An empirical analysis of the adequacy of the infrastructure delivery rate to address poverty in South Africa

by

Adél Bosch
920100281

Submitted in fulfillment of the requirements for the degree of

MAGISTER COMMERCII

in

ECONOMETRICS

at the

UNIVERSITY OF JOHANNESBURG

Supervisor
Dr I Botha

Co-supervisor
Ms M Pretorius

November 2008

Opsomming

Titel:	’n Empiriese analise van die doeltreffentheid van infrastruktuur ontwikkeling om armoede in Suid-Afrika aan te spreek
Kandidaat:	Adél Bosch
Studieleier:	Dr. Ilsé Botha
Mede-studieleier:	Me. Marinda Pretorius
Graad:	Magister Commercii in ekonometrie
Departement:	Ekonomie en Ekonometrie
Taal:	Engels
Universiteit:	Universiteit van Johannesburg

In ’n poging om armoede te verlig, investeer die Suid-Afrikaanse regering ‘n groot deel van die jaarlikse begroting om infrastruktuur aan armes te voorsien. Daarom is dit nie net belangrik om die effektiwiteit van infrastruktuur-voorsiening ter verligting van armoede in Suid-Afrika te verstaan nie, maar ook om belangrike vrae aan te spreek soos hoe om die armes te identifiseer. Meer onlangs het lande begin weg beweeg van die tradisionele breë maatstawwe van armoede soos bruto nasionale inkomste per kapita en die Menslike-ontwikkelingsindeks. Inligting oor armoede en ander eienskappe van huishoudings word al hoe meer deur huishouding-opnames verkry. Hierdie opnames voorsien inligting in terme van monetêre en nie-monetêre maatstawwe wat gebruik kan word om armes te identifiseer. Deur van ’n maatstaf vir monetêre armoede, bv. besteding gebruik te maak, kan per kapita huishoudingbesteding bereken word. Huishoudings kan ingedeel word in kwintiele, gebaseer op hul huishoudelike per kapita besteding, en die onderste 20 en 40 persent word gewoonlik as die norm gebruik om huishoudings as arm te klassifiseer. Die data in terme van die geidentifiseerde behoeftiges kan dan geanaliseer word na aanleiding van hul toegang tot dienste en ander huishoudelike karaktereinskappe.

Kwalitatiewe regressie analise het populariteit verwerf in die ekonometriese veld, veral onder sosiale wetenskappe en mediese navorsing. Inligting van huishoudingopnames is dikwels kwalitatief, of binêr van aard. As gevolg van die nie-liniere eienskappe van

binêre afhanklike veranderlike modelle, is die logit en probit-modelle geskik vir hierdie studie. Die volle-inligtings- maksimum aanneemlikheidsmetode, in die konteks van kwalitatiewe regressie analises, is toegepas om die omvang van die effek wat infrastruktuur besteding en ander huishoudings eienskappe op die armes het te bepaal. Die resultate het empiries bewys hoe infrastruktuur investering die kans dat ‘n huishouding arm sal wees betekenisvol verminder, gegewe sekere karaktereienskappe.

Abstract

Each year, in an attempt to alleviate poverty, government invests large parts of the budget to provide infrastructure to poor households in South Africa. This not only necessitates an understanding of the effectiveness of government's infrastructure delivery rate to address poverty in South Africa, but also raises important questions on how the poor can be identified. In recent years, countries have moved away from traditional broad poverty measures such as gross national income (GNI) per capita and Human Development Index (HDI). Information on poverty and other household information are more often collected through household surveys. From these surveys, monetary and non-monetary poverty measures can be used to identify the poor. By making use of a monetary poverty measure such as expenditure, per capita household expenditure can be calculated. Households are divided into quintiles based on their per capita household expenditure, and the bottom 20 and 40 per cent are usually the benchmark for households to be identified as being poor. This is analysed in terms of the poor's access to services and other household characteristics.

Qualitative regression models have gained more recognition in econometrics, especially in the social sciences field. Information collected from household surveys is often qualitative, or binary in nature. Due to the non-linear nature of binary-dependent variable models, logit and probit models were appropriate for this study. The maximum likelihood method, within the binary choice framework, was employed to determine the extent to which infrastructure delivery and other household characteristics have an impact on poverty. The results provided empirical evidence that infrastructure investment can significantly reduce the likelihood that a household will be poor, given a set of characteristics.

Acknowledgements

If it were not for the love and blessings of Jesus Christ, this study would not have been possible. I would like to thank my Lord for the opportunities He has provided me with over the years. May I always follow the path that He has set out for me.

I would like to thank my best friend Cobus for standing by me all the way to the end of what seemed to be a never-ending journey. All your love and support is greatly appreciated.

To my parents who supported me in so many ways: thank you for your support and motivation, not only during the time spent on this study, but through all the years thus far. Thank you for providing me with the opportunities that we so often take for granted. I would also like to express my thanks and admiration to Dr. Ilsé Botha my supervisor, and Me. Marinda Pretorius, my co-supervisor. You provided me with much more than just advice and feedback. Thank you for all your assistance, words of motivation and support throughout the past few years.

I would also like to thank my colleagues at work, and all my friends for their support and tolerance.

Lastly, I would like to thank Dr. Ingrid Woolard who inspired me to pursue the topic for this study.

Pretoria,
October 2008

Content

Figures

Tables

List of Abbreviations

ATM	automated teller machine
CAPI	computer-assisted personal interviewing
CAPS	Cape Area Panel Study
CDF	cumulative distribution function
Cosatu	Congress of South African Trade Unions
CWSS	Community Water Supply and Sanitation
DFID	Department for International Development
DME	Department of Minerals and Energy
DWAF	Department of Water Affairs and Forestry
EA	enumeration areas
FGT	Foster–Greer–Thobecke
FIFA	Fédération Internationale de Football Association
GCIS	Government Communication and Information System
GHS	General Household Survey
GIC	growth incidence curves
GNI	gross national income
HDI	Human Development Index
HPI	Human Poverty Index
HSRC	Human Sciences Research Council
ICT	information and communications technologies
IDASA	Institute for Democracy in South Africa
IES	Income and Expenditure Survey
INEP	Integrated National Electrification Programme
ILO	International Labour Organisation
KIDS	KwaZulu-Natal Income Dynamics Study
kWh	kilowatt-hour
MLE	maximum likelihood estimation
LFS	Labour Force Survey
LPM	linear probability model

MRC	Medical Research Council of South Africa
MTEF	Medium-Term Expenditure Framework
NCA	National Credit Act
NEF	National Electricity Fund
NEP	National Electrification Programme
NGO	Non-Governmental Organisation
OHRO	Oral History Research Office
OHS	October Household Survey
OLS	ordinary least squares
PPP	purchasing power parity
PSLSD	Project for Statistics on Living Standards and Development
PSU	primary sampling unit
RDP	Reconstruction and Development Programme
SACC	South African Council of Churches
SAIRR	South African Institute of Race Relations
SALDRU	South African Labour Development Research Unit
Samwu	South African Municipal Workers Union
SANGOCO	South African Non-Governmental Organisation Coalition
SHS	Solar home system
Stats SA	Statistics South Africa
UN	United Nations
VAT	value-added tax
WHO	World Health Organisation
WRC	Water Research Commission

Chapter 1: The aim of the study and the method of research

1.1 Problem statement

During the 2008/09 budget speech by the Minister of Finance, South Africa's National Treasury Department committed itself to a poverty reduction strategy which included the provision of services such as "*water, electricity, sanitation, education, health care and public transport*" (National Treasury, 2008). There is an increasing acknowledgement that pro-poor infrastructure delivery in some cases has the potential to have a negative social impact on poor communities in a country. This is especially true where organisational models are imposed on these communities, with little or no consultation between the parties concerned (Masika and Baden, 1997). Pro-poor infrastructure projects in many instances fail to recognise gender, cultural and social diversities within poor communities.

Financial investment in infrastructure is often not representative of the actual provision and quality of service delivery. This is in many instances due to the misallocation of funds, and the lack of adequate project planning and management. It is suggested that non-poor households benefit more from infrastructure investment than poorer households and that poor households in some cases gain no benefit at all from infrastructure delivery. The question that arises is to what extent access to infrastructure benefits the poor and what role infrastructure plays in alleviating poverty in South Africa.

In order to answer this question successfully, a necessary step is to identify a measure that will distinguish the poor from the non-poor. The challenge here is to select a measure that accurately measures poverty, from a range of measures that are usually used by researchers. Information on households is usually collected by surveys that are done on samples from a country's population. These surveys also collect information on poverty, and other labour market and social information, such as access to services. Information gathered from these types of household surveys can be useful in measuring the effect that

infrastructure and other household characteristics have on poverty. In an attempt to explain how changes in these characteristics impact on the poverty status of a household, econometric modelling techniques can be used to estimate the interaction between these characteristics (usually referred to as the 'explanatory variables') and poverty (usually referred to as the 'dependent variable'). Different econometric models can be used; however, because the dependent variable only distinguishes between being poor and being non-poor, a non-linear binary-dependent variable model (logit model) is more appropriate to estimate the likelihood that a household will become less poor, should it have better access to infrastructure and simultaneously to control for other demographic characteristics. By employing the logit model, the likelihood of a household becoming less poor, given the extent to which a household has access to infrastructure can be measured.

1.2 Relevance of the study

The South African government invests a large part of the fiscal budget each year in providing social infrastructure, especially to poor households. It is therefore important to establish whether this infrastructure provision lessens the likelihood for households to be poor. Using information from the South African Labour Force Survey (LFS) a non-linear binary-dependent econometric model can be used to estimate to what extent various forms of access to infrastructure will increase the probability that households will be less poor.

1.3 Methods of research

In Chapter 2 the various methods of identifying the poor are discussed. Firstly an overview of both international and domestic poverty measures is provided, followed by an overview of the current measures researchers apply in South Africa. This section also reviews literature on poverty measures, more specifically poverty measures using data from household surveys in South Africa. The last section in this chapter looks at other

poverty benchmarks and the progress made by South Africa in terms of identifying an official poverty line for the country.

