 1983,

the Second Carnegie Conference was held and it examined poverty amongst South Africa's black population and highlighted the appalling conditions in the rural areas and townships of South Africa (May 1998). As a starting point to address the inadequate information base in South Africa for the measurement of poverty and inequality, the Project for Statistics on Living Standards and Development (PSLSD) was undertaken in 1993 with the objective of providing a quantitative base-line survey. The PSLSD culminated into Living Standards Measurement Survey (LSMS) for South Africa. The survey came to be known as the SALDRU 1993 dataset. This dataset became the first nationally representative micro-dataset. Also in 1993, the then Central Statistical Service (CSS) now known as Statistics South Africa (StatsSA) ran the first October Household Survey and has continued to do so annually.

In 1995, a detailed Income and Expenditure Survey (IES) was conducted in conjunction with the October Household Survey, making this an extremely rich data. The South African Participatory Poverty Assessment (SAPPA) was also conducted in 1995/96 by Southern Africa Labour and Development Research Unit (SALDRU), to provide a fuller and more integrated understanding of poverty from the perspective of those who are poor. The efforts to develop baseline data for poverty measurement continued when the annual October Household Survey was linked to the five year IES. This linkage allowed for the development of a large database by means of which to compare household income and expenditure with living conditions and life circumstances. With the availability of data, it became clear that poverty threshold (in the form of single or multiple poverty lines) was needed to monitor trends in this dimension of poverty.

In 2000, Statistics South Africa published a report titled '*Measuring Poverty in South Africa*' and this was the first official step, under the democratic government, in presenting a multi-dimensional approach to poverty analysis in South Africa. The report gave an analysis of the 1996 Population Census income data in comparison to the 1995 IES income data (StatsSA 2000). The formal proposal of the official poverty lines was done in 2007 when Statistics South Africa and National Treasury released a report which was reviewing all methodological issues and recommendations based on the previous work done. The proposal of official poverty lines for South Africa was to keep up with the practice in many other countries as well as to respond to its international obligations. The proposal was to construct the official poverty line as a measure of the money income required to attain basic minimal standard of living (StatsSA & National Treasury 2007). The technical process on the development of poverty lines in South Africa was concluded in 2008 when Statistics South Africa published a methodological report on the development of poverty lines for statistical reporting. The three pilot poverty lines (PPL) which were to be piloted in the period leading to 2014 were presented

in the report. The three lines are the Food Poverty Line, and the Upper and Lower bound Poverty Lines (StatsSA 2008).

The efforts to respond to demands for regular and quality poverty data continued when StatsSA developed and implemented a multi-topic, user guided poverty survey known as the Living Conditions Survey (LCS). The LCS was first conducted between September 2008 and August 2009 and this survey was the first tool specifically designed to measure the multi-dimensional nature of poverty. Several poverty analysis reports (StatsSA 2012b, StatsSA 2012c and StatsSA 2013) have been published using LCS 2008/2009 data. Most recently, StatsSA published two reports (StatsSA 2014a and StatsSA 2014b). The former report examines poverty from a money-metric perspective of households using the IES 2005/2006, the LCS 2008/2009 and the recent IES 2010/11 against the three poverty lines (Food Poverty line, Lower bound poverty line and the Upper bound poverty line). The report presents individual poverty, household poverty and household expenditure (StatsSA 2014a). The latter provides poverty maps and poverty data at provincial and municipal levels. The report details how the South African Multidimensional Poverty Index (SAMPI) was conceptualised and constructed using data collected by Statistics South Africa (Stats SA) through the censuses of 2001 and 2011. The SAMPI is based on the Alkire-Foster methodology but domesticated to include an additional dimension of economic activity. This is in line with the UN General Assembly [1] call to countries to develop national indices that reflect country-specific conditions and needs (StatsSA 2014b).

The three poverty lines (Food Poverty line, Lower bound poverty line and the Upper bound poverty line) have been used since their first publication in 2008 and updated annually using the consumer price index (CPI) data. In addition to the need for updated national poverty lines, Stats SA continuously experienced increasing demand for provincial poverty lines and poverty lines for rural and urban areas. In 2015, a methodological report on rebasing of poverty lines was released (StatsSA 2015). The need for rebasing the national poverty lines was informed by the fact that spending and consumption patterns change over time, which implies that, the basket of goods and services on which the existing poverty lines are based may have changed. It was therefore necessary to update estimates using recent consumption data in order to make sure that the lines remain relevant and accurate. The methodological report also published pilot poverty lines for provinces, which were derived following the same approach used for the national poverty lines. However, in deriving the pilot poverty lines for provinces, the national reference food basket was subjected to province-

[1] United Nations General Assembly 2013 to report progress towards the current Millennium Development Goals

specific prices for food items. In addition, factors such as consumption patterns, average household size of the province, and household composition were considered in deriving the province-specific food poverty lines. Data from the Income and Expenditure Survey 2011 was used to rebase the national poverty lines as well as to provide pilot poverty lines for all the provinces. Due to lack of data that disaggregate prices for food and non-food items according to rural/urban places, the separate poverty lines for urban and rural areas could not be provided. These are efforts to improve poverty measurement for the country and to align itself with the growing international trend towards measuring poverty beyond the traditional money-metric method.

1.3. The importance of poverty measurement

If poverty reduction and social development programmes are to be well designed and effective, the nature of poverty, vulnerability and income inequality, and their shifts in response to economic trends and policy, need to be better understood. The Poverty and Inequality Report argues that measuring poverty enables reviews of the extent and nature of poverty and provides for an assessment of policy frameworks and programmes for the reduction of both poverty and inequality. The latter provides clear conceptual and practical guidance concerning the issues which need to be taken in consideration in the formulation of policies and programmes and their implementation (May 1998). The World Bank in its World Development Report 2000 (World Bank 2000) gives reasons why it is important to measure poverty. These reasons are also reiterated in Houghton and Khandker (2009). Poverty measurement is essential for a number of reasons including:

- design of poverty reduction policies and other government interventions; Using credible measures of poverty help focus the attention of policy makers on the living conditions of the poor;
- providing statistical standards and a systematic approach to reporting on poverty which includes but are not limited to indicators of poverty levels (e.g., the headcount index poverty depth) and inequality indicators (e.g. the Gini-coefficient and decile shares);
- monitoring and evaluating the collective impact of poverty and inequality reduction programmes. If the country has a poverty measure, the country will be able, at appropriate intervals, to evaluate whether the poverty programmes are being effective and whether the well-being of people is improving, both in the short term and over an extended period of time. This would include the evaluation of the effectiveness of institutions whose goal is to help poor people. Evidence is needed to tell if institutions such as government are doing a good job of combating poverty. This is not possible unless poverty is measured.

- to enable evidence-based decision making in public policy; if the measure of poverty is put in the public domain, it can help in building a national commitment to eradicate poverty that goes beyond government (Studies in Poverty and Inequality Institute (SPII) 2007).

To assess the impact of poverty reduction programmes, it is necessary to capture changes in the circumstances of potential beneficiaries – both the general population and groups that are known to be deprived. This can be achieved by monitoring the poverty status of individuals, vulnerable groups, communities and society as a whole. An appropriate index to assist in measuring and tracking poverty over time is therefore a useful statistical instrument for research and analysis.

Understanding poverty dynamics as well as measuring and tracking poverty over time can contribute to effective poverty eradication. If a country is able to measure poverty it can also begin to map geographically where poverty is more severe and so direct resources accordingly; The understanding of various dimensions of deprivation experienced by people living in poverty assists government to focus its resources on specific programmes such as basic services, etc. The geographic mapping enables identification of the poor and targeting of appropriate interventions; This also helps to keeping poor people on the agenda because poor people are easily ignored if they are statistically invisible.

It is emphasized in Sen (1976) that a particular way of measuring a phenomenon should depend on the purpose to which the resulting measure will be used. In the case of poverty measurement, three possible purposes and applications of measuring poverty are discussed as follows:

- *Assessment:* Poverty may be measured by a government to provide a continuous assessment of how its various policies are affecting the conditions of the poor;
- *Diagnosis:* Poverty can be measured to help uncover the causes and correlates of poverty in order to formulate policies to fight poverty. Diagnosis generally includes more dimensions of poverty than only income;
- *Targeting:* In addition to the above purposes, a standard use for the poverty measurement is to enable governments to identify individuals or families as being in poverty and thereby focus services and policies directly upon them.

As discussed in Section 1.1, the elimination of poverty and reduction of inequality remains the priority of government in South Africa. It is therefore important that poverty is measured in order for government to target the 'social wage' and other development programmes to the relevant beneficiaries. In addition, measuring poverty will assist in monitoring progress and assessing the impact of all poverty alleviation programmes in the country.

1.4. Problem Statement

Different approaches employed in the measurement of poverty in South Africa provide different levels of poverty. Depending on the poverty measurement approach used, the poverty line, the source of data, whether poverty is measured at the household or individual level, it can be observed that the levels of poverty in South Africa differs over a wide range (8% - 45.5%) (StatsSA 2014a; StatsSA 2014b; Leibbrandt et al 2010). The wide gap in these poverty estimates is an area of concern. Furthermore, these poverty estimates are more often presented as point estimates without sampling errors and confidence intervals. As argued by Houghton and Khandker (2009), it is more useful and honest to report poverty estimates with sampling errors and confidence limits than reporting only the point estimate. This is not constantly practiced in South Africa. Theoretically, if all the approaches employed in poverty measurement resulted in the same levels of poverty, then any approach can be used to measure poverty. However, this is not the case depending on the approach adopted. This can be misleading, especially when reporting poverty estimates at country level on different international obligations such as the Millennium Development Goals, the United Nations Development Program's Human Development Report, World Bank's World Development Report or any other international obligations. This can also be misleading locally in several areas of strategic importance such as, formulation of policies to fight poverty, assessment of policy frameworks and programmes for the reduction of both poverty and inequality, and in enabling government to identify individuals or households who are poor and thereby targeting services to them.

Given this situation, this study seeks to give an overview and assessment of the different approaches to poverty measurement currently employed in South Africa. The overview includes different data sources used for poverty measurement in South Africa but focusing in the Income and Expenditure Survey (IES). The overview includes the name of the survey, the custodian, the focus, coverage, variables, methodologies, the frequency or time periods in which the survey was conducted as well as the purpose of each survey. The assessment on the other hand will focus on one dimensional poverty measures. The assessment will include calculating the estimators of poverty measures based on complex sample; calculating standard errors and confidence intervals of estimators of poverty; analysis of in-kind consumption as a measure household well-being as measured by IES 2010/11; sample distribution as presented by means of Cumulative Distribution Function (CDF) and quantile function; calculating poverty measures for the whole sample but also comparing for different groups (Gender, population group, settlement type and province). Multiple tests are conducted using Bonferroni adjusted confidence intervals to test hypothesis of differences in estimated poverty by gender, population group, settlement type and province of the head of household.

1.5. Purpose of the study

The primary objective of this study is to contribute towards a better understanding of different approaches to poverty measurement in South Africa, and, at the same time, assess these approaches. The study presents the estimation of different poverty measures and their standard errors and confidence intervals in the case of complex multi-stage sampling design as opposed to the standard case of simple random sampling. The objectives of this study could be broken down into the following components:

(i). a discussion of the background to poverty measurement in South Africa. In this regard, the attainment of democracy by South Africa will be used as a reference point;
(ii). a discussion of the importance of measuring poverty, as stated in the literature;
(iii). a presentation of aspects to consider in defining poverty;
(iv). a discussion of different approaches to poverty measurement;
(v). presentation of different indices of poverty measurement;
(vi). a discussion of strength and limitations to survey data in poverty analysis;
(vii). estimating and analysing South African household consumption as an indicator of well-being
(viii). how to estimate standard errors of sample statistics in the case of complex multi-stage sampling as opposed to the standard case of simple random sampling;
(ix). how to present poverty profiles and poverty analysis over time;
(x). an overview of key data sources used in measuring poverty in South Africa.

All these objectives will be accomplished with a view to proposing how best to analyse, interpret and report poverty measures from survey data.

1.6. Organisation of the Study

The study is organised into five chapters. The first chapter gives an introduction to the study and discusses background to poverty measurement in South Africa. The chapter establishes the rationale for studying the subject matter and also discusses the importance of measuring poverty in general. The second chapter presents a review of relevant literature on the definition and measurement of poverty. Aspects such as space and time horizon over which poverty is defined, the multidimensionality of poverty, the unit of poverty measurement and different classifications of poverty are discussed. Chapter 3 presents theoretical concepts and methods of measuring poverty. The methodology includes discussion of different poverty indices and their properties, how to estimate poverty from household survey data, strengths and limitations to survey data in poverty

analysis, how to estimate variances of sample statistics in the case of complex multi-stage sampling, how to present poverty profiles and how poverty analysis over time can be done and presented. Chapter 4 analyses poverty based on the South Africa's 2011 Income and Expenditure Survey. Estimators of poverty measures, their standard errors and confidence intervals are calculated based on in-kind consumption as an indicator of well-being. A sampling distribution is presented by means of Cumulative Distribution Function and quantile function. Poverty measures are calculated for the whole sample, but also compared by gender, population group, settlement type and by province. Chapter 5 summarises the findings of the study and gives concluding remarks and recommendation.

Chapter 2: Defining poverty

2.1. Introduction

Chapter 1 presented background to poverty measurement in South African and why it is important to measure poverty. This chapter presents a review of relevant literature on the definition and measurement of poverty which is essential in understanding the remainder of the study. This chapter includes: aspects to consider in defining poverty, space and time horizon over which poverty is defined, the multidimensionality of poverty, the unit of poverty measurement, different classifications of poverty which includes discussions of chronic and temporary poverty, subjective and objective poverty, and absolute and relative poverty. These aspects are critical to discuss as they fundamentally underlie any discussion of definition of poverty. Concepts such as inequality, vulnerability and deprivation which are related to poverty are also discussed. These concepts are sometimes used interchangeably, but it is important to understand what each means and how they complement each other in their definitions. The chapter also discusses different general approaches to poverty measurement. Both one dimensional and multidimensional measurement of poverty are defined as they complement each other. The definition of poverty within the South African context is given, which is the money-metric approach for one dimensional approach and the South Africa Multidimensional Poverty Index (SAMPI) for multidimensional approach. An overview of different approaches to poverty measurement is introduced in this chapter- each worthy to be analysed in more detail. The chapter concludes by selecting a specific approach to poverty measurement which is applied in this study to estimate and do inference on poverty.

2.2. Defining poverty

2.2.1. Aspects to consider in defining poverty

There are a number of general aspects to consider when defining poverty and these apply to all approaches in defining and measuring poverty. It is therefore critical to discuss these aspects in general terms as they fundamentally underlie any discussion of definition of poverty.

The initial aspect to consider in defining poverty is that there is no single universally accepted definition of poverty (Townsend 2004). The definition of poverty has been the subject of debate and discussion among experts and policy analysts (May 1998). There are many difficulties inherent in defining poverty. There are so many different ways to think about what poverty means. This lack of agreed definition leads to conceptual and measurement issues which remains to be addressed or

clarified. Poverty is therefore an unclear concept without a single definition (Naidoo 2007). Magasela (2005) argues that most researchers claim their understanding of poverty as the correct one based on logical argument or scientific research. However, in exploring the problem of understanding poverty, it was evident that there is no one correct, scientific, agreed definition. This is because poverty is a socio-economic concept and thus inherently a contested one and therefore the first thing to understand is that poverty is not a simple phenomenon which we can learn to define by adopting the correct approach. It is a series of contested definitions and complex arguments which overlap and at times contradict each other (Magasela 2005). While the definition of poverty is not agreed, the understanding of the concept of poverty has improved and expanded over the years. Currently, there are analytical tools to identify and locate the poor, to describe their characteristics and to measure the extent of poverty at different levels of aggregation. There is also consensus that poverty is disapproved of and its elimination is regarded as morally good.

The space over which poverty is defined

The definition of poverty depends on the society to which it is to be applied (Magasela 2005). When a definition of poverty is chosen, it denotes specific characteristics with which to identify the poor. These characteristics thus defines what is acceptable (or unacceptable) in a society and says a great deal about the way the society would like things to be. (Stewart et al. 2007) argues by asking whether we expect definitions of poverty applied to one type of society to be transferrable to other societies. This is referred to as the question of *universality* of the definition of poverty (Stewart et al. 2007). It is however concluded that the interpretation of the definitions will differ between societies with radically different characteristics. In both arguments, the state of being in poverty is directly related to lack of an acceptable quality of life. (Noble et al. 2004) argues that the concept and definition of poverty in a society is like a mirror-image of the ideals of that society. It is therefore vital that concepts and definition of poverty, as well as being theoretically robust, is appropriate to the society in which it is to be applied. Having agreed on a definition or definitions, the method of measurement must appropriately operationalise the definition. (Sen 1979) defines poverty according the conventions of the society in which it occurs. Poverty is defined as a situation in which there is lack of essential facilities, resulting from inadequate income. Since there is a socially accepted minimum level of living in every society, those who live below this minimum level are said to live in poverty. (Townsend 1979) in his study on Poverty in the United Kingdom argued that individuals, families and groups in the population can be said to be in poverty when they lack the resources to obtain the type of diet, participate in the activities and have the living conditions and amenities which are customary, or at least widely encouraged or approved, in the societies to which

they belong. Their resources are so seriously below those commanded by the average individual or family that they are effectively excluded from ordinary living patterns, customs or activities.

The time horizon over which poverty is defined

How poverty is experienced is different from time to time (Naidoo 2007). This makes it difficult to formulate its definition because poverty might be experienced differently during different time periods. Poverty experienced in South Africa today is different from the poverty experienced in South Africa 100 years ago. (Stewart et al. 2007) defines this as the *time* horizon over which poverty is defined. Time horizon is viewed as a technical issue since people move in and out of poverty over seasons and years. (Kanbur 2001) argues that according to normal usage, poverty is "The state of one who lacks a usual or socially acceptable amount of money or material possessions." In this definition, it is emphasised that the definition of poverty will be different at different times and in different societies and what is "socially acceptable" in, say, South Africa today may differ from what was socially acceptable 100 years ago.

The multidimensionality of poverty

The understanding of poverty can be broader than the extent of low income or low expenditure. The understanding can include the denial of opportunities and choices most basic to human development to lead a long, healthy, creative life and to enjoy a decent standard of living, freedom, dignity, self-esteem and respect from others (StatsSA 2000). (Stewart et al. 2007) poses the universal question of how to deal with multidimensionality considering that individual well-being (and lack of it) manifests itself in multiple dimensions. Although there is consensus that poverty needs to be understood in a multidimensional manner (World Bank 2000), there is no consensus as to what the dimensions of poverty should be or how many dimensions are adequate. Some examples of dimensions of poverty are: a lack of nutrition, housing, safety, clothing and health, income, education, literacy and clean drinking water. Similarly, each dimension contributes more to poverty than the other, depending on the time and place; this is referred to as the *horizontal vagueness* of poverty (Naidoo 2007). There is no consensus on where or how to distinguish between the poor and those who are not poor in each dimension. There is also no consensus as to what level of each dimension is acceptable, since the requirements of a society may differ from place to place; this is referred to as the *vertical vagueness* of poverty (Naidoo 2007).

The United Nations Development Programme (UNDP) in its 1997 Human Development report defined poverty in the human development perspective. Poverty was defined to reflect poor health and education, deprivation in knowledge and communication, inability to exercise human and

political rights and the absence of dignity, confidence and self-respect. Several measures including the two popular measures of Human Development (HD), the Human Development Index (HDI) and the Human Poverty Index (HPI) were introduced.

The unit of poverty measurement

The unit over which poverty is defined is discussed by (Stewart et al. 2007) as one of the problems encountered in defining and measuring poverty. Poverty can be defined at the individual, household, community, country, region or global level. Although it is individuals who experience poverty or who are defined to be in poverty, data, be it monetary or pertaining to access to services (i.e. water, sanitation, electricity, etc.) is normally at household level. It is often difficult to ascertain the distribution of some of the services to individuals. When defining poverty, the unit of definition is important for a few reasons. Firstly, in identifying the society with respect to which poverty lines are drawn. Secondly, to define boundaries or society which access to services (or lack of) is defined or boundaries of the relevant market, for example, in obtaining prices for valuations. Thirdly, if geographic areas are used for targeting, then how well the areas are defined will affect the efficiency of targeting. The unit over which poverty is defined also informs those in power how to source and allocate resources directed at its eradication (Stewart et al. 2007).

A method commonly used to measure poverty at country level is based on income or consumption levels measured at household level. Personal income or consumption is typically based on the total income of the household divided by the total number of household members sharing that income/consumption. A person is considered to be in poverty if his or her consumption or income level falls below some minimum level necessary to meet basic needs, i.e. poverty line. At the global level, the World Bank uses reference poverty lines set at $1.25 and $2 per person per day (in 2005 Purchasing Power Parity terms) to estimate poverty worldwide, and express it in a common unit across countries. In this study, the unit over which poverty is defined and measured is mostly a household.

Chronic and temporary poverty

Experiences of poverty can be over a long term or for a short period of time. Chronic poverty describes the state of poverty that exists over time. People are “chronic poor” if they experience poverty for an extended period of time or throughout their lives. Chronic poverty is the more difficult to address and is often associated with persistent inter-generational poverty, i.e. chronic poverty is likely to be transferred across generations. In contrast, temporary or transitory poverty refers to a state in which people are able to move out of poverty after a short spell. This may result from a one-time decline in living standards (for example following the loss of a job), from which a

household gradually emerges. This type of poverty may show itself in fluctuations in well-being that result in frequent declines in living standards. For example, natural events such as disaster or major policy changes by governments may plunge a household into poverty. The concepts of chronic and transitory poverty are important because they focus attention on the dynamics associated with escaping poverty or staying poor. The aspect of transitory poverty should be taken into consideration in the development of measurement instruments, for example, by doing cohort studies. The tracking of evolution of poverty over time is useful in evaluating the effects of policies and programs. For example, policy experts might be interested in knowing how interventions to targeted regions have impacted on the poverty rates in those regions. The comparison of poverty rates before the implementation of intervention and after the implementation would therefore be needed in order to evaluate the effect of these interventions

Subjective and objective poverty

Poverty can be defined or measured in subjective or objective terms. The subjective approach to poverty considers the people's perception of what constitutes their wellbeing. The uniqueness of the subjective approach to poverty measurement is that the threshold between poor and non-poor is determined on the basis of people's perception of their own well-being. In the measurement of poverty, the subjective approach can be used either in monetary or non-monetary contexts. A survey to gauge the population's opinion is carried out in order to define the poverty line. The best-known method for measuring subjective poverty is based on a Minimum Income Question (MIQ), such as "what do you, in your circumstances, consider to be an absolute minimum income for your family?" Other known methods are the Income Evaluation Question (IEQ) and the Consumption Adequacy Question (CAQ). A frequently cited advantage of the subjective approach to poverty measurement is that it is free from arbitrariness, since the classification of the poverty line is derived directly from the population itself and not indirectly by comparing to some threshold.

The objective approach to poverty measurement can either be absolute or relative. These approaches are defined as follows:

Absolute and relative poverty

Poverty can be defined in absolute or relative terms. In absolute terms, poverty usually refers to a state of deprivation defined in relation to an objective, invariant and value free external definition of basic human needs. The standard of absolute poverty does not change according to prevailing living standards of a society, or over time, or according to needs of different groups in society. The existing and commonly used approach is through the development of 'poverty lines'. Two approaches, the cost-of-basic-needs approach and the food energy intake method, are commonly used. The cost-of-

basic-needs approach first estimates the cost of acquiring enough food for adequate nutrition, and then adds the cost of other essentials such as clothing and shelter. An alternative food energy intake method is based on calculations of the income needed to secure minimum human calorie requirements, and does not take into account any of their other non-food needs. The World Summit for Social Development in Copenhagen in 1995 defined poverty as

> '*a condition characterized by severe deprivation of basic human needs, including food, safe drinking water, sanitation facilities, health, shelter, education and information. It depends not only on income but also on access to social services*' (World Summit for Social Development, 1995)

This definition suggests absolute poverty measure based on many dimensions.

Relative poverty, in contrast, explicitly relates poverty to a reference group. Rather than referring to an objective standard, this definition links poverty to a particular point in relation to the national distribution of a particular variable such as income/expenditure. People are considered to be poor in comparison to those around them. This definition of poverty is more relevant when debates on societal justice do not merely focus on the needs of bare survival, but also on the question of inequality in society. The definition of poverty in relative terms changes over time and from place to place. In its 2007 Human Development Report (UNDP 2007), the United Nations Development Programme defines absolute poverty as some absolute standard of minimum requirement while relative poverty was referred to as falling behind most others in the community. Using income, a person is defined as absolutely poor if her income is less than the defined income poverty line, while she is relatively poor if she belongs to, say a bottom income group (such as poorest 10% of the population). (Townsend 2004) also gives the distinction between absolute and relative poverty. Absolute poverty is referred to a lack of the needs for physical subsistence, i.e. minimum needs necessary for the 'maintenance of physical health' and 'physical efficiency'. Relative poverty extends the concept of poverty to consider individuals as social beings, who have psychological needs to participate in a society and share in its customs and norms.

Definition of poverty

Poverty can be defined as the condition of living below socially accepted living standards generally connected to suffering and shortage of a wide range of resources. The application of the definition consists of shortage of material need, which includes the lack of necessary goods and services, a variety of deprivations, and a repetitive form of deficiency in a specific time period. Poverty is therefore briefly described as a situation of living without the essential goods and services for a proper well-being such as sufficient shelter, foodstuff, a job, adequate earnings, access to essential public services and societal standing. What is clear in defining poverty is, firstly, that poverty and

the poor are associated with a state of want and deprivation and, secondly, that such deprivation is related to the necessities of life. Therefore the definition or the use of the term "poverty" implies a comparison between the condition of a person, household or a society on the one hand, and the perception of the person who speaks or writes about what is necessary to sustain life on the other hand. This also implies the comparison of the perception of who defines poverty on one hand, and the perception of the society on what is acceptable or the way the society would like things to be, on the other. The emerging consensus on its definition sees poverty as generally characterised by the inability of individuals, households or communities to command sufficient resources to satisfy a socially acceptable minimum standard of living. The above definitions make it clear that poverty may be understood in a narrow or broad sense. In the narrowest sense it means lack of income. In a broader sense poverty can be seen as multidimensional, encompassing other issues such as housing, health, education, access to services and to other avenues of accessing resources.

Discussed above are several aspects to consider when defining poverty. There could be considerable debate as to whether poverty should be regarded as absolute or relative; or whether it should be measured as necessities or capabilities or functions; or whether it is only a monetary phenomenon, whether is one dimensional or multidimensional. In order to identify the poor, what they lack and their location, it is necessary that we have ways of defining and measuring as many aspects of the manifestation of poverty as possible in order to be able to develop appropriate and evidence-based policy interventions.

Concepts related to poverty

The concepts of poverty, inequality, deprivation and vulnerability are related. They are often used interchangeably and require brief elaboration since they are not the same. Although sometimes used interchangeably, it is important to understand what each means and how they complement each other in their definitions. Inequality focuses broadly on the distribution of attributes, such as income or consumption, across the whole population. If it is believed that the welfare of individuals depends on their economic position relative to others in society (relative poverty), then for poverty analysis, inequality should be examined. Vulnerability is applied in understanding the risk of falling into poverty in the future, for a person who is not necessarily in poverty at the current time. It terms of perceptions of their own situations, vulnerability is a key dimension of well-being as if affects individual's behaviours in terms of investment, production patterns, and coping strategies. These concepts and the concept of deprivation are elaborated in detail in the next paragraphs.

Inequality: In defining the meaning of 'inequality' within the social context requires consensus on what is meant by 'equality'. The term 'equality' within this context refers to a state of social organisation that gives equal access to resources and opportunities to all its members (May 1998). Inequality is therefore lack of equality as defined above. Inequality is a characteristic of social power relations. If members of different social groups have highly differential power relations, then inequality is present. Inequality is also closely linked to the notion of social exclusion, in that unequal power relations may be linked to differential access to political or socio-economic rights. Within the quantitative and economic sense, inequality can refer to an imbalance in the distribution of particular resources, such as income, in a specific population. In societies which are well-resourced, the existence of poverty can be said to be a manifestation of inequality. The common measure of distributional inequality in a population is the *Gini coefficient.* Other known measures of inequality are the *Theil Index* and *decile shares.* If inequality is to be addressed, the policies adopted to reduce it must contain elements of redistribution of resources from the wealthier to the poorer members of that society. In the context of poverty analysis, inequality requires examination if one believes that the welfare of individuals depends on their economic position relative to others in society.

Vulnerability: Poverty is not a static condition among individuals, households or communities. Others experience chronic poverty while other experience transitory or temporary poverty. The experiences of transitory poverty may be a result of life-cycle changes, specific events such as the illness of a main income earner, or deterioration in external economic conditions. In trying to understand these processes of change, the concept of vulnerability is applied. Vulnerability refers to the negative outcomes of processes of change. These changes may be economic, social, environmental or political, and may be long-term changes, 'shocks' or recurring processes such as seasonality. Vulnerability can also refer to a state of being that is defenceless to threats to the well-being of people. People are vulnerable when they live in a way that, when a shock that they would recover from with relative ease causes a disastrous and hard to reverse reduction in their well-being or access to resources. When vulnerability is analysed, an assessment of assets that can be called on to withstand or mitigate the impact of the threat in question is done. Individuals, households and communities with more and better managed assets are less vulnerable. Those with lesser assets have greater insecurity and associated poverty. The factor which characterise vulnerability is therefore not only a lack of assets and an inability of the poor to accumulate more of different assets but also an inability to devise appropriate management strategies in times of crisis.

Deprivation: According to the *Oxford Advanced Learners Dictionary (2006)*, the noun 'deprivation' is defined as 'the fact of not having something that you need, like enough food, money or a home'. Within the context of poverty, deprivation refers to the effects of poverty on a person's life. People are defined as deprived if 'they lack the types of diet, clothing, housing, household facilities and fuel and environmental, educational, working and social conditions, activities and facilities which are customary'. Deprivation takes into account how being in poverty or being poor limits what a person "can and cannot do" in terms of both immediate and future actions. While poverty refers to the lack of resources required to meet people's needs, deprivation refers to these unmet needs. In assessing deprivation, different types of indicators that go beyond merely measuring lack of income are used. Deprivation is closely linked to the multidimensional measurement of poverty. The multidimensional measurement of poverty incorporates a wide range of indicators to capture the complexity of poverty, since several factors can contribute to a poor person's experience of deprivation.

2.3. Different approaches to poverty measurement

Before poverty can be measured, it has to be defined. How it is defined is extremely important as different definitions of poverty imply use of different criteria for measurement, potentially the identification of different individuals and groups as poor, and the use of different policy interventions for poverty reduction (Stewart et al. 2007). When the definition of poverty is broader, it is more difficult to measure it. Different approaches to poverty measurement are discussed below.

2.3.1. One dimensional and multi-dimensional measurement of poverty

In a one dimensional measurement of poverty, poverty is defined from income, consumption or expenditure. The poverty line is chosen in such a way that any household whose income (consumption or expenditure) falls below this line is considered to be poor. The poverty line defines the level of income (consumption or expenditure) needed for a household to escape poverty. While a one dimensional measurement of poverty is very useful in measuring poverty, it is limited in capturing the multiple aspects that constitute poverty (StatsSA 2014b). There are several factors that can contribute to a poor person's experience of deprivation. These can include poor health, lack of education, inadequate living standards, lack of income (as one of several factors considered), disempowerment, lack of decent work and threat from violence. The multidimensional measurement of poverty aims at incorporating this wide range of indicators to capture the complexity of poverty. As discussed in Chapter 1, when conceptualising poverty, there is also a need to consider vulnerability, inequality, the poverty of categories of people (women, children, older

people, disabled people), and collective poverty (of regions, nations, groups). In addition, the poor describe their experience of poverty as multi-dimensional. This implies that the more policy relevant information there is available on poverty; the better-equipped policy makers will be to reduce it. Multidimensional methods of poverty measurement and definition can also be used for additional purposes such as targeting or conditional transfers (StatsSA 2014b). When these complementarities are understood, they become essential for designing and implementing programs and projects that help people escape poverty.

2.3.2. General approaches to measurement of poverty

(Stewart et al 2007) classifies general approaches to poverty measurement into four. These include the monetary approach, the social exclusion approach, the participatory approach, and the capability approach. The classification of the poverty measurement is an extension of the work of (Ruggeri et al. 2003). (Naidoo 2007) argues that theoretically, if all the approaches identify the same people as being poor, any one of these approaches can be used to measure poverty. However, empirical evidence shows that poverty rates differ significantly depending on the approach adopted (Stewart et al. 2007; Ruggeri et al. 2003). These approaches are briefly discussed below.

The monetary approach

The monetary approach is the most frequently used approach to define and measure poverty. This approach identifies poverty with a shortfall in income, consumption or expenditure; i.e. a poverty line is defined in terms of the monetary income (expenditure) sufficient for a person to attain a minimal standard of living. A person whose income falls below the poverty line is considered to be poor (Naidoo 2007; Stewart et al 2007). The *absolute*, the *relative* and the *subjective* poverty lines are defined. An *absolute* poverty line is fixed at a value (cut-off point) of income or expenditure that is necessary to acquire goods and services regarded as essential for a minimum standard of living. Using the *relative* poverty line approach, a household is defined as 'poor' relative to others in the same society or economy. This line may be set at a value two-thirds of the mean or at the median value, and any household below this value is then regarded as poor. *Subjective* poverty lines are based on households' perceptions of their needs. The advantage of subjective poverty lines over objective ones (absolute and relative) is that though they use income as a monetary indicator of standard of living, they do not require the use of equivalency scales as the household itself takes size into account when providing the information on income. The World Bank estimate for the absolute poverty line is $2 per person per day for developing countries. In South Africa, the poverty line for households was set at R620 per person per month in 2011 prices (StatsSA 2014b).

The social exclusion approach

In defining the Social exclusion approach, we start by defining the concept of social exclusion. Social exclusion is defined as a process through which individuals or groups are wholly or partially unable to participate fully in the society in which they live (Stewart et al. 2007). The social exclusion approach emphasizes relations between individuals. Through this approach, poverty is a social construct and has little to do with the fulfilment of the individual's minimum needs. This is often a characteristic of groups rather than individuals, for example, the aged, the handicapped or particular racial or ethnic categories. The main way in which social exclusion is different from other approaches of poverty measurement is that it focuses attention on social process and social relations, and emphasises the ways in which adverse power relations, discrimination and identity can help to marginalize and impoverish people.

The participatory approach

In a participatory approach to poverty measurement, people are asked to define what constitutes poverty. This approach takes into account the views of poor people themselves. The people decide what it means to be poor and that determines the magnitude of poverty categories. Participatory poverty assessments are open-ended, interactive and qualitative. People are allowed to describe what constitutes poverty in whatever dimension they choose. One of the key advantages of the participatory approach is that it may be considered to be more democratic.

Usually during participatory poverty assessments, two aspects of poverty will emerge, the first one being vulnerability and the second being lack of voice. With vulnerability, the concern is with risk and volatility of incomes. Poverty is expressed not only a state of having little, but also of being vulnerable to losing the little that one has. Lack of voice and political rights, often described as a sense of powerlessness, is revealed when the poor describe their interactions with government employees and institutions. Where there is interaction between those in power and the poor, it is more one sided – with those in power dominating the discourse (Kanbur 2001).

The capability approach

The capabilities approach defines poverty as a deprivation in the space of capabilities. According to this approach, a person who's capabilities or functioning falls below a minimum acceptable standard is considered to be poor. This approach emphasizes that income is only valuable in so far as it maximises the capabilities of individuals and thereby permits them to function in their society. A person can be capability poor but not be income or monetary poor. This approach is much broader and addresses the omissions of social goods in monetary approach.

2.4. Definition of poverty within South African context

Poverty in South Africa is defined and measured in both one-dimensional and multi-dimensional approaches. The one-dimensional approach to poverty measurement used by Statistics South Africa identifies poverty with a shortfall on consumption expenditure. The traditional approach of developing poverty lines and identifying poor persons as those whose consumption expenditure is below the poverty line is used. This approach is called the money-metric or the monetary approach to poverty measurement as discussed in the general approaches above. A set of three national poverty lines – the food poverty line (FPL), lower-bound poverty line (LBPL) and upper-bound poverty line (UBPL) – are used for poverty measurement in the country. The FPL is the level of consumption below which individuals are unable to purchase sufficient food to provide them with an adequate diet. Those below this line are either consuming insufficient calories for their nourishment, or must change their consumption patterns from those preferred by low income households. The LBPL includes non-food items, but requires that individuals sacrifice food in order to obtain these non-food items, while individuals at the UBPL can purchase both adequate food and non-food items (StatsSA 2014a). These poverty lines are universal in that they apply across all provinces in the country and they are updated annually using the Consumer Price Index data to maintain their relevance in long term statistical use.