Chapter 3 elaborates on the research design and methodology used in collecting data from surveys. The first section highlights the importance of a well-formulated and designed research question. The second section analyses the sample design process, which first looks at the steps taken in selecting a sample. These steps are: defining the target population, selecting a sample from the sampling frame, planning the procedure for selecting sampling units, determining the sample size and determining the sampling unit. The third section provides information regarding the various data collection methods, and also tables the strength and weakness of each of these methods. The fourth section examines data collection instruments, and also discusses the three phases of questionnaire and data collection instrument design. The fifth section explains the actual data collection process, referred to as 'fieldwork'. The last section in this chapter focuses on the South African LFS as a data collection method to provide information on households that can be analysed.

In Chapter 4 data from the LFS are used to calculate per capita household expenditure as a method of distinguishing the poor from the non-poor in South Africa. The first section looks at household expenditure information as collected in the LFS. Second the process of creating a per capita household expenditure indicator is studied in an attempt to identify the poor based on the LFS data. Households are divided into quintiles based on household's per capita expenditure. The last section provides an analysis of how different demographic, labour market, unemployment and other household characteristics differ between expenditure groups. The focus is on the poor, being those in the bottom 20th and 40th quintile of households, and the non-poor, which are those in the top two to three quintiles.

Chapter 5 deals with the infrastructure development in general and in South Africa in particular. The first section looks at the problems in defining and measuring infrastructure development, followed by a discussion on whether infrastructure development does in fact benefit the poor. The third section provides information on the

extent to which the poor have access to infrastructure. Infrastructure refers to housing, water and sanitation, electricity, communication, and financial and transport infrastructure. This section, furthermore, analyse LFS data to determine to what extent the poor have access to these services and, lastly, each of the infrastructure measures is analysed in terms of the policy and financial environment in which they are provided.

Econometrics models for simulating survey data is discussed in Chapter 6. The first section in this chapter looks at the usefulness of binary response modelling approaches in modelling survey data. The first modelling approach that is discussed is the linear probability model, followed by the logit and the probit model. Lastly, this chapter discusses the measures that are used to evaluate these models. They are the pseudo R^2 methods, such as the count R^2 and the McFadden R^2, and the likelihood ratio index to test for overall significance in such a model.

In Chapter 7 the various methods discussed in Chapter 6 are used to analyse LFS data. In the first section the dependent and explanatory variables for estimating a poverty model are explained. The explanatory variables are divided into two parts. The first part looks at demographic and status variables and the second part at infrastructure variables. The second section specifies the three different models. The models that are fitted to the LFS data is the linear probability model, the probit and the logit model. The results, such as the marginal effects and the odds ratio, where appropriate, as well as the goodness of fit of each model are discussed. This model provides the relevant information on how infrastructure development affects the probability of becoming less poor, making it possible to evaluate how infrastructure delivery can alleviate poverty.

Chapter 2: Methods to identify the poor

2.1 Introduction

In 2005 1,4 billion people in the developing world lived below the revised international poverty line of US$1,25 a day in 2005 prices (World Bank, 2008). Because different countries make use of different poverty measures, it is often a challenge to compare poverty across various nations. The World Bank aimed to overcome this by introducing what is considered a benchmark poverty measure. It is not always feasible for individual countries to adopt international poverty lines, as there are various factors that have to be considered when implementing a poverty line. These factors vary between the state of development of a country, the political regime and welfare system, the population demography and location, and the level of infrastructure development.

Poverty can be measured either in monetary terms or non-monetary terms. International benchmarks tend to implement monetary measurements of poverty, such as an income or expenditure value. It is difficult to select a poverty line based on monetary terms due to the subjective nature of distinguishing the poor from the non-poor merely based on a monetary value. It is difficult to argue that a person earning R800 per month is classified as poor, while a person earning R801 per month is not. South Africa is in the process of developing an official poverty line. At present there is no consensus on which rand value will suffice as a poverty line and researchers use a variety of poverty measures, making it difficult to compare results over time.

Poverty can also be measured in non-monetary terms. During the developmental stage of a nation, large numbers of people not only have low income levels, but very often also lack non-monetary wealth, which refers to inadequate infrastructure such as housing, electricity, water and sanitation, deteriorating health conditions, little access to education and overall limited future prospects (Todaro and Smith, 2003). Increasing development is the key to escaping the harsh conditions associated with a *developing* state. Todaro and

Smith (2003) identify three core values of development, which if without, a country's people can easily fall into poverty and a state of impoverishment. These are

1. "*Sustenance*"*:* being able to meet basic needs;
2. "*Self-esteem*"*:* to have a sense of belonging; and
3. "*Freedom from servitude*"*:* enabling broader decision-making.

The above core values only aim to define what is seen as necessary conditions for alleviating poverty. Poverty, in fact, is a relative term and has various definitions. For this reason, the poor are usually identified in terms of a range of characteristics which, depending on the research question, would define a group of people who in some way are excluded from these core values.

Poverty information is collected differently from country to country. Household surveys are commonly used to collect information from households regarding living conditions, employment status and access to service delivery. These surveys also aim to measure income or expenditure for households. Given the above challenges in determining a poverty line, researchers often steer clear of implementing a direct poverty line by dividing households into equal shares, most commonly five or ten equal parts, based on their income or expenditure level. The bottom shares can then be seen as those households whose income or expenditure falls within the bottom 20 or 10 per cent of households. The households can then broadly be classified as poor, compared with households who fall within the top 80 or 90 per cent of households, ranked according to income or expenditure.

This chapter contains an overview of the various domestic and international poverty measures. The second part of this chapter provides an overview of current poverty measures applied by researchers in South Africa, using data collected by means of household surveys. This part will also review literature on poverty measures using household survey data. The last part of this chapter discusses where South Africa is in

terms of identifying a poverty line, together with identifying other benchmarks that could be used instead of those proposed by authorities.

2.2 Overview of measuring poverty in a global and national context

2.2.1 The global context

Gross national income (GNI) per capita as introduced and compiled by the World Bank, has for a long time been seen as the main measure of poverty (Thirlwall, 1999). This measure served as motivation for countries to strive for a high GNI per capita and, in return, be classified in one of four income categories. These categories are classified according to the World Bank Atlas method (World Bank, 2007). The Atlas method smoothes exchange rate fluctuations by using a three-year moving average, price-adjusted conversion factor. The categories are: low income (US$875 or less per month); lower middle income (US$876–$3 465 per month); upper middle income (US$3 466–US$10 725 per month) and high income (US$10 726 or more per month).

The main argument for using this income method specifically to measure *poverty* is that, in theory, the basic purpose of development as argued by Haq (1995) is to enlarge people's choices. However, the increases in income are usually restricted to the top income earners of a country and tend not to trickle down to grass-roots levels, resulting in the enlargement of choices for the already well-off.

Many reasons can be given for why GNI per capita fails to give a true reflection of the poverty profile of a country, one being that a single per capita income figure can be arbitrary when determining the developmental state of a country, as it ignores factors such as distribution of income, differences in development potential and other physical indicators that provide information on a country's quality of life. The second concern, as already mentioned above, is that economic growth and income accumulation in many countries do not trickle down to those living at grass-roots levels. Haq (1995) emphasises

this by explaining that many human choices extend beyond economic well-being, and that the automatic link between expanding income and expanding human choices is not automatic, but rather depends on the quality and distribution of growth, and not only on the quantity of growth.

Owing to the lack of fully functioning markets, growth and growth distribution are often not achieved (May, 2000). Government intervention is in many instances needed to address market failures in an economy (Auerbach and Feldstein, 1987). Ensuring quality growth and more equal distribution thereof are some of the failures that require attention by the so-called political-economic structures of an economy. Government is therefore responsible for improving resource allocation and for creating a sound economic environment that will be inductive for sustainable growth and its redistribution. Development without growth is inconceivable, but growth is possible without development, because often growth does not translate into development, due to the misallocation of resources and the lack of fully functioning markets (Thirlwall, 1999). Conversely growth cannot take place without expanding economic activities and consequently broadening choices.

In summary, the drawback to using a single GNI per capita income figure is that a country may be classified, in World Bank terms, as a high-income country, but in development terms may struggle with unequal income distribution, low human development and restricted choices. According to Thirlwall (1999), development should be a process in which economic and social objectives and values should be drawn together in order to ensure social transformation; which cannot be captured in one single statistic. Because of the limitations inherent in using GNI per capita as a measure of poverty, the United Nations (UN) developed two alternative indices by which the level of development and the progress made can be compared between countries (Thirwall, 1999). These two indices are the Human Development Index (HDI) and the Human Poverty Index (HPI).

The HDI attempts to rank all countries on a scale of 0 (lowest human development) to 1 (highest human development) based on three sets of criteria, which are:

1. longevity, measured by life expectancy at birth
2. knowledge, as measured by a weighted average of adult literacy (two-thirds) and the mean years of schooling (one third)
3. standard of living as measured by real per capita income adjusted for the differing purchasing power parity (PPP) of each country.

The UN released the HPI in 1997. This index measures human poverty as opposed to income poverty. The HPI is calculated in terms of three key deprivations, where a low HPI indicates a lower level of deprivation (Todaro and Smith, 2003). These deprivations are classified in terms of

- *Life:* the average age people are expected to reach
- *Basic education:* measured by the percentage of adults who are illiterate
- *Overall economic provisioning:* measured by the percentage of people without access to health services and safe water plus the percentage of children under the age of five years who are under-weight.

Lastly, poverty is often measured in terms of inequality which, in essence, looks at income distribution. The size distribution of income is the measure most commonly used, and deals with individual persons or households and the total income they receive. In their research, Todaro and Smith (2003) arranged individuals in ascending order based on their personal incomes. Individuals are further divided into distinct groups, such as quintiles (fifths) or deciles (tenths), and they then determined what proportion of the total nation's income was received by each income group. This method can also be used to calculate a poverty gap, which measures the total amount of income necessary to raise everyone who is below the poverty line to above the line.

Alternatively, by plotting the quantitative relationship between these percentages of income recipients and the percentage of the total income they did, in fact, receive during a given period, a Lorenz curve can be constructed (Todaro and Smith, 2003). From the Lorenz curve the Gini coefficient, also a measure of inequality, can be calculated. The Gini coefficient is a number between 0 and 1, where 0 corresponds to perfect equality (i.e. everyone has the same income) and 1 corresponds to perfect inequality (i.e., a small group has all the income, while the rest has zero income).

Despite the potential risks associated with the interpretation of the above poverty measures (GNI per capita, HDI, HPI and various poverty-related inequality measures), these measures are still widely used today to enable comparison between different countries and their state of development. However, evidence shows that countries are changing the way in which they approach poverty research, from a global focus to a more national and internal focus. They exchange broad poverty measures, such as GNI per capita, for more comprehensive measures, which usually involve an analysis of community and household surveys to give researchers a more in-depth picture of how to address poverty challenges adequately and in response tailor policy initiatives.

2.2.2 National context

Apart from international comparable measures of poverty, countries conduct independent research on poverty and other issues such as inequality. This type of research aims to create a poverty profile of a country. Research is also conducted by government or working groups consisting of government and civil society. Research organisations then use various instruments such as surveys to collect information for their research.

It is common for South African literature on social research topics to include discussions on data challenges associated with the apartheid regime (Klasen and Woolard, 2000). The outcome being that South Africa does not have more than ten years of uninterrupted data available for analysis and research, due to the exclusion of the former so-called homelands in surveys; comparability issues due to a change in definitions and questions;

re-estimation and re-weighting of data based on new census releases, and fragmented data, where surveys were discontinued after a short period of time.

Most household- and poverty-related information in South Africa can be gathered from Statistics South Africa's (Stats SA) October Household Survey (OHS) (Stats SA, 1999), the General Household Survey (GHS) (Stats SA, 2007a); the Labour Force Survey (LFS) (Stats SA, 2007b) and the five-yearly Income and Expenditure Survey (IES) (Stats SA, 2000b). These different surveys are in many instances not truly comparable with one another due to differences in sample design. Despite these drawbacks, researchers use these surveys to compare statistics over time, even though sample frames and methodologies between them are different.

These surveys include information on households' labour market status, living conditions and other measures of living satisfaction. Information on household income and expenditure can also be analysed. The most important benefit of working with survey data is that large amounts of descriptive information can be gathered to establish an individual's poverty profile.

Various instruments can be used to measure poverty using a household survey. Woolard and Leibbrandt (1999) refer to the following measures:

- per capita consumption
- household consumption
- per capita income
- per capita food expenditure
- per capita calorie intake
- budget share of food expenditure (food ration)
- average educational level of an adult household member.

2.2.2.1 Poverty research in South Africa using household survey data

Literature on quantitative methods of poverty measures abounds in South Africa, using data from the various surveys available. These sources vary between Statistics South Africa's OHS, the GHS, the LFS and the five-yearly IES. Other independent surveys include the KwaZulu-Natal income dynamics study (KIDS) (KIDS, 2003), Cape Area Panel Study (CAPS) (Lam, Seekings and Sparks, 2006), and the Project for Statistics on Living Standards and Development (PSLSD) (PSLSD, 1994) managed by the Southern Africa Labour Development Research Unit (SALDRU) located at the School of Economics at the University of Cape Town and other smaller surveys run in conjunction with academia and international sponsors. The majority of these studies shows a worsening poverty trend up to 2000, and a more promising trend after 2000 specifically since 2002, when the positive effects from the roll-out of governments' extensive grants programmes can be seen (Woolard, 2003).

The first study to be considered partly summarises the earlier literature on poverty in South Africa. Gyekye and Akinboade (2001) explain how the first poverty research started with the first and second Carnegie inquiries into poverty in South Africa. The first inquiry was initiated in 1929 during the Great Depression, since a large part of the white population was poverty stricken as a result of war, drought and pestilence. This inquiry resulted in funds being allocated to poor white South Africans only. This resulted in the improvement of conditions for whites at the expense of black South Africans. The second inquiry was initiated in 1982 when Alan Pifer, who was president of the Carnegie Corporation until 1982, wanted to deal with the outcomes of the corporation's first study. He felt additional research into poverty in South Africa was necessary; this time including all races. The results were mainly summarised in a "study of black poverty in rural areas" (OHRO, 2005).

The second part of Gyekye and Akinboade's (2001) study constructs a poverty profile of the Northern Province, where around 13 per cent of South Africa's population is situated. They measured poverty by following measures suggested by Deaton and Muelbauer

(1980) cited in (Gyekye and Akinboade, 2001), correcting household expenditure for household size and demographic structure. Apart from selecting a poverty line of 40 per cent of the population based on an ascending expenditure distribution, the authors also select a so called hard-core poverty line where the bottom 20 per cent of the population, based on their expenditure, is classified as poor.

Taking into account the number of individuals who are poor and the depth of their poverty, 41 per cent of the population in the Northern Province live in households with a monthly adult equivalent expenditure of less than R259,11. Using IES data, 537 of the 2 668 households in the province were classified as hard-core poor (Gyekye and Akinboade, 2001).

The expenditure per-adult-equivalent is then used in constructing a poverty line to identify the poor. Although some authors argue that using an adult equivalent approach does not necessarily give better results than using a per capita approach (Pauw, Leibbrandt, Edwards and Dieden, 2006), many authors acknowledge the effect that household size and structure have on measuring poverty. Woolard and Leibbrandt (2006) explain that a household consisting of only adults are "unlikely to have the same consumption requirements" as a household where children are present (Woolard and Leibbrandt, 2006). The result is that poverty measured in terms of monetary value (income or expenditure) is adjusted for household size and composition by using equivalent measures. Woolard and Leibbrandt (1999) cover this topic in detail. Other authors who employ this method are Carter and May (1999), and Meth and Dias (2003).

Meth and Dias (2003) use a variant of the head-count method to establish what happened to the number of people living in poverty between 1999 and 2002. They established a poverty line and then counted the number of individuals whose expenditure or income falls below this line. Data on the distribution of households by expenditure level are used. Nothing is known about the distribution of households by expenditure level. In order to overcome this, it was necessary to construct the relevant distribution by assumption. The study aims to measure the total number of people in poverty in the two bottom

expenditure categories, namely those in households where expenditure per adult equivalent lies between (1) R0 and R399; and (2) R400 and R799 per month respectively. The authors compared LFS 2002 and OHS 1999.

The research adds an additional poverty dimension by incorporating a social wage measure in terms of access to basic services. The result shows an increase in poverty between 1999 and 2002 (assuming the weights are correct) and disseminates the data according to the number of adults, children, pensioners and those unemployed in the household. The results concluded that the number of people in the lowest expenditure category grew by about 2,9 million and in the second lowest category by 1,4 million from 1999 to 2002 (Meth and Dias, 2003).

The number of poor households increased substantially between 1999 and 2002 (assuming that the weightings are correct). Poverty is measured for households with a single adult, two adults and children at various ages. The number of people in the bottom two expenditure classes (R0–R399 and R400–R799 per household per month) increased by approximately 4,2 million between 1999 and 2002. There were 12 million people in the R0–R399 category and 13,6 million people in the R400–R799 category. This was an increase of "approximately 31 per cent and 11 per cent respectively" (Meth and Dias, 2003).

Meth and Dias (2003) employ the poverty line or household subsistence level established by Woolard and Leibbrandt (1999), which amounted to R384 per month in 1999 and R467 per month in 2002. The maximum number of new poor (using the household subsistence level) lies in the region of 4,4 to 4,5 million. In order to calculate the maximum potential per capita expenditure for households with children, Meth and Dias (2003) used a child cost ratio of α=0,5, since it is assumed that children have less expenditure. Households with more than one adult were assigned an economies-of-scale factor of 0,9, since it is assumed that there is a marginal saving in expenditure for more than one adult.

Deaton and Muelbauer (1980) used the following formula to compute household expenditure, correcting for household size and composition:

$$E = (A + \alpha K)^{\theta} \quad (1)$$

Where E = number of adult equivalents

A = number of adults

K = number of children

α = fraction representation of children in adult equivalence i.e., child cost ratio

θ = household economies-of-scale parameter

Hoogeveen and Özler (2004) introduced the various poverty measures used internationally and locally; and comment on other challenges in South Africa such as crime and unemployment. The section of their work dealing with poverty measures used internationally and locally is relevant to this dissertation. They used IES/OHS 1995 and LFS/IES 2000 data. They constructed a poverty line based on "cost of basic needs", which consists of a selected basket of expenditure goods, with a value of between R322 and R593 per capita per month in 2000 prices. The authors also used a lower poverty line of R174 per capita per month (US$2 per day).

To measure head-count rates of poverty and the poverty gap, they employed the Foster–Greer–Thobecke (FGT) measure, which estimates the head-count poverty rate, the poverty gap and the poverty gap squared. The head-count poverty rate represents the total share of the population that is poor, while the poverty gap measures the income amount that is necessary to take people above the poverty line. The poverty gap squared measures the severity of poverty by giving a greater weight to the poorest of the poor. The authors warned about possible shortcomings in the data and dealt with them by doing sensitivity analyses and reported the results for each potential problem.

The results show that roughly one third of the population lived on less than US$2 per day in 2000. For any other poverty line below R322 per capita per month, there is a

significant increase in the gap as well as the severity of poverty between 1995 and 2000. Severe poverty among black Africans has increased using any of the three FGT measures suggested in the paper. Across provinces a large variation in the level and changes in poverty was found, and although some provinces have improved, others have seen increases in the number of the poor.

Dieden (2002) used OHS/IES 1995 data to analyse how income is generated by households in lower-income groups. The way of compensating for household size and structure when comparing various poverty lines is taken from research done by Deaton and Muellbauer (1980), and Dieden (2002). The bottom 40 per cent of a household's cut-off line is used to identify the poor, after ranking households in ascending order based on their income level. Dieden (2002) also looked at the main sources of income of the poor and established that a large part of households only have one main income earner. The study outlines the various linkages between poverty and demographic characteristics, such as gender, population group, location and income. The result is one that confirms South Africa's discriminating heritage. After excluding Indians and the white population group from the analysis, differences were found between black African and coloured population groups based on spatial dimensions. The results show how main income sources vary across demography, especially in the rural–urban divide and across provinces. Multinomial logit regressions for each of the poverty lines were constructed.

The results show that the variation in poverty risk associated with certain income sources changes as the severity of poverty changes. In fact, the variation in poverty risk increases the more severe poverty becomes. Income sources strongly associated with severe poverty were pensions, public transfers, primary sector wages and wages from several sectors.