Statistics South Africa also uses subjective approaches to measures poverty. Three of the most widely used subjective poverty measures are used in order to estimate subjective levels of poverty in South Africa (StatsSA 2012b, StatsSA 2012c). These approaches are the self-perceived wealth question (SPWQ), minimum income question (MIQ) and the income evaluation question (IEQ). The self-perceived wealth question asks respondents to select the category which best describes their households according to an ordinal scale ranging from ‘very poor’ to ‘wealthy’. Households who responds ‘poor’ or ‘very poor’ are identified as ‘poor’ and all other categories are classified as ‘non-poor’. The minimum income question asks respondents to select the smallest level of income with which their household could make ends meet. If the reported per capita household consumption falls below this minimum income level, then the household is identified as poor. The income evaluation question is linked to the minimum income question. Respondents are asked whether or not their household’s actual level of income is above or below the minimum level reported in the minimum income question. Responses are presented in an ordinal scale ranging from ‘much lower’ to ‘much higher’. Households are therefore identified as ‘poor’ if their income is described as ‘lower’ or ‘much lower’ than the minimum required income. All other responses are identified as ‘non-poor’.

Poverty is defined and measured multidimensionally using the South African Multidimensional Poverty Index (SAMPI) (StatsSA 2014b). The SAMPI was conceptualised and constructed based on the global Multidimensional Poverty Index (MPI) (Alkire and Santos 2010). The MPI has three dimensions namely; health, education, and standard of living. The SAMPI adjusted the indicators of the MPI and included an additional dimension of economic activity to reflect country specific conditions and needs. SAMPI complements the consumption expenditure-based poverty measures by capturing the severe deprivations that each person or household faces. It reveals a different pattern of poverty than expenditure based poverty, as it illuminates a different set of deprivations. Poor households are identified and an aggregate measure constructed based on a weighted average of the deprivations they experience using a nested weight structure (equal weight across dimension and equal weight for each indicator within dimensions), i.e. each dimension is equally weighted; each indicator within a dimension is also equally weighted. A household is identified as multidimensionally poor if, and only if, it is deprived in some combination of indicators whose weighted sum is thirty percent or more of the dimensions (StatsSA 2014b).

South Africa also uses Living Standard Measure (LSM) to define and measure poverty (The Presidency 2011). Living Standard Measure divides the population into ten living standard measure groups, where 1 is the lowest and 10 is the highest. It is based on access to services and durables, and geographic indicators as determinants of standards of living. The LSMs are calculated using 29 variables from the South African Audience Research Foundation (SAARF) All Media Products Survey (AMPS). It calculates an imputed average monthly income and national poverty lines are used to estimate poverty levels.

The National Income Dynamics Study (NIDS) conducted by the South African Labour Development Research Unit (SALDRU) measures poverty using income and thus approaches deprivation from a money-metric perspective (The Presidency 2013). Households are defined as poor if their monthly household income per capita is below specified periodic poverty lines. This study also uses multidimensional poverty index (MPI) containing nine indicators spread over the three dimensions of education, health and living standards as defined by Alkire and Foster (2011).

2.5. Conclusion

This chapter covered a review of literature on the definition and measurement of poverty, an analysis that will prove beneficial in understanding the remainder of the study. Some of the important aspects that were covered in this chapter include the following: aspects to consider in

defining poverty, concepts related to poverty, different approaches to poverty measurement and the definition of poverty within the South African context.

With regard to key issues for consideration, it was highlighted that poverty is a phenomenon without a single definition. Furthermore, it has also been established that the space and time over which poverty is defined and the unit of measurement determines how poverty is measured. It has also been established that poverty is a multidimensional phenomenon with different meanings to different people hence, the importance of identifying concepts that are closely related to poverty, as well as different types of poverty.

With regards to measurement issues, it was noted that, just like defining poverty, there is no unique way to measure poverty in a society. Four common approaches to poverty measurement were discussed. The first approach, monetary approach, measures poverty with regard to the shortfall in income, consumption or expenditure. The second approach, the social exclusion approach, focuses attention on social process and social relations, and emphasises the ways in which adverse power relations, discrimination and identity can help to marginalize and impoverish people. The third approach, the participatory approach, takes into account the views of poor people themselves. The fourth and last approach, the capabilities approach, defines poverty as a deprivation in the space of capabilities. The definition of poverty within South African context was discussed. Poverty in South Africa is defined and measured using two approaches, the one dimensional approach to poverty measurement, which identifies poverty with a shortfall on consumption expenditure, and the multidimensional approach, which identifies poor households if and only they are deprived in some combination of indicators.

This study does not focus on the *relative approach to measurement of poverty*. This approach is not useful for the monitoring of progress in the reduction of poverty over time or space; because people will always be poor in comparison to those around them. For example, if the poor are defined as the bottom 10 percent of the population(in terms of welfare distribution), there will always be the bottom 10 percent of the population even if the living standards for the whole population have risen over time. This approach also does not allow comparisons of poverty across regions. It is also quite arbitrary in that it defines poverty in terms of a specific predetermined point. The study also does not consider *subjective approach to measurement of poverty* since the currently available source has only one data point and will not allow comparisons.

This study focuses on objective, absolute one-dimensional poverty measurement. In the South African public policy context these measures are highly relevant. This approach is well suited for long term statistical use because it is easy to maintain. This is possible because base period values

can be projected retrospectively or prospectively depending on analytical purpose. The objective absolute approach also allows comparisons over time and across groups. This approach allows the evolution of poverty over time to be tracked, and is very useful when evaluating the effects of policies and programs aimed at alleviating poverty. The table below summarises different approaches to poverty measurement:

Table 2.1: Summary of different approaches to poverty measurement

Poverty measurement approach / **Dimension**		**One-dimensional poverty measures**	**Multi-dimensional poverty measures**
Objective	**Absolute**	**(Focus of this study)** The poverty measures are determined in relation to an objective, invariant and value free external definition of basic human needs. The commonly used approach is the development of a poverty line. Some of the poverty measure are: • Poverty Headcount Index • Poverty Gap • Poverty Gap Index • Squared Poverty Gap • Sen-Shorrocks-Then Index • Watts Index	Poor households are identified and an aggregate measure constructed based on a weighted average of the deprivations they experience using a nested weight structure. A common multidimensional poverty measures is the Multidimensional Poverty Index (MPI). The MPI reveals the combination of deprivations that batter a household at the same time.
	Relative	Poverty measures are determined from a cut-off point in the welfare distribution such as income or consumption. That is, poverty is linked to a particular point in relation to the national distribution of a particular variable such as income or expenditure	Poverty is linked to a particular point in relation to a national distribution of a combination of variables such as income, access to services, etc.
Subjective	**Subjective**	Poverty is measured by considering the people's perception of what constitutes minimum wellbeing. People are asked to define a poverty line and it is used to measure the extent of poverty.	People define their well-being not only from income or expenditure perspective but also using other welfare variables such as availability and access to services such as water, electricity, sanitation, etc.

Chapter 3: Theoretical concepts and methods of measuring poverty

3.1. Introduction

In Chapter 2, review of relevant literature on the definition of poverty was discussed. This chapter presents theoretical concepts and methods of measuring poverty. The methodology includes discussion of general primary steps on how to compute one-dimensional and multidimensional poverty measures. The South African specific approaches to computing one-dimensional and multidimensional poverty measures are also discussed to give the South African context. Different poverty indices which include the poverty headcount, poverty gap index, squared poverty index, Watts index and the time taken to exit poverty are discussed. These are common one-dimensional measures of poverty used internationally and are critical to understand as they are applied in this study (Coudouel, Hentschel & Wodon 2002), (Houghton and Khandker 2009). A brief overview of the application of these absolute poverty measures in South Africa is given.

General approaches on how to estimate poverty from household survey data and the strength and limitation to survey data in poverty analysis are presented. These discussions are particularly important in recognising the strength and limitations of survey data as well as to set up and interpret such data correctly. When discussing how to estimate poverty from household survey data, the chapter discusses several dimensions that encompass quality of surveys. The dimension of accuracy is discussed in detail with emphasis on survey design and properties of estimators for two design components, namely, Simple Random Sampling and Stratified Sampling.

The chapter concludes by discussing how to estimate variances and standard errors of sample statistics in the case of complex multi-stage sampling. Two methods of estimating variances in the case of complex surveys are discussed in detail. These methods are the Taylor Series method and the Jackknife method. The discussion of these methods is critical because the data used in this study is the household data from the 2011 Income and Expenditure Survey which follows a multi-stage sampling design and not simple random sampling design. Other methods of variance estimation in case of a complex survey which are the Balanced Repeated Replication (BRR) method and the Bootstrap method are mentioned briefly.

3.2. Poverty measures

Both one-dimensional and multidimensional approaches to the measurement of poverty were discussed in Chapter 2. Poverty is often defined by one-dimensional measures, such as income or expenditure. While this provides us with very useful way of measuring absolute poverty, it does not capture the multiple aspects that constitute poverty (StatsSA 2014b). The common one-dimensional measures of poverty used internationally are the poverty headcount, the poverty gap, the poverty gap index and the squared poverty gap index (Coudouel, Hentschel & Wodon 2002), (Houghton and Khandker 2009). South Africa also uses these one-dimensional absolute measures of poverty. In contrast, since there are several factors that can contribute to a poor person's experience of deprivation, the multidimensional measurement of poverty aims at incorporating this wide range of factors to capture the complexity of poverty. The common multidimensional measures of poverty used internationally include the Human Development Index (HDI), Human Poverty Index (HDI) both developed by the UNDP, and the Multidimensional Poverty Index (MPI) developed by Alkire and Foster for the UNDP (Alkire and Santos 2010). South Africa constructed and uses the South African Multidimensional Poverty Index (SAMPI) based on the international MPI. The one-dimensional poverty measures are discussed in detail in this chapter whereas the multidimensional poverty index is only introduced.

Computing one dimensional poverty measures

There are three primary steps required in computing a one-dimensional absolute poverty measure. First, one has to define and choose the relevant dimension and indicator of well-being. Second, one has to select a poverty line, that is, a threshold below which a given household or individual will be classified as poor. Finally, one has to select a poverty measure to be used for reporting for the population as a whole or for a population subgroup only. The usual choice of the relevant dimension and indicator of well-being is income or consumption expenditure. Consumption is preferred over income for a number of reasons including that it is less understated than income, because expenditure is easier to recall. Consumption is preferred also because it comes closer to permanent income, however, it requires the survey to value durable goods (by assessing the implicit rental cost) and housing (by estimating what it would have cost to rent), as well as accounting for household composition differences. The identification of the relevant indicator of well-being is followed by two distinct problems: (1) the specification of the poverty line; and (2) once the poverty line is determined, construction of an index to measure the intensity of poverty suffered by those below that line. The value assigned to the poverty line is a critical parameter in the poverty indices (Klugman, 2002). The traditional approach in measuring poverty involves establishing a threshold

and calculating how many individuals, families or households fall below it. This threshold is often referred to as a Poverty Line. The question is how to establish the poverty line. There is no single correct approach; a wide range of methods has been used in different countries and at different times. The three common approaches are the *cost of basic needs approach*, the *food energy intake approach* and *subjective evaluations*. The cost of basic needs approach is the most commonly used. In this approach, the cost of acquiring enough food for adequate nutrition is first estimated, and then the cost of other essentials such as clothing and shelter are added. The food energy intake method is used when price information is unavailable. This method plots expenditure (or income) per capita against food consumption (in calories per person per day) to determine the expenditure (or income) level at which a household acquires enough food. In the third method, subjective poverty lines are based on asking people what minimum income level is needed just to make ends meet. The construction of a poverty line is the most difficult step in the practical measurement of poverty. It is also said in Woolard and Leibbrandt (1999 p.9) that a 'A poverty line will always be an imperfect construct, but for purposes of analysis one has to draw the line in order to go forward in understanding the nature of poverty.' It is common practice to use a measure of household well-being, say income or consumption expenditure to obtain a poverty line (Klugman 2002).

In developing its poverty lines, South Africa followed the three-step approach as discussed above. The indicator of well-being chosen is money income required to attain basic minimal standard of living. In defining the poverty line, the "cost-of-basic-needs" approach was used. In this case, the first step is the setting of the food poverty line and the other steps are the addition of non-food expenditure to obtain two additional poverty lines –a lower bound poverty line and the upper bound poverty line. This approach estimates the cost of a minimum basket of goods that would satisfy the necessary daily energy requirement per person over a period of a month. This cost defines the food poverty line. In deriving the upper bound and the lower bound poverty lines, the same cost-of-basic-needs approach was followed as in the food poverty line. Two different sets of non-food expenditure was obtained from separate reference households and added to the poverty line to yield the two sets of poverty lines: an upper bound and a lower bound poverty line. The assumption is that in cases where food expenditure is equivalent to the food poverty line, households are considered able to meet basic food and basic non-food needs. Therefore by adding the non-food expenditure of such households to the food poverty line, an upper bound poverty line is obtained. A similar approach was used to obtain the lower poverty line but with a different set of households. The assumption in this case is that households whose total expenditure is close to the food poverty line subsist on "survival food needs" and therefore sacrifice fulfilment of basic food-needs in order to meet their non-food needs; this implies that those non-food items typically purchased by

households can be regarded as essential, as such households sacrifice spending on food to buy these non-food items. Here as well, by adding the non-food expenditure of such households to the food poverty line, a lower bound poverty line is obtained.

The methodology followed in constructing these poverty lines produced period specific poverty lines which require regular updates to maintain relevance in long term statistical use. The last step is to maintain the integrity of absolute poverty lines. The two ways to maintain the integrity of absolute poverty lines is adjustment by means of inflation index or construction of new poverty lines. South Africa follows the inflation index approach. This approach uses the changes in the cost of living to make updates to the poverty line annually based on Consumer Price Index (CPI) series. The approach inflates each portion (food and non-food) of the poverty line by the relevant CPI component. This is to accommodate different direction in which food and non-food prices move and to ensure that these price movements in each component are adequately measured.

At the time of finalisation of this study in 2015, the three poverty lines (Food Poverty line, Lower bound poverty line and the Upper bound poverty line) were rebased using data from the 2011 Income and Expenditure Survey (StatsSA 2015). The need for rebasing the national poverty lines was informed by the fact that spending and consumption patterns change over time, which implies that, the basket of goods and services on which the existing poverty lines are based may have changed. It was therefore necessary to update estimates using recent consumption data in order to make sure that the lines remain relevant and accurate. This is another step to maintain the integrity of absolute poverty lines. During rebasing of national poverty lines, pilot poverty lines for provinces which were derived following the same approach used for the national poverty lines were published (StatsSA 2015). This study however uses the current poverty lines and not the rebased poverty lines. The values of the three current poverty lines in 2011 prices are presented in Chapter 4.

A poverty measure, whether one-dimensional or multidimensional, is a function that translates the comparison of the indicator(s) of household well-being and the chosen poverty line/deprivation cut-offs into one aggregate number for the population as a whole or a population subgroup (Coudouel, Hentschel & Wodon 2002). The most commonly used one-dimensional absolute poverty measures (Houghton and Khandker 2009) are discussed in the next sections.

3.2.1. Definition of one-dimensional absolute poverty measures

Let $\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_\mathrm{N}$ be a standard of living variable (e.g. income or consumption or expenditure) of a population of N households. Let z be a predetermined poverty line and suppose that $z > 0$ is a known constant. Let G_i be the income shortfall of household (hh) i.

$$G_i = \begin{cases} z - Y_i \ \ if\ 0 \le Y_i \ < z \\ \\ 0\ if\ Y_i \ \ge z \end{cases}, \text{ where } 0 \le G_i \le z.$$

Let Q be the number of poor households and N be the total number of households.

$$Q = \sum_{i=1}^{N} I_i,$$

where

$$I_i = \begin{cases} 1 \ \ if \ \ Y_i < z \\ 0 \ \ if\ Y_i \ \ge z \end{cases}$$

Using the notation above, the following absolute poverty measures are defined.

3.2.1.1. Poverty Headcount Index

Poverty headcount index is the number of poor households relative to the total number of households. It is the proportion of households in the population for whom income (or other measures of living standard) is less than the poverty line. The poverty headcount is defined as

$$P_{HC} = \frac{1}{N}\sum_{i=1}^{N} I_i = \frac{Q}{N} \tag{3.1}$$

It follows that $0 \le \mathrm{P_{HC}} \le 1$.

The advantage of the Poverty Headcount Index is that it is easy to construct and to understand. There are however a few disadvantages to this poverty measure. The Poverty Headcount Index is insensitive to differences in well-being between different poor households. It assumes all poor are in the same situation. It does not take the intensity of poverty into account. It is insensitive to differences in the depth of poverty of the poor. If households below the poverty line become poorer or richer, as long as they remain below the line, the index does not change.

In terms of policy, a transfer of income to a very poor household would probably leave the Poverty Headcount Index unchanged (if poor remains below the line) even though poverty has overall lessened. Therefore the easiest way to reduce the Poverty Headcount Index is to target benefits to people just below the poverty line, because they are the ones who are cheapest to move across the line. Policies based on the Poverty Headcount Index might be sub-optimal. In order to ensure rigorous analysis, however, it is important to carry out sensitivity analysis for instance, by calculating the measure for different poverty lines.

3.2.1.2. Poverty Gap

The poverty gap is the average, over all households, of the gaps between households' living standards and the poverty line. It indicates the average extent to which households fall below the poverty line (if they do). The poverty gap is formally defined as the average difference between households expenditure or income and the poverty line. Using the same notation, the poverty gap is defined as

$$P_{PG} = \frac{1}{N}\sum_{i=1}^{N} G_i \ , \tag{3.2}$$

It follows that $0 \leq P_{PG} \leq z$.

If the poverty gap is close to the poverty line, a high proportion of the total population's income or consumption is close to zero. A poverty gap of zero implies either that the proportion of households below the poverty line is low or the income or expenditure fall just below or on the poverty line.

3.2.1.3. Poverty Gap Index

The poverty gap index is the poverty gap relative to the poverty line. It is formally defined as the ratio of the Poverty Gap to the poverty line. Using the same notation, the poverty gap index is defined as

$$P_{PGI} = \frac{1}{z} P_{PG}, \tag{3.3}$$

It follows that $0 \leq P_{PGI} \leq 1$.

The poverty gap shows how much would have to be transferred to the poor to bring their expenditure up to the poverty line. The poverty gap and the poverty gap index are proportional to the "minimum" cost for eliminating poverty with transfers (the cost to eliminate poverty with perfect targeting of the poor and no targeting costs or distortion effects). The poverty gap has the virtue that it does not imply that there is a discontinuity ("jump") at the poverty line.

These poverty measures also have a few disadvantages. They do not capture differences in the severity of poverty amongst the poor and ignore "inequality among the poor". They are therefore insensitive to transfers among the poor; i.e. there will be no change in poverty gap or poverty gap index if there is a transfer of income from a poor household to an even poorer household. Although they complement the headcount index, they might not be sufficient for analysis.

3.2.1.4. Squared Poverty Gap Index

The squared poverty gap index is a weighted sum of poverty gaps (as a proportion of the poverty line), where the weights are the proportionate poverty gaps themselves (Like the poverty gap index, but with weights given to each observation). The squared poverty gap index is formally defined as the average of the square relative poverty gap of the poor. Using the same notation, the squared poverty gap index is defined as

$$P_{SPGI} = \frac{1}{Nz^2}\sum_{i=1}^{N} G_i^2 \ , \tag{3.4}$$

It follows that $0 \leq P_{SPGI} \leq 1$.

The squared poverty gap index takes inequality among the poor into account. A transfer of income from a poor to an even poorer household would reduce the index; a transfer of income from a very poor to a less poor household would increase the index.

3.2.1.5. Foster-Greer-Thorbecke (FGT) class of poverty measures

The Poverty Headcount, the Poverty Gap, the Poverty Gap Index and the Squared Poverty Gap Index all belong to the Foster-Greer-Thorbecke class of measures. Using similar notation,

$$P_{\propto} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{G_i}{z}\right)^{\propto}, \text{with } \alpha \geq 0 \tag{3.5}$$

It follows that $0 \leq P_{\propto} \leq 1$.

The measures are defined for $\propto \geq 0$, and $\propto$ is a measure of the sensitivity of the index to poverty.

When $\propto = 0$, the measure is defined to be the poverty headcount

When $\propto = 1$, the measure is defined to be the poverty gap index

When $\propto = 2$, the measure is defined to be the squared poverty gap index

3.2.1.6. The Sen Index and the Sen-Shorrocks-Thon Index

The *Sen Index* seeks to combine the effects of the number of poor, the depth of their poverty, and the distribution of poverty within the group (Houghton and Khandker 2009). The index is given by

$$P_S = P_{HC}\left(1 - (1 - G^P)\frac{\bar{y}^P}{z}\right) \tag{3.6}$$

where P_{HC} is the poverty headcount, $\bar{y}^P$ is the mean income (or expenditure) of the poor, and G^P is the Gini coefficient of inequality among the poor. It also follows that $0 \leq P_S \leq 1$. The Gini coefficient ranges from 0 (perfect equality) to 1 (perfect inequality). An alternative Sen Index is given by the average of the poverty headcount and poverty gap index, weighted by the Gini coefficient of the poor. It is given by

$$P_S = P_{HC} G^P + P_{PGI}(1 - G^P). \tag{3.7}$$

It follows that $0 \leq P_S \leq 1$. This index has been widely discussed and has the virtue of taking the income distribution among the poor into account. However, the index is seldom used outside of the academic literature. This is perhaps because it lacks the intuitive appeal of some of the simpler measures of poverty.

The *Sen-Shorrocks-Thon Index (SST),* a modified *Sen Index*, combines measures of the proportion of poor people, the depth of their poverty, and the distribution of welfare among the poor. It is the product of the headcount index, the poverty gap index (applied to the poor only), and a term with the Gini coefficient of the poverty gaps. This Gini coefficient is usually close to 1 which indicates great inequality in the incidence of poverty gaps. The *Sen-Shorrocks-Thon Index* is given by

$$P_{SST} = P_{HC} P_{PGI}^P \left(1 + \hat{G}^P\right). \tag{3.8}$$

One of the strengths of the SST index is that it helps give a good sense of the sources of change in poverty over time. The index allows one to decompose poverty into three components and to ask whether there are more poor, whether the poor are more poorer, and if there is higher inequality among the poor? This decomposition is given by

$$\Delta \ln P_{SST} = \Delta \ln P_{HC} + \Delta \ln P_{PGI}^P + \Delta \ln(1 + G^P) \tag{3.9}$$

The decomposition may be interpreted as, percentage change in SST index = percentage change in headcount index + percentage change in poverty gap index (among poor) + percentage change in (1 + Gini coefficient of poverty gaps).

3.2.1.7. The Watts Index

The Watts Index is the first distribution-sensitive poverty measure (Houghton and Khandker 2009). The measure is "distributionally-sensitive" by virtue of its use of logarithms. The Watts index is more sensitive to changes in the lowest incomes than it is to changes for those with higher incomes; that is, transferring R10 to a very poor person counts as a larger contribution to poverty reduction than transferring R10 to a richer (but still poor) neighbour. The Watts index is defined as

$$P_W = \frac{1}{Q}\sum_{i=1}^{n} I_i \ln\left(\frac{z}{y_i}\right) = \frac{1}{Q}\sum_{i=1}^{n} I_i[\ln(z) - \ln(y_i)], \tag{3.10}$$

where $y_1, y_2, \ldots y_n$ indicate the income or consumption of households. The Watts index is increasingly used by researchers because it satisfies all the theoretical properties that one would want in a poverty index. Houghton and Khandker (2009) argue that three axioms are essential to any good measure of poverty. Under the *focus axiom*, the measure should not vary if the income of the non-poor varies; under the *monotonicity axiom*, any income gain for the poor should reduce poverty; and under the *transfer axiom*, inequality-reducing transfers among the poor should reduce poverty. When allocating anti-poverty resources to minimize the Watts index, the effort would tilt towards the poorest due to its distribution sensitiveness. This feature of the index is found by many analysts to be appealing.

3.2.1.8. Time taken to exit poverty

The expected time to exit poverty (that is, to reach the poverty line), for the j_{th} person below the poverty line, if consumption per capita grows at positive rate g per year, is given by

$$t_g^j = \begin{cases} \dfrac{\ln(z) - \ln(y_j)}{g} & if\ 0 \le y_j < z \\ 0 & \text{otherwise} \end{cases} \tag{3.11}$$

The average exit time is then t_g^j averaged over the whole population, including the non-poor for whom $t_g^j = 0$.

$$T_g = \frac{1}{N}\sum_{j=1}^{N} t_g^j = \frac{1}{N}\sum_{j=1}^{N} \frac{\ln(z) - \ln(y_j)}{g}. \tag{3.12}$$

This poverty statistic is useful when policy makers develop poverty reduction strategies. This is so because it may be useful to show how long it would take the average poor person to exit poverty, at different potential economic growth rates. The time taken to exit poverty is decomposable by population subgroups and is also sensitive to the distribution of expenditure (or income) among the poor. The higher the expected growth rate per year, the shorter time it takes for a poor person to exit poverty. Hence, economic growth that acts to raise the real consumption levels of the poor can have a powerful effect on the elimination of poverty. It should be noted however that despite the strength of economic growth, it generally takes more than just growth to improve the lives of the poor.

3.2.1.9. Multidimensional Poverty Index (MPI)

The Multidimensional Poverty Index was developed by the Oxford Poverty and Human Development Initiative (OPHI) for the United Nations Development Programme's Human Development Report (Alkire and Santos 2010). It is an international index of acute multidimensional poverty and it reflects deprivations in very rudimentary services and core human functionings. The MPI has three dimensions namely; health, education, and standard of living. These dimensions are measured using ten indicators. Poor households are identified and an aggregate measure constructed based on a weighted average of the deprivations they experience using a nested weight structure (equal weight across dimension and equal weight for each indicator within dimensions), i.e. each dimension is equally weighted; each indicator within a dimension is also equally weighted. The MPI reveals the combination of deprivations that batter a household at the same time. It is a linear combination of poverty headcounts in different dimensions of poverty. A household is identified as multidimensionally poor if, and only if, it is deprived in some combination of indicators whose weighted sum is 30 percent or more of the dimensions.

The South African Multidimensional Poverty Index (SAMPI) was conceptualised and constructed based on the global Multidimensional Poverty Index (StatsSA 2014b). The SAMPI adjusted the indicators of the MPI and included an additional dimension of economic activity to reflect country specific conditions and needs. Just like the Multidimensional Poverty Index, each indicator within the South African Multidimensional Poverty Index has a defined deprivation cut-off and weights allocated to it. The same method of aggregation of deprivation cut-offs is used at indicator and dimension level to determine the multidimensionally poor households.

The measurement of multidimensional poverty requires a single survey which effectively measures all the dimensions of the index. As this is not available, the practical implementation is problematic. In South Africa, it will only be possible to calculate SAMPI with Census information, which is in any way limited in measuring to the level of detail required by the index. The multidimensional poverty measures are not discussed in detail as they are not the focus of this study. This study focuses on one-dimensional absolute poverty measures (see Table 2.1)

3.3. Absolute poverty measurement in South Africa

Measuring poverty enables review of the extent and nature of poverty and provides for an assessment of policy frameworks and programmes for the reduction of both poverty and inequality (May 1998). There are a number of key researchers and institutions working on poverty measurement and analysis in South Africa. Some of the poverty measures discussed in Section 3.2 above are currently being applied and used by these researchers and analysts locally. Below is a list of some applications of these measures:

Statistics South Africa (StatsSA) uses some of the poverty measures discussed above when reporting the country's progress with regard to Millennium Development Goals (MDGs) (StatsSA 2005, 2010b and 2013b). The poverty measures were used in developing the poverty profile of South Africa in 2012 (StatsSA 2012b). An examination of absolute poverty between 2006 and 2011 was conducted and the poverty measures were used (StatsSA 2014a). SAMPI was first officially reported by StatsSA using census data (StatsSA 2014b).

The University of Stellenbosch under its Bureau for Economic Research has produced a number of working papers since 2005. Some of the papers in this series applies and uses the poverty measures discussed in Section 3.2 above (Van der Berg *et al.* 2005), (Van der Berg 2010), (Yu 2010 and 2013), (Van der berg, Louw & Yu 2007).

The University of Cape Town under its Development Policy Research Unit (DPRU) and the Centre for Social Science Research (CSSR) have also produced a number of working papers in their working paper series since 1995 and 2001 respectively. Some of the papers in the series apply absolute poverty measures (Woolard and Leibbrandt 1999), (Leibbrandt et al. 2004 and 2005), (Van der Berg *et al.* 2006), (Bhorat and Westhuizen 2012).

Other application of the absolute measures in South Africa includes Hoogeveen and Ozler 2004, Klasen 2000, Leibbrandt M *et al.* 2010 to mention a few. It is evident from the research reports by Statistics South Africa, research centres within academic institutions as well as individual researchers that Poverty Headcount, Poverty Gap Index, and Squared Poverty Gap Index are widely used in

South Africa. There is however no visible evidence which indicates the use of Sen Index, Sen-Shorrocks Index, Watts Index, and the time taken to exit poverty.

3.4. Estimating poverty from household survey data

3.4.1. Strength and limitations to survey data in poverty analysis

Measures of poverty are usually calculated using data from national household surveys. It is therefore important to recognize the strengths and limitations of such data and to set up and interpret them correctly. Houghton and Khandker (2009) lists several issues which require attention before undertaking a survey to measure poverty or before poverty is analysed using survey data. These issues are discussed below.

- *The relevance of sample frame*: The appropriateness of a survey's particular sample frame will depend on the inferences one wants to draw from it. The survey may represent a whole country's population, or some more narrowly defined subset, such as workers or residents of one region. Thus, a survey of urban households would allow one to measure urban poverty, but not poverty in the country as a whole.
- *The unit of observation*: In surveys measuring poverty, the unit of observation is typically the household or occasionally the individuals within the household.
- *The number of observations over time*: Most surveys are single cross-sections, covering a sample of households just once. *Longitudinal surveys*, are surveys in which the same households or individuals are resurveyed one or more times. Its data is called panel data and these are not collected often. Panel data or longitudinal surveys are critical in analysis of poverty over time (This is discussed further in Section 4.2.2).
- *The principal living standard indicator collected*: It is common practice to use household consumption expenditure or household income as the measures of welfare. Although some surveys collect both, it typically requires at least two interviews per household. Some surveys collect data on either income or expenditure to reduce the cost. When the questionnaire is more detailed and complex, it takes longer to administer. This result in the sample size having to be smaller, which reduces the precision of the statistics and limits the possible amount of disaggregation of data (for example, to the provincial level).

In addition, (Houghton and Khandker, 2009) lists common problems encountered when interpreting data from surveys pertaining to the survey design, sampling methodology, variability and time

period of measurement and other factors which are common survey problems. These are discussed below:

- *Survey design*: If the sample on which a survey is based is not a probability sample, the interpretation of the resulting estimates of poverty becomes complex. Due to sampling, the ideal of each person or household having an equal chance of being selected is always not achieved. This is due to reasons such as people or households which may be difficult to reach for interviewing; measuring poverty using surveys such as Labour Force Surveys which were not designed for this purpose in their sampling frame. Their sampling frames focuses on economically active population and exclude certain population subgroups of the poor such as children and the pensioners. Surveys very often oversample some small groups (for example, minority households in remote areas) and undersample large and homogeneous groups.
- *Sampling*: Since the measures of poverty and inequality are based on survey data, it means that they are *sample statistics*, and so estimate the true population parameters with some error. This also implies that it is essential to know how the sampling was done and to use the appropriate weights and sample design to calculate poverty measures and their standard errors.
- *Goods coverage and valuation*: It has been widely observed that when questions about income and expenditure are more detailed, the resulting reported levels of income and expenditure are higher. Therefore if the variable of interest (e.g. economic welfare) is to be measured satisfactorily, these questions must be comprehensive. In addition, the questions should be consistent over time to ensure comparability.
- *Variability and the time period of measurement*: The levels of both households' income and consumption vary from month to month, year to year, and over a lifetime. However, income levels normally vary more than consumption levels. The reason being that households try to smooth their consumption over time, for instance by managing their savings. Analysts in developing countries prefer to use current consumption rather than current income as an indicator of living standards, because consumption reflects more accurately the resources that households control and it reveals information about incomes at other dates, in the past and future.
- *Comparisons across households at similar consumption levels*: Characteristics of households vary. These might include their income or expenditure levels, household size, the prices they face, the publicly provided goods (such as roads and schools) to which they have access, the amount of leisure time they enjoy, and in the agreeableness of the environment in which they live (some areas are too hot or too cold or too dry or too flood-prone). These

characteristics make comparison of household welfare difficult. Although it is not possible in practice to take all such factors fully into account, some corrections can be made to correct some of the differences, if data permits. The corrections include applying equivalence scales, setting poverty lines for different areas such as urban or rural, as well as measuring comprehensively indicators of welfare which might include measuring durable good, value of housing services, weddings and funerals etc.

It is critical to know survey design and sampling methodologies prior to interpreting and analysing survey data to improve accuracy and comparability over time. Houghton and Khandker (2009) further provide some cautionary notes on using survey data. These cautionary notes include caution to do the sampling right, use consistent recall method, use the correct price indexes, use consistent questions, adjust for nonresponse bias if possible, define living standards consistently, value own production, distinguish between values that are zero and those that are missing, use expenditure per capita and not per household as well as to use weights where they are needed.

The quality of estimates calculated from survey data is affected by a number of factors. When making poverty comparisons or even when reporting on poverty at a single time, it is important to examine the robustness of the results since the estimated poverty rates are sometimes fragile. (Lohr 2010) considers quality of surveys to encompass several dimensions as follows:

- *Relevance of statistical concept*: The statistics collected must meet user needs. It is emphasized here that the identification of the users and their expectation is necessary. This dimension relates to the strength of the sampling frame as discussed above.
- *Accuracy of estimates*: Estimates should be close to the true values of population quantities. Assessment of accuracy involves analysis of the total error associated with the estimate. This study explores further this dimension relating to accuracy estimates of poverty.
- *Timeliness*: Results need to be disseminated quickly to be useful. This is also the reason for taking surveys rather than conducting censuses since surveys can be completed much more rapidly.
- *Accessibility and clarity of information*: Statistical data and data products, particularly in official statistics, must be accessible to users, and sufficient documentation should be provided to enable users to interpret the results.
- *Comparability*: Many surveys have a purpose of comparing estimates over time; such surveys must be conducted so that these comparisons are meaningful. This dimension relates to the practice of keeping the survey methodologies consistent to allow comparisons over time.

- *Coherence:* Common definitions and standards should be used when data come from several sources.
- *Completeness:* Domains for which statistics are available should reflect the needs and priorities expressed by the community of data users. The data collector should be able to provide statistics for all domains identified by the community of data users.

Although not discussed as a dimension of quality, the resources available for the production of statistics act as a constraint on quality. There is clearly a strong link between the quality of statistics and the resources available to produce them.

Given the above considerations, measures of poverty are usually calculated using data from national household surveys. The sample design of official household surveys is typically complex, involving multi-stage stratified cluster sampling. The result is that observations are weighted, to ensure unbiased estimates of the population characteristics. Furthermore, variance estimation should take into consideration the specific sample design (UNDESA 2005a).

Each household in a household sample is assigned a weight which is the reciprocal of the probability of including the household in the sample. In the case of a simple random sample each unit has the same inclusion probability and consequently, all sampling weights are the same. In all other survey designs, the weights are typically not the same and should be taken into consideration to ensure unbiased estimation of the population statistics (in this case, the poverty measures). Suppose $\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_n$ represent a sample of income or consumption expenditure drawn from a population with corresponding weights $\mathrm{w}_1, w_2, \ldots, w_n$. An estimator for the poverty headcount index (3.1) is given by

$$\hat{P}_{HC} = \frac{1}{\sum_{i=1}^{n} w_i} \sum_{i=1}^{n} w_i I_i \qquad (3.13)$$

An estimator for the poverty gap (3.2) is given by

$$\hat{P}_{PG} = \frac{1}{\sum_{i=1}^{n} w_i} \sum_{i=1}^{n} w_i G_i \qquad (3.14)$$

An estimator for the poverty gap index (3.3) is given by

$$\hat{P}_{PGI} = \frac{1}{\sum_{i=1}^{n} w_i z} \sum_{i=1}^{n} w_i G_i \qquad (3.15)$$

An estimator for the squared poverty gap index (3.4) is given by

$$\hat{P}_{SPGI} = \frac{1}{\sum_{i=1}^{n} w_i z^2} \sum_{i=1}^{n} w_i {G_i}^2 \tag{3.16}$$

An estimator for Foster-Greer-Thorbecke (FGT) class of poverty measures (3.5) is given by

$$\hat{P}_{\propto} = \frac{1}{\sum_{i=1}^{n} w_i} \sum_{i=1}^{n} w_i \left(\frac{G_i}{z}\right)^{\propto} \tag{3.17}$$

An estimator of the Sen Index is given, using (3.13) by

$$P_S = \hat{P}_{HC}\left(1 - (1 - G^P)\frac{\bar{y}^P}{z}\right) \tag{3.18}$$

An estimator for the Sen-Shorrocks-Thon Index uses results from (3.13) and (3.15) but with poverty gap index applied to the poor only. It is given by

$$P_{SST} = \hat{P}_{HC}\hat{P}^P_{PGI}\left(1 + \hat{G}^P\right). \tag{3.19}$$

An estimator for the Watts Index is given by

$$P_W = \frac{1}{\sum_{i=1}^{n} I_i w_i} \sum_{i=1}^{n} I_i w_i \, ln\left(\frac{z}{y_i}\right) = \frac{1}{\sum_{i=1}^{n} I_i w_i} \sum_{i=1}^{n} I_i w_i [ln(z) - ln(y_i)], \tag{3.20}$$

where $y_i, y_2, \ldots y_n$ in this case denotes a sample of households income or consumption.