Van der Berg et al. (2006) cover a large part of the literature and descriptions of the deficiencies in the data sets, and cites authors such as Simkins (2004), Ardington et al. (2005), Vermaak (2005) and Hoogeveen and Özler (2004) who actually re-weighted the IES 2000 to the 2001 census. Problems associated with the data are the many zero and

missing incomes reported, and incomes reported in bands, which make it difficult to determine the true distribution. Other problems are comparability and sampling error, in terms of over-estimating black Africans and under-estimating whites (Hoogeveen and Özler, 2004). This research focused on income distribution from 1993 to 2004 using various surveys such as LFS and IES. An important variation that it employs is based on work done by Leibbrandt, Levinsohn and McCrary (2005) (cited in van der Berg et al. (2006)), where the national accounts series for aggregated household income were used to adjust the income distribution estimates of the household surveys.

The research analyses per capita income tend to establish whether poverty has worsened or improved since 1994. In identifying the poor, they set one poverty line at R250 income per household following van der Berg and Louw (2004) (cited in Van der Berg et al. 2006), and a second one following Bhorat, Naidoo and van der Westhiuzen (2004) (in van der Berg et al. (2006)) at R281 per month. In addition, they test the sensitivity of their poverty findings against alternative poverty lines.

The results show that for the above poverty lines "poverty seems to have decreased sharply since about 2002 while there was some indication of a slight increase over the period 1993 to 2000". Giving four theoretical explanations with a view to dismiss the decrease as data failure, the authors, however, do warn that they might have over-estimated poverty in the analyses. The authors also note that when comparing IES 1995 and IES 2000, one must be aware of the sharp decrease in income levels. The research further analyses growth in the middle class in South Africa, and adds that it is the increase in black African people in the middle and upper classes that has contributed favourably towards reducing poverty. The results are tested by estimating cumulative density functions, which emphasises that poverty is not dependent on the poverty line chosen or the poverty measure used.

Bhorat et al. (2006) use South Africa's 1993 integrated household survey from the PSLSD, the 1999 OHS and the 2004 GHS. The 1993 integrated household survey was

conducted by SALDRU at the University of Cape Town's School of Economics and was South Africa's first representative multipurpose household survey (Bhorat et al., 2006).

The shifts in non-income welfare are analysed by using a factor analysis technique to construct an asset index. They analyse the provision of free basic services (i.e., electrification, water, sanitation and formal housing). Growth incidence curves (GICs) were employed, developed by Ravallion (2004), and Ravallion and Chen (2003) (cited in Bhorat et al. (2006)). GIC measures growth rates in per capita income or expenditure over a selected period at each percentile of the distribution, ranked from low to high.

Results show an increase in service delivery between 1993 and 1999. It is found that a large proportion of poor households still do not have access to basic services, especially in terms of toilets and water in the bottom deciles of expenditure. An asset index was derived (which is a non-monetary measure of poverty) and a scoring method was implemented for a selected basket of assets, such as type of dwelling material and asset ownership.

By making use of a kernel density estimate, the results show an improvement in access to assets and services at the bottom end of the distribution, together with an increase in the top end, which could be a result of economic growth rather than service delivery. Poverty lines are then calculated based on the distribution of the asset index. Bhorat et al. (2006) use the 20th and 40th percentile of the 1993 asset index, after which the FGT measures of poverty is employed to determine the depth of poverty and the poverty gap. Results based on the 40th percentile head-count rates showed poverty almost halved between 1994 and 2004, from 40 per cent to 21,6 per cent. Most of the decrease took place between 1994 and 1999. Adding all demographic factors showed the same results.

Pauw et al. (2006) draw linkages between poor households and trade, employment and earnings in South Africa. Although the research only highlights selected poverty characteristics of the poor in South Africa, it still draws on important methodology. The data used for the analysis is the merged IES/LFS 2000. Concerns were raised pertaining

to merged data, and in terms of their reliability and comparability. See Poswell (2003), Pauw (2005) and Simkins (2003) (cited in Pauw et al. (2006)).

Poverty was then defined based on the World Bank's definition as stated by Woolard and Leibbrandt (1999) (cited in Pauw et al. (2006)) the "inability to attain a minimal standard of living". Measured in terms of annualised per capita expenditure of households, calculated by taking a household's total annual expenditure divided by its number of members, household quintiles are derived by dividing household's per capita expenditure into percentiles. The bottom two quintiles are then classified as being 'poor'. The ultra poor refer to those in the bottom quintile.

The results show that 13 million people lived in households who were in the bottom 20 per cent of households based on per capita household expenditure and 10 million in the second quintile. Together, 56 per cent of individuals lived in poor households in 2000. Woolard and Leibbrandt (1999) summarise the steps taken in analysing poverty starting with the ranking of households or individuals on the basis of some monetary variable such as income or consumption expenditure. Material wellbeing is then measured based on the amount that an individual consumes, ignoring public consumption. The data used for this analysis were data taken from PSLSD. Woolard and Leibbrandt (1999) identify alternative poverty measures and compare the number of poor in the bottom 40th percentile for the various poverty measures. The results show that "some definitions of poverty are much more strongly correlated with per capita expenditure than others" (Woolard and Leibbrandt, 1999).

Woolard and Leibbrandt (1999) then aim to select a poverty line to distinguish between the poor and the non-poor. The research also reviews the difference between an absolute and relative poverty line, and the various poverty lines used in South Africa. Consumption patterns are adjusted for household size and structure, by computing a per capita consumption and per capita adult equivalent poverty measure. A demand model is derived where the share of the food budget is regressed on the log of per capita expenditure, using different equivalent measures. The results show that the choice of

equivalent scale does not make a big difference in estimating the number of households that are poor. Finally, by making use of household survey information, a poverty profile is drawn up according to demographic and labour market characteristics. While using a head-count index and a poverty gap index, the depth of poverty is measured.

The results show the severity of rural poverty in South Africa. Of those living in rural areas, 63 per cent of people are poor at a poverty line of R3 509 per adult equivalent per annum. The Eastern Cape and Northern provinces had the largest number of poor people when using the head-count ratio and poverty depth measure. Racial poverty shows that black African and coloured population groups experience greater poverty than whites and Indians. Poverty among gender groups was higher among female-headed households than male-headed households, using the higher poverty line. The results also show a relationship between education and poverty. Certain health issues associated with nutrition were found to be more likely among the poor. Poverty was also highly linked to unemployment, income sources and access to services.

2.2.3 A South African poverty line

A poverty line is constructed based on a selected poverty measure. The measure also depends on the type of poverty that is to be measured. It can be a monetary measure, such as income or consumption expenditure; or a subjective measure, such as children's calorie food intake; social exclusion; or measuring asset-based poverty. A poverty line divides the population into two groups based on the selected measure (Woolard and Leibbrandt, 1999). Below this line, individuals or households are classified as poor, and those that are above the line can be classified as those who are better off. A poverty line is a measure of the minimum resources that one require for a household or individual to survive (Stats SA/Treasury, 2007).

Although the above statement aims to define a poverty line, it is only a description. There can never really be only one definition of a poverty line, because of the subjective nature of a poverty measure. It is also for this reason that some countries or researchers use two

or more poverty lines, making it possible to distinguish between various levels of poverty (Frye, 2005).

There are two main approaches in identifying a poverty line, namely (1) a direct approach and (2) an income approach. The direct approach identifies individuals or households who cannot meet a selected minimum level of basic resources, and the income method identifies those who are not capable of affording a satisfactory basket of consumption goods. Woolard and Leibbrandt (1999) also distinguish between absolute and relative poverty lines. An absolute poverty line does not change over time, with the exception of inflation adjustment, whereas a relative poverty line changes as the standard of living changes, for example, a change in the average income of households or individuals will increase the average 40th percentile cut off-point for classifying the poor.

South Africa is currently in the process of developing an official poverty line, which will enable poverty researchers to compare results, and the time spent on debating the correct poverty line can be put to more productive use. The South African Medical Research Council (MRC) recommended that the daily energy requirement, which is 2 261 kilocalories per person per month, should be used as a poverty measure (Stats SA/Treasury, 2007). According to the income and expenditure survey, in 2000 prices, to meet this requirement it will cost a person R211 per month.

Statistics South Africa estimated a further non-food component to determine the poverty line (Stats SA/Treasury, 2007). This amount comes to around R111 per person per month. Added together, the poverty line amounts to approximately R322 per person per month in 2000 prices.

Other researchers identify various other measures in which a poverty line can be defined. Poverty can be benchmarked in the following ways:

- The number of people who live in households who fall within the 20th, 40th or 50th percentile cut-off of household per capita expenditure (Pauw et al., 2006)

- Utilising expenditure brackets as used in Statistics South Africa's questionnaires of R0–399, and R400–R799 (Meth and Dias, 2003; Bhorat et al., 2006)
- Amount required for a certain minimum calorie intake required per day or per month or per year (Woolard and Leibbrandt, 1999)
- Minimum supplement levels as set by the Bureau of Market Research as reported in Woolard and Leibbrandt (1999)
- International poverty lines set by the United Nations Development Programme of US$1 and US$2 a day (Van der Berg et al., 2006; Hoogeveen and Özler, 2004)
- Poverty index, based on non-monetary values such as level of access to basic services, health care and access to finance (Bhorat et al., 2006).

2.3 Conclusion

This chapter takes an extensive look at the literature on poverty, and how poverty is measured in a global and national context. It was established that with a view to establishing an international poverty measure, countries used to make use of broad poverty measures, such as GNI per capita, HDI and a HPI. These measures, however, display various drawbacks, one being that the GNI per capita does not provide any information on the distribution of wealth in a country. More recently, countries have changed the way in which they approach poverty measurements, from a global to a more national and internal focus. This comprehensive method usually involves analysing community and household surveys. South Africa, and many research and statistical organisations use instruments such as surveys to collect information on poverty. A number of surveys are introduced in this chapter. These include the OHS, GHS, LFS, IES, KIDS and CAPS. Information on households' labour market status, living conditions and other measures of living satisfaction are usually collected in these types of surveys.