An estimator for the expected time to exit poverty is given by

$$T_g = \frac{1}{\sum_{i=1}^{n} w_i} \sum_{i=1}^{n} w_i \, t_g^j \tag{3.21}$$

where t_g^j is given by (3.11)

All these estimators take weights into consideration to ensure unbiased estimation of poverty measures.

3.5. Precision of poverty estimates

In practice, there are two reasons for the difference between estimates and true values of population parameters. These are *sampling* and *non-sampling errors*. *Sampling error* is the error that arises as a result of taking a sample from a population rather than using the whole population. An estimate of a population parameter, such as a sample mean or sample proportion, is likely to be different for different samples (of the same size) taken from the same population and each estimate is likely to be different from the true population parameter. There is only one solution to eliminate

sampling error which is to eliminate the concept of a sample, and to test the entire population. Testing the entire population is in most cases not possible. The best possible way is to minimise this sampling error. This can be achieved by a proper and unbiased probability sampling and by using a large sample size. In contrast, *non-sampling* errors are errors that cannot be attributed to the sample-to-sample variability. Examples of such errors include non-response error, coverage error, model assumption error, processing error, and measurement error, to mention a few. The non-sampling errors would occur even if the entire population was surveyed, which is why they are distinct from sampling error (which would disappear if everyone were included in the survey).

Lohr 2010 argues that in many surveys, the margin of error reported is based entirely on the sampling error. Non-sampling errors are sometimes acknowledged in the text, but generally are not included in the reported measures of uncertainty (Lohr 2010, 527). Lohr further argues that until now, statisticians have focused mainly on sampling errors:

"*Estimates of accuracy published by statistical agencies usually cover only sampling errors. The estimated variance depends in practice on response rates and certain non-sampling errors - e.g. random measurement errors - but it does not account for all non-sampling errors. In particular, most systematic errors are not included in the measure of accuracy." (Lohr 2010, p527)*

Accuracy of estimates from survey data is the most important aspect of data quality as per the dimensions discussed above. Timely, coherent, comparable statistics are of little use if they are inaccurate. The concept of probability sampling is a possible solution to minimising sampling error. In addition, total survey design is emphasized in that the survey should be designed and conducted to reduce errors in general and not only sampling errors. This can be achieved if major error components are known; for example, if errors in survey estimates are caused by coverage problems, then coverage should be improved.

When calculating poverty measures from a sample survey, it is not sufficient to simply report the point estimate such as the sample proportion of households below the poverty line. To do any inference, it is critical to give an indication of how accurate your estimates are by reporting standard errors and confidence intervals. In this study, the focus is on the accuracy of the poverty estimates calculated from the survey data and specifically on the sampling error component. Sampling errors of the poverty estimates will be derived for complex surveys using Taylor Series and Jackknife methods.

3.5.1. Sample design of household surveys

In developing countries, National Statistics Offices (NSOs) are usually the main providers of national official statistics. In executing this responsibility, consideration must be made for a broad scope of information needs in the areas of demographic, social and economic statistics. Different data sources and methods of data collection are used by NSOs. Some of the surveys such as the Demographic and Health Survey are standardized in design while others are tailor-made to fit specific national demands. Standardisation of surveys programmes allows for integration of survey design whereby the same concepts and definitions can be used for variables occurring in several surveys.

Sample designs for household surveys in developing and transition countries have common features (UNDESA 2005a). Most of the surveys are based on multistage stratified area probability sample designs. The sampling units used at the first stage are called primary sampling units (PSUs). The second stage consists of a sample of secondary sampling units (SSUs) selected within the selected PSUs. The last-stage sampling units in a multistage sample are called ultimate sampling units (USUs). The PSU are selected from the master sample. Most NSOs develop a master sample to serve the needs of their household surveys. Independent samples for different surveys are selected from this master sampling frame.

In multi-stage sample designs, survey weights are needed for each unit of analysis in order to produce valid estimates of parameters of the survey population. The weights are needed to compensate for unequal selection probabilities, non-response, and non-coverage. If weights are not used, the analyses will result in distorted estimates of population values. Another consideration is the computation of variances and sampling errors for survey estimates in a way that takes account of the survey's complex design. Standard methods assume simple random sampling whereas household surveys use stratified multistage sampling (UNDESA 2005b). In general, variances and sampling errors for estimates from a stratified multistage sample are larger than those from a simple random sample of the same size, so that the application of the standard methods will overstate the precision of the estimates. The next section discusses variance estimation in the case of simple survey designs.

3.6. Variance estimation in the case of simple survey designs

In order to derive the variance of estimators of poverty measures based on household samples with complex designs, the results for simple random sample and stratified random samples (Lohr, 2010) are briefly stated.

In the case of a simple random sample $\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_\mathrm{n}$, the population mean is estimated as

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i. \tag{3.22}$$

The population variance S^2 is estimated using the sample variance given by

$$s^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \bar{y})^2. \tag{3.23}$$

To calculate the standard error of the sample mean, the variance of the sample mean is required. The variance of the sample mean is given by

$$V(\bar{y}) = \frac{S^2}{n}\left(1 - \frac{n}{N}\right) \tag{3.24}$$

Since S^2 is unknown and estimated by s^2, the variance of the sample mean, $(\bar{y})$, is estimated by

$$\hat{V}(\bar{y}) = \frac{s^2}{n}\left(1 - \frac{n}{N}\right) \tag{3.25}$$

The *standard error* (SE) is the square root of the estimated variance of $\bar{y}$. It is given by

$$SE(\bar{y}) = \sqrt{\hat{V}(\bar{y})} = \sqrt{\frac{s^2}{n}\left(1 - \frac{n}{N}\right)} \tag{3.26}$$

All the above results apply in estimating the population total. The population total is given by $t = \sum_{i=1}^{N} y_i = N\bar{Y}$. To estimate the total, $\bar{y}$ is used, such that

$$\hat{t} = N\bar{y} \tag{3.27}$$

and from (3.24) and (3.27) above

$$V(\hat{t}) = N^2 V(\bar{y}) = N^2\left(1 - \frac{n}{N}\right)\frac{S^2}{n} \tag{3.28}$$

and

$$\hat{V}(\hat{t}) = N^2\left(1-\frac{n}{N}\right)\frac{s^2}{n} \tag{3.29}$$

As the sample becomes larger, the standard error of the sample mean decreases. The estimate becomes more precise and sampling error decreases.

A large sample $100(1-\propto)\%$ confidence interval for the sample mean is given by

$$\left[(\bar{y}) \pm Z_{\frac{\propto}{2}}\, SE(\bar{y})\right] \tag{3.30}$$

where $Z_{\frac{\propto}{2}}$ is the $(1-\propto/2)th$ percentile of the standard normal distribution. $Z_{\frac{\propto}{2}}$ is often substituted by $t_{\frac{\propto}{2},n-1}$, $(1-\propto/2)th$ percentile of t distribution with $n-1$ degrees of freedom. However, for large samples, $t_{\frac{\propto}{2},n-1} \approx Z_{\frac{\propto}{2}}$. In smaller samples, using $t_{\frac{\propto}{2},n-1}$ instead of $Z_{\frac{\propto}{2}}$ produces a wider CI.

This theory can be applied to Foster-Greer-Thorbecke (FGT) class of poverty measures, as they are linear functions of the sample. In practice, however, the computation of the standard error of the sample mean is complicated by the fact that most household surveys are not simple random samples.

In a *stratified random sample*, the population is divided into subgroups called strata. The strata do not overlap, and they constitute the whole population so that each sampling unit belongs to exactly one stratum. A Simple Random Sample is selected from each stratum independently. Elements in the same stratum are more often similar than randomly selected elements from the whole population, so stratification often increases precision.

Using similar notation, the population of N sampling units is divided into H strata, with N_h sampling units in stratum h. Stratified sampling is only possible if the values of $N_1, N_2, N_3, \ldots, N_H$ are known and they sum to N (i.e. each sampling unit belongs to exactly one stratum). The simplest form of stratified sampling is stratified random sampling where Simple Random Sample is taken independently from each stratum so that n_h observations are randomly selected from the N_H population units in stratum h. Use the notation S_h to be a set of n_h observations in the SRS from stratum h. Then the total sample size is $n = n_1 + n_2 + \ldots + n_H$. Let

- y_{hj} be the value of the j_{th} unit in stratum h.
- $t_h = \sum_{j=1}^{N_h} y_{hj}$ be the population total in stratum h
- $t = \sum_{h=1}^{H} t_h$ be the population total

- $\bar{Y}_h = \frac{\sum_{j=1}^{N_h} y_{hj}}{N_h}$ be the population mean in stratum h
- $\bar{Y} = \frac{t}{N} = \frac{\sum_{h=1}^{H}\sum_{j=1}^{N_h} y_{hj}}{N}$ be the overall population mean
- $S_h^2 = \sum_{j=1}^{N_h} \frac{(y_{hj}-\bar{Y}_h)^2}{N_h-1}$ be the population variance in stratum h

When using the SRS estimators within each stratum, the corresponding quantities for the sample are:

- $\bar{y}_h = \frac{\sum_{j=1}^{n_h} y_{hj}}{n_h}$ is the mean of the sample in stratum h.
- $\hat{t}_h = \frac{N_h}{n_h}\sum_{j=1}^{n_h} y_{hj} = N_h\bar{y}_h$ is the sample total in stratum h.
- $s_h^2 = \sum_{j=1}^{n_h} \frac{(y_{hj}-\bar{y}_h)^2}{n_h-1}$ is the sample variance in stratum h.

Assuming that we only sampled the h^{th} stratum, we will have a population of N_h and we take a SRS of n_h units. Then we can estimate the population mean $\bar{Y}_h$ by $\bar{y}_h$ and t_h by $\hat{t}_h = N_h\bar{y}_h$. The population total $t = \sum_{h=1}^{H} t_h$, so we estimate t by

$$\hat{t}_{str} = \sum_{h=1}^{H} \hat{t}_h = \sum_{h=1}^{H} N_h\bar{y}_h \tag{3.31}$$

Then the overall population mean, $\bar{Y}$ is estimated by

$$\bar{y}_{str} = \frac{\hat{t}_{str}}{N} = \sum_{h=1}^{H} \frac{N_h}{N}\bar{y}_h \tag{3.32}$$

The overall population mean is the weighted average of the sample stratum averages. $\bar{y}_h$ is multiplied by $\frac{N_h}{N}$, the proportion of the population units in stratum h. As mentioned above, stratified sampling is only possible if the sizes or relative sizes of the strata are known. The properties of the estimators follow directly from the properties of SRS estimators discussed above. Based on the assumption that we sampled independently from the strata, it follows from (3.28) that

$$V(\hat{t}_{str}) = \sum_{h=1}^{H} V(\hat{t}_h) = \sum_{h=1}^{H} {N_h}^2 \left(1 - \frac{n_h}{\mathrm{N_h}}\right)\frac{S_h^2}{n_h} \tag{3.33}$$

The estimator of $V(\hat{t}_{str})$ is obtained from (3.33) by substituting the sample estimator s_h^2 for the population parameter S_h^2.

$$\hat{V}(\hat{t}_{str}) = \sum_{h=1}^{H} \hat{V}(\hat{t}_h) = \sum_{h=1}^{H} {N_h}^2 \left(1 - \frac{n_h}{\mathrm{N_h}}\right) \frac{s_h^2}{n_h} \tag{3.34}$$

The estimator for the variance of the population mean follows from (3.32) and (3.34) and it is given by

$$\hat{V}(\bar{y}_{str}) = \hat{V}\left(\frac{\hat{t}_{str}}{N}\right) = \frac{1}{N^2}\hat{V}(\hat{t}_h) = \sum_{h=1}^{H} \left(\frac{N_h}{N}\right)^2 \left(1 - \frac{n_h}{\mathrm{N_h}}\right) \frac{s_h^2}{n_h} \tag{3.35}$$

The standard error is therefore given by the square root of the estimated variance,

$$SE(\bar{y}_{str}) = \sqrt{\hat{V}(\bar{y}_{str})} \tag{3.36}$$

If either the sample sizes within each stratum are large, or the sampling design has a large number of strata, an approximate $100(1-\propto)\%$ confidence interval for the population mean $\bar{Y}$ is

$$\left[\bar{y}_{str} \pm Z_{\frac{\propto}{2}} SE(\bar{y}_{str})\right] \tag{3.37}$$

In stratified sampling, inclusion probabilities in different strata may have different weights and as a result the weights may be unequal for some stratified sampling designs. The stratified sampling estimator $\hat{t}_{str}$ can be expressed as a weighted sum of the individual sampling units.

The estimator of the population total in stratified sampling may thus be written as

$$\hat{t}_{str} = \sum_{h=1}^{H} \sum_{j=1}^{n_h} w_{hj}\, y_{hj} \tag{3.38}$$

where the sampling weight for unit j of stratum h is $w_{hj} = \frac{N_h}{n_h}$. The sampling weight can again be interpreted as the number of units in the population represented by the sample member y_{hj}. The probability of including unit j of stratum h in the sample is $\pi_i = \frac{n_h}{N_h}$, the sampling fraction of stratum h. The sum of the sampling weights in stratified random sampling, as was the case in simple random sampling, is equals the population size N, that is $\sum_{h=1}^{H} \sum_{j=1}^{n_h} w_{hj} = N$. The population mean can thus be estimated by

$$\bar{y}_{str} = \frac{\sum_{h=1}^{H} \sum_{j=1}^{n_h} w_{hj}\, y_{hj}}{\sum_{h=1}^{H} \sum_{j=1}^{n_h} w_{hj}} \tag{3.39}$$

Unlike in the case of simple random sampling, in stratified sampling, we cannot simply divide the estimate of the population variance by N to obtain $\hat{V}(\bar{y})$. If this is done, incorrect standard error

will be obtained and it would give a spurious sense of precision. Hence the total sampling weight is applied.

Statistics South Africa uses two stage sampling design in the Income and Expenditure Survey. The first stage is the selection of PSU's to form the master sample and the second stage is the selection of dwelling units by means of a systematic random sample from the PSU's. The notation used for stratified random sampling can be expanded to incorporate the second stage of sampling as follows:

Let y_{hij} be the value of the j_{th} unit in the i_{th} PSU of stratum h, where

$h = 1, 2, \dots, H$ is the stratum index
$i = 1,2, \dots, n_h$ is the PSU index within stratum h
$j = 1,2, \dots, m_{hi}$ is the household index within PSUi of stratum h
Suppose w_{hij} denotes the sampling weight for unit i in PSU j of stratum h.

The estimate of the mean is given by

$$\hat{\bar{y}} = \left(\sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij} y_{hij} \right) / w \dots \qquad (3.40)$$

where $w \dots = \sum_{h=1}^{H} \sum_{i=1}^{n_h} \sum_{j=1}^{m_{hi}} w_{hij}$ is the sum of weights over all observations in the sample.

The sampling rate f_h for stratum h is the fraction of first-stage units (PSUs) selected for the sample.

The variance of the sample mean is estimated by

$$\hat{V}(\hat{\bar{y}}) = \sum_{h=1}^{H} \hat{V}_h(\hat{\bar{y}})$$

where, if $n_h > 1$, then

$$\hat{V}_h(\hat{\bar{y}}) = \frac{n_h(1 - f_h)}{n_h - 1} \sum_{i=1}^{n_h} (e_{hi.} - \bar{e}_{h..})^2 \qquad (3.41)$$

$$e_{hi.} = \left(\sum_{j=1}^{m_{hi}} w_{hij} \left(y_{hij} - \hat{\bar{y}} \right) \right) / w \dots$$

$$\bar{e}_{h..} = \left(\sum_{i=1}^{n_h} e_{hi.} \right) / n_h$$

and if $n_h = 1$, then

$$\hat{V}_h(\hat{\bar{y}}) = \begin{cases} \text{missing} & \text{if } n_{h'} = 1 \text{ for } h' = 1,2, \dots, H \\ 0 & \text{if } n_{h'} > 1 \text{ for some } 1 \leq h' \leq H \end{cases} \qquad (3.42)$$

The results from simple random sampling and stratified random sampling show that estimation of means and totals is not as complicated as estimation of variances. It is however noted that in a complex survey with several levels stratification and clustering, variances for estimated means and totals are calculated at each level and then combined as the survey is ascended.

Lohr (2010) argues that most surveys sample clusters of households because it is less expensive. The sampling of cluster of households does not affect the estimate of the sample mean; however it does increase the variance, relative to simple random sampling. When sampling errors of sample means are calculated, it is vital to correct for clustering which in the context of this study implies PSU. To compute the standard error of the estimated mean in the case of clustering, it is helpful to use methods of variance estimation developed for complex surveys. When there is stratification, the computation of the variance of population mean is further complicated because it is required that each observation be weighted using sampling weights. The next section discusses methods of estimating standard errors of the poverty estimators in the case of complex surveys.

3.7. Standard Error estimation in the case of complex surveys

In the previous section (Section 3.6), we presented and discussed how to estimate means, variances and standard errors for different sampling designs. Some of the variance formulas such as those for simple random samples (SRSs) are simple to derive and apply whereas other formulas such as, $\widehat{\mathrm{V}}(\hat{\mathrm{t}}_{\mathrm{str}})$ from stratified sampling are more complicated. This is because in complex surveys with several levels of stratification and clustering, variances for estimated means and totals are calculated at each level and then combined as the survey design is ascended. While these formulas work for estimating variances of estimated means, it is often required to estimate variances of other quantities such as the variance for a ratio of two means. The traditional analytic approach for estimating variances of such quantities is more complicated, especially when the survey is not a SRS. There are several methods for estimating variances of estimated totals and other statistics from complex surveys (Lohr 2010). These methods are categorised as Linearization methods, Random Group methods and Resampling and Replication methods. In this section, we discuss Taylor series which is within the Linearization methods and the Jackknife method which is within the Resampling methods. We also briefly discuss the Balanced Repeated Replication (BRR) and the Bootstrap which are within the Resampling and Replication methods respectively. Some of the poverty measures considered in this study are linear functions and consequently exact standard error can be calculated using (3.41) and (3.42). In addition to the exact calculation, the Jackknife method was used to estimate standard errors and confidence intervals of different poverty measures using the IES 2011 data.

As discussed in Chapter 1, the objective of this study is to estimate different poverty measures and their standard errors in the case of complex multi-stage sampling design as opposed to the standard case of simple random sampling. The absolute poverty measures discussed in Section 3.2.1 are averages, and therefore these poverty measures will be calculated as sample means to get the point estimates. As discussed in the previous section (Section 3.5), when taking a sample survey, it is not sufficient to simply report the point estimates such as the sample proportion of households below the poverty line. It is useful to give an indication of how accurate your estimates are. In statistics, the standard error of the estimator is used to indicate the accuracy of an estimate. This associated confidence interval will provide a range within which the population mean is likely to fall with a specified level of confidence (e.g. 95%). In this case, the population mean is the unknown population poverty measure.

3.7.1. Linearization (Taylor Series) methods

Linearization is used for estimating variances of non-linear functions by using Taylor series expansions, e.g. the ration of two means. Let y_{ij} be the response for unit i to the item j, say, the income or expenditure for household i. Suppose $\hat{t}_1,\ldots,\hat{t}_k$ are unbiased estimators of the k population totals with $\hat{t}_j = \sum_{i\epsilon S} w_i\, y_{ij}$. Then, for any constants $a_1,\ldots, a_k$, we can define a new variable

$$q_i = \sum_{j=1}^{k} a_j\, y_{ij}$$

so that

$$\hat{t}_q = \sum_{i\epsilon S} w_i\, q_i = \sum_{j=1}^{k} a_j\, \hat{t}_j$$

and

$$V\left(\sum_{j=1}^{k} a_j\, \hat{t}_j\right) = V(\,\hat{t}_q) = \sum_{j=1}^{k} a^2 V(\hat{t}_j) + 2\sum_{j=1}^{k-1}\sum_{l=j+1}^{k} a_j a_l\, Cov(\,\hat{t}_j \hat{t}_l) \qquad (3.43)$$

The variance formula in (3.43) can be used to estimate the variance of any quantity that can be expressed in the form $a_1t_1+,\ldots,+a_qt_q$ for constants $a_1,\ldots a_q$. However, quantities such as proportions given by, say, t_1/t_q, cannot be expressed in the form $a_1t_1 + a_2t_q$. The Taylor's theorem allows linearization of smooth nonlinear functions such as $h(t_1, t_2,\ldots, t_k)$ of population totals (Lohr 2010). Taylor's theorem gives the constants $a_0, a_1,\ldots, a_k$ so that

$$h(t_1,\ldots,t_k) \approx a_0 + \sum_{j=1}^{k} a_j\, t_{j.}$$

Then the variance of the nonlinear function $h(t_1, t_2,\ldots, t_k)$ given by $V\left(h(\hat{t}_1,\ldots,\hat{t}_k)\right)$ can be approximated by

$$V\left(\sum_{j=1}^{k} a_j\, \hat{t}_j\right)$$

which is calculated using (3.43).

One key advantage of the Linearization method is that if the partial derivatives are known, this method almost always gives a variance estimate for a statistic and can be applied in general sampling designs. Some of the quantities of interests such as the median and other quantiles cannot be expressed as smooth functions of the population totals and as such do not fit within the framework of Linearization. The accuracy of the linearization approximation depends on the size of the sample. If the sample is not large enough, the variance estimator is often biased downwards (underestimated). Taylor series method is given as a variance estimation method in the SAS software. SAS actually applies the Taylor series method by default if no variance estimation method is specified. The method will be applied to the Income and Expenditure Survey 2011 data in estimating variances of the weighted poverty measures discussed in Section 3.4.1 which are to be calculated in Chapter 4.

The Taylor series approximation for the variance of a linear function is equal to the exact formula. Many of the poverty measures are linear functions and consequently the variance need not be estimated, but the exact value can be calculated using the Taylor series method.

3.7.2. Jackknife

Jackknife is a resampling method for variance estimation. The estimator of a parameter, say variance, is found by systematically deleting each observation from the data and calculating the estimate, and then finding the average of these calculations. Suppose we have a sample of size n, the jackknife estimate is found by aggregating the estimates of each $n-1$ estimates in the sample. In this section, we discuss the *delete-1 jackknife,* so called because for each replicate, one observation or PSU is deleted.

In case of *Simple Random Sampling*, let $\hat{\theta}_{(j)}$ be the estimator of the same form as $\hat{\theta}$, but not using observation j. If $\hat{\theta} = \bar{y}$ then $\hat{\theta}_{(j)} = \bar{y}_{(j)} = \sum_{i\neq j}^{k} y_i /(n-1)$. The jackknife estimator is then defined as

$$\hat{V}_{JK}(\hat{\theta}) = \frac{n-1}{n}\sum_{j=1}^{n}(\hat{\theta}_{(j)} - \hat{\theta})^2 \tag{3.44}$$

The multiplier $(n-1)/n$ is used here because for $\hat{V}_{JK}(\hat{\theta})$ when $\hat{\theta} = \bar{y}$, we have

$$\bar{y}_{(j)} = \frac{1}{n-1}\sum_{i\neq j} y_i = \frac{1}{n-1}\left(\sum_{i=1}^{n} y_i - y_j\right) = \bar{y} - \frac{1}{n-1}(y_j - \bar{y})$$

Then

$$\sum_{j=1}^{n}(\bar{y}_{(j)} - \bar{y})^2 = \frac{1}{(n-1)^2}\sum_{j=1}^{n}(y_j - \bar{y})^2 = \frac{1}{n-1}s_y^2$$

So, $\hat{V}_{JK}(\bar{y}) = \frac{s_y^2}{n}$, which is the with-replacement estimator of the variance of $\bar{y}$ given in (3.22), but ignoring the finite population correction. For a *cluster sample*, the jackknife variance estimator in (3.43) is applied but letting n be the number of Primary Sampling Units (PSUs), and letting $\hat{\theta}_{(j)}$ be the estimate of θ that we would obtain by deleting all the observations in PSU j.

In a *stratified multistage cluster sample*, the same jackknife variance estimator given by (3.44) is applied but separately in each stratum at the first stage of sampling, with one PSU deleted at a time. Suppose there are H strata, and n_h PSUs are chosen for the sample from stratum h. It is assumed these PSUs are chosen with replacement. Jackknife is applied by deleting one PSU at a time. Let $\hat{\theta}_{(hj)}$ be the estimator of the same form as $\hat{\theta}$, but not using PSU j from stratum h. New weight variable is defined to calculate $\hat{\theta}_{(hj)}$. Let

$$w_{i(hj)} = \begin{cases} w_i & \text{if observation unit } i \text{ is not in stratum } h \\ 0 & \text{if observation unit } i \text{ is in psu } j \text{ of stratum } h \\ \dfrac{n_h}{n_h - 1} w_i & \text{if observation unit } i \text{ is in stratum } h \text{ but not in psu } j \end{cases}$$

Then the weights are used to calculate $\hat{\theta}_{(hj)}$, and

$$\hat{V}_{JK}(\hat{\theta}) = \sum_{h=1}^{H} \frac{n_h - 1}{n_h} \sum_{j=1}^{n} (\hat{\theta}_{(hj)} - \hat{\theta})^2 \tag{3.45}$$

The Jackknife resampling method for variance estimation is an all-purpose method because it uses the same procedure to estimate the variance for every statistic for which jackknife can be used. This method provides a consistent estimator of the variance when θ is a smooth function of population totals. The jackknife method however requires a large amount of computation for some sampling designs. The jackknife performs poorly for estimating the variances of some statistics that are not smooth functions of population totals. For example, the jackknife does not give a consistent estimator of the variance of quantiles in a SRS. The Jackknife method is given as a variance estimation method in the SAS software. The method will be applied to the Income and Expenditure Survey 2011 data in estimating variances of the weighted poverty measures discussed in Section 3.4.1 which are to be calculated in Chapter 4.

3.7.3. Other variance estimation methods

Other methods for estimating variances from complex surveys are Balanced Repeated Replication (BRR) and Bootstrap. These methods fall within the family of resampling methods. Resampling methods treat the sample as if it were itself a population; these methods take different samples from this new "population" and use the subsamples to estimate the variance. Balanced Repeated Replication (BRR) is a resampling method for variance estimation used when there are two PSUs sampled per stratum. Bootstrap is a resampling method for variance estimation in which samples of PSUs with replacement are taken within each stratum. Both BRR and Bootstrap can be applied in case of Stratified Random Sample or Stratified Multistage Survey. These methods are not discussed any further as they are not applied in this study.

3.8. Conclusion

Theoretical concepts and methods of measuring poverty were covered in this chapter. The methods discussed included discussion of primary steps on how to compute one-dimensional and multidimensional poverty measures, the procedure followed in developing poverty lines for South Africa as well as different poverty indices and their application within the South African context. In this chapter it was also explained how to estimate poverty from household survey data, strengths and limitation to survey data in poverty analysis and how to estimate variances of sample statistics in the case of complex multi-stage sampling.

With regard to poverty indices, the poverty headcount, poverty gap, poverty gap index, squared poverty gap index, Sen-Shorrocks-Thon Index, watts index and the time taken to exit poverty were discussed. The indices are important as they translate the comparison of the indicators of household well-being and the chosen poverty line into one aggregate measure for the whole population. A brief overview of the commonly used indices was made.

With regard to estimating poverty from household survey data, several dimensions that encompass quality of surveys were discussed. The dimension of accuracy was discussed in detail with emphasis on survey design and properties of estimators for two sample designs components, namely, Simple Random Sampling and Stratified Sampling. The notation for stratified sampling was expanded to present the estimation of mean and variance in case of a two stage sampling design such as the Income and Expenditure Survey. Two methods of estimating variances in case of complex surveys were presented. These are the Taylor Series method and the jackknife method. It was shown how standard errors of the mean and confidence intervals will be used to indicate the precisions of the poverty estimates. In the next chapter (Chapter 4), all these theoretical concepts will be applied in the analysis of poverty based on South Africa's 2011 Income and Expenditure Survey.

Chapter 4: Analysis of poverty based on South Africa's 2011 Income and Expenditure Survey

4.1. Introduction

The previous chapter (Chapter 3) discussed theoretical concepts and different methods of measuring poverty. This chapter is an application of all the theoretical concepts discussed in Chapter 3 through the analysis of poverty based on South Africa's 2011 Income and Expenditure Survey. The chapter starts by presenting in general how analysis of poverty can be made through poverty profiles. These poverty profiles are comprehensive poverty comparisons showing how poverty varies across different subgroups. The importance of tracking the evolution of poverty over time is also discussed and the three main ways of measuring poverty over time are presented. These are through a *single cross-section survey*, conducting *repeated cross-sectional surveys*, or through *a longitudinal survey* (*panel data)*. A suggestion is made on how poverty estimates should be reported, which includes reporting poverty with sampling errors and confidence intervals. The chapter continues by giving an overview of South Africa's 2011 Income and Expenditure Survey sample design. The IES 2011 provides data on household income, expenditure and living conditions. This data is necessary for calculating poverty estimates. As a way of introducing estimation of poverty from the IES 2011, the distribution of a key variable for analysis, which is *In-kind Consumption per capita*, is presented. A brief motivation of why specific domains or subgroups were selected for analysis is made using the sample quantile function of the per capita in-kind consumption. This is critical in explaining the reasons behind selecting the four domains of analysis which are gender of head of household, population group of head of household, settlement type and the province where the household is located. This chapter also presents inflation adjusted poverty lines applied to the IES data to distinguish between poor and the non-poor households when calculating poverty estimates. Different weighted poverty measures discussed in Section 3.4.1, their standard errors, and confidence intervals, are estimated using IES 2011 household data. The standard errors and confidence intervals are estimated using the exact method and a Jackknife method. Multiple tests are performed using Bonferroni adjusted confidence intervals of estimated poverty measures by the four domains mentioned above. This chapter concludes by giving an overview of different data sources for poverty measurement in South Africa. In this study, the primary tool used to calculate poverty estimates is SAS software. SPSS and Microsoft Excel were also used to generate other necessary outputs.

4.2. Analysing poverty

4.2.1. Poverty profiles

Analysis of poverty can be done through poverty profiles. A poverty profile sets out the major facts on poverty (and typically inequality), and then examines the pattern of poverty to see how it varies by geography, by community characteristics and by household or individual characteristics. Variation by geography includes by region, urban or rural, mountain or plain, and so on. Community characteristics can include for example, in communities with and without a school, and household characteristics include for example, education of household head or household size. A poverty profile is therefore a comprehensive poverty comparison, showing how poverty varies across subgroups of society. Houghton and Khandker (2009) present a set of key questions that may be asked when presenting a poverty profile:

- Whether poverty varies widely between different areas in the country?
- Whether the most populated areas are also the areas where most of the poor live?
- How income poverty is correlated with characteristics such as gender, age, urban and rural, race or ethnicity?
- What are the main sources of income for the poor?
- How large a factor is unemployment or underemployment?
- Which public services do the poor have access to and what is the quality of these services?
- What assets, such as land, housing, and financial, do the poor own?
- How variable are the incomes of the poor and what risks do they face?
- What is the level of vulnerability of certain population groups in society?

These questions are not a comprehensive list. A poverty profile which clearly responds to the above questions is considered helpful. However, the level of detail a poverty profile can be constructed depends on the availability of data.

There are two main ways in which a poverty profile can be presented. The first approach divides the sample by some characteristic such as region of residence, or education level of head of household, and shows the poverty rate for each component. In the second approach, the sample is divided by poverty status (for example, poor against non-poor), then summarizes the rate of characteristics, such as population group, educational level, or access to services, for each group. The value of each approach depends on the use to which the poverty profile will be put. In this study, poverty profiles are presented using the first approach.

4.2.2. Poverty analysis over time

Tracking the evolution of poverty over time is critical for a number of reasons. Understanding the dynamics of poverty is useful as this information is needed for a number of purposes, including the following:

- To be able to distinguish households that are poor occasionally (in temporary or transient poverty) from those that are poor all the time (in chronic poverty). The practical importance of this distinction is beneficial in that the types of interventions relevant for dealing with persistent poverty are likely to be different from those relevant for addressing transient poverty.
- To inform adjustment of the way in which targeting is done for poverty alleviation.

There are three main ways to measure what happens to poverty over time. These can be achieved through appropriate questions in a *single cross-section survey*, conducting *repeated cross-sectional surveys*, or through *a longitudinal survey* (*panel data)*. Each of these approaches is briefly discussed below.

In a ***single cross-sectional survey,*** information on a sample of households or individuals is available at a single point in time. To track evolution of poverty over time, information about current living standards is asked, but in addition, some questions about living standards in the past are also asked. Examples of single cross sectional surveys in South Africa are the Community Survey 2007 and the Living Conditions Survey (LCS) conducted in 2007 and 2008/09 respectively by Statistics South Africa. ***Repeated cross-sectional surveys*** is the most common way that poverty is tracked by using the results from two or more household surveys over time. Poverty measures from repeated cross-sectional surveys should ideally use data from comparable questionnaires that use a similar sampling frame and research protocol and the same definitions of living standards measures such as income or consumption. Such poverty measures should be comparable with and reflective of differences over time in the cost of living across regions. Examples of repeated cross-section surveys are the Income and Expenditure Survey (IES) and the General Household Survey (GHS). These surveys are also conducted by Statistics South Africa. ***Longitudinal survey*** is when a survey is repeated, and we have multiple observations for the same person (or household, or firm, or community) and data gathered in this way is called ***panel data***. The most common panel design is when the sampling unit is the household and subsequent rounds of the survey return to the same households each time. An example of a longitudinal survey in South Africa is the National Income Dynamics Study (NIDS) conducted by the South African Labour Development Research Unit (SALDRU) at the University of Cape Town. All these examples of surveys are discussed further in Section 4.5 of this report.

4.2.3. Reporting poverty estimates

In reporting poverty estimates, statistical agencies usually report point estimates without mentioning sampling and non-sampling errors. Houghton and Khandker (2009), however, argue that it is more useful and honest to report poverty estimates with the sampling errors than reporting only point estimates. In addition, it is argued that it is more acceptable to present the confidence interval which provides a range within which the population poverty estimate is likely to fall with a specified level of confidence. In this study, as part of the assessment of the different approaches to poverty measurement in South Africa, standard errors of the poverty estimates are calculated and the results are presented with confidence intervals.

4.3. Income and Expenditure Survey sample design

The Income and Expenditure Survey (IES) has been conducted every five years since 1995 by Statistics South Africa. The most recent IES was conducted between September 2010 and August 2011. The survey is designed to gather information about the income and expenditure from South African households. Its main purpose is to determine the average expenditure patterns of households in different areas in the country. The survey identifies goods and services purchased by these households. The data collected through this survey form the basis for the determination of the basket of goods and services used to calculate and update the Consumer Price Index (CPI). The IES data is collected through a household questionnaire, weekly diaries and the summary questionnaire. A specific household is in a sample for a period of four weeks. The data collection instruments are administered in stages at different visits during the duration of data collection. The data collected through the household questionnaire includes household expenditure on education, health, dwellings and services, clothing, footwear, expenditure when away from home, domestic workers, furniture, computers and telecommunications equipment and transport. Household income, finances and banking are also collected by means of the household questionnaire. Households use weekly diaries to track their expenditure on a weekly basis for the survey period. Data is collected at household level and the unit of analysis is a household. The response rate for the 2010/11 survey was 91.6%.

The sampling frame for the IES was obtained from Statistics South Africa's Master Sample (MS) based on the 2001 Population Census enumeration areas (EAs) (StatsSA 2012d). The scope of the Master Sample (MS) is national coverage of all households in South Africa. The survey covered private dwellings, workers' hostels, residential hotels, and nurses' and doctors' quarters. It did not

include institutions such as hospitals and clinics, hotels and guest houses, prisons, schools and student hostels and old-age homes.

The sample for the IES used a two-stage stratified design with probability-proportional-to-size (PPS) sampling of primary sampling units (PSUs) from strata in the first stage, and systematic sampling (SYS) of dwelling units (DUs) from the sampled PSUs. The Master Sample stratification was divided into two levels: First, the primary stratification was defined by metropolitan and non-metropolitan geographic area type; and during the second stratification, the Census 2001 data were summarised at PSU level using the following variables: household size, education of head of household, occupancy status, gender of head of household, industry and household income. A PPS sample of PSUs was drawn in each stratum with the measure of size (MOS) being the number of DUs in the PSU. In each selected PSU, a systematic sample of DUs was drawn. The number of selected DUs per PSU depends on the PSU sampling rate and the number of dwelling units the PSU had. The sample for 2010/2011 IES is based on an extended sample of 3 254 PSUs and the sample file contained 31 419 sampled dwelling units (DUs). In the case of multiple households at a sampled DU, all households in the DU were included.

This study uses data from the 2010/11 IES. The data has been annualised to ensure comparability of diary data and COICOP (Classification of Individual Consumption by Purpose) data defined as monthly and annually. In annualising, diary data was multiplied by 26 weeks to get an annual value. The questionnaire data was multiplied by 12 months if the COICOP was defined as monthly. The reported expenditure data was benchmarked to March 2011, which was the midway point of the survey year. All expenditures incurred before the end of February 2011 was inflated to March 2011 prices and all expenditure after March 2011 were deflated back to March 2011 prices using Consumer Price Index (CPI) data. Statistics South Africa generated and released four primary data files for public consumption. These data files are *Person_Info*, which contains particulars on each person in household, *House_Info*, which contains households' characteristics, *Total_IES*, which contains the income and expenditure per item per household and *Person_Income*, which contains data on individual income. This study uses the *House_Info* data which has all information collected about household characteristics, including household assets, income and consumption. The household data files have a total of 145 variables with 25 328 cases (households).

The 2010/11 Income and Expenditure Survey (IES) discussed above provides data on household income and expenditure for poverty measurement. This study analyses poverty using the variable called Household in-kind consumption derived from the Income and Expenditure Survey. The household in-kind consumption is total household consumption plus consumption in-kind (e.g.

subsidies from employer, non-refundable bursaries, etc.) un-weighted but annualised and inflated/deflated to March 2011. In order for the annual total household consumption to be comparable to the poverty lines, consumption per capita was derived by dividing the annual total household consumption by the number of persons in the household (household size). This annual per capita consumption was further divided by 12 to get monthly per capita consumption. That is,

$$Conspumtion\ per\ capita\ per\ month = annual\ total\ household\ consumption/\ (household\ size * 12).$$

This gives consumption per capita per month in 2011 rand value. The IES 2011 data has 25 328 observations (households). Each household has a weight (full_calwgt). These weights were constructed by calibrating the non-response-adjusted (editing and imputation) design weights to the known population estimates using the 'Integrated Household Weighting' method. The final weights were benchmarked to the known mid-March 2011 population estimates of Age-Group by Race by Gender population groups at the national level, and Age-Groups at the province level. The data has a unique identifier for each household (UQNo). The (UQNo) is an eighteen-digit number, which is made up of a primary sampling unit (PSU) number (11 digits), dwelling units (DU) number (5 digits) and household (HH) number (2 digits).

4.4. Poverty estimation

Section 3.2 of Chapter 3 presented definitions and notation of one-dimensional absolute poverty measures. Section 3.4 of the same chapter presented how to estimate poverty from household survey data. Strengths and limitations to survey data in poverty analysis as well as estimating the precisions of poverty estimates were discussed. This was followed by a discussion of how to calculate variances, standard errors and confidence intervals under different sampling designs. Two methods of estimating standard errors in complex surveys were also discussed.

One of the common errors made by researchers and analysts using survey data is to analyse a cluster sample as if it were a Simple Random Sample (SRS) (Haughton and Khandker 2009). As discussed in Chapter 3, this error leads to reporting standard errors that are much smaller than they should be which gives the impression that the survey results are much more precise than they really are.

In this section, poverty measures, their standard errors and confidence intervals are estimated using the two methods of estimating variances in complex surveys, which are Linearization and Jackknife methods. In addition, since IES used stratification and clustering based on the systematic random sample, the relevant strata and cluster variables as well as weights are used when calculating poverty measures, their standard errors and confidence intervals. The distribution and properties of the per capita in-kind consumption are presented. Poverty lines used in this study are presented as

the poverty estimates are based on the household per capita consumption relative to the poverty line.

4.4.1. Distribution of total in-kind consumption per capita per month and disaggregation by domain of analysis

In Chapter 3 (Section 3.1), it was discussed that the first step in computing one dimensional poverty measures is to define and choose the relevant dimension and indicators of well-being. In Section 4.2, it was presented that the indicator of well-being to be used in this study to analyse poverty is the variable called household in-kind consumption derived from the Income and Expenditure Survey. The household in-kind consumption is total household consumption plus consumption in-kind (e.g. subsidies from employer, non-refundable bursaries, etc.) un-weighted but annualised and inflated/deflated to March 2011. It is important to understand the distribution of this variable before any further analysis can be done, as some of the poverty measures include the Gini coefficient of inequality. The high level of inequality in well-being (Income/Consumption) is one of the key challenges in the South African context.

4.4.1.1. Distribution of the in-kind consumption

Table 4.1: Descriptive statistics of the in-kind consumption per capita

Descriptive statistics of monthly per capita in-kind consumption										
Number of households	Mean	Percentiles					95% CL for Mean		95% CL for Mean (Jackknife)	
		0% (Min)	25% (Q1)	50% (Median)	75% (Q3)	100% (Max)				
25 328	2 991.65	8.08	547.58	1135.53	2957.48	101 803.17	2831.38	3151.91	2831.30	3151.99

Table 4.1 above shows that there are 25 328 households in the IES 2011 data. The sample mean is the weighted average of in-kind consumption. Each household was weighted with the variable *full_calwgt* which was provided with the sample. The average monthly per capita in-kind consumption was R2 991.65. This implies that on average, each household consumed a total amount equals to approximately R3000.00 per person per month. A household which consumed the lowest amount spent just above R8.00 per person per month whereas the highest consuming household spent in excess of R100 000 per person per month. It is not surprising that the mean consumption per capita is higher than the median consumption per capita because of the highly positive skewed distribution. The histogram of the total monthly per capita in-kind consumption is given by Figure 4.1 where the sum of weights of households in different intervals of inkind consumption are presented relative to the total sum of weights of all households.

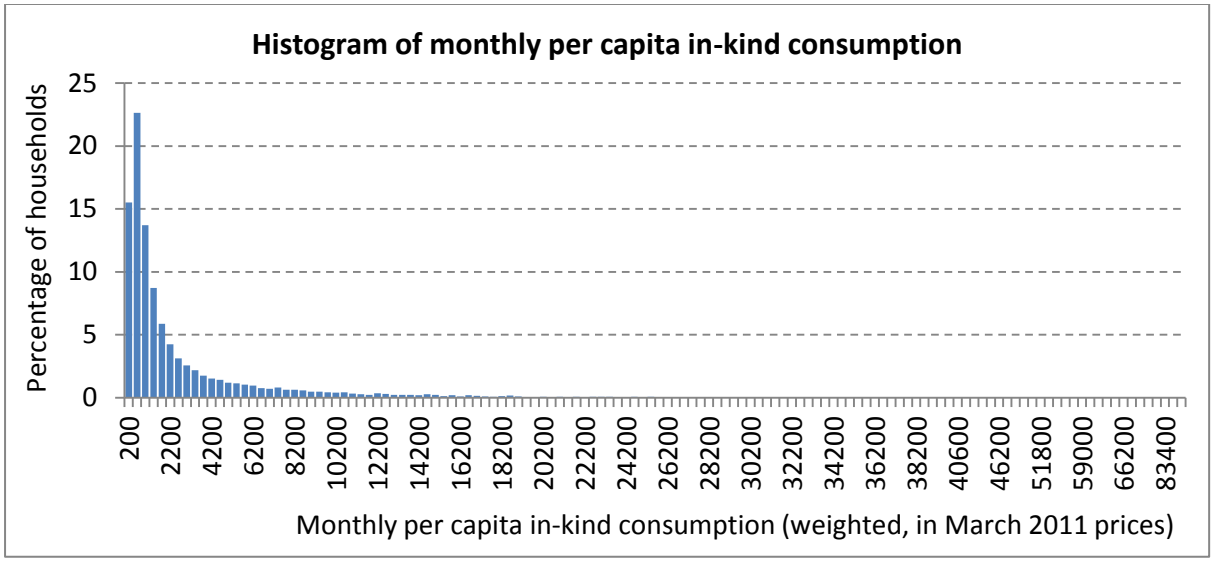

Figure 4.1: Histogram of monthly per capita in-kind consumption (Total)

The distribution of per capita per month in-kind consumption is highly skewed to the right. Higher percentages of households have lower per capita per month in-kind consumption. There are, however, unusually high monthly per capita consumptions.

The empirical cumulative distribution function (CDF) was calculated by sorting the in-kind consumption observations in ascending order and using the cumulative sum of the weights as a proportion of the sum of the weights as an estimate of the cumulative probability. If $Y_1, Y_2, \ldots, Y_n$ are the in-kind consumptions sorted in ascending order with corresponding weights $w_1, w_2, \ldots, w_n$, then the empirical CDF is

Y_i	Relative cumulative weight
Y_1	$w_1 / \sum_{j=1}^{n} w_j$
Y_2	$\sum_{j=1}^{2} w_j \,/\, \sum_{j=1}^{n} w_j$
.	.
.	.
.	.
Y_i	$\sum_{j=1}^{i} w_j \,/\, \sum_{j=1}^{n} w_j$
.	.
.	.
Y_n	$\sum_{j=1}^{n} w_j \,/\, \sum_{j=1}^{n} w_j$

The empirical CDF of the In-kind consumption from IES 2011 is given by Figure 4.2 below. The CDF shows that about 90 percent of households had a monthly per capita in-kind consumption of less

than R8 000. Consider an entry-level state salary for a newly graduated teacher. According to the salary scales from Department of Basic Education (DBE), which took effect from April 2011, the entry salary was R162 354 per annum, excluding benefits such as medical aid, pension and housing subsidy contributions, which would be made by the DBE in addition to this amount. The basic conditions of service recommend adding an additional 37% for these benefits. This would bring the starting salary of an entry-level state schoolteacher to above R222 424 per annum, or R18 535 a month. Assuming equivalence between income and inkind consumption, this will put the teacher in the 95th percentile of the population. If the teacher is the only earning adult from a family of three, the monthly per capita income of this household will be R6 178. Using this household's income as a proxy for the in-kind consumption, the household will be among the 90% whose monthly per capita income falls below R8000 despite a professional working adult.

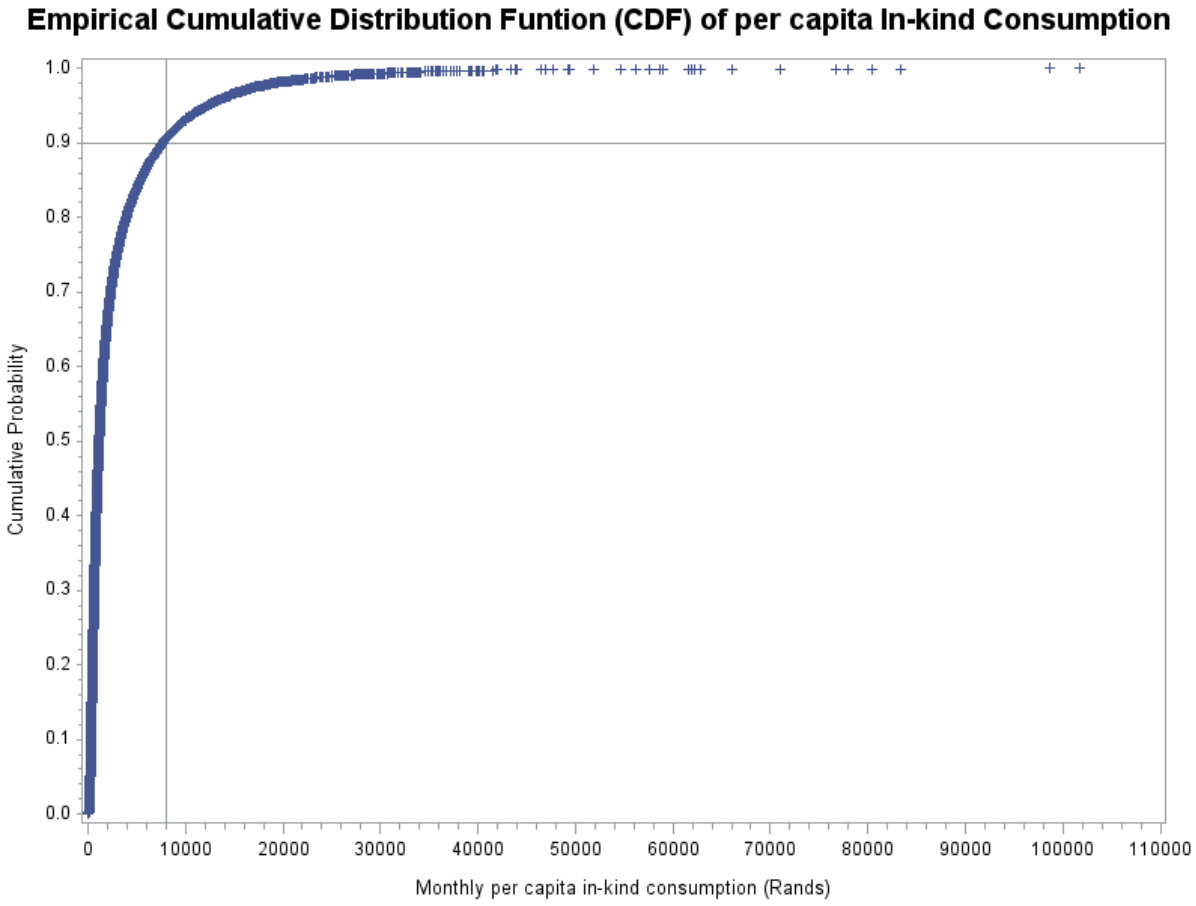

Figure 4.2: Empirical Cumulative Distribution Function (CDF) of monthly per capita in-kind consumption

The quantile function

The quantile function $Q(p)$ is obtained by switching the x and the y axes from the CDF. This function has argument p and returns the in-kind consumption such that $100p\%$ of the households consumes less than that amount per person per month.

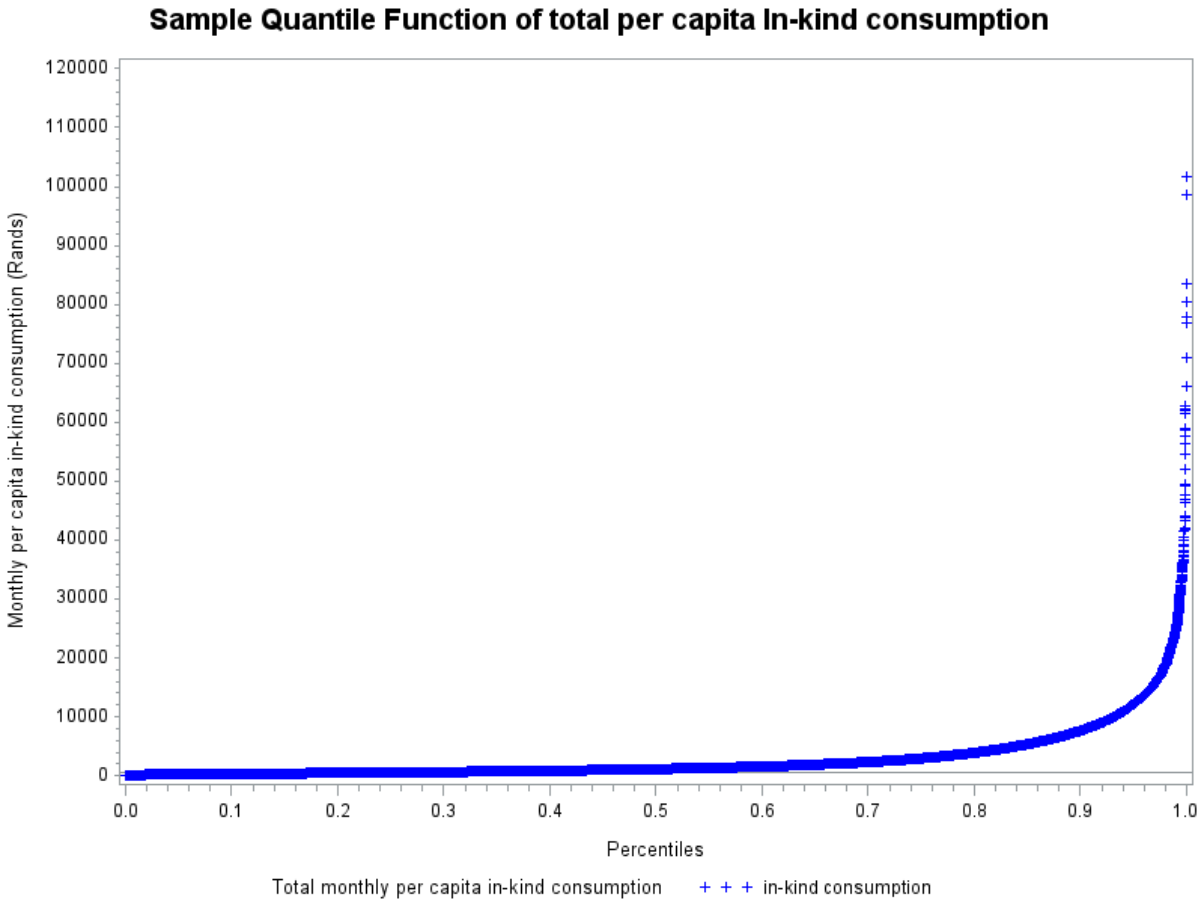

Figure 4.3: Sample Quantile Function of monthly per capita in-kind consumption

The sample quantile function of total monthly per capita in-kind consumption is given in Figure 4.3 above. The quantile function shows that the monthly per capita consumption is highly skewed. The upper bound poverty line is illustrated to indicate the percentage of households consuming below the poverty line. This is better presented in the magnified sample quantile function in Figure 4.3a. As can be seen from Figure 4.3a, the quantile function shows that around 29% of households' monthly per capita consumption is below the upper bound poverty line of R620 (Poverty Headcount). Figure 4.3b indicates that the inkind consumption is smooth around the poverty line, indicating that changes in the poverty line will not have a disproportionate effect on the poverty headcount.

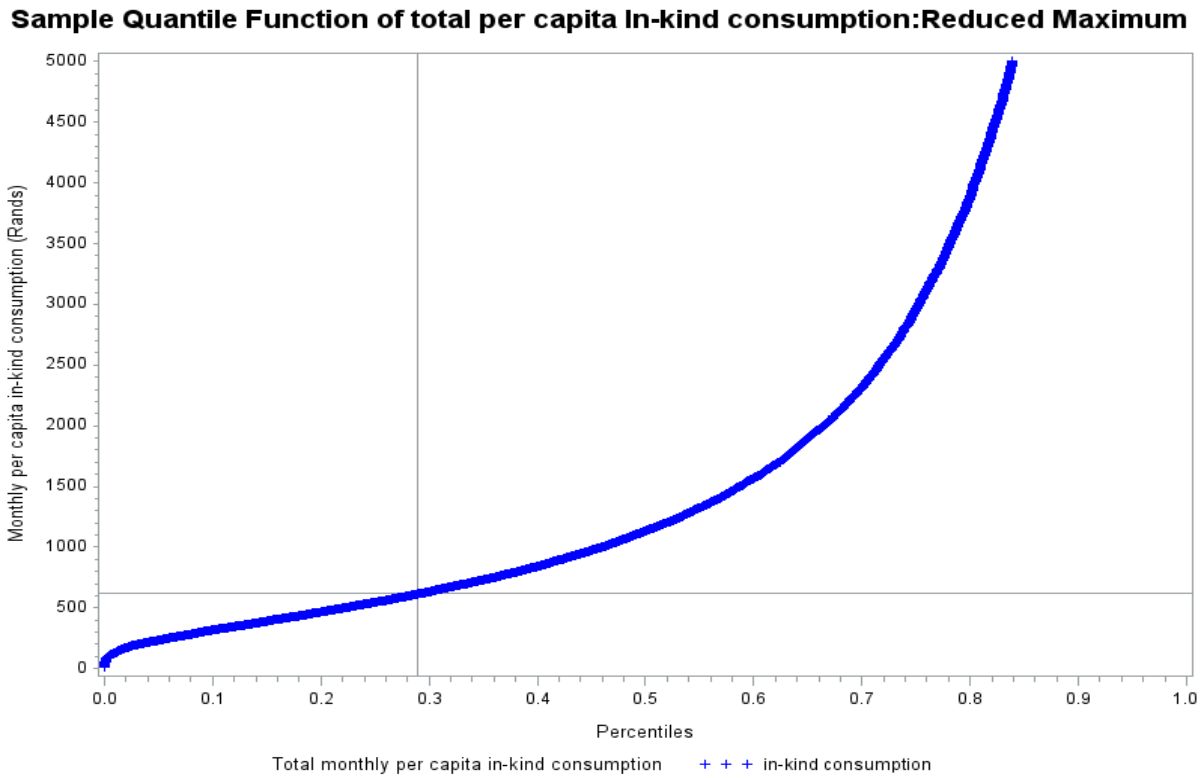

Figure 4.3a: Sample Quantile Function of monthly per capita in-kind consumption: Referencing of the upper bound poverty line.

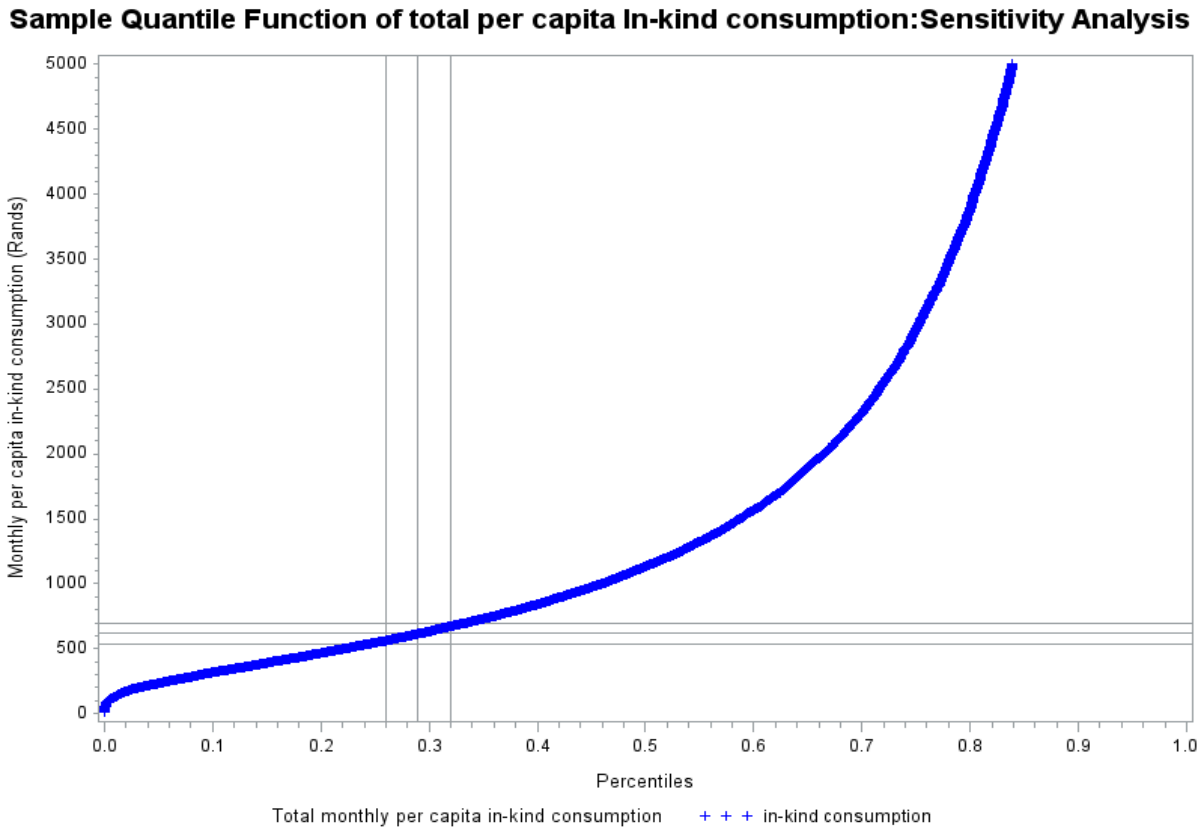

Figure 4.3b: Sample Quantile Function of monthly per capita in-kind consumption: Sensitivity analysis.

4.4.1.2. Domains of analysis

Chapter 1 emphasized the importance of measuring poverty to assess the impact of poverty reduction programmes whereby it is necessary to capture changes in the circumstances of potential beneficiaries – both the general population and groups that are known to be deprived. The World Bank in its World Development Report 2000 (World Bank, 2000) also states that measuring poverty is important to enable the identification of poor people and to be able to target appropriate interventions. Poverty profiles were suggested as a useful way of reporting comprehensive poverty comparison, showing how poverty varies across subgroups of society.

This study estimates poverty for different domains (disaggregation variables), namely gender of head of household, population group of head of household, settlement type and province. It is however critical to indicate the reasons behind selecting these domains. The motivation is made using the historical background of South Africa in relation to these domains as well as the sample quantile function of monthly per capita in-kind consumption for all the domains.

Historical background of South Africa in relation to the use of different domains

This study aims to disaggregate estimated poverty measures by gender, population group, settlement type and province, to quantify the nature and extent of poverty. It is well known that South Africa's Apartheid past and years of discriminatory legislation had left the country with a deeply divided socio-economic structure. The Apartheid government polarised the country into urban and non-urban, black and white and privileged and oppressed. In addition, Apartheid further entrenched discrimination based on gender and introduced laws that oppressed women (The Presidency 2014). Although the formal policies of spatial separation by race are long gone, a lingering legacy remains in the rural-urban marker of inequality and poverty (Leibbrandt et al. 2010). Since attainment of democracy in 1994, the state has set out systems to dismantle Apartheid social relations and create a democratic society based on principles of equity, non-racialism and non-sexism (The Presidency 2003). These include the quest to achieve gender equality and to create a free, open and equal society where the rights of all people are promoted irrespective of attributes such as race, gender, age, or disability (StatsSA 2010). The constitution of the Republic of South Africa (Act 108 of 1996) also established provinces which are granted legislative powers ranging from service delivery functions such as education and health to more facilitative functions such as town, regional and development planning.

Although South Africa has undergone a dramatic economic, social and political transformation in the last two decades, many of the distortions and dynamics introduced by Apartheid continue to reproduce poverty and perpetuate inequality (May 1998). It is therefore critical that poverty is

analysed at disaggregated level to assess the impact of the legacy of Apartheid. In addition, domain specific estimates are important as they may be used individually for specific policy purposes where overall estimates might be more useful if analysed over time or benchmarked against other countries. The sample quantile functions of monthly per capita in-kind consumption disaggregated by gender, population groups, settlement type and province confirm the differences. These functions are presented below.

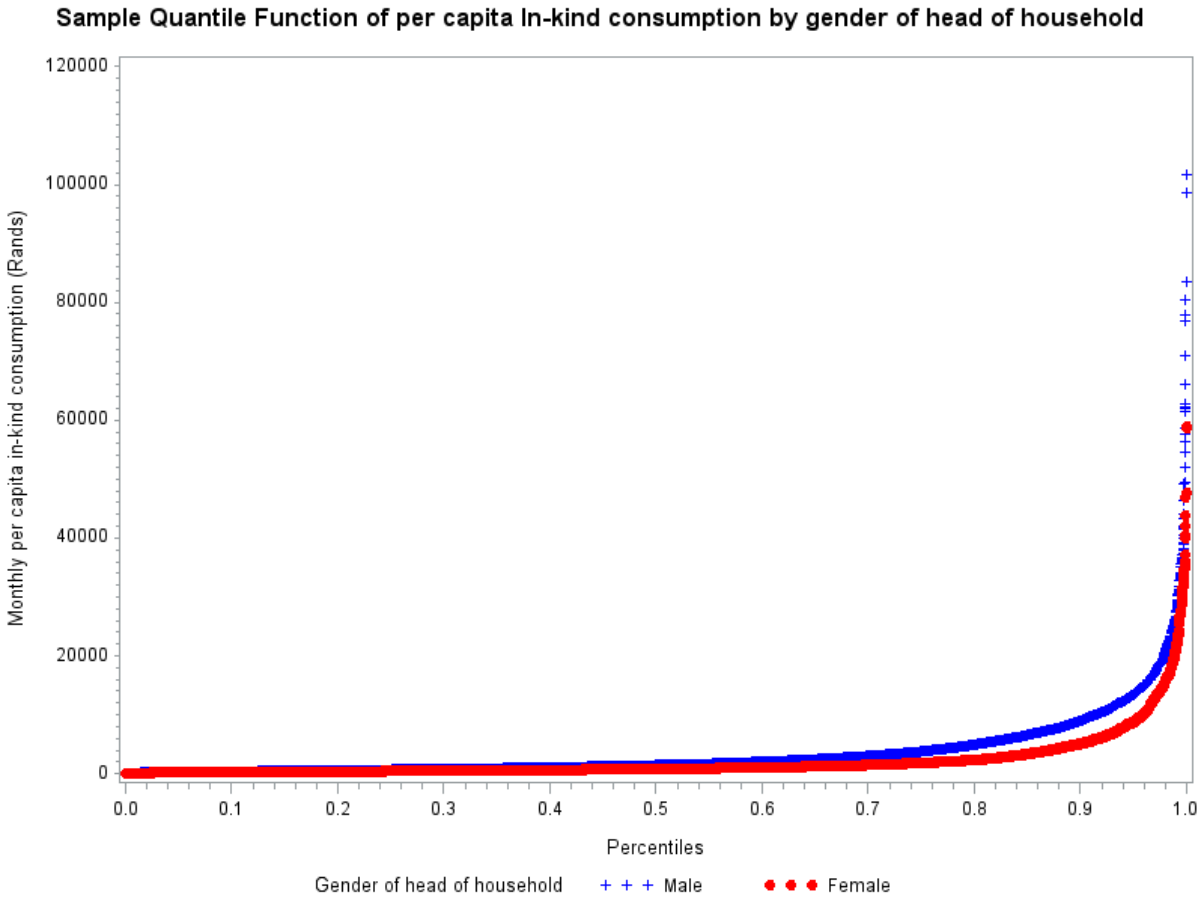

Figure 4.4: Sample Quantile Function of monthly per capita in-kind consumption by gender of head of household

The sample quantile function of monthly per capita in-kind consumption by gender of head of household is given in Figure 4.4 above. It suggests that the monthly per capita consumption in male headed households is generally higher than that in households headed by females and this tendency is more prominent in households with higher consumption. It is therefore necessary that when analysing poverty, it is disaggregated by gender of head of household to assess the statistical significance of the difference suggested by the quantile function.

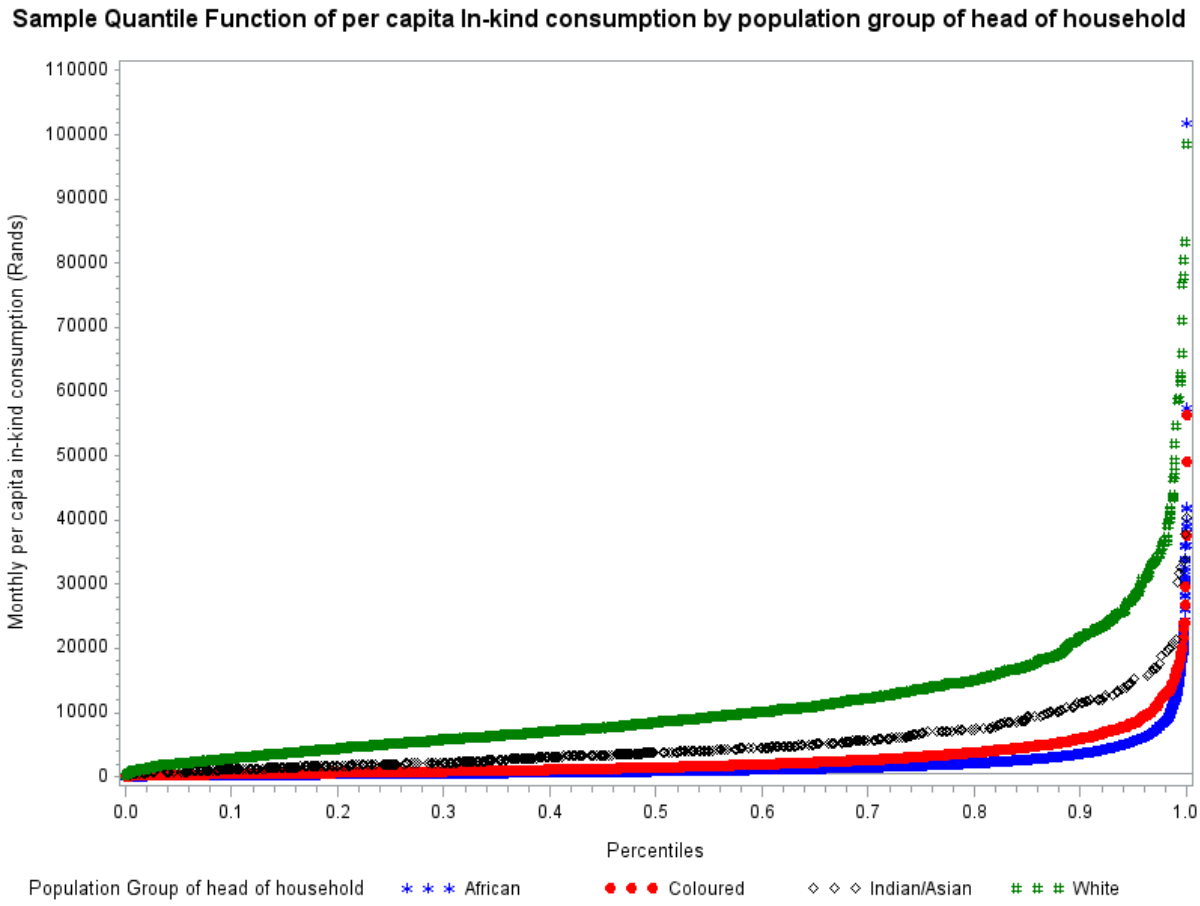

Figure 4.5: Sample Quantile Function of monthly per capita in-kind consumption by population group of head of household

The sample quantile function of monthly per capita in-kind consumption by population group of head of household is given in Figure 4.5. It suggests that the monthly per capita consumption in households headed by Whites is generally higher as compared to other population groups. In contrast, the monthly per capita consumption in households headed by Africans is generally lower. It is however worth noting that the highest monthly per capita consumption in this sample is that of a household headed by an African. The picture given by Figure 4.5 necessitates the estimation of poverty by population group of head of household to assess whether there is difference in poverty amongst households headed by different population groups and if there is, whether it is statistically significant or not.