The next challenge that arose was whether a monetary poverty measure or a non-monetary poverty measure should be used to distinguish between the poor and the non-poor. When using a monetary poverty measure, a suitable cut off-point for those who can be defined as poor must be determined. This chapter has established that this process can

be very subjective. To give an overview of how other researchers approached the process of determining who the poor are, various examples in the literature on poverty research were reviewed.

Among others, Meth and Dias (2003) determined that when using a monetary poverty measure the number of people who fall within the lowest category of R0–R399 per month increased by approximately 1,4 million between 1999 and 2002. Hoogeveen and Özler (2004) showed that roughly one third of the population lived on less than US$2 per day in 2000. When determining these measures, household size and composition can also be taken into account and adjusted for by using an expenditure-per-adult equivalent measure, which adjust for the number of adults and children in the household. The FGT method can also be used to measure the severity of poverty, by calculating the poverty head-count, poverty gap and the poverty gap squared.

Bhorat et al. (2006) implemented a non-monetary poverty measure by analysing the provision of free basic services. The results showed that even though progress was made during 1993 and 1999, little headway was made thereafter, resulting in a large proportion of the poor still not having access to basic services.

The last part of this chapter investigates the process of establishing an official poverty line for South Africa. In consultation with the MRC, it was recommended that the official poverty line should reflect the minimum daily energy requirement of 2 261 kilocalories per person per month. It is estimated that meeting this kilocalorie intake requirement will cost an individual around R211 per month in 2000 prices. When adding a further non-food component to this, this value increases to R322 per person per month.

Chapter 3: Research design and methodology: overview of the data collection process

3.1 Introduction

> The research design is the blueprint for fulfilling objectives and answering questions.
>
> — D. R. Cooper and P. S. Schindler. Business Research Methods (p.81)

In Chapter 2 of this dissertation it was established that researchers rely extensively on data and information collected from communities to measure poverty and other development characteristics. In order for researchers to collect accurate information, the purpose for which the information will be used must be clear and concise. Results from the information collected are often used in policy making, therefore sampling methods and design provide the foundation for accurate results and interpretation.

In this chapter first the importance of a well-thought-through and proper research question design is discussed. Second, the sample design process is analysed. This process looks at the steps taken in selecting a sample, which include defining the target population, selecting a sample from a sample frame, planning procedure for selecting sampling units, determining the sampling size and, determining the sampling unit. Third, the different data collection methods, including each method's strengths and weaknesses, are set out. Fourth, an outline of the various data collection instruments are provided, followed by a discussion on the three phases of questionnaire and data collection instruments. Next, the actual collection of the information through fieldwork is discussed. Owing to the importance of the LFS for this particular dissertation, the LFS as a data collection tool will be discussed in detail.

3.2 Define the research question

Before attempting to collect information, a clear purpose of the study should be identified. The clearer the goal of the research project, the easier the collection process

will be, and the easier the results will be to interpret. A 'research question' can be defined as the hypothesis of choice that best states the objective of the research study (Cooper and Schindler, 2003). The research question will further also determine the research design that will be used to collect the information. If the research question is more explorative and descriptive in nature, rather than testing a hypothesis, surveys are usually used to collect data. The researcher should start off by getting acquainted with the facts that are required to meet the purpose of the study and knowing in what way the information will be used (Taylor-Powell, 1998). To make the process of defining the research question simpler, the following is suggested by Taylor-Powell (1998):

- List the information that is required.
- Find out whether the required information is already available to avoid duplicating work.
- Plan how the information gathered can be optimally used.

After establishing the research goal or question, the researcher must decide from whom the information needs to be collected. The researcher also has to determine how the sample will have to be designed to collect optimal information and results.

3.3 Designing the sample from which data will be collected

The sample design plays an important role in determining the statistical properties of the data, the sample size and the way in which the information should be analysed. Most household surveys make use of complex sample design because of the precision that is required to make it possible to make accurate inferences about the population from the sample (Chromy and Abeyasekera, 2005). In most cases it is not possible to survey each person in the population and it is therefore important to design a sample as sufficiently as possible to represent the population accurately (Madow, Nisselson and Olkin, 1983).

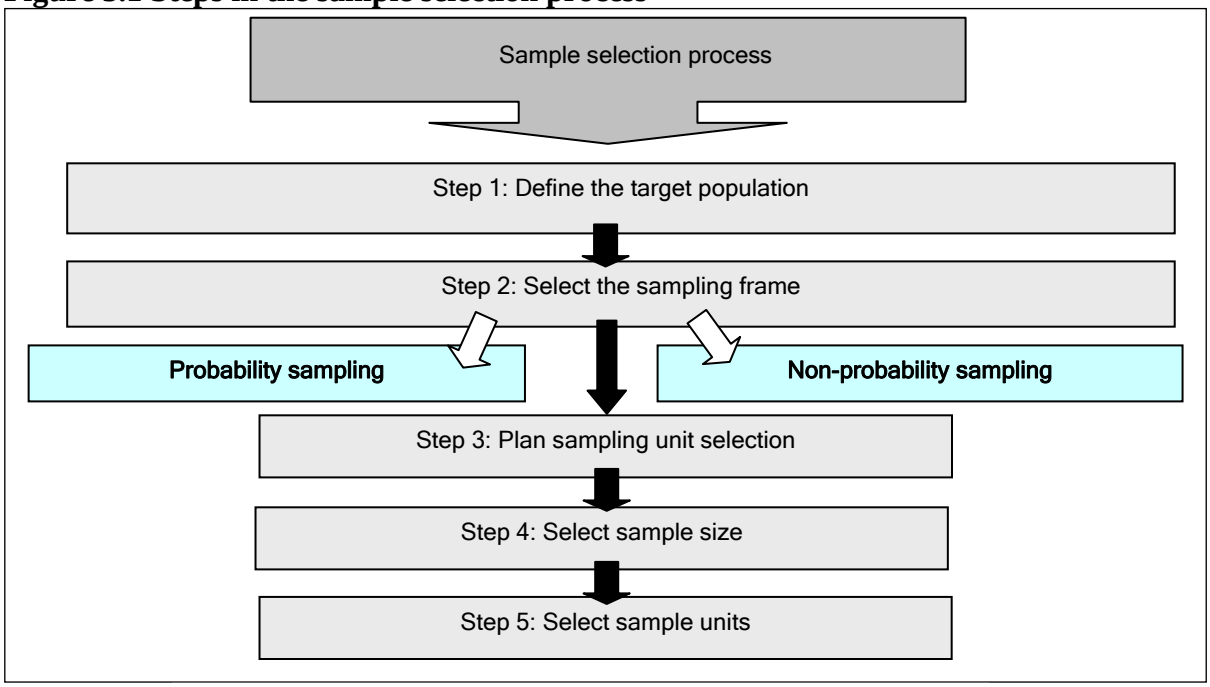
Nassirpour (2004) proposes that the following criteria should be considered to ensure a successful survey:

- *The resources available:* this includes all costs and human resources especially for face-to-face interview surveys (Owens, 2002).
- *Degree of accuracy required:* which includes balancing accuracy with efficiency. Here the designer should decide to what degree is he or she is willing to make an error (Lehtonen and Pahkinen, 1995).
- *Time requirement:* the time needed to write and programme the questionnaire, as well as the time needed for implementation and quality checking (Owens, 2002).
- *A certain level of knowledge of the population is essential:* specifically when conducting complex surveys. It is necessary for the designer to have some initial information on the spread and natural groupings of the population to be surveyed. This will assist when a sample design employs stratification and cluster sampling methods.
- *Before the survey is designed, it should be determined for which purpose the survey and information will be used:* which will enable the design of a survey that will reach its objectives and that is tailor made in terms of its design.

3.3.1 Steps in selecting a sample

Nassirpour (2004) suggests a five-step sample design process to ensure a high-quality sample design. The sample design covers the rules and methods by means of which the researcher selects sampling units from the population. This is usually done randomly. These steps are described in the following diagram followed by a description of each step:

Figure 3.1 Steps in the sample selection process



*Source: A*dapted from Cooper and Schindler (2003)

3.3.1.1 Step 1: Defining the target population

The specification of elements or groups of individuals, from whom information will be collected needs to be defined as clearly as possible. There are no strict rules to follow, because sometimes it could be that the population size is small enough and the entire population can be included in the study. However, usually the population is too large to attempt to survey every element in the group (Statpac inc., 2007). A subset or representation of the population is then selected as a sample. This should reflect the characteristics of the population from which it is drawn. This population from which the sample is selected is usually referred to as the 'sampling frame'. There are different methods in which the sample can be selected from the sampling frame. The next step will be to determine which sample selection method will more accurately collect the required information.

3.3.1.2 Step 2: Selecting a sample from the sample frame

A sampling frame is a register or units of the population elements from which the sample is drawn and often includes additional information on the structure of the population (Lehtonen and Pahkinen, 1995). A sampling frame is thus a physical representation of the target population (Nassirpour, 2004). Selecting a sample from the sampling frame can be done in two ways:

i. Probability selection

When a sample is drawn from the sample frame using probability selection, each member of the population has a known probability of being selected. This probability lies between 0 and 1. The probability selection method can be divided into different approaches (Trochim, 2006):

- *Simple random sampling,* where the sample is drawn not based on any information, but simply randomly.
- *Systematic random sampling,* where the units to be sampled are randomly ordered, and then based on the interval size (i), every i^{th} unit is then selected as part of the sample.
- *Stratified random sampling,* which is also sometimes referred to as 'proportional' or 'quota random sampling', is often preferred to simple random sampling because it reduces sampling error (Statpac inc, 2007). In stratified sampling the target population is divided into non-overlapping sub-populations called 'strata'. Regional, demographic and socio-economic variables are often used as the stratifying variables (Lehtonen and Pahkinen, 1995). Examples of stratums might be males and females, or black Africans and whites.
- *Cluster random sampling,* helps to overcome the challenge of having to travel across the entire country, trying to survey everyone in every province. With random sampling methods it becomes challenging to sample a whole population that is distributed across a wide geographic region; for instance, conducting an interview

survey for all the residents in South Africa. With cluster sampling the population is divided into clusters, which could be based on any type of information, but is usually determined in terms of geographic boundaries. These clusters are then randomly sampled and the units within each cluster are determined (Trochim, 2006).