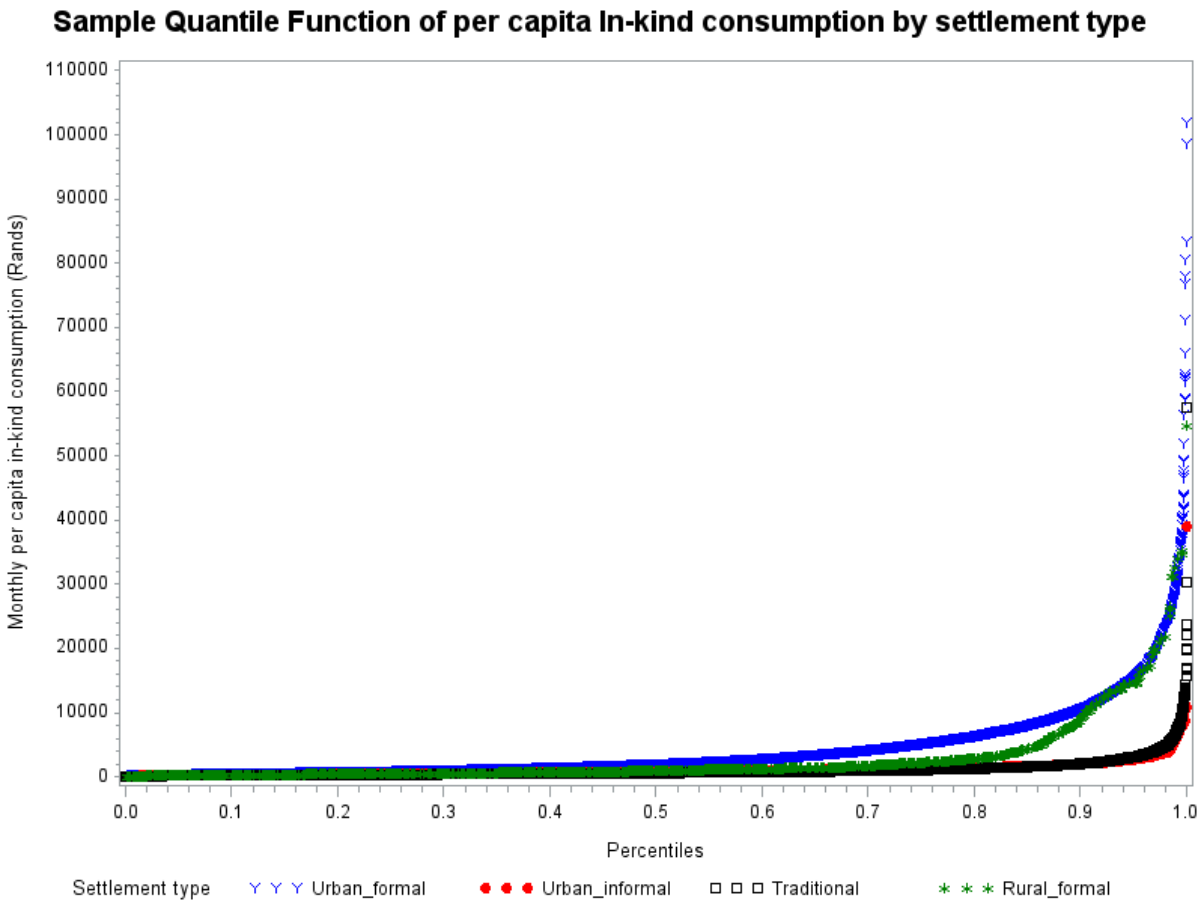

Figure 4.6: Sample Quantile Function of monthly per capita in-kind consumption by settlement type

The sample quantile function of monthly per capita in-kind consumption is given by settlement type in Figure 4.6. It suggests that the monthly per capita consumption amongst households in urban formal areas is generally higher as compared to other settlement types. The monthly per capita consumption amongst households in urban informal areas and in traditional areas has similar distribution. There are visible differences in upper quartile especially in urban formal and rural formal settlement types. It is worth noting that although monthly per capita in-kind consumption of households in rural formal areas is generally higher than that of households in urban informal and traditional areas, there are outlying households in urban informal and traditional areas in the upper quartile with monthly per capita in-kind consumption higher than that of households in rural formal areas. This interesting distribution of in-kind consumption per settlement type also makes it necessary to estimate poverty by settlement type.

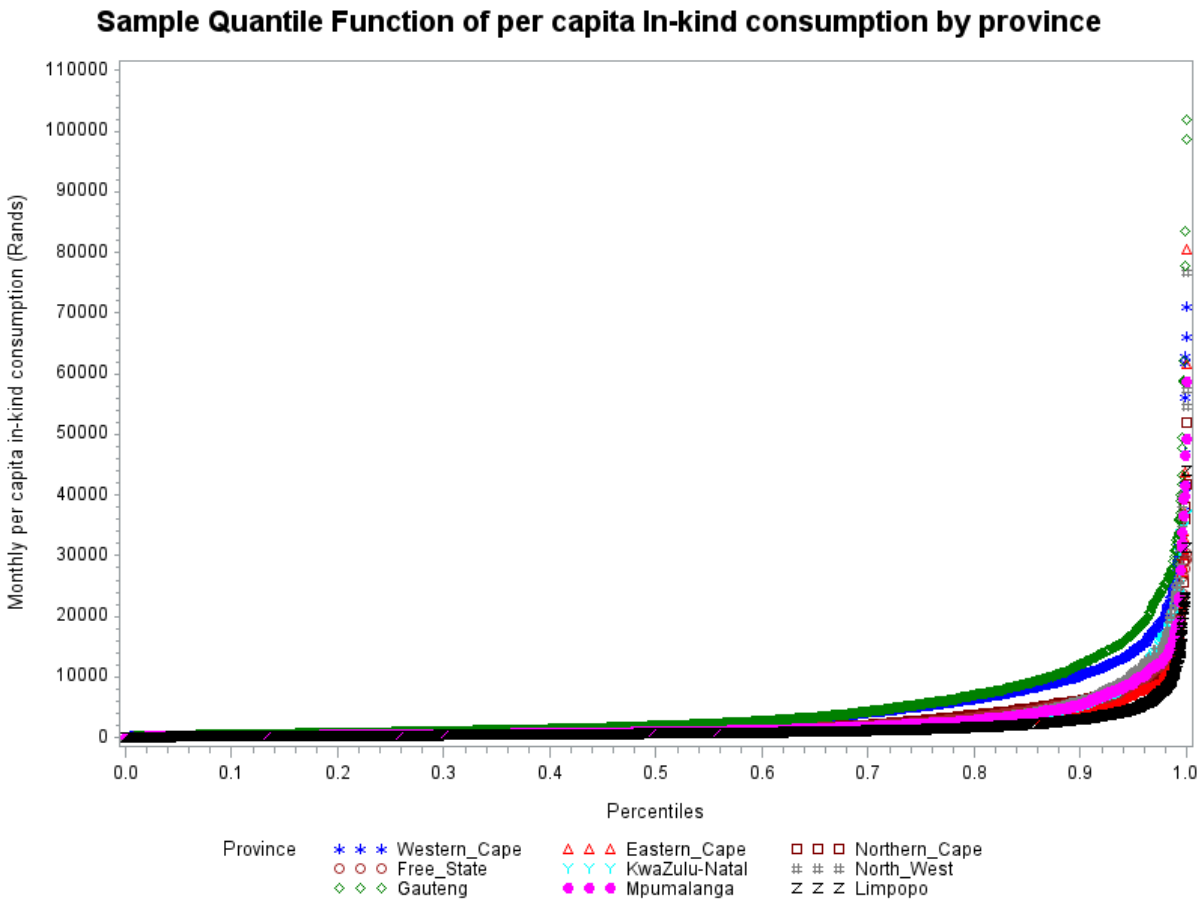

Figure 4.7: Sample Quantile Function of monthly per capita in-kind consumption by province

The sample quantile function of monthly per capita in-kind consumption by province is given in Figure 4.7 above. Monthly per capita consumption amongst households in Gauteng is generally higher as compared to other provinces. The monthly per capita consumption amongst households in Limpopo seems to be the lowest. It is interesting to note that though monthly per capita in-kind consumption of households in Western Cape is generally higher than that of other provinces except Gauteng, there are outlying households in North West and Eastern Cape in the upper quartile with monthly per capita in-kind consumption higher than that of households in Western Cape.

4.4.2. Poverty Lines

In order to derive or calculate poverty estimates, a poverty line must be selected to distinguish between the poor and the non-poor households. The methodology followed in developing poverty lines in South Africa was discussed in Section 3.2. As discussed, the poverty lines constructed are period specific and require regular updates to maintain relevance in long term statistical use. Table 4.2 shows inflation adjusted poverty lines used to calculate poverty estimates using the 2010/11 IES household data. These poverty lines have been benchmarked to March 2011 prices as these

represent the mid-term point of the survey. The values for the poverty lines are **R321** for Food Poverty Line, **R443** for Lower-bound poverty line and **R620** for the Upper-bound poverty line. The upper-bound poverty line is used specifically in this study to derive poverty estimates.

Table 4.2: Inflation adjusted poverty lines

Inflation Adjusted Poverty lines (per capita per month in Rands)				
	Year	Food Poverty Line (FPL)	Lower-bound poverty line(LBPL)	Upper-bound poverty line (UBPL)
Amount	2011 (March)	321	443	620

Source: Poverty Trends in South Africa: An examination of absolute poverty between 2006 and 2011 / Statistics South Africa, p8

It is however worth mentioning that in 2015, a methodological report on rebasing the national poverty lines and development of provincial poverty lines was released by Statistics South Africa (StatsSA 2015). The rebasing of poverty lines is another way of maintaining the integrity of absolute poverty lines. This rebasing was informed by the fact that spending and consumption patterns change over time, which implies that, the basket of goods and services on which the existing poverty lines are based may have changed. Data from the Income and Expenditure Survey 2011 was used to rebase the national poverty lines as well as to provide pilot poverty lines for all the provinces. The rebased poverty lines are presented in Table 4.3. Because of the late release of the new poverty lines, this study uses the poverty lines provided in Table 4.2. Given the smoothness of the in-kind consumption distribution as indicated by the sensitivity analysis, many of the conclusions would not be seriously affected by a slightly changed poverty line.

Table 4.3: Rebased poverty lines using IES 2011

Rebased Poverty lines (per capita per month in Rands)				
	Year	Food Poverty Line (FPL)	Lower-bound poverty line(LBPL)	Upper-bound poverty line (UBPL)
Amount	2011 (March)	335	501	779

Source: Methodological report on rebasing of national poverty lines and development of pilot provincial poverty lines / Statistics South Africa, p11

4.4.3. Bonferroni Adjustment

In this study, inference is done using hypothesis testing and confidence intervals. Multiple tests are performed using confidence intervals of estimated poverty measures. Bonferroni adjustment is a type of multiple comparison test used in statistical analysis. When several dependent or

independent statistical tests are being performed simultaneously on a single dataset, the probability of obtaining false-positive results (Type I error) due to chance increases as more hypotheses are tested. Bonferroni adjustment is used to reduce the chances of obtaining false-positive results (Type I errors) when these multiple pair wise tests are performed on a single set of data. It is based on the idea that if an experimenter is testing n dependent or independent hypotheses on a set of data, the probability of type I error is offset by testing each hypothesis at a statistical significance level $1/n$ times what it would be if only one hypothesis were tested; i.e. Bonferroni adjustment suggests that the value of the level of significance for each test must be equal to $\propto$ divided by the number of tests, where $\propto$ is the level of significance (Miller 1981).

The multiple tests are conducted using confidence intervals of estimated poverty measures by population group, settlement type and by province. When testing hypotheses for different population groups and types of settlements, four combination two ($_4C_2$) =6 tests or comparisons are made. Similarly, when testing hypothesis for different provinces, ($_9C_2$) =36 comparisons are made. This therefore implies that when using Bonferroni Adjustment, the level of significance ($\propto = 0.05$) will be divided by $n = 6$ for the cases of population groups and settlement types and by $n = 36$ for the case of comparisons for different provinces. The Bonferroni Adjustments made are summarised in Table 4.4.

Table 4.4: Adjusted level of significance using the Bonferroni Adjustments for confidence intervals of poverty measures

Domain of analysis	**Original level of significance**	**Number of comparisons performed (n)**	**Adjusted level of significance ($\propto$/n)**
Population group of head of household	$\propto = 0.05$	6	0.00833
Settlement type	$\propto = 0.05$	6	0.00833
Province	$\propto = 0.05$	36	0.00139

4.4.4. Estimated poverty measures using exact and Jackknife Method

Linearization is a method for estimating the variance of a nonlinear function of estimated population totals by using a Taylor series expansion. See Section 3.7.1 for full discussion. The Taylor Series method is given as one of the variance estimation methods in SAS Survey Means Procedure. SAS applies the Taylor series method by default if no variance estimation method is specified. If the statistics (poverty measure) is a liner function, the Taylor series option calculates the standard error

using (3.41) and (3.42). The Jackknife method is a resampling method for variance estimation. The estimator of a parameter, say variance, is found by systematically deleting each observation from the data and calculating the estimate, and then finding the average of these calculations. See Section 3.7.2 for full discussion. The Jackknife method is given as a variance estimation method in the SAS software. In Tables 4.6 to 4.10, both exact and Jackknife methods are applied to the Income and Expenditure Survey 2011 household data in estimating standard errors of the weighted poverty measures discussed in Section 3.4.1. These measures include the poverty headcount, poverty gap index, squared poverty gap index and the Watts index. Standard errors and 95% confidence intervals are presented for all the poverty measures. All these measures are weighted and estimated by gender of household head, population group of household head, settlement type and by province. The upper bound poverty line of R620 is applied as discussed above. The poverty headcount is calculated from the in-kind consumption per person per month. If the in-kind consumption per person per month is below the upper bound poverty line (R620), then the household is considered poor. Since the data is weighted, the number of these poor households relative to the total number of households which gives the poverty headcount is estimated by (3.13). Other measures including the poverty gap, the poverty gap index, the squared poverty gap index and the Watts index are estimated using equations (3.14) to (3.21).

Table 4.5: Properties of the IES 2011 sample

Data Summary	
Number of Strata	363
Number of PSUs	3 062
Number of Observations	33 080
Number of Observations Used	25 328
Number of Observations with Non-positive Weights	7 752
Sum of Weights	13 112 216

Table 4.5 displays the data summary produced when estimating standard errors of different weighted poverty measures. The summary shows that 7 752 observations with non-positive weights are not used for analysis. These are households from the Master Sample which are excluded from analysis. The number of strata was 363 and the number of PSUs was 3 062.

Table 4.6: Poverty estimates, standard errors and 95% confidence intervals

Disaggregation variable	Group/ Domain	Sample size (n)	Poverty measure	Poverty estimate	Standard error	Standard error (Jackknife)	95% confidence interval		95% confidence interval (Jackknife)	
Total	Total	25 328	P_{HC}	0.29295	0.00453	0.00453	0.28408	0.30183	0.28407	0.30183
	Total		P_{PGI}	0.11234	0.00212	0.00212	0.10818	0.11650	0.10818	0.11650
	Total		P_{SPGI}	0.05759	0.00130	0.00130	0.05505	0.06014	0.05505	0.06014
	Total		P_{W}	0.16589	0.00343	0.00343	0.15916	0.17261	0.15916	0.17261

The poverty headcount, which is given by (3.13), was 0.293 in 2011. This implies that approximately 29.3% of households in South Africa were living below the upper bound poverty line of R620 per capita per month. The poverty gap index, squared poverty gap index and the watts index are given by 0.11, 0.06 and 0.17 respectively. The squared poverty gap index and the watts index takes inequality amongst the poor into account. The watts index is even more sensitive to changes in the lower consumption than it is to changes for those with higher consumption. The squared poverty gap index and the watts index can be used to target poverty alleviation resources. Due to their distributional sensitivity, allocation of poverty alleviation resources based on the latter will tilt towards the poor and will result in successful eradication of poverty. The standard errors and confidence intervals based on the jackknife method were very similar to the exact method.

Table 4.6a: Estimation of Sen Index, Alternative Sen Index and Sen-Shorrocks-Thon Index

Variable	P_{HC}	P_{PGI}	P^{P}_{PGI}	G^{P}	$\widehat{G}^{P}$	$\bar{y}^{P}$	z	P_{S}	$P_{S(alt)}$	P_{SST}
Value	0.29295	0.11234	0.38347	0.20785	0.33447	382.24486	620	0.14988	0.14988	0.14991

$P_{S(alt)}$ = alternative Sen Index given by (3.7)

Table 4.6a gives values of variables required in calculating the Sen Index (P_S), Alternative Sen Index ($P_{S(alt)}$) and the Sen-Shorrocks-Thon Index (P_{SST}). The values for poverty headcount index and poverty gap index are taken from Table 4.6 above. The value for the upper bound poverty line is taken from Table 4.2 above. Values for mean of the poor ($\bar{y}^{P}$), poverty gap index for the poor (P^{P}_{PGI}) and Gini coefficient for the poor (G^{P}) and Gini coefficient of the gaps ($\widehat{G}^{P}$) are estimated from the IES 2011 data using in-kind consumption for poor households only, i.e. households whose income is below the upper bound poverty line of R620 per capita per month. This allows the estimation of the Sen Index, Alternative Sen Index and the Sen-Shorrocks-Thon Index.

The standard errors for Sen Index, Alternative Sen Index and the Sen-Shorrocks-Thon can be calculated by means of the Jackknife method. The calculation requires non-standard software and was not included in this analysis.

Table 4.7: Poverty estimates, standard errors and 95% confidence intervals by gender of head of household

Disaggregation variable	Group/ Domain	Sample size (n)	Poverty measure	Poverty estimate	Standard error	Standard error (Jackknife)	95% confidence interval		95% confidence interval (Jackknife)	
Gender of head of household	**Male**	14 283	P_{HC}	0.22590	0.00511	0.00512	0.21587	0.23592	0.21587	0.23593
			P_{PGI}	0.08390	0.00230	0.00230	0.07940	0.08840	0.07940	0.08840
			P_{SPGI}	0.04263	0.00142	0.00142	0.03985	0.04541	0.03985	0.04541
			P_W	0.12353	0.00376	0.00376	0.11616	0.13089	0.11616	0.13090
	Female	11 045	P_{HC}	0.39617	0.00667	0.00667	0.38309	0.40925	0.38309	0.40926
			P_{PGI}	0.15612	0.00332	0.00332	0.14961	0.16262	0.14961	0.16262
			P_{SPGI}	0.08062	0.00209	0.00209	0.07652	0.08472	0.07652	0.08472
			P_W	0.23109	0.00545	0.00545	0.22041	0.24177	0.22040	0.24177

The poverty measures are estimated by gender of head of household as shown in Table 4.7. All poverty measures show that poverty was high among female-headed households as compared to households headed by males. All the poverty estimates are almost twice as high for female headed households compared to male headed households. The standard errors for all estimated poverty measures were higher for females headed households than those for male headed households.

The hypothesis is,

$H_0: \pi_{male} = \pi_{female}$

$H_1: \pi_{male} < \pi_{female}$,

where π is equal to P_{HC} or P_{PGI} or P_{SPGI} or P_W.

The test statistic given by $Z = \frac{P_{male} - P_{female}}{\sqrt{S^2_{male} + S^2_{female}}}$ was used to calculate the critical value for this one-sided test provided in Table 4.7a. The null hypotheses for all poverty measures are rejected and it follows that the poverty measures were significantly lower for male than female headed households.

Table 4.7a: Test statistics for the comparison of the poverty measures of male and female headed households

Poverty measure	z-value for $H_0: \pi_{male} = \pi_{female}$
P_{HC}	-20.26 (p-value< 0.00001)
P_{PGI}	-17.91 (p-value< 0.00001)
P_{SPGI}	-15.05 (p-value< 0.00001)
P_W	-16.26 (p-value< 0.00001)

All p-values are far less than 0.001.

Figure 4.8 shows graphically that the 95% confidence intervals of all poverty measures do not overlap. This confirms the rejection of the null hypotheses and it is concluded that poverty measures were significantly lower for male than female headed households.

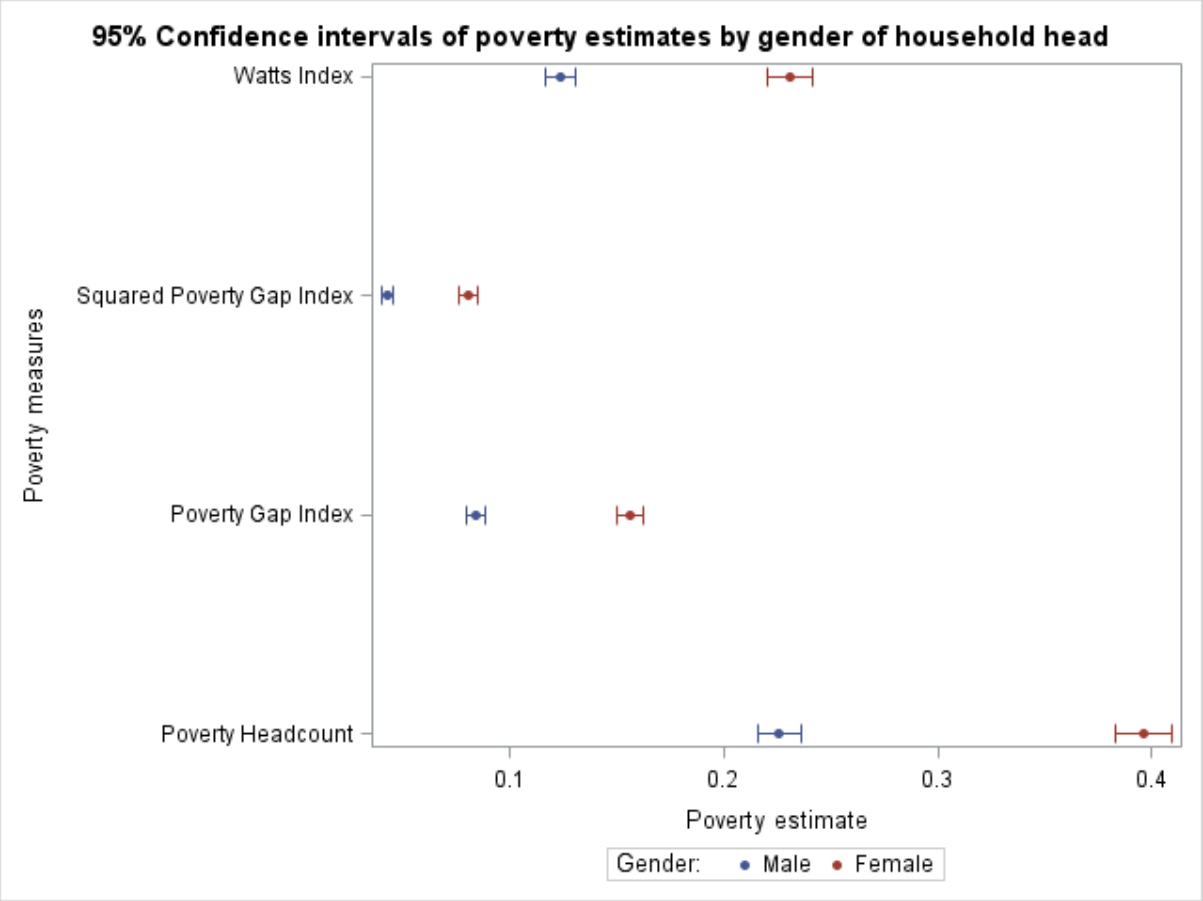

Figure 4.8: Confidence intervals of poverty estimates by gender of household head

Table 4.8: Poverty estimates, standard errors and 95% confidence intervals by population group of head of household

Disaggregation variable	Group/ Domain	Sample size (n)	Poverty measure	Poverty estimate	Standard error	Standard error (Jackknife)	95% confidence interval		95% confidence interval (Jackknife)	
Population group of head of household	African/Black	20 083	P_{HC}	0.3598	0.0055	0.0055	0.3490	0.3706	0.3490	0.3706
			P_{PGI}	0.1394	0.0027	0.0027	0.1341	0.1446	0.1341	0.1446
			P_{SPGI}	0.0720	0.0016	0.0016	0.0687	0.0752	0.0687	0.0752
			P_W	0.2066	0.0043	0.0043	0.1981	0.2151	0.1981	0.2151
	Coloured	2 723	P_{HC}	0.1941	0.0123	0.0123	0.1700	0.2182	0.1699	0.2182
			P_{PGI}	0.0631	0.0045	0.0045	0.0543	0.0720	0.0543	0.0720
			P_{SPGI}	0.0280	0.0023	0.0023	0.0234	0.0326	0.0234	0.0326
			P_W	0.0864	0.0065	0.0065	0.0736	0.0992	0.0735	0.0992
	Indian/Asian	468	P_{HC}	0.0179	0.0079	0.0079	0.0024	0.0333	0.0024	0.0334
			P_{PGI}	0.0052	0.0026	0.0027	0.0001	0.0104	0.0001	0.0104
			P_{SPGI}	0.0023	0.0013	0.0013	0*	0.0049	0*	0.0049
			P_W	0.0071	0.0038	0.0038	0*	0.0145	0*	0.0146
	White	2 054	P_{HC}	0.0044	0.0015	0.0015	0.0016	0.0073	0.0016	0.0073
			P_{PGI}	0.0011	0.0004	0.0004	0.0003	0.0019	0.0003	0.0019
			P_{SPGI}	0.0004	0.0002	0.0002	0*	0.0009	0*	0.0009
			P_W	0.0015	0.0006	0.0006	0.0002	0.0027	0.0002	0.0027

** = truncated at zero*

Table 4.8 shows that the highest proportion of households living below the upper bound poverty line were households headed by Black Africans whereas the poverty headcounts for households headed by both Indians and Whites were very low. The standard errors for all poverty measures were relatively high amongst households headed by Coloureds, followed by those headed by Indians. This might be due to high variations in in-kind consumptions in these population groups because of relatively smaller sample size.

The hypothesis is

$H_0: \pi_{African} = \pi_{Coloured} = \pi_{White} = \pi_{Indian/Asian}$

H_1: At least one of the four population groups' poverty measures differs from the others

where π is equal to P_{HC} or P_{PGI} or P_{SPGI} or P_W.

Figure 4.9 shows the Bonferroni adjusted confidence intervals for different poverty measures by population group of household head. There is an overlap of confidence intervals for different poverty measures between Indian and White headed households which implies that the difference between estimated poverty in Indian and White headed households is not statistically significant. This is confirmed by the z-test in Table 4.8a. However, confidence intervals for poverty measures of households headed by Africans and Coloureds do not overlap with confidence intervals of households headed by other population groups. This implies that there is a statistically significant difference in estimated poverty between households headed by Africans and Coloureds as compared to households headed by other population groups. Poverty amongst households headed by Africans is significantly higher than poverty in households headed by other population groups.

Table 4.8a: Test statistics for the comparison of the poverty measures of Indian and White headed households

Poverty measure	Z-statistic for Indian/Asian-White (critical value: + or -2.63826)	p-value	Reject null hypothesis
P_{HC}	1.68323	0.04616	No
P_{PGI}	1.55269	0.06025	No
P_{SPGI}	1.33749	0.09053	No
P_{W}	1.48460	0.06882	No

Since this is one of a set of 6 multiple comparisons, the significance level for this two-sided test is given by 0.05/12=0.004167.

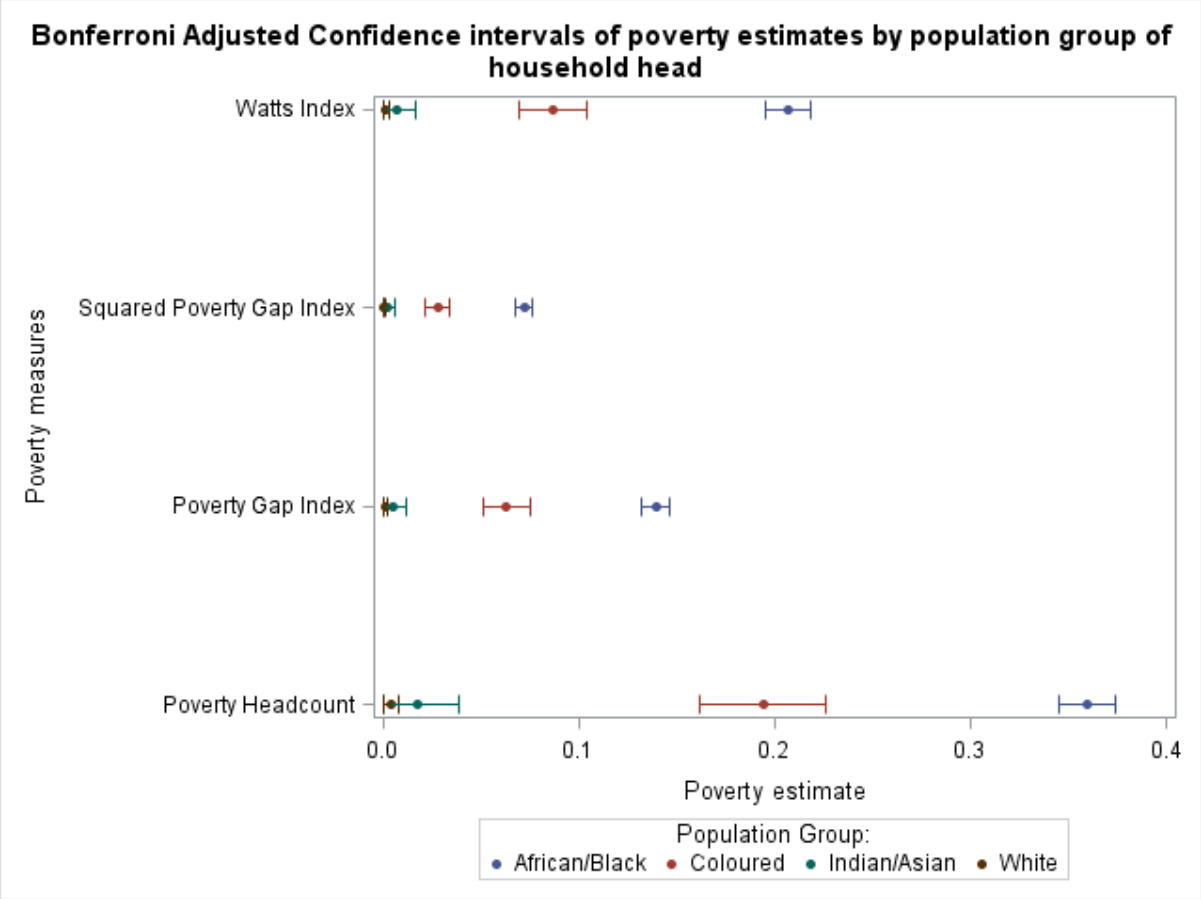

Figure 4.9: Bonferroni adjusted confidence intervals of poverty estimates by population group of household head

Table 4.9: Poverty estimates, standard errors and 95% confidence intervals by type of settlement

Disaggregation variable	Group/ Domain	Sample size (n)	Poverty measure	Poverty estimate	Standard error	Standard error (Jackknife)	95% confidence interval		95% confidence interval (Jackknife)	
Settlement type	**Urban formal**	14 583	P_{HC}	0.1635	0.0059	0.0059	0.1520	0.1751	0.1520	0.1751
			P_{PGI}	0.0549	0.0025	0.0025	0.0501	0.0598	0.0501	0.0598
			P_{SPGI}	0.0261	0.0014	0.0014	0.0233	0.0290	0.0233	0.0290
			P_W	0.0784	0.0039	0.0039	0.0708	0.0861	0.0708	0.0861
	Urban informal	1 712	P_{HC}	0.3878	0.0192	0.0192	0.3502	0.4254	0.3501	0.4255
			P_{PGI}	0.1475	0.0091	0.0091	0.1296	0.1653	0.1296	0.1654
			P_{SPGI}	0.0731	0.0054	0.0054	0.0626	0.0836	0.0626	0.0837
			P_W	0.2124	0.0141	0.0142	0.1847	0.2402	0.1847	0.2402
	Traditional area	8 063	P_{HC}	0.5213	0.0075	0.0075	0.5066	0.5361	0.5066	0.5361
			P_{PGI}	0.2173	0.0043	0.0043	0.2089	0.2258	0.2089	0.2258
			P_{SPGI}	0.1166	0.0029	0.0029	0.1109	0.1224	0.1109	0.1224
			P_W	0.3284	0.0075	0.0075	0.3137	0.3431	0.3137	0.3431
	Rural formal	970	P_{HC}	0.3499	0.0308	0.0312	0.2895	0.4103	0.2888	0.4110
			P_{PGI}	0.1278	0.0125	0.0126	0.1033	0.1523	0.1030	0.1525
			P_{SPGI}	0.0639	0.0071	0.0072	0.0500	0.0779	0.0499	0.0780
			P_W	0.1873	0.0193	0.0195	0.1494	0.2252	0.1491	0.2255

Table 4.9 shows that poverty headcount was the highest in traditional areas at 0.52. In contrast, poverty headcount was the lowest in urban formal areas at 0.16. All poverty measures ranked poverty high in traditional areas and low in urban formal areas. The standard errors for all poverty measures were relatively high amongst households in rural formal areas. This can be attributed to high variations in in-kind consumptions in this settlement type because of relatively smaller sample size.

The hypothesis is

$H_0: \pi_{Urban\ formal} = \pi_{Urban\ informal} = \pi_{Traditional} = \pi_{Rural\ formal}$

H_1: At least one of the four settlement types' poverty measures differs from the others

where π is equal to P_{HC} or P_{PGI} or P_{SPGI} or P_W.

Figure 4.10 shows that there is an overlap of confidence intervals for different poverty measures between households in urban informal and rural formal areas. This overall implies that there is no statistically significant difference in estimated poverty amongst households in urban informal and rural formal areas. This is confirmed by the z-test in Table 4.9a. However, confidence intervals for poverty measures of households in urban formal and traditional areas do not overlap with confidence intervals of households in other settlement types. This implies that there is a statistically significant difference between estimated poverty in households in urban formal and traditional areas as compared to households in other settlement types. Urban formal poverty measures are significantly lower than poverty measures in all other settlement types and rural formal poverty measures are significantly higher than poverty measures all other settlement types.

Table 4.9a: Test statistics for the comparison of the poverty measures of households in urban informal and rural formal areas

Poverty measure	Z-statistic for Rural formal – urban informal (critical value: + or -2.63826)	p-value	Reject null hypothesis
P_{HC}	1.04507	0.147996	No
P_{PGI}	1.27436	0.101269	No
P_{SPGI}	1.03267	0.150879	No
P_{W}	1.05032	0.146786	No

Since this is one of a set of 6 multiple comparisons, the significance level for this two-sided test is given by 0.05/12=0.004167.