- *Multi-stage sampling,* involves a combination of the above sampling methods. Should the residents of a country be sampled by means of face-to-face interviews, cluster sampling as the first stage of the process will be very useful. However, it will still be difficult to measure each person in each cluster, and therefore in the second stage the units in the clusters might be stratified. The combination of different sampling methods enables researchers to make use of all the benefits that each sampling method provides.

ii. Non-probability selection

When using non-probability sampling selection, population elements are non-randomly selected (Statpac inc., 2007). This does not necessarily mean that non-probability sampling is not truly representative of the population. It does, however, mean that the assumptions and underlying principle of probability theory can not be applied. With probability sampling the odds of the sample presenting the population successfully can be determined, or at least be estimated within a selected confidence interval.

With non-probability sampling selection, it is difficult to presume with a certain level of precision whether the sample is representative of the population or not. Non-probability sampling methods can be divided into accidental or purposive sampling. The latter being used more frequently because the selection method is based on some set of pre-determined criteria (Trochim, 2006):

- *Accidental sampling* is when sampling is done unintentionally or even only for the purpose of convenience. Some researchers, such as Trochim (2006) refer to this method as "man on the street" interviews. Journalists usually use this type of sampling because the person is available and easy to interview, even though the

opinion of that person is not representative of the population as whole. It is clear that this type of sampling cannot be assumed to be representative of the population and therefore is open to criticism and misinterpretation of results.

- *Purposive sampling* is when the researcher is sampling specific subjects relevant to the research problem. For instance, conducting a survey at a shopping complex or school for a specific group of people, say to research the credit behaviour of females. The fieldworkers will first find a female respondent and then enquire whether she buys on credit. With this type of sampling it is likely that the sample will over-represent a certain group, especially in respect of availability and accessibility.

3.3.1.3 Step 3: Plan procedures for selecting sampling units

When using stratification and the strata sizes are very different, small strata may have a very small sample size. Therefore, a power allocation with a selected power value is usually used to guarantee a sufficient sample size (Lehtonen and Pahkinen, 1995). A power value of 1 gives proportional allocation, a power value of 0 gives equal size allocation, and a power value between 0 and 1 gives an allocation between proportional allocation and equal size allocation, thus being a compromise between the two (i.e., disproportionate) (Macro International Inc., 2005).

3.3.1.4 Step 4: Determine the sample size

It is important to consider the sample size carefully, because it has statistical implications (Asraf and Brewer, 2004). This decision is usually made after the background information has been collected and the sample design has been completed. The desirable sample size depends on the size of the sampling error the researcher is willing to make. The sample size will depend on the following factors (NAO, 2000):

- The more variation in the data, the larger the sample needs to be to obtain an adequate level of accuracy in generalising the results to the population

- To what extent the researcher is willing to make an error
- The precision to which the researcher wishes to measure his or her estimates based on the selected confidence level
- The population size and proportion of the population relevant to the research question plays important roles in determining the sample size

Statistic experts recommend having 1 000 participants to ensure a statistically significant number of responses depending on the type of questionnaire, although this is not always possible, given time frames and resources (di Martino, 2002). The following calculation is usually used to determine the sample size, given the margin of error the researcher is willing to make; the standard error formula is (Keller and Warrack, 2000):

$$\sigma_p = \sqrt{\frac{pq}{n}} \quad (1)$$

Where:

p = the probability of a success; and

q = the probability of a failure (1-p);

n = the sample size that should be determined, based on the selected confidence interval.

For a given sample the standard error is at its maximum when $p = 0,5$. If a coin is flipped, then there is a 50 per cent chance of it falling on heads and a 50 per cent chance of it falling on tails. The same holds true for a random sample, namely if the zero hypothesis states that "black African females live in poverty" then the black African female who is selected should have a fifty-fifty chance of living in poverty. It could also be that the probability is higher when black African females are sampled in rural areas, and the p value will then perhaps be 0,8 and the alternative $q = 0,2$. It is assumed that no information is available on population and the probability is 0,5. The 95 per cent confidence interval is represented by z, which covers a distance of $\pm 1,96$ standard errors under the normal curve, and whilst allowing for a 1 per cent error on either side of the estimated proportion. By applying the following formula the sample size can be determined (Keller and Warrack, 2000):

$$n = \frac{(z)^2 (p)(q)}{(0,01)^2} \qquad (2)$$

$$n = \frac{1,96^2 (0,5)(0,5)}{(0,01)^2}$$

$$n = 9,604$$

It can now be said that on a 95 per cent confidence level of prediction, within 1 per cent error on either side of the estimated proportion, a random sample of 9,604 is required.

3.3.1.5 Step 5: Select sample units

After the sample size has been determined, the population needs to be subdivided into primary sampling units. Some of these units are then randomly selected whereafter they are subdivided into a series of secondary sampling units. Some of these secondary units are then randomly selected as the sample of the second stage. This process can be repeated (Brack and Wood, 1998).

3.4 Methods of data collection

Data collection can be classified in four ways, namely (1) focus groups, (2) interviews, (3) monitoring, and (4) interrogation/communication, which consists of surveys and follow-up interviews (Cooper and Schindler, 2003). When using the interrogation/communication method of data collection, information is obtained by questioning subjects and documenting their responses. This can be done in three ways: (1) by means of an interview or telephone conversation; (2) by means of surveys sent through the mail, e-mail or left at strategic locations; and (3) information collected before and after an experiment has been conducted (Cooper and Schindler, 2003). Each method presents its own set of strengths and weaknesses. Before a researcher embarks on collecting information, he or she must clearly define the research question. By setting

clear goals, the task of deciding on a method of data collection should become obvious. What follows is a description of each of the data collection methods, together with its strengths and weaknesses:

3.4.1 Focus groups

Focus groups consist of a panel of typically six to ten people who meet for 90 minutes to 2 hours. The group is led by a trained moderator or facilitator who guides the group in an exchange of ideas, feelings and experiences pertaining to a specific topic. This method of data collection became popular in marketing research during the 1980s. Although this method is used in many diverse ways, a common purpose continues to be in the consumer arena, where new products or product concepts are usually developed. Participants of a focus group do not have to be literate, since the technique relies on oral communication (Durrance, Fisher and Hinton, 2005).

3.4.2 Interviews

Interviews can be divided into three groups (Cooper and Schindler, 2003):

1. *A personal interview* (i.e., face-to-face communication) is a conversation between an interviewer and a participant with the aim of the interviewer being able to collect information from the participant. The interviewer and the participant are usually strangers, and the interviewer generally controls the topics and patterns of discussion. Highly trained interviewers usually gather better-quality information. The participant is required to give information and usually does not receive any compensation for his or her co-operation.
2. *Telephone interviewing* is when questionnaire data is gathered from people by trained interviewers over the telephone.
3. *Computer-assisted personal interviewing* (CAPI) is a method of personal interviewing, where information is gathered from a respondent and directly imputed into a laptop or hand-held computer.

3.4.3 Monitoring

Monitoring involves the researcher observing the actions of the subject without expecting a direct response. Examples would include counting traffic at an intersection and searching through books in a library. In this type of data collection the researcher makes notes and records the information obtained from observation. This is an excellent method to discover the behavioural patterns of subjects (Durrance et al., 2005).

3.4.4 Surveys

Surveys involve recording people's responses for analysis. Surveys can include both close-ended and open-ended questions, and can be administered either in written form or online. Surveys usually consist of questionnaires that are completed by the respondent. A well-designed survey contains different types of survey questions. Open-ended questions are questions to which there is not one definite answer. Open-ended questions may be a good ice-breaker because it gives the respondent the opportunity to answer in his or her own words. However, close-ended questions have a pre-defined set of answers from which the respondent can choose. Usually an option for "other" is included, and sometimes an extra space is allowed for the respondent to write in a response, should he or she choose this option (Trochim, 2006).

Conceptually, sample surveys can be divided into two broad categories namely, (1) descriptive surveys and (2) analytical surveys. In a descriptive survey certain population characteristics need to be precisely and efficiently estimated. This type of survey usually focuses on estimating the state of a fixed population at a specific time, and examines means and proportions. A large sample is usually necessary. An analytical survey, in turn, is often multi-purposed so that a variety of information can be gathered from it. This type of survey usually gives attention to relationships between variables for an infinite population and is usually used in regression analysis (Caldwell, 2007). Countries usually use the following types of surveys to collect information from households:

- *Population census:* A population census not only counts the elements in a population, but also collects information on the demographic status of a country (Scott, 2008). A typical census collects information on age, gender, family and household composition, basic education levels and basic access to infrastructure. Because of the extensive coverage of this type of survey it is very costly. According to international recommendations, a population census should be done once every ten years (Scott, 2008). For these reasons the census scope is limited for poverty measurement in the short term. It does, however, generate the frame from which household samples are drawn for household surveying.

- *Labour Force Survey:* The LFS is designed to collect information from the working age population (15–65 years) on key labour market variables such as labour force participation, unemployment, industry segregation and occupation profiles. LFSs are usually carried out systematically over time and in some cases can be used to measure welfare and monitor poverty. LFSs also collect information on aggregate income and expenditure of the household. A problem usually associated with this type of information is that incomes and expenditure of some households could be underestimated more than others, which could lead to misidentification of the poor (Scott, 2008).

- *Social surveys:* Social surveys focus specifically on household demographics, welfare status and service delivery. They usually contain multi-topic information to provide a comprehensive picture of the population and how it lives (Scott, 2008). These surveys often have the same challenges associated with the LFS. The frequency of social surveys is usually one to five years.

- *Income and expenditure surveys:* Most countries have implemented an IES (Scott, 2008). These surveys are also sometimes referred to as 'family' or 'household budget surveys'. According to the International Labour Organisation's (ILO) resolution concerning household income and expenditure statistics, adopted in December 2003,

a survey of household expenditure should be done every five years, and preferably not longer than that. The ILO further recommends that countries that experience fast-changing political, economic and welfare conditions embark on these types of surveys more frequently (ILO, 2003). For poverty research, the IES provide good information on consumption patterns of households. The goal of the IES, however, is not to measure households' welfare, but to measure the expenditure on certain goods and services. Owing to the short reference periods usually used in IES, the measurement of individual household expenditure can be problematic.

3.4.5 Follow-up surveys/interviews

Follow-up surveys/interviews have become synonymous with modern-day living because businesses are continuously trying to improve their service delivery and therefore rely on feedback from consumers on a frequent basis. These types of surveys usually take up the form of a short questionnaire that is left behind at restaurants, hotels and so forth. Participants are then required to evaluate service delivery and consumer satisfaction (Cooper and Schindler, 2003).

Table 3.1 Strengths and Weaknesses of data collection methods

Data collection method	Strengths	Weaknesses
Focus Groups	▪ Data collection is fast. ▪ Respondents can respond in their own words. ▪ Free flow of ideas and discussion.	▪ Limited sampling accuracy. ▪ Time consuming. ▪ Lack of generalisation.
Interviews	▪ As the respondent starts to trust the interviewer, it is more likely that all the questions will be answered. ▪ Overcome literacy problem. ▪ More psychological depth.	▪ Does not allow the survey to be completely anonymous. ▪ Interviewer can influence the answers intentionally or unintentionally. ▪ Time consuming and hard to control.

Data collection method	Strengths	Weaknesses
Monitoring	▪ Capture actual behaviour or process. ▪ Fairly easy to accomplish. ▪ Provide rich information.	▪ Can seem intrusive. ▪ Hard to capture everything, due to note taking. ▪ Difficult to know what to observe: information overload.
Surveys	▪ Efficient and less expensive. ▪ Less time consuming for participants. ▪ Easy to distribute, and large samples are feasible.	▪ Difficult to design a good survey. ▪ Possible to ask biased or leading questions. ▪ Accuracy is questionable due to self-reporting of information.
Follow-up surveys/interviews	▪ Good way to follow up on changes in feelings and experiences. ▪ Decrease the non-response rate.	▪ Costs of follow-up usually outweigh the benefits. ▪ Require additional resources (i.e. staff).

Source: Meyers (2006); Cooper and Schindler (2003).

Now the researcher must decide what the time dimension of the survey should be. A survey can be conducted at one point, or can be repeated over an extended period. The former is referred to as a 'cross-sectional survey', while the latter is referred to as a 'longitudinal survey'. Longitudinal surveys can be divided into trend surveys, cohort surveys and panel surveys (Cooper and Schindler, 2003). Trend surveys are surveys that repeatedly monitor the same sample population at different times. The sample taken from the population is the same each time, however, it usually consists of different people. Cohort surveys have a different focus, even though they also survey particular populations, at different times. A researcher would typically survey a graduate class of the year 2000 in one period and a different graduate class of the same year five years later, whereas, a panel survey is done to determine why certain changes are happening in the population (Palmquist, 1997). The same sample of people is used every time.

3.5 Questionnaires as Data Collection Instruments

Questionnaires play a central role in the data collection process and are an important data collection instrument. When choosing an instrument, the following should be considered (Cochran, 2007):

- appropriateness
- reliability
- validity
- responsiveness
- precision
- interpretability
- acceptability
- feasibility.

When a questionnaire is well designed and planned, the required information can be collected efficiently with the minimum number of errors. The instrument design can be divided into two phases (Adapted from Cooper and Schindler, 2003):

3.5.1 Phase 1

During the first phase, it must be decided what type of data needs to be collected. This can be divided into four data types, namely nominal, ordinal, interval and ratio.

1. *Nominal data* can be grouped into two or more categories that are mutually exclusive and collectively exhaustive (Cooper and Schindler, 2003). Nominal data have no specific order or distance relationships, and no specific origin. Types of questions using this data type are usually questions around marital status, gender, and so forth.
2. *Ordinal data* are the same as nominal data; however, information is given in a specific order. An example of ordinal data is when the respondent can choose on a scale of 1–5 whether he or she is very happy (1), or very sad (5). The numbers in between represent the different places between very happy and very sad, however the following condition should hold: $1 > 2 > 3 > 4 > 5$.

3. *Interval data* combine nominal and ordinal data, and in addition adds the concept of equal intervals. An example would be calendar times and temperature scales.
4. *Ratio data* have all the concepts of the above plus provide for an absolute zero origin. Measures of ratio data would include weight, height and distance.

The next step in the first phase is to identify the question structure. The researcher can decide between administrative, classification and target questions (Cooper and Schindler, 2003). Administrative questions identify who the respondent is, who the interviewer is, and what the location and conditions where the interview takes place are. Classification questions are usually questions that cover the social aspects of the respondent, where target questions can be defined as those questions that address the specific research question.

3.5.2 Phase 2

During phase two the questions must be developed and designed. Shorter and well-laid-out questionnaires are more likely to be completed. General principles of questionnaire design are (Cochran, 2007):

- All questions and sub-questions need to be clearly labelled.
- Steer clear of interesting questions and concentrate on purposeful ones.
- Demographic questions should usually be placed at the end of the questionnaire.
- Use familiar words and avoid technical jargon.
- The participant's level of education and knowledge should be considered when designing the questions.
- Leading questions should be avoided.
- Be careful of relying too heavily on recall or memory. People usually do not remember what happened in the past unless it was dramatic (Cooper and Schindler, 2003).
- A good mixture of multiple choice, open-ended, close- ended, opinion, rating and ranking questions should be included in the questionnaire.

3.5.3 Phase 3

During this phase the instrument is presented to the respondent and tested by conducting a test interview or pilot study. The following three concepts are important when conducting a test interview or pilot study (Cooper and Schindler, 2003):

1. *Introduction and screening:* It is important that the interviewer inform the respondent upfront about the topics that will be covered in the interview, as well as how long it will take to complete the interview. The interviewer should also state the name of the company for whom the interview is intended, as well as the purpose for which the information will be used.
2. *Sequencing and measurement questions:* Questions should be formulated and asked in such a way that they are interesting and motivate participation. Furthermore, personal questions should only be asked later in the interview. Simple questions should be asked first and then move on towards more difficult questions.
3. *Refining the instrument:* A pilot study can also be called a 'feasibility study'. The purpose of a pilot study is to resemble a smaller version of the full-scale study. During this phase in the research, the aim is to test the research instrument and smooth out any possible problems which were not considered during the planning phases (van Teijlingen and Hundley, 2001). The size of the pilot study can differ from data collection method to data collection method. Usually the size ranges from between 25 and 100 individuals (Cooper and Schindler, 2003). A pilot study also permits the researcher to refine and make changes before the full study is started.

3.6 Implementing the Survey: Conducting fieldwork

Fieldwork refers to the process that interviewers follow to conduct interviews. When planning the fieldwork, the following aspects should be considered: arranging the resources required, including recruiting and training of personnel; as well as preparing instructions; and planning visits to the various clusters and households (Audience Dialogue, 2006). Fieldwork usually takes place at respondent's homes, public places or even a call centre where they do telephone interviewing. The interviewer usually starts

off by establishing a connection with the respondent, after which the interviewer will start gathering technical information (Cooper and Schindler, 2003). Interviewers must be trained and well acquainted with the survey instrument. The survey must be presented exactly the way it has been drafted, this means that wording and sequences should not be altered. In some instances the respondent might need to be prompted. However, the respondent should not be led in any way (Cooper and Schindler, 2003). Although methods of recording the information differ, interviewers will mostly write down the answers that the participant provides. Answers should ideally be recorded immediately, and it could be useful to repeat the answers as they are written down, for accuracy and clarity purposes. Once all the information is accurately collected, the information will be converted into statistical data, after which analysis can start.

3.7 The LFS as data collection method in South Africa

The LFS represents one of the core household surveys conducted by Stats SA every year. Stats SA also conducts other surveys on an ad hoc basis or when commissioned by other departments. In February 2000 Stats SA launched the LFS, which is a biannual survey conducted in March and September of every year. A typical household survey randomly selects households from a list of all households in the population known as the 'sampling frame'. The sampling frame is often the most recent national census. Before revision in 2004 the LFS for 2000, 2001 and 2002 used the 1996 census as the sampling frame. In 2004, Stats SA rebased all the surveys from September 2000 on the 2001 census.

The LFS covers some areas previously covered by the OHS but not all, since it is a specialised survey designed to measure the dynamics in the labour market (Stats SA, 2004a). Until September 2005, a section that was designed to measure social indicators such as access to infrastructure was included in the LFS. However, it was later taken out after it was identified that there was a need for a more regular survey that was specifically designed to measure the level of development and performance of government. This gave rise to the General Household Survey (GHS). The first round of the GHS was conducted in July/August 2002 and is not part of the scope of this study.

3.7.1 First round of the LFS in February 2000

During the pilot round of the first LFS, fieldwork took place in February 2000, and a probability sample of 10 000 dwelling units was used (Stats SA, 2000a). Six months later the September 2000 survey was conducted, using a larger probability sample of 30 000 dwelling units. Among the 10 000 households visited in February 2000, approximately 40 per cent were re-visited during the September 2000 survey round. According to Stats SA (2000a), the fieldworkers had difficulty identifying certain dwelling units in the sample, particularly in those areas with a lack of visible addresses. Therefore, survey results were based on a cross-sectional analysis (Stats SA, 2000a).

3.7.2 LFS sample design

The master sample for the 2000 to 2003 LFS was based on the 1996 population census of enumeration areas (EA) and estimated number of dwelling units. In the master sample 3 000 primary sampling units (PSUs) were included. The PSU included either one EA or more in a case where the number of dwelling units within an EA was less than 100. Dwelling units such as prisons, hospitals, boarding houses, hotels, guest houses, schools and churches were excluded from the sample (Stats SA, 2003). The sample procedure for the master sample involved explicit stratification by province and within each province by urban and non-urban areas. Within each explicit stratum the PSUs were implicitly stratified by district council, magisterial district and within the magisterial district, by average household income or EA. In explicit stratification, the population is divided into strata and a separate sample is chosen from each stratum. In implicit stratification, the population is first sorted by a chosen characteristic. Next, the sample is selected from this sorted list using a random starting point and a fixed sampling interval (Sandene, et al., 2005). An independent sample of PSUs was drawn for each stratum within each province. The smaller provinces were given a disproportionately larger number of PSUs compared with bigger provinces. A simple random sample of 10 dwelling units was selected to visit in each PSU. If there were more than one household in a dwelling unit, all households were interviewed. A number of 3 000 households was drawn from the master sample by probability proportionate to size in each stratum.