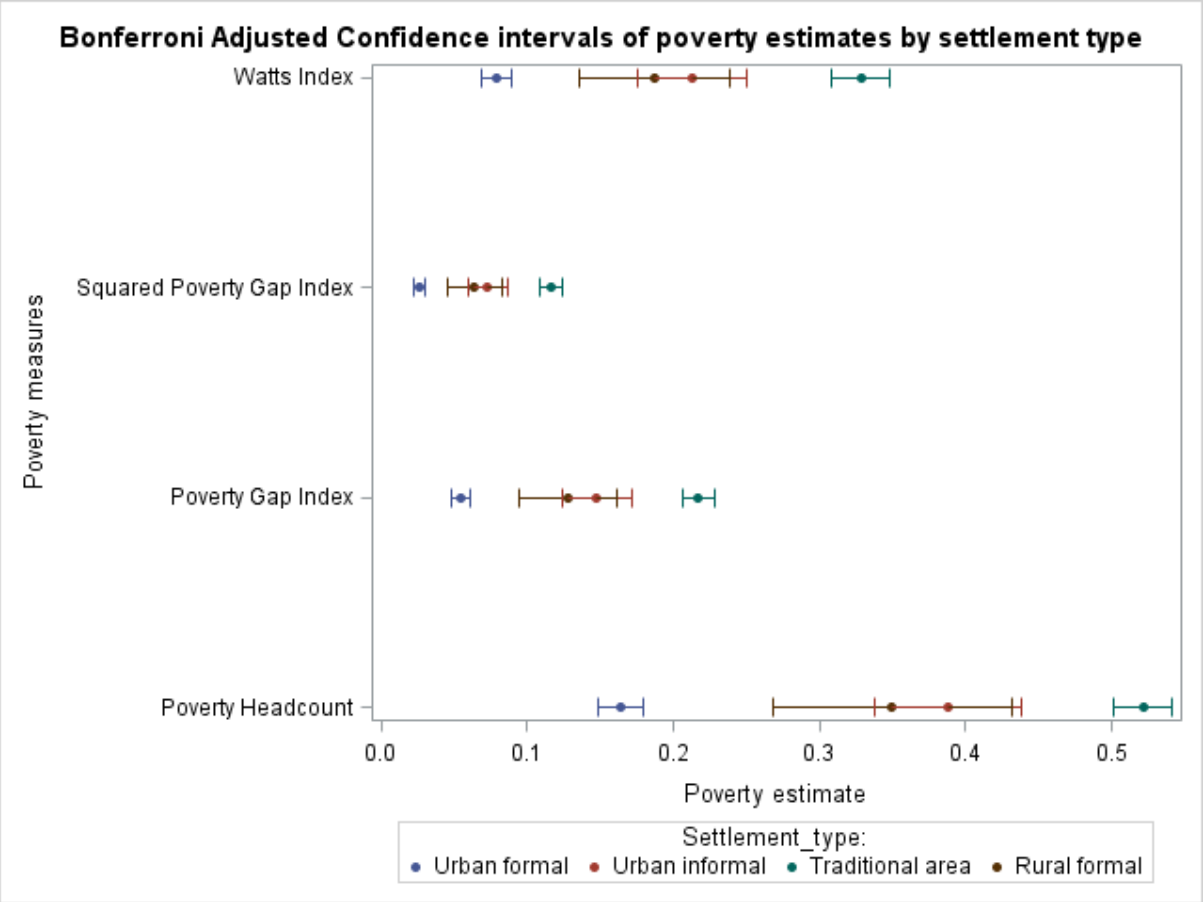

Figure 4.10: Bonferroni adjusted confidence intervals of poverty estimates by settlement type

Table 4.10: Poverty estimates, standard errors and 95% confidence intervals by province

Disaggregation variable	Group/ Domain	Sample size (n)	Poverty measure	Poverty estimate	Standard error (Taylor)	Standard error (Jackknife)	95% confidence interval (Taylor)		95% confidence interval (Jackknife)	
Province	**Western Cape**	2 970	P_{HC}	0.15440	0.01204	0.01205	0.13079	0.17801	0.13078	0.17802
			P_{PGI}	0.04785	0.00466	0.00466	0.03871	0.05699	0.03870	0.05699
			P_{SPGI}	0.02096	0.00232	0.00232	0.01641	0.02550	0.01641	0.02550
			P_{W}	0.06550	0.00663	0.00663	0.05249	0.07850	0.05249	0.07850
	Eastern Cape	3 333	P_{HC}	0.42754	0.01391	0.01397	0.40028	0.45481	0.40015	0.45494
			P_{PGI}	0.17125	0.00722	0.00726	0.15709	0.18542	0.15702	0.18548
			P_{SPGI}	0.08957	0.00459	0.00461	0.08057	0.09856	0.08054	0.09860
			P_{W}	0.25507	0.01188	0.01193	0.23178	0.27837	0.23169	0.27846
	Northern Cape	1 205	P_{HC}	0.33312	0.01766	0.01770	0.29850	0.36774	0.29841	0.36782
			P_{PGI}	0.12284	0.00742	0.00743	0.10829	0.13738	0.10827	0.13740
			P_{SPGI}	0.06018	0.00429	0.00430	0.05176	0.06860	0.05175	0.06861
			P_{W}	0.17562	0.01156	0.01158	0.15295	0.19828	0.15291	0.19832
	Free State	2 172	P_{HC}	0.25997	0.01232	0.01234	0.23581	0.28413	0.23577	0.28417
			P_{PGI}	0.09208	0.00545	0.00546	0.08139	0.10277	0.08137	0.10279
			P_{SPGI}	0.04427	0.00336	0.00337	0.03768	0.05087	0.03767	0.05088
			P_{W}	0.13059	0.00878	0.00879	0.11339	0.14780	0.11337	0.14782
	KwaZulu-Natal	3 625	P_{HC}	0.38265	0.01290	0.01291	0.35736	0.40794	0.35734	0.40796
			P_{PGI}	0.15322	0.00598	0.00598	0.14150	0.16494	0.14148	0.16495
			P_{SPGI}	0.08059	0.00367	0.00368	0.07339	0.08779	0.07338	0.08779
			P_{W}	0.22853	0.00967	0.00968	0.20957	0.24750	0.20955	0.24751
	North West	2 524	P_{HC}	0.33513	0.01615	0.01625	0.30347	0.36679	0.30327	0.36699
			P_{PGI}	0.13153	0.00718	0.00722	0.11745	0.14562	0.11738	0.14569
			P_{SPGI}	0.06863	0.00427	0.00429	0.06025	0.07700	0.06022	0.07703
			P_{W}	0.19649	0.01151	0.01155	0.17393	0.21906	0.17384	0.21915
	Gauteng	3 887	P_{HC}	0.13968	0.00840	0.00844	0.12321	0.15614	0.12313	0.15622
			P_{PGI}	0.04633	0.00372	0.00374	0.03903	0.05363	0.03900	0.05366
			P_{SPGI}	0.02201	0.00230	0.00231	0.01751	0.02651	0.01749	0.02653
			P_{W}	0.06657	0.00620	0.00623	0.05440	0.07873	0.05434	0.07878
	Mpumalanga	2 306	P_{HC}	0.33296	0.01489	0.01494	0.30376	0.36216	0.30366	0.36226
			P_{PGI}	0.12173	0.00663	0.00665	0.10873	0.13474	0.10870	0.13477
			P_{SPGI}	0.06047	0.00414	0.00415	0.05235	0.06859	0.05233	0.06861
			P_{W}	0.17677	0.01075	0.01078	0.15569	0.19786	0.15564	0.19790
	Limpopo	3 306	P_{HC}	0.45647	0.01173	0.01174	0.43348	0.47946	0.43346	0.47948
			P_{PGI}	0.19154	0.00675	0.00675	0.17831	0.20477	0.17829	0.20478
			P_{SPGI}	0.10307	0.00463	0.00464	0.09400	0.11215	0.09398	0.11216
			P_{W}	0.29048	0.01184	0.01185	0.26727	0.31369	0.26724	0.31372

Table 4.10 shows the poverty measures estimated by province. The province with high presence of poverty is Limpopo and poverty was the lowest in Gauteng. The standard errors for all the poverty measures were high amongst households in Northern Cape and very low amongst households in Gauteng. This is attributed to small sample size in Northern Cape and high sample size in Gauteng.

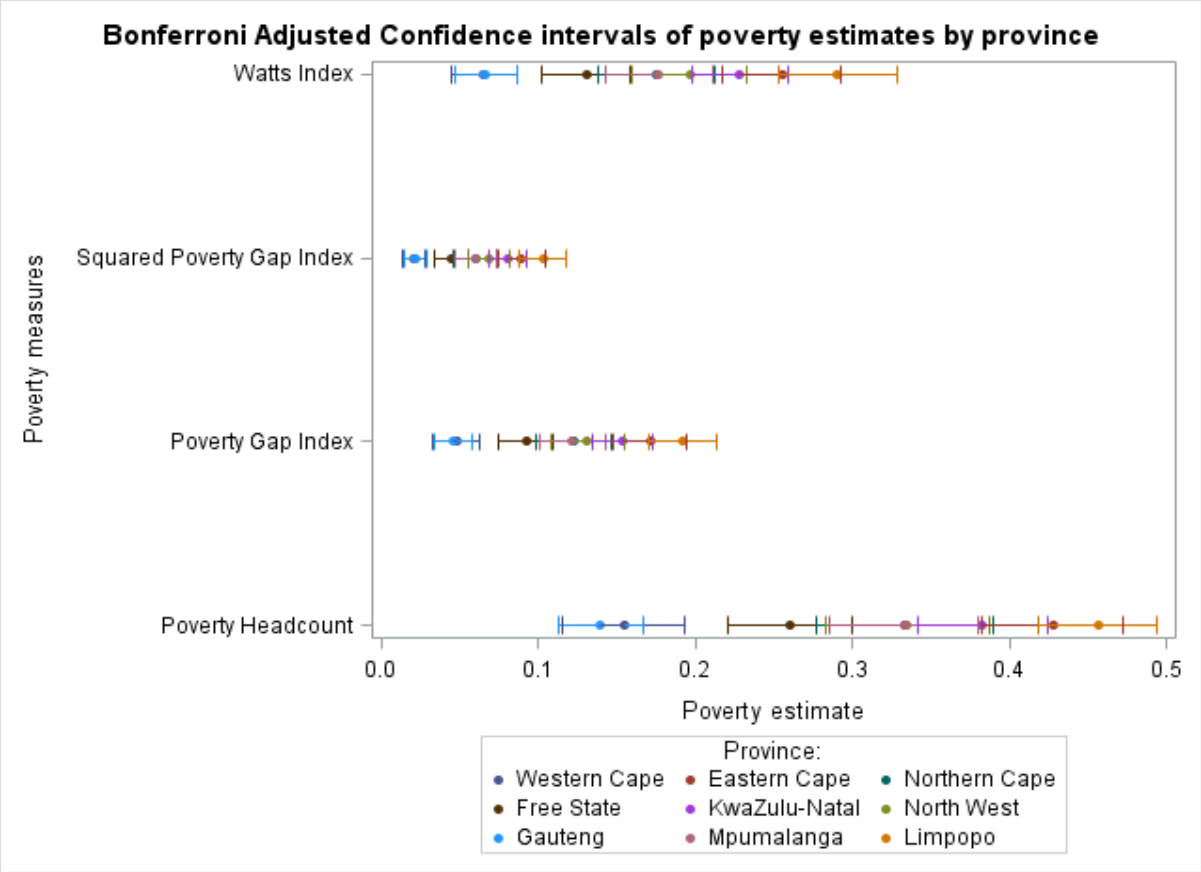

Figure 4.11: Bonferroni adjusted confidence intervals of poverty estimates by province

The hypothesis is

H_0: $\pi_{WC} = \pi_{EC} = \pi_{NC} = \pi_{FS} = \pi_{KZN} = \pi_{NW} = \pi_{GP} = \pi_{MP} = \pi_{LP}$

H_1: At least one of the nine provinces' poverty measures differs from the others

where π is equal to P_{HC} or P_{PGI} or P_{SPGI} or P_W.

Figure 4.11 shows that there is an overlap of confidence intervals for different poverty measures between households in different provinces. The confidence intervals for poverty measures in Gauteng overlap with those of poverty measures in Western Cape. This implies that the difference in estimated poverty amongst households in these two provinces is not statistically significant. However the estimated poverty in these two provinces is statistically different from estimated poverty in all other provinces. For ease of analysis, pairwise comparisons per poverty measure are conducted. The test statistics for these pairwise comparisons of the different poverty measures by province are given in Tables 4.10a – 4.10d. The pairwise comparisons are sets of 36 multiple comparisons, therefore the significance level for these pairwise comparisons is given by 0.05/72=0.000694. The z-statistics for all pairwise comparisons is -3.19695. The bold and shaded z values indicate differences that are not statistically significant.

Table 4.10a: Test statistics for the pairwise comparison of the poverty headcount (P_{HC}) for different provinces

Province	Western Cape	Eastern Cape	Northern Cape	Free State	KwaZulu-Natal	North West	Gauteng	Mpumalanga
Eastern Cape	14.84845							
Northern Cape	8.36287	-4.20152						
Free State	6.12832	-9.01981	-3.39780					
KwaZulu-Natal	12.93656	**-2.36723**	**2.26544**	6.87874				
North West	8.97293	-4.33678	**0.08410**	3.70084	**-2.29956**			
Gauteng	**-1.00315**	-17.71999	-9.89465	-8.06875	-15.78884	-10.73983		
Mpumalanga	9.32336	-4.64184	**-0.00684**	3.77645	**-2.52217**	**-0.09879**	11.30507	
Limpopo	17.97307	**1.59004**	5.82011	11.55404	4.23538	6.08083	21.96635	6.51599

Table 4.10b: Test statistics for the pairwise comparison of the poverty gap index (P_{PGI}) for different provinces

Province	Western Cape	Eastern Cape	Northern Cape	Free State	KwaZulu-Natal	North West	Gauteng	Mpumalanga
Eastern Cape	14.35417							
Northern Cape	8.56126	-4.67655						
Free State	6.16638	-8.74706	-3.34101					
KwaZulu-Natal	13.90058	**-1.92328**	**3.18957**	7.55562				
North West	9.77232	-3.89859	**0.84232**	4.37414	**-2.32026**			
Gauteng	**-0.25397**	-15.37113	-9.21951	-6.92896	-15.17722	-10.52984		
Mpumalanga	9.11607	-5.04970	**-0.11067**	3.45397	-3.52638	**-1.00220**	9.91536	
Limpopo	17.52415	**2.05239**	6.85290	11.46558	4.25135	6.08887	18.84533	7.37929

All poverty measures show no significance difference in poverty in Western Cape and Gauteng which confirms the conclusion using Bonferroni adjusted confidence intervals. All poverty measures show poverty in Free State is significantly different from poverty in all other provinces except when measured by Squared Poverty Gap Index. The Squared Poverty Gap Index shows poverty in Free State is not significantly different from poverty in Northern Cape and in Mpumalanga. Another interesting observation is that all poverty measures show that poverty in Limpopo significantly different from poverty in all provinces except Eastern Cape. All poverty measures show that poverty is Eastern Cape is significantly different from poverty in all other provinces except KwaZulu-Natal and Limpopo.

Table 4.10c: Test statistics for the pairwise comparison of the squared poverty gap index (P_{SPGI}) for different provinces

Province	Western Cape	Eastern Cape	Northern Cape	Free State	KwaZulu-Natal	North West	Gauteng	Mpumalanga
Eastern Cape	13.34840							
Northern Cape	8.04042	-4.67670						
Free State	5.70809	-7.96132	**-2.91675**					
KwaZulu-Natal	13.73363	**-1.52795**	3.61219	7.29201				
North West	9.81400	-3.34136	**1.39488**	4.48065	**-2.12414**			
Gauteng	**0.32244**	-13.16674	-7.83999	-5.46569	-13.52448	-9.61624		
Mpumalanga	8.32857	-4.70828	**0.04846**	**3.03632**	-3.63528	**-1.37150**	8.12350	
Limpopo	15.86540	**2.07220**	6.79425	10.27661	3.80565	5.47056	15.68918	6.86045

Table 4.10d: Test statistics for the pairwise comparison of the watts index (P_W) for different provinces

Province	Western Cape	Eastern Cape	Northern Cape	Free State	KwaZulu-Natal	North West	Gauteng	Mpumalanga
Eastern Cape	13.93501							
Northern Cape	8.26243	-4.79359						
Free State	5.91862	-8.42877	**-3.10196**					
KwaZulu-Natal	13.90384	**-1.73274**	3.51079	7.49992				
North West	9.86431	-3.54219	**1.28007**	4.55423	**-2.13153**			
Gauteng	**0.11772**	-14.06626	-8.31145	-5.95796	-14.09646	-9.93948		
Mpumalanga	8.80891	-4.88698	**0.07335**	3.32751	-3.57910	**-1.25230**	8.87807	
Limpopo	16.58116	**2.11121**	6.94185	10.85037	4.05256	5.69323	16.75362	7.11025

Comparison of poverty measures and their confidence intervals

A question that may be asked is whether the different measures of poverty tell a consistent story or not. The poverty headcount index, poverty gap index, squared poverty gap index and the watts index ranked poverty amongst female headed households to be higher than amongst male headed households. When estimating poverty per population group of household head, all these measures ranked African headed households high, followed by coloured headed household and then by Indian headed households. The white headed households were ranked lowest by all the measures. These measures were also consistent when estimating poverty levels by settlement type. For all the measures, households within traditional areas was ranked high, followed by households in urban informal areas, and then by households in rural formal areas. Poverty levels were estimated low within households in urban formal areas. A slightly different observation is made when estimating poverty by province. The poverty headcount and the poverty gap index both rank poverty estimates by provinces consistently with the highest being Limpopo and the lowest being Gauteng. The

squared poverty gap index and the watts index, however, rank the provinces differently from the poverty headcount index and the poverty gap index. Both these measures ranks poverty in Mpumalanga to be higher than in Northern Cape and also ranks poverty in Gauteng to be higher than in Western Cape. Although these two measures still ranks Limpopo as the province with the highest presence of poverty, they are different in that they rank Western Cape as the province with the lowest presence of poverty, unlike poverty headcount and poverty gap index which ranks Gauteng as the province with the lowest presence of poverty.

4.5. Overview of different data sources for poverty measurement in South Africa

Different sources of information have been used at different time periods by different researchers to measure levels of poverty in South Africa. These sources differ in focus, coverage, time periods, frequency, variables and methodologies. Since 1995, there has been a steady flow of micro-datasets, differing in focus and coverage. Some of these sources are the Income and Expenditure Survey (IES) introduced in 1995, the population Census first undertaken in 1996 and subsequently in 2001 and 2011, the Labour Force Survey (LFS) introduced in 2000 which has since 2008 became the Quarterly Labour Force Survey (QLFS), the General Household Survey (GHS) introduced in 2002, which replaced the OHS since 2000, the 2007 Community Survey (CS) which was conducted to fill the information gap between the 2001 and 2011 censuses and the Living Conditions Survey (LCS) which was designed specifically to measure the living conditions of South African households. 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.

- The Income and Expenditure Survey (IES), conducted by Statistics South Africa, gathers information about the income and expenditure from South African households. The survey also identifies goods and services purchased by these households. Although the main purpose of IES is to collect data to update the Consumer Price Index (CPI) basket of goods and services, this survey have become an important source of information for poverty and inequality analysis.

- The National Income Dynamics Study (NIDS) is the first national household panel study in South Africa. The NIDS is conducted by the Southern Africa Labour and Development

Research Unit (SALDRU) based at the University of Cape Town. It examines the livelihoods of individuals and households over time and provides information about how they cope with positive or negative shocks, such as a death in the family or an unemployed relative obtaining a job. This survey is relevant for multidimensional poverty measurement as it contains information relevant for common measures of multidimensional poverty.

- The Living Conditions Survey, first conducted in 2008/09 by Statistics South Africa, is aimed to provide for the first time in South Africa, data with which to measure poverty levels. It provides data that can be used for measuring poverty using objective, subjective and multi-dimensional measures, and covers a wide range of indicators. It is very relevant in poverty measurement as it was designed specifically to measure the living conditions of South African households. Though this survey is relevant, it has got only one data point and does not allow poverty comparisons over time.

- The General Household Survey (GHS), also conducted annually by Statistic South Africa, is designed to measure 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. It is a useful instrument for poverty analysis from a multidimensional perspective, but it is not designed to measure household income or consumption.

- The Census and Community Survey, also conducted by Statistics South Africa, provide important baseline data on population characteristics. However, the surveys collect no consumption data and very limited income information necessary for poverty analysis.

- The Quarterly Labour Force Survey, also conducted by Statistics South Africa, is designed exclusively for labour force measurement at a certain point in time; however, for poverty measurement, employment information over a longer period of time is required to get a better understanding and measurement of livelihoods.

In this study, the source of data used is the Income and Expenditure Survey (IES) 2011 household data. The IES is used mainly because it provides information relevant in calculating absolute monetary measures of poverty. An overview of some of these data sources, along with their uses and shortfalls, is given in the Appendix D. The overview includes the name of the survey, the custodian, the frequency or time periods in which the survey was conducted as well as the purpose of each survey.

4.6. Conclusion

Poverty measures were estimated in this chapter. The chapter gave an overview of South Africa's 2011 Income and Expenditure Survey sample design and the IES 2011 household consumption data was used for calculating poverty estimates. The variable used in particular was *In-kind Consumption per capita.* Using the sample quantile function, it was found that the distribution of *In-kind Consumption per capita* varies for different subgroups and therefore it was critical for poverty to be estimated for each subgroup. Inflation adjusted poverty lines were used to distinguish between poor and non-poor households when calculating poverty estimates. The poverty measures estimated followed the first approach of presenting poverty profiles, where the sample was divided into subgroups namely, gender, population group, settlement type and province and the poverty rates were calculated for each subgroup.

The Poverty Headcount (P_{HC}), Poverty Gap Index (P_{PGI}), Squared Poverty Gap Index (P_{SPGI}), Sen Index (P_S), Sen-Shorrocks-Thon Index (P_{SST}) and Watts Index (P_W) were estimated and standard errors and confidence intervals were calculated. The standard errors found using the Jackknife and the exact method were very similar. Multiple comparisons were performed using the z-test and 95% Bonferroni adjusted confidence intervals of estimated poverty measures by the four subgroups. It was found that poverty was significantly higher among female-headed households as compared to households headed by males. Although the difference in estimated poverty between households headed by Black Africans and Coloureds was statistically different, this was not the case between households headed by Indians and Whites. It was found also that the difference in estimated poverty between households in urban formal and traditional areas was statistically significant, but this was not the case between households in urban informal and rural formal areas. Poverty was also not statistically different in Gauteng and Western Cape, however estimated poverty in these two provinces was found to be statistically different from estimated poverty in other provinces. The Free State province's poverty measures were statistically significant from other provinces except when compared using Squared Poverty Gap Index. This chapter concluded by giving an overview of different data sources for poverty measurement in South Africa.

Chapter 5: Summary and conclusions

5.1. Summary

The purpose of this study was to do an overview and assessment of different approaches to poverty measurement in South Africa. The primary objective of the study was to research and present the estimation of different poverty measures and their standard errors and to do basic inference in the case of complex multi-stage sampling as opposed to the standard case of simple random sampling.

The main focus of this chapter is to give summary and conclusions of the study. The chapter commences by giving synopsis of previous four chapters by highlighting most crucial issues presented in these chapters. The next section presents key findings, conclusions and recommendations of this study. The chapter finally concludes by stating study limitations as well as future research and possibilities

Synopsis of previous chapters

Chapter One gave an introduction to the study and discussed background to poverty measurement in South Africa. The chapter discussed the importance of measuring poverty in general and established the rationale for studying the subject matter by stating the problem statement and the purpose of the study.

Chapter Two was on review of literature on the definition of poverty in general and within the South African context. The literature review was an important analysis that proved beneficial in understanding the remainder of the study. Some of the important aspects that were covered in this chapter include the following: aspects to consider in defining poverty, concepts related to poverty, different approaches to poverty measurement and the definition of poverty within the South African context.

Chapter Three covered theoretical concepts and methods of measuring poverty. Different poverty indices including the Poverty Headcount (P_{HC}), Poverty Gap (P_{PG}), Poverty Gap Index (P_{PGI}), Squared Poverty Gap Index (P_{SPGI}), Sen Index, (P_S) Sen-Shorrocks-Thon Index (P_{SST}) and Watts Index (P_W) and the time taken to exit poverty were discussed. The procedure followed in developing poverty lines for South Africa as well as how absolute poverty measures are used in the country were discussed to give the South African context of the application of these absolute poverty measures. In addition, the chapter also explained how to estimate poverty from household survey data as well as strengths and limitation to survey data in poverty analysis. Two methods of estimating variances in

case of complex surveys were presented. These are the Taylor Series method and the jackknife method. It was shown how standard errors of the mean and confidence intervals will be used to indicate the precision of the poverty estimates.

Chapter Four presented the estimated poverty measures using South Africa's 2011 Income and Expenditure Survey data. This was preceded by an overview of the sample design of the IES 2011. The variable of analysis used in estimating poverty was *In-kind Consumption per capita.* Using the sample quantile function, it was found that the distribution of *In-kind Consumption per capita* varies for different domains and therefore it was critical for poverty to be estimated for each domain. The poverty measures estimated followed the first approach of presenting poverty profiles, where the sample was divided into domains such as gender, population group, settlement type and province. The poverty rate was shown for each domain. Poverty Headcount (P_{HC}), Poverty Gap Index (P_{PGI}), Squared Poverty Gap Index (P_{SPGI}), and Watts Index (P_W) were estimated. The standard errors and confidence intervals of these poverty measures were calculated using the exact and Jackknife Method. Multiple pairwise comparisons were performed using hypothesis tests and Bonferroni-adjusted 95% confidence intervals of estimated poverty measures by the four domains mentioned above. The chapter concluded by giving an overview of different data sources for poverty measurement in South Africa.

5.2. Findings and conclusions

The conclusions of this study are that South Africa has significantly improved since the fall of Apartheid in addressing the inadequate information base for the measurement of poverty and inequality. There are now three national poverty lines, the Food Poverty Line and the Upper and Lower bound Poverty Lines which were developed and published by Statistics South Africa to be used when reporting poverty levels in the country. These improvements in measuring poverty are complemented by surveys such as the multi-topic, user guided survey known as the Living Conditions Survey which was specifically designed to measure poverty, and the National Income Dynamics Study which examines the livelihoods of individuals and households over time and provides information about how they cope with positive or negative shocks.

Poverty measurement is essential in providing statistical standards and systematic approach to reporting on levels. The measurement of poverty not only contributes to evidence-based decision making in public policy but also help in assessing the impact of poverty reduction programmes. Regular and appropriate reporting on poverty assists in identifying poor people as well as targeting appropriate interventions.

The study found that there is no single universally accepted definition of poverty as there are so many ways to think about what poverty means and there are several aspects to consider in defining poverty.

Poverty in South Africa is defined and measured in both one-dimensional and multi-dimensional approaches in line with international practices. The one-dimensional approach to poverty measurement uses monetary approach which identifies poverty with a shortfall on consumption expenditure. In multidimensional terms, South Africa uses the South African Multidimensional Poverty Index (SAMPI) which was conceptualised and constructed based on the global Multidimensional Poverty Index (MPI). SAMPI complements the consumption expenditure-based poverty measures by capturing the severe deprivations that each person or household faces. However, there are other poverty measures such as the Sen Index and the Watts Index which are not often utilised.

Poverty measures in South Africa are often not studied and reported with standard errors. This study recognises the importance of reporting poverty estimates with standard errors and confidence intervals as these allows inference when subsequent poverty levels are published. The distributional properties of a variable used to measure poverty, *in-kind consumption*, were studied using the empirical cumulative distribution functions and the sample quantile functions for different subgroups. The subgroups include gender of head of household, population group of head of household, settlement type and province. The distribution of the in-kind consumption was found to vary within and amongst different subgroups. This variation concludes that in South Africa, it is critical that poverty is analysed at disaggregated level.

This research required detailed meta-data on the Income and Expenditure Survey. Access to metadata would be critical for expanding this research to other surveys.

The study found that all poverty measures were consistent and complimentary in estimating poverty levels in South Africa. The Poverty Headcount Index and Poverty Gap Index ranked poverty consistently amongst all subgroups. The Squared Poverty Index and Watts index ranked poverty consistently amongst three subgroups except province. Both these measures ranks poverty in Mpumalanga to be higher than in Northern Cape and also ranks poverty in Gauteng to be higher than in Western Cape. Although these two measures still ranks Limpopo as a province with the highest presence of poverty, they are different in that they rank Western Cape as a province with the lowest presence of poverty, unlike poverty headcount and poverty gap index which ranks Gauteng as the province with the lowest presence of poverty.

The squared poverty gap index and the watts index are considered to be best in that they take inequality amongst the poor into account. The watts index is found to be even more appealing since it is more sensitive to changes in the lowest incomes than it is to changes for those with higher incomes. When allocating anti-poverty resources to minimize the watts index, the effort would tilt towards the poorest due to its distributional sensitivity.

The analysis of in-kind consumption as measured by the IES 2011 indicated that for all domains analysed (gender, population group, settlement type and province), the null hypothesis that poverty measures are the same for all subgroups could be rejected which highlights the value of disaggregated analysis for improved targeting of poverty interventions.

5.3. Study limitations

This study analysed only money metric poverty using the variable called *in-kind consumption.* However, in-kind-consumption expenditure is only one dimension of wellbeing and poverty often involves deprivation on a number of levels. In addition, only data from IES 2011 was used.

5.4. Recommendations

It is common practice in South Africa that poverty estimates are reported without standard errors or confidence intervals. In order to allow inference, particularly interpretation of the statistical significance of shifts in poverty between different reporting times, it is recommended that poverty estimates be reported with standard errors and confidence intervals.

Given the differences in living conditions in South Africa in relation to its Apartheid historical background, as well as the uneven distribution of poverty among the provinces, it is recommended that provincial specific poverty interventions be explored.

It is recommended that when allocating anti-poverty resources, the target should be to minimize the squared poverty gap index and the watts index due to their appealing feature of taking inequality amongst the poor into account. In this way, the poorest will benefit more from the allocated resources.

5.5. Future work and possibilities

This study focused on one-dimensional absolute poverty measures. South Africa has, however, developed the South African Multidimensional Poverty Index (SAMPI) which reveals the combination of deprivations that batter a household at the same time. Future studies can be conducted with the focus on SAMPI.

Efforts should be made, particularly by Statistics South Africa as the National Statistics Office, to ensure promotion of application of poverty thresholds. This will ensure alignment, by different surveys, to poverty thresholds used by Statistics South Africa. This is in line with one key importance of measuring poverty which is to ensure that measures of poverty are put in public domain to help in building a national commitment to eradicate poverty that goes beyond government.

The study used data for one survey period, the 2011 IES. Comparisons of poverty measures over time can be made using data from earlier survey periods such as IES 2000 and IES 2005.

The study disaggregated estimates of poverty by four subgroups namely, gender, population group, settlement type and province. Other determinants of poverty can be studied and poverty can be analysed by such determinants. Such determinants can include analysis of poverty by different age groups, level of education of the household head. etc. or a combination of different subgroups or variables.

This study focused on poverty profiles on key intrinsic aspects such as gender and population group. It did not analyse the causes of poverty or tried to prioritise any factor that would be the main contributor to poverty. The key variable to consider in such analysis is the level of education of head of household even though this will not be a simple cause and effect analysis within the South African context.

This study applied the Jackknife as an alternative to the exact method. The Jackknife method can be used to estimate the standard errors for the Sen Index and the Sen-Shorrocks-Thon Index. There are also other methods such as the Balanced Repeated Replication and Bootstrap which can be used to estimate standard errors and confidence intervals of poverty estimates in case of complex surveys.

The effect of different provincial poverty lines as published in 2015 can be explored. It will be of interest to compare the overall poverty measures based on a single poverty line as opposed to applying different provincial poverty lines.

Appendices

A. Appendix A: Key variables for analysis of poverty from the Income and Expenditure Survey 2011

The following variables are key variables from IES 2011 used for analysis of poverty in this study. The variable names given in brackets are variables names used in the *House_Info* dataset.

- *Province (province)* is the names of the nine South African provinces. These are Western Cape, Eastern Cape, Northern Cape, Free State, KwaZulu-Natal, North West, Gauteng, Mpumalanga, and Limpopo.
- *Type of settlement (settlement_ type)* is the type of settlement where the dwelling unit is situated. The four categories are urban formal, urban informal, traditional area and rural formal.
- *Household size (hsize)* is the total number of persons per household. The range for household sizes in 2011 was 1- 21.
- *Gender of head of the household (GenderOfHead)* asks the gender of the head of the household. Gender was either male or female.
- *Population group of head of the household (PopGrpOfHead)* asks the population group of the head of the household. The four categories for population group are African/Black, Coloured, Indian/Asian and White.
- *Household consumption (Consumptions)* is total un-weighted household consumption, annualised and inflated/deflated to March 2011 prices using the Consumer Price Index (CPI).
- *Household in-kind consumption (In-KindConsumptions)* is the total household consumption plus consumption in-kind (e.g. subsidies from employer, non-refundable bursaries, etc.) un-weighted but annualised and inflated/deflated to March 2011 prices using the CPI.
- *Weight of the household (Full_calwgt)* is the weight of the household.

These variables were used to calculate all the one dimensional poverty measures in this study.

B. Appendix B: Composition and characteristics of households from the IES 2011

In this study, poverty is analysed at household level. It is therefore critical to understand the composition of households as found by the 2010/11 IES. The next few tables present the distribution of households by different critical variables.

Table B.1: Percentage distribution of households by population group and gender of household head

Population group of household head	Gender of household head		Total (%)
	Male (%)	Female (%)	
African/Black	57.06	42.94	76.58
Coloured	64.52	35.48	8.48
Indian/Asian	79.81	20.19	2.55
White	76.02	23.98	12.40
Total	60.62	39.38	100

The Income and Expenditure Survey 2011 found that a little more than three quarters (76.58%) of the households were headed by Black Africans. A total of 12.4% households were headed by whites with less than one in ten (8.48%) households headed by coloureds. Only 2.55% were headed by Indians/Asians. The majority of households in South Africa are headed by men, with six out of every ten (60.62%) households being male-headed as opposed to four tenths (39,38%) being female-headed. More than two fifths (42.94) of households were headed by female Black Africans while more than a third (35.48%) were headed by female Coloureds. On the other hand, less than a quarter of white (23.98%) and Indian/Asian (20.19%) households were headed by women.

Table B.2: Percentage distribution of households by settlement type and gender of household head

Settlement type	Gender of household head		Total (%)
	Male (%)	Female (%)	
Urban formal	65.7	34.3	58.42
Urban informal	64.79	35.21	8.86
Traditional area	46.5	53.5	28.33
Rural formal	75.7	24.3	4.38
Total	60.62	39.38	100

Table B.2 shows that in rural formal areas, three quarters (75.7%) of all households were headed by men. Also, approximately two-thirds of households in urban formal areas (65.7%) and urban informal areas (64.79%) were also found to be headed by men. The majority of female-headed households were found in traditional areas, where more than half (53.5%) of all households were headed by women. Approximately a third (34.3%) of households in urban formal and just above a third (35.21%) of households in urban informal areas were headed by females.

Table B.3: Percentage distribution of households by province

Province	Estimated number of households	Percent
Western Cape	1 430 924	10.91
Eastern Cape	1 660 125	12.66
Northern Cape	230 673	1.76
Free State	788 743	6.02
KwaZulu-Natal	2 383 424	18.18
North West	991 725	7.56
Gauteng	3 404 189	25.96
Mpumalanga	855 572	6.53
Limpopo	1 366 841	10.42

Table B.3 gives percentage distribution of households per province. According to the IES 2011, the province with the highest percentage of households was Gauteng province which accounted for more than a quarter of all households in the country. It was followed by KwaZulu-Natal where more than 18% of the households were found. The third province which accounted for more than 10% o of the households was Limpopo. Other provinces accounted for less than 10% of households each, with the lowest being Northern Cape at 1.8 percent.

C. Adjusted confidence intervals using Bonferroni Adjustment

In chapter, only figures showing the adjusted Bonferroni confidence intervals were presented. This appendix presents the tables showing actual values of the confidence limits.

Table C1: Poverty estimates, standard errors and Bonferroni adjusted confidence intervals by population group of head of household

Disaggregation variable	Group/ Domain	Sample size (n)	Poverty measure	Poverty estimate	Standard error	Standard error (Jackknife)	Bonferroni adjusted confidence interval		Bonferroni adjusted confidence interval (Jackknife)	
Population group of head of household	African/Black	20 083	P_{HC}	0.359769	0.005513	0.005516	0.3452	0.3743	0.3452	0.3743
			P_{PGI}	0.139361	0.002660	0.002661	0.1323	0.1464	0.1323	0.1464
			P_{SPGI}	0.071960	0.001642	0.001643	0.0676	0.0763	0.0676	0.0763
			P_W	0.206592	0.004327	0.004329	0.1952	0.2180	0.1952	0.2180
	Coloured	2 723	P_{HC}	0.194076	0.012278	0.012308	0.1616	0.2265	0.1616	0.2266
			P_{PGI}	0.063141	0.004503	0.004512	0.0512	0.0750	0.0512	0.0751
			P_{SPGI}	0.028016	0.002339	0.002343	0.0218	0.0342	0.0218	0.0342
			P_W	0.086387	0.006535	0.006548	0.0691	0.1036	0.0691	0.1037
	Indian/Asian	468	P_{HC}	0.017870	0.007866	0.007913	0*	0.0386	0*	0.0388
			P_{PGI}	0.005247	0.002636	0.002650	0*	0.0122	0*	0.0122
			P_{SPGI}	0.002264	0.001342	0.001349	0*	0.0058	0*	0.0058
			P_W	0.007132	0.003772	0.003791	0*	0.0171	0*	0.0171
	White	2 054	P_{HC}	0.004406	0.001452	0.001453	0.0006	0.0082	0.0006	0.0082
			P_{PGI}	0.001101	0.000426	0.000426	0.0000	0.0022	0.0000	0.0022
			P_{SPGI}	0.000442	0.000234	0.000234	0*	0.0011	0*	0.0011
			P_W	0.001455	0.000628	0.000629	0*	0.0031	0*	0.0031

** = truncated at zero*

Table C2: Poverty estimates, standard errors and Bonferroni adjusted confidence intervals by type of settlement

Disaggregation variable	Group/ Domain	Sample size (n)	Poverty measure	Poverty estimate	Standard error	Standard error (Jackknife)	Bonferroni adjusted confidence interval		Bonferroni adjusted confidence interval (Jackknife)	
Settlement type	Urban formal	14 583	P_{HC}	0.163544	0.005905	0.005913	0.1479	0.1791	0.1479	0.1792
			P_{PGI}	0.054934	0.002479	0.002483	0.0484	0.0615	0.0484	0.0615
			P_{SPGI}	0.026131	0.001437	0.001440	0.0223	0.0299	0.0223	0.0299
			P_W	0.078416	0.003902	0.003909	0.0681	0.0887	0.0681	0.0887
	Urban informal	1 712	P_{HC}	0.387801	0.019185	0.019206	0.3371	0.4385	0.3371	0.4385
			P_{PGI}	0.147488	0.009107	0.009119	0.1234	0.1715	0.1234	0.1716
			P_{SPGI}	0.073137	0.005351	0.005359	0.0590	0.0873	0.0590	0.0873
			P_W	0.212434	0.014146	0.014167	0.1751	0.2498	0.1750	0.2499
	Traditional area	8 063	P_{HC}	0.521344	0.007506	0.007509	0.5015	0.5412	0.5015	0.5412
			P_{PGI}	0.217343	0.004315	0.004317	0.2059	0.2287	0.2059	0.2287
			P_{SPGI}	0.116626	0.002923	0.002924	0.1089	0.1243	0.1089	0.1244
			P_W	0.328398	0.007493	0.007496	0.3086	0.3482	0.3086	0.3482
	Rural formal	970	P_{HC}	0.349881	0.030798	0.031175	0.2685	0.4312	0.2675	0.4322
			P_{PGI}	0.127772	0.012507	0.012619	0.0947	0.1608	0.0944	0.1611
			P_{SPGI}	0.063946	0.007112	0.007158	0.0452	0.0827	0.0450	0.0829
			P_W	0.187278	0.019327	0.019467	0.1362	0.2383	0.1359	0.2387

Table C3: Poverty estimates, standard errors and Bonferroni adjusted confidence intervals by province

Disaggregation variable	Group/ Domain	Sample size (n)	Poverty measure	Poverty estimate	Standard error	Standard error (Jackknife)	Bonferroni adjusted confidence interval		Bonferroni adjusted confidence interval (Jackknife)	
Province	**Western Cape**	2 970	P_{HC}	0.154401	0.012041	0.012046	0.1159	0.1929	0.1159	0.1929
			P_{PGI}	0.047846	0.004661	0.004662	0.0329	0.0628	0.0329	0.0628
			P_{SPGI}	0.020957	0.002317	0.002318	0.0135	0.0284	0.0135	0.0284
			P_{W}	0.065496	0.006631	0.006632	0.0443	0.0867	0.0443	0.0867
	Eastern Cape	3 333	P_{HC}	0.427544	0.013907	0.013970	0.3830	0.4720	0.3828	0.4722
			P_{PGI}	0.171251	0.007224	0.007257	0.1481	0.1944	0.1480	0.1945
			P_{SPGI}	0.089566	0.004588	0.004606	0.0749	0.1042	0.0748	0.1043
			P_{W}	0.255074	0.011879	0.011925	0.2171	0.2931	0.2169	0.2932
	Northern Cape	1 205	P_{HC}	0.333117	0.017655	0.017698	0.2766	0.3896	0.2765	0.3898
			P_{PGI}	0.122835	0.007416	0.007428	0.0991	0.1466	0.0991	0.1466
			P_{SPGI}	0.060181	0.004293	0.004299	0.0464	0.0739	0.0464	0.0739
			P_{W}	0.175615	0.011561	0.011577	0.1386	0.2126	0.1386	0.2127
	Free State	2 172	P_{HC}	0.259969	0.012319	0.012344	0.2205	0.2994	0.2205	0.2995
			P_{PGI}	0.092081	0.005453	0.005461	0.0746	0.1095	0.0746	0.1096
			P_{SPGI}	0.044273	0.003364	0.003368	0.0335	0.0550	0.0335	0.0551
			P_{W}	0.130593	0.008775	0.008786	0.1025	0.1587	0.1025	0.1587
	KwaZulu-Natal	3 625	P_{HC}	0.382647	0.012896	0.012909	0.3414	0.4239	0.3413	0.4240
			P_{PGI}	0.153217	0.005978	0.005984	0.1341	0.1723	0.1341	0.1724
			P_{SPGI}	0.080587	0.003672	0.003675	0.0688	0.0923	0.0688	0.0923
			P_{W}	0.228532	0.009671	0.009679	0.1976	0.2595	0.1976	0.2595
	North West	2 524	P_{HC}	0.335129	0.016146	0.016248	0.2835	0.3868	0.2831	0.3871
			P_{PGI}	0.131532	0.007184	0.007218	0.1085	0.1545	0.1084	0.1546
			P_{SPGI}	0.068626	0.004269	0.004285	0.0550	0.0823	0.0549	0.0823
			P_{W}	0.196494	0.011506	0.011553	0.1597	0.2333	0.1595	0.2335
	Gauteng	3 887	P_{HC}	0.139675	0.008397	0.008436	0.1128	0.1665	0.1127	0.1667
			P_{PGI}	0.046331	0.003723	0.003737	0.0344	0.0582	0.0344	0.0583
			P_{SPGI}	0.022009	0.002297	0.002307	0.0147	0.0294	0.0146	0.0294
			P_{W}	0.066565	0.006204	0.006232	0.0467	0.0864	0.0466	0.0865
	Mpumalanga	2 306	P_{HC}	0.332959	0.014893	0.014943	0.2853	0.3806	0.2851	0.3808
			P_{PGI}	0.121734	0.006631	0.006646	0.1005	0.1430	0.1005	0.1430
			P_{SPGI}	0.060470	0.004140	0.004150	0.0472	0.0737	0.0472	0.0738
			P_{W}	0.176773	0.010752	0.010777	0.1424	0.2112	0.1423	0.2113
	Limpopo	3 306	P_{HC}	0.456467	0.011725	0.011735	0.4189	0.4940	0.4189	0.4940
			P_{PGI}	0.191537	0.006746	0.006754	0.1699	0.2131	0.1699	0.2131
			P_{SPGI}	0.103070	0.004628	0.004635	0.0883	0.1179	0.0882	0.1179
			P_{W}	0.290480	0.011838	0.011854	0.2526	0.3284	0.2525	0.3284

D. Overview of different data source for poverty measurement

This appendix gives an overview of some of the poverty data sources, along with their uses and shortfalls. The overview includes the name of the survey, the custodian, the frequency or time periods in which the survey was conducted as well as a brief purpose of each survey.

Table D1: Overview of the Income and Expenditure Survey

Survey Name	**Income and Expenditure Survey (IES)**
Custodian	**Statistics South Africa**
Year(s)	**1995, 2000, 2005/06, 2010/11**
Purpose of the survey	Is conducted every 5 years and is designed to gather information about the income and expenditure information from South African households. The survey also identifies goods and services purchased by these households. The primary purpose is to collect data to update the Consumer Price Index (CPI) basket of goods and services.

Table D2: Overview of the Living Conditions Survey

Survey Name	**Living Conditions Survey**
Custodian	**Statistics South Africa**
Year(s)	**2008/09**
Purpose of the survey	Aims to provide for the first time in South Africa, data with which to measure poverty levels. The survey provides data that can be used for measuring poverty using objective, subjective and multi-dimensional measures, and covers a wide range of indicators such as household income, household expenditure, ownership of assets, minimum income, household's self-perceived poverty status, and access to facilities and services.

Table D3: Overview of the National Income Dynamics Study

Survey Name	**National Income Dynamics Study (NIDS)**
Custodian	**University of Cape Town, Southern Africa Labour and Development Research Unit (SALDRU)**
Year(s)	**2008, 2010, 2012**
Purpose of the survey	NIDS is the first national household panel study in South Africa and is implemented by the Southern Africa Labour and Development Research Unit (SALDRU) based at the University of Cape Town. The study began in 2008 and survey continues to be repeated with these same household members every two years. NIDS examines the livelihoods of individuals and households over time. It also provides information about how households cope with positive or negative shocks, such as a death in the family or an unemployed relative obtaining a job. Other themes include changes in poverty and well-being; household composition and structure; fertility and mortality; migration; labour market participation and economic activity; human capital formation, health and education; vulnerability and social capital.

Table D4: Overview of the population Census

Survey Name	**CENSUS**
Custodian	**Statistics South Africa**
Year(s)	**1996, 2001 and 2011**
Purpose of the survey	Census is conducted after every 10 years. It is the principal means of collecting basic population and housing statistics required for social and economic development, policy interventions, their implementation and evaluation. A number of population and household attributes are measured with a variety of indicators generated.

Table D5: Overview of the Labour Force Survey

Survey Name	**Labour Force Survey (LFS)**[2]
Custodian	**Statistics South Africa**
Year(s)	**2000 - 2007**
Purpose of the survey	The Labour Force Survey has since become the Quarterly Labour Force Survey (See overview of Quarterly Labour Force Survey in the table below).

Table D6: Overview of the Quarterly Labour Force Survey

Survey Name	**Quarterly Labour Force Survey (QLFS)**
Custodian	**Statistics South Africa**
Year(s)	**2008 - 2014**
Purpose of the survey	The QLFS is conducted quarterly. The survey is specifically designed to measure the dynamics of the South African labour market, producing indicators such as employment, unemployment and inactivity. It measures a variety of issues related to the labour market, including the official unemployment rate.

Table D7: Overview of the October Household Survey

Survey Name	**October Household Survey (OHS)**[3]
Custodian	**Statistics South Africa**
Year(s)	**1993 - 2001**
Purpose of the survey	The October Household Survey has since become the General Household Survey (See overview of General Household Survey in the table below)

Table D8: Overview of the General Household Survey

Survey Name	**General Household Survey (GHS)**
Custodian	**Statistics South Africa**
Year(s)	**2002- 2012**
Purpose of the survey	The GHS is conducted annually. It is designed to measure 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. The survey collects a variety of household information, such as housing types and access to services, as well as person-level data about, for example, education, health and work status**.**

[2] The LFS has since become the QLFS
[3] The OHS has since become the GHS

Table D9: Overview of the Community Survey

Survey Name	**Community Survey**
Custodian	**Statistics South Africa**
Year(s)	**2007**
Purpose of the survey	The main objectives of the Community Survey conducted in February 2007 were to: • provide data at lower levels of geography (at district and municipal levels) in addition to national and provincial levels, • build human, management and logistical capacity for Census 2011; provide the primary data as a base for population projections.

Table D10: Overview of the KwaZulu-Natal Income Dynamics Study

Survey Name	**KwaZulu-Natal Income Dynamics Study (KIDS)**
Custodians	**University of KwaZulu-Natal (UKZN), the University of Wisconsin-Madison and the International Food Policy Research Institute (IFPRI).**
Year(s)	**1993, 1998, 2004**
Purpose of the survey	This panel study offers unique insight through the collection of data over period that spans South Africa's political and demographic transition, the introduction of many policies intended to reduce poverty, as well as the era of rapid HIV/AIDS infection. The data collected is used to provide an analysis of the impact of the HIV/AIDS epidemic as well as of the impact of poverty reduction interventions undertaken since 1993.

Table D11: Overview of the Project for Statistics on Living Standards and Development

Survey Name	**Project for Statistics on Living Standards and Development (PSLSD)**
Custodians	**University of Cape Town, Southern Africa Labour and Development Research Unit (SALDRU)**
Year(s)	**1993**
Purpose of the survey	The PSLSD Integrated Household Survey is a nationally representative, multi-purpose household survey which collected information on a wide range of indicators of standard of living including household composition, education, health, fertility, expenditures, employment and other income earning activities. The survey was undertaken prior to South Africa's first democratic elections in 1994. The principal purpose of the survey was to collect data on living standards in order to provide policy makers with the data required for planning strategies to implement such goals as those outlined in the Government of National Unity's Reconstruction and Development Programme.

Table D12: Overview of the Living Standards Measurement Survey

Survey Name	**Living Standards Measurement Survey (LSMS)**
Custodian	**University of Cape Town, Southern Africa Labour and Development Research Unit (SALDRU)**
Year(s)	**1993**
Purpose of the survey	LSMS was established by the Development Research Group (DECRG) to explore ways of improving the type and quality of household data collected by statistical offices in developing countries. The goal is to foster increased use of household data as a basis for policy decision making. The objectives of the LSMS are to • improve the quality of household survey data • increase the capacity of statistical institutes to perform household surveys • improve the ability of statistical institutes to analyse household survey data for policy needs • provide policy makers with data that can be used to understand the determinants of observed social and economic outcomes The 1993 LSMS was undertaken by the Southern Africa Labour & Development Research Unit (SALDRU) and the survey came to be known as the SALDRU 1993 dataset.

Table D13: Overview of the South African Participatory Poverty Assessment

Survey Name	**South African Participatory Poverty Assessment (SA-PPA)**
Custodian	**University of Cape Town, Southern Africa Labour and Development Research Unit (SALDRU),**
Year(s)	**1995**
Purpose of the survey	The purpose of this survey was to provide a fuller and more integrated understanding of poverty from the perspective of those who are poor. The SA-PPA included fifteen linked studies and involved some 45 researchers from 20 organisations. Work was undertaken with 25 communities and the study covered all provinces of South Africa with the exception of Gauteng and Free State.

E. SAS codes for estimating poverty

This appendix provides SAS program statements used to produce different outputs presented in this study.

E.1. Appendix E1: SAS code for data preparation and formatting

```
/*This programme generates the variable PSUNo_seg1 and save the data as
household*/
data finald.household;
set finald.house_info_1;
PSUNo_seg1=substr(uqno,1,16);
run;
/*This programme sorts the household data by the variable PSUNo_seg1*/
data finald.householdsorted;
set finald.household;
run;
proc sort data = finald.householdsorted;
by PSUNo_seg1;
run;
/*This programme prints the contents of the sampling data created from a
dataset called Ies2014_finak_sample*/
data finald.sampling;
set finald.Ies2014_final_sample;
proc contents data = finald.sampling;
run;
/*This programme sorts the sampling data by the variable PSUNo_seg1*/
proc sort data= finald.sampling;
by PSUNo_seg1;
run;
/*This programme merges the household data and the sampling data by
variable PSUNo_reg1*/
data finald.household_sampling;
merge finald.household finald.sampling;
by PSUNo_seg1;
run;
/*-----------------------------------------------------------------------
Formatting the codes in the IES 2011 dataset into labels to give more
descriptive names to the codes of the variables
-----------------------------------------------------------------------*/
proc format;
   value PROVINCE
      1 = 'Western Cape'
      2 = 'Eastern Cape'
      3 = 'Northern Cape'
      4 = 'Free State'
      5 = 'KwaZulu-Natal'
      6 = 'North West'
      7 = 'Gauteng'
      8 = 'Mpumalanga'
      9 = 'Limpopo' ;
   value SETTLEMENT_TYPE
      1 = 'Urban formal'
      2 = 'Urban informal'
      4 = 'Traditional area'
      5 = 'Rural formal' ;
   value GENDEROFHEAD
      1 = ' Male'
      2 = ' Female' ;
   value POPGRPOFHEAD
```

```
   1 = ' African/Black'
   2 = ' Coloured'
   3 = ' Indian/Asian'
   4 = ' White' ;
value Q110SUPPORTS
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q111PERSSUPPORTED
   8 = ' Not applicable' ;
value Q112A1FINSUPPORT
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q112A2FOOD
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q112A3ALCOHOL
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q112A4CLOTHING
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q112A5FURNITURE
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q112A6TRANSPORT
   1 = ' Yes'
   2 = ' No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q112A7ACCOMODATION
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q112A8PERSONALCARE
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q112A9OTHER
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q113A1FOOD
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q113A2ALCOHOL
   1 = ' Yes'
```

```
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q113A3CLOTHING
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q113A4FURNITURE
   1 = 'Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q113A5TRANSPORT
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q113A6ACCOMMODATION
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q113A7PERSONALCARE
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q113A8POCKETMONEY
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q113A9OTHER
   1 = 'Yes'
   2 = 'No'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q3101MEDSACQUIRED
   1 = 'Yes'
   2 = 'No'
   3 = 'Don't Know'
   9 = 'Unspecified' ;
value Q31021MEDAID
   1 = ' Yes'
   2 = ' No'
   3 = ' Don't Know'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q31022CASH
   1 = ' Yes'
   2 = ' No'
   3 = ' Don't Know'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q30123FLATRATE
   1 = ' Yes'
   2 = ' No'
   3 = ' Don't Know'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q31024FRIEND
```

```
   1 = ' Yes'
   2 = ' No'
   3 = ' Don't Know'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q31025NATURE
   1 = ' Yes'
   2 = ' No'
   3 = ' Don't Know'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q31026OTHER
   1 = ' Yes'
   2 = ' No'
   3 = ' Don't Know'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q31031DOCTOR
   1 = ' Yes'
   2 = ' No'
   3 = ' Don't Know'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q31032TRADITIONALHEALER
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q31033SELFPRESCRIPTION
   1 = ' Yes'
   2 = ' No'
   3 = ' Don't Know'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q3104MEDSNOTBOUGHT
   1 = 'Yes'
   2 = 'No'
   3 = 'Don't Know'
   8 = 'Not applicable'
   9 = 'Unspecified' ;
value Q41MAINDWELLING
   1 = ' DWELLING/HOUSE OR BRICK/CONCRETE BLOCK STRUCTURE ON A SEPARATE
       STAND OR YARD OR ON A FARM'
   2 = ' TRADITIONAL DWELLING/HUT/STRUCTURE MADE OF TRADITIONAL
       MATERIALS'
   3 = ' FLAT OR APARTMENT IN A BLOCK OF FLATS'
   4 = ' CLUSTER HOUSE IN SECURITY COMPLEX'
   5 = ' TOWN HOUSE/SEMI-DETACHED HOUSE'
   6 = ' DWELLING/HOUSE/ FLAT/ROOM IN BACKYARD'
   7 = ' INFORMAL DWELLING/SHACK IN BACKYARD'
   8 = ' INFORMAL DWELLING/SHACK NOT IN BACKYARD, E.G IN AN
       INFORMAL/SQUATTER SETTLEMENT OR ON FARM'
   9 = ' ROOM/FLATLET ON A PROPERTY OR A LARGER DWELLING, SERVANTS
       QUARTERS/GRANNY'S FLAT'
  10 = ' CARAVAN/TENT'
  11 = 'OTHER, specify in the box below' ;
value Q41OTHERDWELLING
   1 = ' DWELLING/HOUSE OR BRICK/CONCRETE BLOCK STRUCTURE ON A SEPARATE
       STAND OR YARD OR ON A FARM'
   2 = ' TRADITIONAL DWELLING/HUT/STRUCTURE MADE OF TRADITIONAL
       MATERIALS'
   3 = ' FLAT OR APARTMENT IN A BLOCK OF FLATS'
```

```
   4 = ' CLUSTER HOUSE IN SECURITY COMPLEX'
   5 = ' TOWN HOUSE/SEMI-DETACHED HOUSE'
   6 = ' DWELLING/HOUSE/ FLAT/ROOM IN BACKYARD'
   7 = ' INFORMAL DWELLING/SHACK IN BACKYARD'
   8 = ' INFORMAL DWELLING/SHACK NOT IN BACKYARD, E.G IN AN
         INFORMAL/SQUATTER SETTLEMENT OR ON FARM'
   9 = ' ROOM/FLATLET ON A PROPERTY OR A LARGER DWELLING, SERVANTS
         QUARTERS/GRANNY'S FLAT'
   10 = ' CARAVAN/TENT'
   11 = 'OTHER, specify in the box below'
   88 = ' Not applicable' ;
value Q42WALLS
   1 = ' BRICKS'
   2 = ' CEMENT BLOCK/CONCRETE'
   3 = ' CORRUGATED IRON/ZINC'
   4 = ' WOOD'
   5 = ' PLASTIC'
   6 = ' CARDBOARD'
   7 = ' MIXTURE OF MUD AND CEMENT'
   8 = ' WATTLE AND DAUB'
   9 = ' TILE'
   10 = ' MUD'
   11 = ' THATCHING'
   12 = ' ASBESTOS'
   13 = ' OTHER, specify in the box below'
   99 = ' UNSPECIFIED' ;
value Q42ROOF
   1 = ' BRICKS'
   2 = ' CEMENT BLOCK/CONCRETE'
   3 = ' CORRUGATED IRON/ZINC'
   4 = ' WOOD'
   5 = ' PLASTIC'
   6 = ' CARDBOARD'
   7 = ' MIXTURE OF MUD AND CEMENT'
   8 = ' WATTLE AND DAUB'
   9 = ' TILE'
   10 = ' MUD'
   11 = ' THATCHING'
   12 = ' ASBESTOS'
   13 = ' OTHER, specify in the box below'
   99 = ' UNSPECIFIED' ;
value Q43WALLS
   1 = 'Weak, needs major repairs (e.g. not windproof, leaking)'
   2 = 'Needs minor repairs'
   3 = 'Good'
   9 = 'Unspecified' ;
value Q43ROOF
   1 = ' Weak, needs major repairs (e.g. not windproof, leaking)'
   2 = ' Needs minor repairs'
   3 = ' Good'
   9 = ' Unspecified' ;
value Q44A1MAINDWELLING
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q44A2DWELLINGSE
   1 = ' Yes'
   2 = ' No'
   3 = ' Don't know'
   9 = ' Unspecified' ;
value Q44BSUBSIDISEDDU
   1 = ' Yes'
```

```
    2 = ' No'
    3 = ' Don't know'
    9 = ' Unspecified' ;
value Q45DRINKINGWATER
    1 = ' PIPED (TAP) WATER IN DWELLING'
    2 = ' PIPED (TAP) WATER ON-SITE OR IN YARD'
    3 = ' BOREHOLE ON-SIT'
    4 = ' RAIN-WATER TANK ON-SITE'
    5 = ' NEIGHBOUR'S TAP'
    6 = ' PUBLIC TAP'
    7 = 'WATER-CARRIER/TANKER'
    8 = ' BOREHOLE OFF-SITE/COMMUNAL'
    9 = ' FLOWING WATER/STREAM/RIVER'
    10 = ' STAGNANT WATER/DAM/POOL'
    11 = 'WELL'
    12 = ' SPRING'
    13 = ' OTHER, specify in the box'
    99 = ' UNSPECIFIED' ;
value Q45OTHERUSEWATER
    1 = ' PIPED (TAP) WATER IN DWELLING'
    2 = ' PIPED (TAP) WATER ON-SITE OR IN YARD'
    3 = ' BOREHOLE ON-SIT'
    4 = ' RAIN-WATER TANK ON-SITE'
    5 = ' NEIGHBOUR'S TAP'
    6 = ' PUBLIC TAP'
    7 = 'WATER-CARRIER/TANKER'
    8 = ' BOREHOLE OFF-SITE/COMMUNAL'
    9 = ' FLOWING WATER/STREAM/RIVER'
    10 = ' STAGNANT WATER/DAM/POOL'
    11 = 'WELL'
    12 = ' SPRING'
    13 = ' OTHER, specify in the box'
    99 = ' UNSPECIFIED' ;
value Q46PIPEDWATER
    1 = ' Yes'
    2 = ' No'
    9 = ' Unspecified' ;
value Q47PAYWATER
    1 = ' Yes'
    2 = ' No'
    8 = ' Not applicable'
    9 = ' Unspecified' ;
value Q48TOILETTYPE
    11 = ' FLUSH TOILET WITH OFF-SITE DISPOSAL IN-DWELLING'
    12 = ' FLUSH TOILET WITH OFF-SITE DISPOSAL ON-SITE'
    13 = ' FLUSH TOILET WITH OFF-SITE DISPOSAL OFF-SITE'
    21 = ' FLUSH TOILET WITH ON-SITE DISPOSAL (SEPTIC TANK) IN-DWELLING'
    22 = ' FLUSH TOILET WITH ON-SITE DISPOSAL (SEPTIC TANK)'
    32 = ' CHEMICAL TOILET ON-SITE'
    33 = ' CHEMICAL TOILET OFF-SITE'
    42 = ' PIT LATRINE WITH VENTILATION PIPE ON-SITE'
    43 = ' PIT LATRINE WITH VENTILATION PIPE OFF-SITE'
    52 = ' PIT LATRINE WITHOUT VENTILATION PIPE ON-SITE'
    53 = ' PIT LATRINE WITHOUT VENTILATION PIPE OFF-SITE'
    62 = ' BUCKET TOILET ON-SITE'
    63 = ' BUCKET TOILET OFF-SITE'
    73 = ' NONE'
    99 = ' UNSPECIFIED' ;
value Q49MAINSCONNECTION
    1 = 'Yes'
    2 = 'No'
    9 = 'Unspecified' ;
```

```
value Q410FREEELECTRICITY
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q52MAINDWELLING
   1 = ' Owned and fully paid off'
   2 = ' Owned, but not yet fully paid off, financed by a mortgage bond'
   3 = ' Owned, but not yet fully paid off, financed by another type of
         loan'
   4 = ' Rented as part of employment contract of household member'
   5 = ' Rented not as part of employment contract of household member'
   6 = ' Occupied rent-free as part of employment contract of household
         member'
   7 = ' Occupied rent-free not as part of employment contract of
         household member'
   8 = ' Occupied as a border or lodger'
   9 = ' Other, specify' ;
value Q531FRENTINCLUDEWATER
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5352SUBSIDY
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5353INCLUDESUBSIDY
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q552GARAGE
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q552NOCARS
   0 = ' Unspecified' ;
value Q553AESTIMATEDAREA
   1 = ' Very Small less than 30 m^2'
   2 = ' Small, between 30 and 59 m^2'
   3 = ' Medium, between 60 and 119 m^2'
   4 = ' Large, between 120 and 239 m^2'
   5 = ' Very large, 240 m^2 or more'
   9 = ' Unspecified' ;
value Q553CSKETCH
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q555ESTIMATEDVALUE
   1 = ' LESS THAN R50 000'
   2 = ' R50 001 - R250 000'
   3 = ' R250 001 - R500 000'
   4 = ' R500 001 - R1 000 000'
   5 = ' R1 000 001 - R1 500 000'
   6 = ' R1 500 001 - R2 000 000'
   7 = ' R2 000 001 - R3 000 000'
   8 = ' MORE THAN R3 000 000'
   9 = ' DON'T KNOW'
   99 = ' UNSPECIFIED' ;
value Q583REPAIRS
```

```
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101A01RADIO
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B0112MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = 'Not applicable'
   9 = ' Unspecified' ;
value Q5101A02STEREO
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B0212MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = 'Not applicable'
   9 = ' Unspecified' ;
value Q5101A03TELEVISION
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B0312MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = 'Not applicable'
   9 = ' Unspecified' ;
value Q5101A04DVDPLAYER
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B0412MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A05REFRIGERATOR
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B0512MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A06STOVE
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B0612MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A07MICROWAVE
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B0712MONTHS
```

```
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A08WASHINGMACHINE
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B0812MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A09MOTORVEHICLE
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B0912MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A10MOTORCYCLE
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B1012MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A11COMPUTER
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B1112MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A12GENERATOR
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B1212MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A13CAMERA
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B1312MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A14CELLULARPHONE
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
```

```
value Q5101B1412MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A15LANDLINEPHONE
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B1512MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A16DSTV
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B1612MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A17INTERNET
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B1712MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A18POWERTOOLS
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B1812MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A19KITCHENFURNITURE
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B1912MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A20DININGFURNITURE
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B2012MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A21LOUNGEFURNITURE
   1 = ' Yes'
   2 = ' No'
```

```
   9 = ' Unspecified' ;
value Q5101B2112MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q5101A22BEDROOMFURNITURE
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q5101B2212MONTHS
   1 = ' Yes'
   2 = ' No'
   8 = ' Not applicable'
   9 = ' Unspecified' ;
value Q611SWIMPOOL
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q711TRIPSAWAY
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q721TRANSPORT
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q731TIMESHARE
   1 = 'Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q81DOMESTIC
   1 = 'Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q91OWNPRODUCTION
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q111HHTEXTILES
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q1231REPAIRAPP
   1 = ' Yes'
   9 = ' Unspecified' ;
value Q1331KEEPACQUIREPETS
   1 = ' Yes'
   2 = ' No' ;
value Q151VEHICLEPRIVUSE
   1 = ' Yes, new only'
   2 = ' Yes, used only'
   3 = ' Yes, both new and used'
   4 = ' No'
   9 = ' Unspecified' ;
value Q15141RUNCOSTSPRIV
   1 = ' Yes'
   2 = ' No'
   9 = ' Unspecified' ;
value Q1516DRIVLESSTESTSLIC
   1 = ' Yes'
   2 = ' No'
```

```
      9 = ' Unspecified' ;
/*---------------------------------------------------------------------------
Formatting the variables in dataset household_sampling and creating a new
formatted dataset named household_sampling_2
---------------------------------------------------------------------------*/
data finald.household_sampling_2;
set finald.household_sampling;

      format province PROVINCE. settlement_type SETTLEMENT_TYPE.
GenderOfHead GENDEROFHEAD. PopGrpOfHead POPGRPOFHEAD.;
      format Q110SUPPORTS Q110SUPPORTS. Q111PERSSUPPORTED
Q111PERSSUPPORTED. Q112A1FINSUPPORT Q112A1FINSUPPORT.;
      format Q112A2FOOD Q112A2FOOD. Q112A3ALCOHOL Q112A3ALCOHOL.
Q112A4CLOTHING Q112A4CLOTHING. Q112A5FURNITURE Q112A5FURNITURE.;
      format Q112A6TRANSPORT Q112A6TRANSPORT. Q112A7ACCOMODATION
Q112A7ACCOMODATION. Q112A8PERSONALCARE Q112A8PERSONALCARE.;
      format Q112A9OTHER Q112A9OTHER. Q113A1FOOD Q113A1FOOD. Q113A2ALCOHOL
Q113A2ALCOHOL. Q113A3CLOTHING Q113A3CLOTHING.;
      format Q113A4FURNITURE Q113A4FURNITURE. Q113A5TRANSPORT
Q113A5TRANSPORT. Q113A6ACCOMMODATION Q113A6ACCOMMODATION.;
      format Q113A7PERSONALCARE Q113A7PERSONALCARE. Q113A8POCKETMONEY
Q113A8POCKETMONEY. Q113A9OTHER Q113A9OTHER.;
      format Q3101MEDSACQUIRED Q3101MEDSACQUIRED. Q31021MEDAID
Q31021MEDAID. Q31022CASH Q31022CASH. Q30123FLATRATE Q30123FLATRATE.;
      format Q31024FRIEND Q31024FRIEND. Q31025NATURE Q31025NATURE.
Q31026OTHER Q31026OTHER. Q31031DOCTOR Q31031DOCTOR.;
      format Q31032TRADITIONALHEALER Q31032TRADITIONALHEALER.
Q31033SELFPRESCRIPTION Q31033SELFPRESCRIPTION.;
      format Q3104MEDSNOTBOUGHT Q3104MEDSNOTBOUGHT. Q41MAINDWELLING
Q41MAINDWELLING. Q41OTHERDWELLING Q41OTHERDWELLING.;
      format  Q42WALLS Q42WALLS. Q42ROOF Q42ROOF. Q43WALLS Q43WALLS.
Q43ROOF Q43ROOF. Q44A1MAINDWELLING Q44A1MAINDWELLING.;
      format Q44A2DWELLINGSE Q44A2DWELLINGSE. Q44BSUBSIDISEDDU
Q44BSUBSIDISEDDU. Q45DRINKINGWATER Q45DRINKINGWATER.;
      format Q45OTHERUSEWATER Q45OTHERUSEWATER. Q46PIPEDWATER
Q46PIPEDWATER. Q47PAYWATER Q47PAYWATER. Q48TOILETTYPE Q48TOILETTYPE.;
      format Q49MAINSCONNECTION Q49MAINSCONNECTION. Q410FREEELECTRICITY
Q410FREEELECTRICITY. Q52MAINDWELLING Q52MAINDWELLING.;
      format Q531FRENTINCLUDEWATER Q531FRENTINCLUDEWATER. Q5352SUBSIDY
Q5352SUBSIDY. Q5353INCLUDESUBSIDY Q5353INCLUDESUBSIDY.;
      format Q552GARAGE Q552GARAGE. Q552NOCARS Q552NOCARS.
Q553AESTIMATEDAREA Q553AESTIMATEDAREA. Q553CSKETCH Q553CSKETCH.;
      format Q555ESTIMATEDVALUE Q555ESTIMATEDVALUE. Q583REPAIRS
Q583REPAIRS. Q5101A01RADIO Q5101A01RADIO.;
      format Q5101B0112MONTHS Q5101B0112MONTHS. Q5101A02STEREO
Q5101A02STEREO. Q5101B0212MONTHS Q5101B0212MONTHS.;
      format Q5101A03TELEVISION Q5101A03TELEVISION. Q5101B0312MONTHS
Q5101B0312MONTHS. Q5101A04DVDPLAYER Q5101A04DVDPLAYER.;
      format Q5101B0412MONTHS Q5101B0412MONTHS. Q5101A05REFRIGERATOR
Q5101A05REFRIGERATOR. Q5101B0512MONTHS Q5101B0512MONTHS.;
      format Q5101A06STOVE Q5101A06STOVE. Q5101B0612MONTHS
Q5101B0612MONTHS. Q5101A07MICROWAVE Q5101A07MICROWAVE.;
      format Q5101B0712MONTHS Q5101B0712MONTHS. Q5101A08WASHINGMACHINE
Q5101A08WASHINGMACHINE. Q5101B0812MONTHS Q5101B0812MONTHS.;
      format Q5101A09MOTORVEHICLE Q5101A09MOTORVEHICLE. Q5101B0912MONTHS
Q5101B0912MONTHS. Q5101A10MOTORCYCLE Q5101A10MOTORCYCLE.;
      format Q5101B1012MONTHS Q5101B1012MONTHS. Q5101A11COMPUTER
Q5101A11COMPUTER. Q5101B1112MONTHS Q5101B1112MONTHS.;
      format Q5101A12GENERATOR Q5101A12GENERATOR. Q5101B1212MONTHS
Q5101B1212MONTHS. Q5101A13CAMERA Q5101A13CAMERA.;
      format Q5101B1312MONTHS Q5101B1312MONTHS. Q5101A14CELLULARPHONE
Q5101A14CELLULARPHONE. Q5101B1412MONTHS Q5101B1412MONTHS.;
```

```
      format Q5101A15LANDLINEPHONE Q5101A15LANDLINEPHONE. Q5101B1512MONTHS
Q5101B1512MONTHS. Q5101A16DSTV Q5101A16DSTV.;
      format Q5101B1612MONTHS Q5101B1612MONTHS. Q5101A17INTERNET
Q5101A17INTERNET. Q5101B1712MONTHS Q5101B1712MONTHS.;
      format Q5101A18POWERTOOLS Q5101A18POWERTOOLS. Q5101B1812MONTHS
Q5101B1812MONTHS. Q5101A19KITCHENFURNITURE Q5101A19KITCHENFURNITURE.;
      format Q5101B1912MONTHS Q5101B1912MONTHS. Q5101A20DININGFURNITURE
Q5101A20DININGFURNITURE. Q5101B2012MONTHS Q5101B2012MONTHS.;
      format Q5101A21LOUNGEFURNITURE Q5101A21LOUNGEFURNITURE.
Q5101B2112MONTHS Q5101B2112MONTHS. Q5101A22BEDROOMFURNITURE
Q5101A22BEDROOMFURNITURE.;
      format Q5101B2212MONTHS Q5101B2212MONTHS. Q611SWIMPOOL Q611SWIMPOOL.
Q711TRIPSAWAY Q711TRIPSAWAY. Q721TRANSPORT Q721TRANSPORT.;
      format Q731TIMESHARE Q731TIMESHARE. Q81DOMESTIC Q81DOMESTIC.
Q91OWNPRODUCTION Q91OWNPRODUCTION. Q111HHTEXTILES Q111HHTEXTILES.;
      format Q1231REPAIRAPP Q1231REPAIRAPP. Q1331KEEPACQUIREPETS
Q1331KEEPACQUIREPETS. Q151VEHICLEPRIVUSE Q151VEHICLEPRIVUSE.;
      format Q15141RUNCOSTSPRIV Q15141RUNCOSTSPRIV. Q1516DRIVLESSTESTSLIC
Q1516DRIVLESSTESTSLIC.;;

    /*-------------------------------------------------------------------
    | This program generates new variables used to calculate poverty
      measures|
    -------------------------------------------------------------------*/

      inkindcons_pp = inkindconsumptions/(hsize*12);
      UBPL=620;
      cons_poor=0;
      if inkindcons_pp<UBPL then cons_poor=1;
      inkindcons_shortfall= 0;
      if inkindcons_pp < UBPL then inkindcons_shortfall= UBPL -
      inkindcons_pp;
      pov_gap_index= inkindcons_shortfall/UBPL;
      squared_pov_gap_index = pov_gap_index**2;
      watts_index = log(UBPL/inkindcons_pp);
      if watts_index < 0 then watts_index =0;
        run;
```

E.2. Appendix E2: SAS code for generating CDF of monthly per capita in-kind consumption

```
*Plot Sample Cumulative Distribution Function of monthly per capita In-Kind
Consumption;
proc freq data=finald.household_sampling_2 noprint;
tables inkindcons_pp/out=cumfreq outcum;
weight full_calwgt;
run;
data cdf;
set cumfreq;
cdf= cum_pct/100;
run;
Title "Sample Cumulative Distribution Funtion (CDF) of per capita In-kind
Consumption";
axis1 label=(angle=90 'Cumulative Probability') ;
axis2 label=('Monthly per capita in-kind consumption (Rands)');
proc gplot data=cdf;
plot cdf*inkindcons_pp /vaxis=axis1 haxis= axis2 href= 8000 vref=0.9;
run;
quit;
```

E.3. Appendix E3: SAS code for generating sample quantile functions for different subgroups

```
data genderquantile2;
set work.genderquantile;
goptions reset = all;
Title "Sample Quantile Function of per capita In-kind consumption by gender
of head of household";
axis1 label=('Percentiles');
axis2 label=(angle=90'Monthly per capita inkind consumption (Rands)')order
= 0 to 120000 by 20000;
symbol1 color =blue value=plus  line=1width=0.001;
symbol2 color =red value=dot  line=1;
legend label=('Gender of head of household');
*Plot Sample Quantile Function by gender of head of household;
proc gplot data = genderquantile2;
plot Male*Quantile_M Female*Quantile_F/ overlay vaxis = axis2 haxis=axis1
legend= legend;
run;

data popgrpquantile2;
set work.popgrpquantile;
goptions reset = all;
Title "Sample Quantile Function of per capita In-kind consumption by
population group of head of household";
axis1 label=('Percentiles');
axis2 label=(angle=90'Monthly per capita inkind consumption (Rands)');
symbol1 color =blue value=star  line=5;
symbol2 color =red value=dot        line=10;
symbol3 color =black value=diamond  line=15;
symbol4 color =brown value=hash     line=20;
legend label=('Population Group of head of household');
*Plot Sample Quantile Function by population group of head of household;
proc gplot data=popgrpquantile2;
plot African*Quantile_A Coloured*Quantile_C Indian_Asian*Quantile_I
White*Quantile_W /haxis= axis1 vaxis=axis2 legend = legend overlay
vref=620;
run;
data settlementquantile2;
set work.settlementquantile;
goptions reset = all;
Title "Sample Quantile Function of per capita In-kind consumption by
settlement type";
axis1 label=('Percentiles') ;
axis2 label=(angle=90'Monthly per capita inkind consumption (Rands)');
symbol1 color =blue value=y  line=5;
symbol2 color =red value=dot line=10;
symbol3 color =black value=square line=15;
symbol4 color =brown value=star line=20;
legend label=('Settlement type');
*Plot Sample Quantile Function by settlement type;
proc gplot data=settlementquantile2;
plot Urban_formal*Quantile_UF Urban_informal*Quantile_UI
Traditional*Quantile_T Rural_formal*Quantile_RF/haxis= axis1 vaxis=axis2
legend = legend overlay;
run;

data provincequantile2;
set work.provincequantile;
goptions reset = all;
```

```
Title "Sample Quantile Function of per capita In-kind consumption by
province";
axis1 label=('Percentiles') ;
axis2 label=(angle=90'Monthly per capita inkind consumption (Rands)');
symbol1 color =blue value=star  line=5;
symbol2 color =red value=triangle   line=10;
symbol3 color =maroon value=square  line=15;
symbol4 color =brown value=circle   line=20;
symbol5 color =cyan value=y  line=25;
symbol6 color =grey value=hash      line=30;
symbol7 color =green value=diamond  line=35;
symbol8 color =magenta value=dot    line=40;
symbol9 color =black value=z  line=45;
legend label=('Province');
*Plot Sample Quantile Function by province;
proc gplot data=provincequantile2;
plot Western_Cape*Quantile_WC Eastern_Cape*Quantile_EC
Northern_Cape*Quantile_NC Free_State*Quantile_FS KwaZulu_Natal*Quantile_KZN
North_West*Quantile_NW Gauteng*Quantile_GP Mpumalanga*Quantile_MP
Limpopo*Quantile_LP/haxis= axis1 vaxis=axis2 legend=legend overlay;
run;
quit;
```

E.4. Appendix E4: SAS code for estimating poverty measures, their standard errors and 95% confidence intervals using exact method

```
/*-----------------------------------------------------------------------------
|This program calculates the descriptive statistics of the Inkind
Consumption variable using Taylor Series Method|
-----------------------------------------------------------------------------*/
title 'Disriptive statistics of monthly per-capita Inkind Consumption using
the IES 2011 dataset';
 proc surveymeans data=finald.household_sampling_2 mean nobs stderr var clm
        varmethod=taylor;
        strata stratum;
        cluster PsuNo_M;
        var inkindcons_pp;
        weight full_calwgt;
   run;

/*-----------------------------------------------------------------------------|
This program calculates the poverty measures, their standard errors and 95%
confidence intervals|
-----------------------------------------------------------------------------*/
title 'Estimation of Standard Error of different Poverty Measures using the
IES 2011 dataset';
 proc surveymeans data=finald.household_sampling_2 mean nobs stderr var clm
        varmethod=taylor;
        strata stratum;
        cluster PsuNo_M;
        ods output Statistics = finald.boxwhiskerstats;
        ods output Domain = finald.boxwhiskerdomain;
        var cons_poor pov_gap_index squared_pov_gap_index watts_index;
        label cons_poor = 'Poverty Headcount'
        pov_gap_index = 'Poverty Gap Index'
        squared_pov_gap_index = 'Squared Poverty Gap Index'
        watts_index = 'Watts Index';
        weight full_calwgt;
        domain genderofhead PopGrpOfHead settlement_type province;
   run;
```

E.5. Appendix E5: SAS code for estimating poverty measures, their standard errors and 95% confidence intervals using Jackknife method

```
data finald.household_sampling_2;
set finald.household_sampling;
/*---------------------------------------------------------------------------*/
/* The below section is to modify the dataset to allow for domain matches
when using the jackknife method in proc surveymeans for proper domain
matches */
/*---------------------------------------------------------------------------*/

/* Format to convert domain variables to character*/
      format PopGrpOfHead1  $10.;
      format GenderOfHead1 $10.;
      format settlement_type1 $10.;
      format province1 $10.;

/*****************************************************************************/
/* Set the Length of the domain variables to character*/
/*****************************************************************************/

      length PopGrpOfHead1  $10;
      length GenderOfHead1 $10;
      length settlement_type1 $10;
      length province1 $10;

/* Modify variable types for use in the domain matches for proc surveymeans
in stratified_household_sampling dataset*/

      PopGrpOfHead1 =PopGrpOfHead;
      GenderOfHead1 =GenderOfHead;
      settlement_type1 =settlement_type;
      province1 =province;

/* Change missing values from "." to " " in the new variables to be used in
the domain matches for proc surveymeans in stratified_household_sampling
dataset*/

      if PopGrpOfHead1= "." then PopGrpOfHead1="";
      if GenderOfHead1= "." then GenderOfHead1="";
      if settlement_type1= "." then settlement_type1="";
      if province1= "." then province1="";

 /*---------------------------------------------------------------------------
    | This program generates new variables used to calculate poverty
      measures|
    ---------------------------------------------------------------------------*/

      inkindcons_pp = inkindconsumptions/(hsize*12);
      UBPL=620;
      cons_poor=0;
      if inkindcons_pp<UBPL then cons_poor=1;
      inkindcons_shortfall= 0;
      if inkindcons_pp < UBPL then inkindcons_shortfall= UBPL -
      inkindcons_pp;
      pov_gap_index= inkindcons_shortfall/UBPL;
      squared_pov_gap_index = pov_gap_index**2;
      watts_index = log(UBPL/inkindcons_pp);
      if watts_index < 0 then watts_index =0;
  run;
```

```
/*-------------------------------------------------------------------------|
This program calculates the descriptive statistics of the Inkind
Consumption variable usnig Jackknife method|
-------------------------------------------------------------------------*/
title 'Disriptive statistics of monthly per-capita Inkind Consumption using
the IES 2011 dataset';
 proc surveymeans data=finald.household_sampling_2 mean nobs stderr var clm
        varmethod=jackknife;
        strata stratum;
        cluster PsuNo_M;
        var inkindcons_pp;
        weight full_calwgt;
   run;
/*-------------------------------------------------------------------------|
This program calculates the poverty measures, their standard errors and 95%
confidence intevals using Jackknife method|
-------------------------------------------------------------------------*/
title 'Variance Estimation of Poverty measures using the IES 2011 dataset:
Jackknife Method';
 proc surveymeans data= finald.household_sampling_2 mean var clm
        varmethod=jackknife;
        strata stratum;
        cluster PsuNo_M;
        var cons_poor pov_gap_index squared_pov_gap_index watts_index;
        label cons_poor = 'Poverty Headcount'
        pov_gap_index = 'Poverty Gap Index'
        squared_pov_gap_index = 'Squared Poverty Gap Index'
        watts_index = 'Watts Index';
        weight full_calwgt;
        domain genderofhead1 PopGrpOfHead1 settlement_type1 province1;
   run;
```

E.6. Appendix E6: SAS code for calculating Bonferroni adjusted confidence intervals using exact method

```
/*-------------------------------------------------------------------------|
This program calculates the poverty measures, their standard errors and
Bonferroni adjusted confidence intervals by population group and
settlement_type|
-------------------------------------------------------------------------*/
title 'Bonferroni Adjusted confidence intervals of Poverty Estimates by
population group and settlement_type';
 proc surveymeans data=finald.household_sampling_2 mean nobs stderr var clm
        alpha= 0.0083 varmethod=taylor;
        strata stratum;
        cluster PsuNo_M;
        ods output Statistics = finald.boxwhiskerstats;
        ods output Domain = finald.boxwhiskerdomain;
        var cons_poor pov_gap_index squared_pov_gap_index watts_index;
        label cons_poor = 'Poverty Headcount'
        pov_gap_index = 'Poverty Gap Index'
        squared_pov_gap_index = 'Squared Poverty Gap Index'
        watts_index = 'Watts Index';
        weight full_calwgt;
        domain PopGrpOfHead settlement_type;
   run;
```

```
/*-------------------------------------------------------------------------|
This program calculates the poverty measures, their standard errors and
Bonferroni adjusted confidence intervals by province|
-------------------------------------------------------------------------*/
title 'Bonferroni Adjusted confidence intervals of Poverty Estimates by
province';
 proc surveymeans data=finald.household_sampling_2 mean nobs stderr var clm
        alpha= 0.00139 varmethod=taylor;
        strata stratum;
        cluster PsuNo_M;
        ods output Statistics = finald.boxwhiskerstats;
        ods output Domain = finald.boxwhiskerdomain;
        var cons_poor pov_gap_index squared_pov_gap_index watts_index;
        label cons_poor = 'Poverty Headcount'
        pov_gap_index = 'Poverty Gap Index'
        squared_pov_gap_index = 'Squared Poverty Gap Index'
        watts_index = 'Watts Index';
        weight full_calwgt;
        domain province;
   run;
```

E.7. Appendix E7: SAS code for calculating Bonferroni adjusted confidence intervals using Jackknife method

```
/*-------------------------------------------------------------------------|
This program calculates the poverty measures, their standard errors and
Bonferroni adjusted confidence intervals using Jackknife method|
-------------------------------------------------------------------------*/
title 'Bonferroni Adjusted confidence intervals of Poverty Estimates using
Jackknife method by population group and settlement type';
proc surveymeans data= finald.household_sampling_2 mean var clm
        alpha= 0.0083 varmethod=jackknife;
        strata stratum;
        cluster PsuNo_M;
        var cons_poor pov_gap_index squared_pov_gap_index watts_index;
        label cons_poor = 'Poverty Headcount'
        pov_gap_index = 'Poverty Gap Index'
        squared_pov_gap_index = 'Squared Poverty Gap Index'
        watts_index = 'Watts Index';
        weight full_calwgt;
        domain PopGrpOfHead1 settlement_type1;
   run;
/*-------------------------------------------------------------------------|
This program calculates the poverty measures, their standard errors and
Bonferroni adjusted confidence intervals using Jackknife method|
-------------------------------------------------------------------------*/
title 'Bonferroni Adjusted confidence intervals of Poverty Estimates using
Jackknife method by province';
 proc surveymeans data= finald.household_sampling_2 mean var clm
        alpha= 0.00139 varmethod=jackknife;
        strata stratum;
        cluster PsuNo_M;
        var cons_poor pov_gap_index squared_pov_gap_index watts_index;
        label cons_poor = 'Poverty Headcount'
        pov_gap_index = 'Poverty Gap Index'
        squared_pov_gap_index = 'Squared Poverty Gap Index'
        watts_index = 'Watts Index';
        weight full_calwgt;
        domain province1;
   run;
```

E.8. Appendix E8: SAS code for calculating Gini coefficient for the poor

```
/*This SAS code is adapted to calculate the mean in-kind consumption,
poverty gap index and Gini coefficient for the poor*/

data poor1; set finald.household_sampling;

inkindcons_pp=InKindConsumptions/hsize/12;
UBPL = 620;
cons_poor=0;
if inkindcons_pp<UBPL;
if inkindcons_pp < UBPL then inkindcons_shortfall= UBPL - inkindcons_pp;
      pov_gap_index_poor= inkindcons_shortfall/UBPL;
run;

title 'Mean inkind consumption and poverty gap index for the poor';
proc surveymeans  data=poor1   mean;
                  strata stratum;
                  cluster PsuNo_M;
                  var inkindcons_pp pov_gap_index_poor;
                  weight full_calwgt;
   run;

/*                        GINI CODE
                          =========

This SAS code was written by Philip N. Cohen. It is meant to be adaptable
to various units of analysis and measures of interest. The Gini coefficient
can be calculated for lots of different distributions, although it is most
often used for income.

The formula used here is from _The methods and materials of demography_,by
Henry S. Shryock, Jacob S. Siegel, and associates. Orlanda, FL: Academic
Press, 1976 (p. 98).

(The author of the code can take no responsibility for its reliability or
accuracy, or for the results obtained with its use; but he would be glad to
take partial credit for it successful use or adaptation.)
*/

/* The variable I use is CAPINC and the weight is CAPWGT.
   Substitute these for your own measure and population weight.
   Those are the only variable names you have to change to suit
   your data.
 */

/* This creates a table with one line for each level of income,
   the number of (weighted) people with that income, and the
   percent with that income. */

title 'In-kind consumption distribution';
proc freq data=poor1;
      tables inkindcons_pp / noprint out = table1;
      format inkindcons_pp 7.0;
      weight full_calwgt;
run;

/* this data step creates cumulative income and population columns */
```

```
data table1;
set table1;

retain suminc perpop;

suminc + (inkindcons_pp * count);
perpop + percent;

/* suminc is the cumulative income at each point in the distribution.
   perpop is the cumulative population at each point in the distribution.
   Note that PERCENT and COUNT are variables created by PROC FREQ.
 */

run;

/* This sort and data step takes the last value of suminc, which is the
total income, and adds it onto every record in the table as totalinc. Then
it divides suminc by totalinc for each line to create the percent of income
below that point in the distribution */

proc sort data=table1;
      by descending suminc ;
run;

data table1;
      set table1;
      by descending suminc;

      if _n_=1 then do;
      totalinc=suminc;
      end;

      retain totalinc;

      perinc = (suminc/totalinc) * 100;

run;

/* this sort just puts it back in order from low to high */

proc sort data=table1;
      by perpop;
run;

/* To calculate Gini: sum[Xsub(i) * Ysub(i+1)] - sum[Xsub(i+1) * Ysub(i)]
   where X is the proportion of population column and Y is the proportion
   of income column.*/

data ginidat1;
      set table1;

      xlag = lag(perpop);
      xlag = xlag / 100;

      ylag = lag(perinc);
      ylag = ylag / 100;

      columna = (perinc/100) * xlag;
      columnb = (perpop/100) * ylag;

      retain suma sumb;
```

```
      suma + columna;
      sumb + columnb;

      gini = suma - sumb;

run;

title 'Gini coefficient for the poor';
      proc print data=ginidat1;
      var gini;
      where perinc = 100;
run;
```

E.9. Appendix E9: SAS code for calculating Gini coefficient of the poverty gaps

```
*This SAS code is adapted to calculate gini coefficient of the poverty
gaps*/

data poor2; set finald.household_sampling;

      inkindcons_pp=InKindConsumptions/hsize/12;
      UBPL = 620;
      if inkindcons_pp < UBPL then inkindcons_shortfall= UBPL -
      inkindcons_pp;

run;

title 'Poverty Gap distribution';
      proc freq data=poor2;
      tables inkindcons_shortfall / noprint out = table2;
      format inkindcons_shortfall 7.0;
      weight full_calwgt;
run;

/* this data step creates cumulative income and population
   columns */

data table2;
      set table2;

      retain suminc perpop;

      suminc + (inkindcons_shortfall * count);
      perpop + percent;

/* suminc is the cumulative income at each point in the distribution.
   perpop is the cumulative population at each point in the distribution.
   Note that PERCENT and COUNT are variables created by PROC FREQ.
 */

run;

/* This sort and data step takes the last value of suminc, which is the
total income, and adds it onto every record in the table as totalinc. Then
it divides suminc by totalinc for each line to create the percent of income
below that point in the distribution */
```

```
proc sort data=table2;
      by descending suminc ;
run;

data table2;
      set table2;
      by descending suminc;

      if _n_=1 then do;
      totalinc=suminc;
      end;

      retain totalinc;

      perinc = (suminc/totalinc) * 100;

      run;

/* this sort just puts it back in order from low to high */

proc sort data=table2;
      by perpop;
run;

/* To calculate Gini: sum[Xsub(i) * Ysub(i+1)] - sum[Xsub(i+1) * Ysub(i)]
   where X is the proportion of population column and Y is the proportion
   of income column. */

data ginidat2;
      set table2;

      xlag = lag(perpop);
      xlag = xlag / 100;

      ylag = lag(perinc);
      ylag = ylag / 100;

      columna = (perinc/100) * xlag;
      columnb = (perpop/100) * ylag;

      retain suma sumb;

      suma + columna;
      sumb + columnb;

      gini = suma - sumb;

run;

title 'Gini coefficient for the poverty gaps';
      proc print data=ginidat2;
      var gini;
      where perinc = 100;
run;
```

E.10. Appendix E10: SAS code for calculating Sen Index, Alternative Sen Index and Sen-Shorrocks-Thon

```
/*Calculating the Sen Index, Alternative Sen Index and the Sen-Shorrocks-
Thon Index using estimated values;*/

data Sen;
input upperbound PH PGI PGI_poor gini_poor gini_gaps mean_poor  @@;
      datalines;
      620 0.29 0.11 0.38 0.21 0.33 382.24
      ;
run;
data SenIndices;
      set Sen;

      Sen_index=PH*(1-(1-gini_poor)*(mean_poor/upperbound));

      Sen_alternative = PH*gini_poor + PGI*(1-gini_poor);

      SST = PH*PGI_poor*(1+gini_gaps);

run;

/* Print the values of Sen Index, Alternative Sen Index and the Sen-
Shorrocks-Thon Index*/
title 'Sen Index, Alternative Sen Index and the Sen-Shorrocks-Thon Index';
proc print data = SenIndices;
      var Sen_index Sen_alternative SST;
run;
```

E.11. Appendix E11: SAS code for calculating test statistics for Sen Index, Alternative Sen Index and Sen-Shorrocks-Thon Index

```
/* This program calculates the pairwise test statistics for Sen Index,
Alternative Sen Index and the Sen-Shorrocks-Thon Index*/

data zstatistics;
input province type mean stderr;
cards;
1      1      0.154401     0.012041
1      2      0.047846     0.004661
1      3      0.020957     0.002317
1      4      0.065496     0.006631
2      1      0.427544     0.013907
2      2      0.171251     0.007224
2      3      0.089566     0.004588
2      4      0.255074     0.011879
3      1      0.333117     0.017655
3      2      0.122835     0.007416
3      3      0.060181     0.004293
3      4      0.175615     0.011561
4      1      0.259969     0.012319
4      2      0.092081     0.005453
4      3      0.044273     0.003364
4      4      0.130593     0.008775
5      1      0.382647     0.012896
5      2      0.153217     0.005978
5      3      0.080587     0.003672
```

```
5       4       0.228532        0.009671
6       1       0.335129        0.016146
6       2       0.131532        0.007184
6       3       0.068626        0.004269
6       4       0.196494        0.011506
7       1       0.139675        0.008397
7       2       0.046331        0.003723
7       3       0.022009        0.002297
7       4       0.066565        0.006204
8       1       0.332959        0.014893
8       2       0.121734        0.006631
8       3       0.06047         0.00414
8       4       0.176773        0.010752
9       1       0.456467        0.011725
9       2       0.191537        0.006746
9       3       0.10307         0.004628
9       4       0.29048         0.011838
;
proc sort by type province; run;proc print; run;
proc iml;
use zstatistics;
read all into x;
province=x[,1];
type=x[,2];
mean=x[,3];
stderr=x[,4];
do it=1 to 4;
 y=x[loc(type=it),];
 mat=j(9,9,0);
 ztest=j(9,9,"          ");
 do ip1=1 to 9;
  do ip2=1 to ip1-1;
   mat[ip1,ip2]=(y[ip1,3]-y[ip2,3])/SQRT(y[ip1,4]##2+y[ip2,4]##2);
   if ABS(mat[ip1,ip2])> 3.196950229 then ztest[ip1,ip2]="Diff";
   if ABS(mat[ip1,ip2])<= 3.196950229 then ztest[ip1,ip2]="No Diff";
end;
 end;
 print it mat ztest;
end;
run;
```

E.12. Appendix E12: SAS code for calculating composition of households as found by IES 2011 for different domains

```
/*-------------------------------------------------------------------------
| This program generates distribution of population by gender and
population group of head of household
-------------------------------------------------------------------------*/
Proc freq data = finald.household_sampling_2;
title 'Percentage distribution of households by population group and sex of
household head';
weight full_calwgt;
tables POPGRPOFHEAD*GENDEROFHEAD/ nofreq nocol;
run;
```

```
/*-------------------------------------------------------------------------
| This program generates distribution of population by gender and
settlement type
-------------------------------------------------------------------------*/
Proc freq data = finald.household_sampling_2;
title 'Percentage distribution of households by settlement type and sex of
household head';
weight full_calwgt;
tables SETTLEMENT_TYPE*GENDEROFHEAD/ nofreq nocol;
run;
/*-----------------------------------------------------------
| This program gives distribution of population by province
-----------------------------------------------------------*/
Proc freq data = finald.household_sampling_2;
title 'Percentage distribution of households by porvince';
weight full_calwgt;
tables PROVINCE/ nofreq nocol;
run;
quit;
```

References

African National Congress 1994, *The reconstruction and development programme: A policy framework*. Johannesburg: African National Congress by Umanyano Publications.

Aliber, M 2003, 'Chronic Poverty in South Africa: Incidence, Causes and Policies', *World Development,* vol. 31, no. 3, pp. 473-490.

Alkire, S & Foster, J 2011, 'Counting and Multidimensional Poverty Measurement', *Journal of Public Economics*, 95(7–8), 476–487.

Alkire, S & Santos, M 2010, 'Acute Multidimensional Poverty: A New Index for Developing Countries', *OPHI Working Paper 38, Oxford Poverty and Human Development Initiative*, University of Oxford.

Alkire, S , Conconi, A & Roche, JM 2013*, 'Multidimensional Poverty Index 2013: Brief Methodological Note and Results'* Also available at: www.ophi.org.uk/multidimensional-poverty-index.

Angus, D & Case, A 1998, 'Large Cash Transfers to the Elderly in South Africa', *The Economic Journal*, vol. 108, no. 450, pp. 1330-1361.

Bhorat, H & Kanbur, R (eds.) 2006b, *Poverty and Policy in Post-Apartheid South Africa*, HSRC Press, Pretoria.

Bhorat, H & Kanbur, R 2006a, 'Poverty and Well-being in Post-Apartheid South Africa: An Overview of Data, Outcomes and Policy', *Development Policy Research Unit*, Working Paper 05/101.

Bhorat, H & Westhuizen, C 2012, 'Poverty, inequality and the nature of economic growth in South Africa', *Development Policy Research Unit*, Working Paper 12/151.

Bhorat, H, Oosthuizen, M & Van der Westhuizen, C 2012, 'Estimating a poverty line: An application to free basic municipal services in South Africa', *Development Southern Africa*, vol. 29, no.1, pp. 77-96.

Central Bureau of Statistics 2005, *Geographic dimensions of well-being in Kenya: Who and where are the poor?* A constituency level profile, Vol.II, Ministry of Planning and National Development, Nairobi.

Coudouel, A, Hentschel,J & Wodon, Q 2002, 'Poverty Measurement and Analysis' In Klugman,J ed., A Sourcebook for Poverty Reduction Strategies, vol. 1: Core Techniques and Cross-Cutting Issues, The World Bank, Washington, D.C.

Cumming, G 2012, *Understanding the new statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*, Routledge, Taylor and Francis Group.

Eurostat 2000, Assessment of the quality in statistics, Luxembourg.

Frye, I & Magasela, W 2005, 'Constructing and adopting an official poverty line for South Africa: Some issues for consideration', National Labour & Economic Development Institute (NALEDI): A discussion document.

Haughton, J, & Khandker, S 2009, Handbook on Poverty and Inequality, The World Bank, Washington DC.

Hirschowitz, R 2000, Measuring Poverty in South Africa, *Statistics South Africa*, Pretoria,

Hirschowitz, R, Orkin, M & Alberts, P 2000, 'Key baseline statistics for poverty measurement', in Measuring Poverty in South Africa, Statistics South Africa, Pretoria.

Hoogeveen, G & Özler B 2006, 'Poverty and inequality in post-apartheid South Africa: 1995–2000',in *Poverty and Policy in post-apartheid South Africa*, HSRC Press, Cape Town.

Hoogeveen, J & Özler B 2004, 'Not Separate, Not Equal: Poverty and Inequality in Post-Apartheid South Africa', *World Bank*, Washington DC.

Hunter, N, May, J & Padayachee, V 2003, 'Lessons for Poverty Reduction Strategy Papers (PRSP) from Poverty Reduction Strategies in South Africa', *Third Meeting of the African Learning Group on the Poverty Reduction Strategy Papers*, Addis Ababa: Ethiopia.

Kanbur, R & Lyn, S 2001, 'The Evolution of Thinking about Poverty: Exploring the Interactions', pp. 183-226, Oxford University Press.

Kingdom G & Knight J 2005, 'Subjective Well-being Poverty versus Income Poverty and Capabilities Poverty? , *Development Policy Research Unit*, Working Paper 05/096, University of Cape Town.

Klasen, S 2000, 'Measuring Poverty and Deprivation in South Africa', *Review of Income and Wealth*, Series 46, Number 1.

Klugman, J 2002, 'A sourcebook for poverty reduction strategies', Vol 1: Core techniques and cross-cutting issues, The World Bank.

Leibbrandt, M 2005, 'Asset-based versus Money Metric Poverty Indices in South Africa', *Centre for Social Science Research (CSSR)*, Working Paper No. 109.

Leibbrandt, M 2005, 'The sensitivity of estimates of post-apartheid changes in South African poverty and inequality to key data imputations', *Centre for Social Science Research (CSSR),* Working Paper No. 106.

Leibbrandt, M, Bhorat H, & Woolard, I 1999, 'Understanding contemporary household inequality in South Africa', *Development Policy Research Unit,* Working Paper 99/25, University of Cape Town.

Leibbrandt, M, Poswell, L, Naidoo, P & Welch, M 2006, 'Measuring recent changes in South African Inequality and Poverty using 1996 and 2001 census data', *Poverty and Policy in post-apartheid South Africa*, HSRC Press, Cape Town.

Leibbrandt, M, Poswell, L, Naidoo, P, Welch, M 2006 & Woolard, I 2005, 'Measuring recent Changes in South African Inequality and Poverty using 1996 and 2001 Census Data' *Development Policy Research Unit,* Working Paper 05/094, University of Cape Town.

Leibbrandt, M, Poswell, L, Naidoo, P, Welch, M 2006 & Woolard, I 2005, 'Measuring recent Changes in South African Inequality and Poverty using 1996 and 2001 Census Data', *Centre for Social Science Research (CSSR),* Working Paper No. 84, South African Labour Development Research Unit, University of Cape Town.

Leibbrandt, M, Woolard, I, Finn, A & Argent, A 2010, 'Trends in South African Income Distribution and Poverty since the fall of apartheid', *OECD Social, Employment and Migration*, Working Papers, No. 101, OECD Publishing.

Lohr, LS 2010, *Sampling: Design and Analysis*, Second Edition. Brooks/Cole, Cengage Learning.

Magasela, W 2005, 'Towards a constitution-base definition of poverty in post-apartheid South Africa', *State of the Nation, South Africa, 2005-2006*, Cape Town, HSRC Press.

May, J & Carter, M 1999, 'Poverty, livelihoods and class in rural South Africa', *World Development*, vol. 27, no. 1, pp. 1-20.

May, J 1998, 'Poverty and Inequality in South Africa', Report prepared for the Office of the Executive Deputy President and Inter-Ministerial Committee for Poverty and Inequality.

May, J 2000, 'Poverty and Inequality in South Africa: Meeting the Challenge', David Phillip and Zed Press, Cape Town and London.

Mbeki, T 2004, *Address by the President of South Africa, Mr Thabo Mbeki, on the occasion of his inauguration and the 10th anniversary of freedom*, Pretoria, 27 April.

Miller, R G 1981, *Simultaneous statistical inference*, Second edition. New York, Springer.

Morduch, J 1998, 'Poverty, economic growth, and average exit time', *Economic Letters,* no. 59, pp. 385-390, ELSEVIER.

Naidoo, A 2007, A multidimensional measurement of poverty in South Africa, doctoral thesis, University of Pretoria, Pretoria.

Naidoo, A G, Yadavalli, V S & Crowther, N A 2005, 'A multi-dimensional measure of poverty using the totally fuzzy and relative approach', *Studies in Economics and Econometrics, vol.* 29, no. 3, pp. 67-80.

Ngwane, A K, Yadavalli, VSS & Steffens, FE 1995, 'Poverty in South Africa in 1995: A totally fuzzy and relative approach', *Journal for Studies in Economic and Econometrics*, vol. 25, no. 1, pp77-87.

Ngwane, A K, Yadavalli , VS S & Steffens, FE 2001, 'Poverty in South Africa: a statistical analysis', *Development Southern Africa*, vol. 18, no. 2, pp. 201-215.

Ngwane, A K, Yadavalli , VS S & Steffens, FE 2002, 'Poverty: Deprivations in terms of basic needs', *Development Southern Africa*, vol. 19, no. 4, pp. 545-560.

Ngwane, A K, Yadavalli , VS S & Steffens, FE 2003, 'Poverty in South Africa: statistical inference for decomposable poverty measures', *Development Southern Africa* vol. 20, no. 2, pp. 283-291.

Noble, M, Babita, M, Barnes, H, Dibben, C, Magasela, W, Noble, S, Ntshongwana, P, Phillips, H, Rama, S, Roberts, B, Wright, G & Zungu, S 2006, 'The Provincial Indices of Multiple Deprivation for South Africa 2001: Technical report', *University of Oxford*, UK.

Noble, M, Ratcliffe, A & Wright, G 2004, 'Conceptualising, Defining and Measuring Poverty in South Africa: An Argument for a Consensual Approach', Centre for the Analysis of South African Social Policy, *University of Oxford*, UK.

Noble, M, Zembe, W, Wright, G & Avenell, D 2013, 'A Diagnostic Evaluation of Poverty and Relative Deprivation at small area level for South Africa', South African Social Policy Research Institute, *University of Oxford*, UK.

Ravallion, M & Pradhan, M 1998, 'Measuring Poverty Using Qualitative Perceptions of Welfare', *The World Bank*, Washington DC.

Ravallion, M 1992, 'Poverty Comparisons: A Guide to Concepts and Methods' Living Standards Measurement Surveys, Working Paper No. 88, World Bank, Washington, DC.

Ruggeri-Laderchi, C, Saith, R, & Stewart, F 2003, "Does it matter that we don't agree on the definition of poverty? A comparison of four approaches", Queen Elizabeth House Working Paper Series, Number 107, University of Oxford.

Sahn, D & Steifel, D 2003, 'Exploring Alternative Measures of Poverty in the Absence of Expenditure Data', *Review of Income and Wealth*, vol. 49, no. 3, pp. 463-489.

Sen, A 1979, 'Issues in the measurement of poverty', *Scand Journal of Economics.*

Southern Africa Labour and Development Research Unit 2013, *National Income Dynamics Study 2012*, Wave 3 dataset, Version 1.2, DataFirst.

Statistics South Africa (StatsSA) 2000, *Measuring poverty in South Africa,* Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2005, *Millennium Development Goals Country Report 2005*, Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2006, *Measuring Poverty in South Africa,* Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2008, *Measuring poverty in South Africa*, *A methodological report on the development of poverty lines for statistical reporting:* Technical Report, Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2010a, *Social Profile of South Africa 2002–2009*, Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2010b, *Millennium Development Goals Country Report* 2010, Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2011, *General Household Survey, Statistical release P0318,* Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2012a, *Income and Expenditure Survey, Statistical release P0100*, Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2012b*, Poverty profile of South Africa*, *Application of the poverty lines on the Living Conditions Survey 2008/2009*, Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2012c, *Subjective Poverty in South Africa*, *Findings of the Living Conditions Survey 2008/2009,* Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2012d, *Income and Expenditure of Households, 2010/2011: Metadata*, Report No. P01000, Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2013a, *Men, Women and Children, Findings of the Living Conditions Survey 2008/2009*), Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2013b, *Millennium Development Goals Country Report 2013*, Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2014a, *Poverty Trends in South Africa: An examination of absolute poverty between 2006 and 2011*, Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2014b, *The South African Multidimensional Poverty Index: Creating a multidimensional poverty index using Census data*, Statistics South Africa, Pretoria.

Statistics South Africa (StatsSA) 2015, *Methodological report on rebasing of national poverty lines and development on pilot provincial poverty lines* – Technical Report, Statistics South Africa, Pretoria.

Statistics South Africa and National Treasury 2007, *A National Poverty Line for South Africa.*

Stewart, F, Saith R, & Harris-White, B 2007, 'Defining poverty in the developing world', *Palgrave Macmillan.*

Studies in Poverty and Inequality Institute (SPII) 2007, *The Measurement of Poverty in South Africa Project: Key Issues*, Richmond: Johannesburg.

The Presidency 2003, *Towards a ten year review, Synthesis report on implementation of government programmes.*

The Presidency 2008, *Towards a fifteen year review, Synthesis report on implementation of government programmes.*

The Presidency 2011, *Development Indicators.*

The Presidency 2013, National Planning Commission, *National Income Dynamics Study*, Wave 3 overview.

The Presidency 2014, *Twenty year review*, South Africa 1994 – 2014.

The Presidency, Republic of South Africa 2011, National Development Plan 2030, *Our future, make it work*, Executive Summary, National Planning Commission.

Thorbecke, E 2003, Conceptual and Measurement Issues in Poverty Analysis, The challenge of growth and poverty: The South African economy since democracy, TIPS/DPRU Forum 2003.

Townsend I, 2004, 'Poverty: Measures and Targets', *House of Commons Library*, Research paper No. 04/23.

Townsend P, 1962, 'The Meaning of Poverty', *The British Journal of Sociology*, vol. 13, no. 3, pp. 210-227.

Townsend P, 1979, 'Poverty in the United Kingdom, A study of household resources and living standards', *Penguin Books Ltd*.

United Nations Department of Economic and Social Affairs (UNDESA) 2005a, Statistics Division, *Household Sample Surveys in Developing and Transition Countries,* Studies in methods, Series F No. 96, New York.

United Nations Department of Economic and Social Affairs (UNDESA) 2005b, Statistics Division, *Designing Household Survey Samples: Practical Guidelines,* Studies in methods, Series F No. 98, New York.

United Nations Development Programme (UNDP) 1997, *Human Development Report*, Oxford University Press, New York.

United Nations Development Programme (UNDP) 2007, *Human Development Report, Fighting climate change: Human solidarity in a divided world*, Oxford University Press, New York.

Van der Berg , S, Burger, R (Ronelle), Burger, R (Rudolf), Louw, M & Yu, D 2005, 'Trends in poverty and inequality since the political transition', *Stellenbosch Economic Working Paper Series*: 1/2005.

Van der Berg, S ,Burger, R (Ronelle), Burger, R (Rudolf), Louw, M & Yu, D 2006, 'Trends in poverty and inequality since the political transition', *Development Policy Research Unit,* Working Paper 06/104, University of Cape Town.

Van der Berg, S 2010, 'Current poverty and income distribution in the context of South African history', *Stellenbosch Economic Working Paper Series*: 22/10.

Van der Berg, S, Louw, M & Yu, D 2007, 'Post-transition poverty trends based on an alternative data source', *Stellenbosch Economic Working Paper Series*: 08/07.

Welman, C, Kruger, F, and Michell, B 2005, Research Methodology, 3rd edition, Oxford University Press, Republic of South Africa.

Woolard, I & Leibbrandt, M 1999, 'Measuring Poverty in South Africa', *Development Policy Research Unit,* Working Paper 99/33.

Woolard, I & Leibbrandt, M 2006, 'Towards a Poverty Line for South Africa: A Background Note', Southern African Labour and Development Research Unit, University of Cape Town

World Bank 1990, *World Development Report 1990, Poverty, World Development Indicators*, World Bank, Washington, Oxford University Press.

World Bank 2000, *World Development Report 2000/2001, Attacking Poverty,* World Bank, Washington, Oxford University Press.

World Bank 2001, *The World Development Indicators 1996-1999*, Washington DC.

Yu, D 2010, 'Poverty and inequality trends in South Africa using different survey data', *Stellenbosch Economic Working Paper Series*: 04/10.

Yu, D 2013, 'Some factors influencing the comparability and reliability of poverty estimates across household surveys', *Stellenbosch Economic Working Paper Series*: 03/13.