During the September 2004 LFS round, the census frame that was used to draw the sample of PSU excluded EAs where the number of households exceeded 25. As in the previous rounds, worker's hostels, convents and monasteries were excluded from the sample. EAs in the census database were pooled in two stages: before and after sampling. The master sample is a multi-stage stratified sample with an overall sample size of PSUs of 3 000. Explicit strata were the 53 district councils. The 3 000 PSUs were allocated to them using the power allocation method. The PSUs were sampled using the same probability proportionate to size as previous survey years. Independent samples of PSUs were drawn for each stratum within each province. The sampled PSUs were listed with the dwelling unit as the listing unit. From these listings systematic samples of dwelling units per PSU were drawn. These samples of dwelling units formed clusters. The LFS used only the clusters that contained ten dwelling units.

3.7.3 LFS rotating panel methodology

Stats SA uses a rotating panel methodology ultimately to obtain a better picture of movements into and out of the labour market over time (Stats SA, 2004b). The rotating panel methodology involves visiting the same dwelling units on a number of occasions. Stats SA visits these dwellings at most five times. After the panel has been established, 20 per cent of the dwelling units are replaced in each round. New dwelling units are then selected to replace those that are taken out.

3.8 Conclusion

In this chapter the importance of a well-formulated research question was established. It was shown that a researcher should be well acquainted with the facts required to meet the purpose of the study and should know how the information will be used. After the research question has been defined, the researcher must design a sample from where the information will be collected. The process of collecting a sample can be divided into five

steps. The first step is to define the target population. It was concluded that there are no strict rules. However, the sample should be a good representation of the population. The second step is to choose a sampling frame using either probability sampling or non-probability sampling selection. During the third step of this process, the procedure for sampling unit selection is planned, and a power allocation method is usually applied to select sampling units. Step four determines the sample size, which mainly depends on the size of the sampling error that the researcher is willing to make. During step five, the population is subdivided into primary sampling units from which the sample will be selected.

The third part of this chapter looked extensively at data collection methods. The main methods that were distinguished were focus groups, interviews, monitoring, surveys and follow-up surveys/interviews. The strengths and weaknesses of the aforementioned methods were also discussed. The next section in this chapter looked at questionnaire design and testing (pilot study), which was divided into three phases. Phase 1 looks at the type of data required to answer the research question, while Phase 2 discusses the questionnaire design. During Phase 3 the instrument is tested, that is, a pilot study is done to determine the feasibility of the survey instrument. The next section looks at conducting fieldwork, where it was determined that interviewers need to be trained and well acquainted with the survey instrument. The last section builds on the foundation set by the first part of this chapter, and looks at the LFS and its design as a method of data collection in South Africa. It is concluded that the LFS uses a combination of the methods discussed in this chapter to collect accurate and useful information.

Chapter 4: Analysing LFS data according to the per capita expenditure method to identify the poor in South Africa

4.1 Introduction

> If Government is to address poverty and inequality, it requires reliable data on the extent and nature of the problem.
>
> — I. Woolard and M. Leibbrandt Measuring poverty in South Africa (p.4)

It is for this reason that the methods of identifying the poor are so relevant. As discussed in Chapter 2, poverty can be measured in many different ways; however, it can broadly be divided into monetary and non-monetary poverty. Non-monetary measures are often difficult to measure in value terms and therefore even though both measures are important for identifying the poor, this dissertation will only use a monetary value to define the poor (Woolard and Leibbrandt, 1999). This chapter starts off by considering the monetary measure of poverty, by making use of household survey data, more specifically the LFS. The LFS, amongst a large range of other information, also collects information on household's monetary income and expenditure. Household expenditure is often used instead of income as a poverty measure, because it is more reliable and stable (Ravallion, 1992).

In the LFS, expenditure data are collected for the household as a whole and not on an individual level. However, to calculate expenditure per person in each household, certain assumptions must be made about how expenditure is shared between members of the household. Household expenditure can either be divided equally among members of the household (i.e., per capita expenditure) or can be adjusted for structure and size of the household (i.e., per capita equivalent expenditure). These methods were discussed in detail in Chapter 2 and it was concluded that because of the complexity of latter method, and with little proof of better results, per capita household expenditure will be used (Pauw et al., 2006).

Furthermore, owing to the nature of the household expenditure question in the LFS it is difficult to provide a point estimate for household expenditure because results are provided in a range. To overcome this, a randomising method was used.

This chapter first give attention to household expenditure information as reported in the LFS. The second focal point is on how to create a per capita expenditure indicator and how the poor will be identified using this method. Thereafter the number of people who live in each quintile are identified, based on household per capita expenditure. This also provides information on how income is distributed between different population groups. Lastly, the LFS results will be analysed and the focus will be on how various demographic, labour market characteristics, unemployment and household livelihoods differ between expenditure groups. The groups are divided into 'poor' (being those in the bottom 20 and 40 per cent) and 'non-poor' (those in the top two to three quintiles).

4.2 Household expenditure information as reported in the LFS

In the LFS information is collected from people present in the household at the time of the survey. Individuals are asked personal and worker-related questions, and the main respondent or head of the household is asked household-related questions. In this dissertation the monetary measure used to distinguish between the poor and non-poor was a household-level variable where respondents were asked what their monthly expenditure level was. The question asked the respondent to report the household's total expenditure for the month (including everything on which the household and its members spent money: food, clothing, transport, rent and rates, alcohol and tobacco, school fees, entertainment and any other expenses). The options that were available in the questionnaire for the respondents to select from are shown in Table 4.1.

Table 4.1 Expenditure question in the LFS

7.30	**What was the total household expenditure in the last month?** *Include everything that the household and its members spent money on, including food, clothing, transport, rent and rates, alcohol and tobacco, school fees, entertainment and any other expenses.* 01 = R 0–R 399 02 = R 400–R 799 03 = R 800–R 1 199 04 = R 1 200–R 1 799 05 = R 1 800–R 2 499 06 = R 2 500–R 4 999 07 = R 5 000–R 9 999 08 = 10 000 or more 09 = Don't know 10 = Refuse	☐ 01 ☐ 02 ☐ 03 ☐ 04 ☐ 05 ☐ 06 ☐ 07 ☐ 08 ☐ 09 ☐ 10

Source: Stats SA, 2004a

The number of respondents for the periods 2001 and 2004, unweighted, is presented in Table 4.2. When analysing the results from the LFS it was found that there was a decrease in the number of respondents who reported an expenditure level of R0–R399 from 9 013 to 6 575 (Stats SA, 2001 and 2004). All other categories showed an increase in the numbers, apart from the 'don't know'; 'refuse' and 'unspecified'. A challenge in making use of these expenditure categories lies in the fact that since the introduction of the LFS in 2000, these categories have not been adjusted for inflation. The result is that the number of households in the lowest expenditure brackets decreased over time, not due to their moving out of poverty, but due to inflation increases. Households could have higher expenditure, without being better off in real terms (Casale and Desmond, 2007). It was also found that there was a decrease in the number of 'don't know' and 'unspecified', which shows that more valid expenditure information was captured in the survey.

Table 4.2 Un-weighted expenditure categories, 2001 and 2004

Category	Description: Rand value category	Number of respondents	
		2001	2004
1	R0–R399	9 013	6 575
2	R400–R799	7 545	8 998
3	R800–R1 199	3 332	4 187
4	R1 200–R1 799	1 953	2 469
5	R1 800–R2 499	1 396	1 719
6	R2 500–R4 999	1 815	2 116
7	R5 000–R9 999	997	1 344
8	R10 000 or more	319	553
9	Don't know	668	344
10	Refuse	220	204
99	Unspecified	57	29
	Total	27 315	28 538

Sources: Stats SA, 2001; 2004a. Own calculations.

The results, after applying Stats SA's population census weights, are presented in Table 4.3. The results show the same patterns as the unweighted results, as shown in Table 4.2. Of the 12 million households in 2004, 3,6 million were in the expenditure category R400–R799; and 2,7 million in the R0–R399 category. This is in line with the findings of Meth and Dias (2003), who showed that in 2002, 3,3 million people were in the expenditure category R0–R399, and 2,6 million in 1999. It is evident that the majority of households was found in the bottom two expenditure categories.

Table 4.3 Weighted expenditure categories, 2001 and 2004

Category	Description: Rand value category	Number of respondents	
		2001	2004
1	R0–R399	3 631 023	2 745 608
2	R400–R799	2 983 364	3 593 837
3	R800 – R1 199	1 324 803	1 769 209
4	R1 200–R1 799	819 522	1 063 532
5	R1 800–R2 499	619 654	827 486
6	R2 500–R4 999	924 224	1 039 197
7	R5 000–R9 999	525 752	826 836
8	R10 000 or more	194 367	386 936
9	Don't know	286 287	195 630
10	Refuse	102 715	122 563
99	Unspecified	28 895	17 126
	Total	11 440 606	12 587 961

Sources: Stats SA, 2001; 2004a. Own calculations.

4.3 Per capita expenditure measures

As discussed in Chapter 2, the following methods are examples of poverty measures: per capita expenditure, per capita equivalent expenditure, US$1 per day method and other non-monetary measures. In this study, as adopted from Pauw et al. (2006), an individual's welfare is measured by the annual per capita expenditure of the household in which he or she lives. The response option in the questionnaire is formulated in such a way that one can link individuals to households and, consequently, to a specific expenditure band. Analysis is made difficult because individuals in the household have a total household expenditure. No information about the distribution of data within the band can be obtained without making serious assumptions about the normality of the data (Meth and Dias, 2003). To work around this, the following procedure was employed using *Stata* software:

A random number between 0 and 1 was generated and then assigned to each household (Ardington et al., 2006). This value was then multiplied by the upper limit of the expenditure categories and the lower limit was added to this. The household expenditure was then divided by the number of people in the household. This dissertation gives equal weight to each individual in the household, whether it is a child or an adult (Woolard and Leibbrandt, 2006). This method thus draws on the assumption that household income or expenditure is shared equally between members (Pauw et al., 2006). Although some researchers prefer to adjust for household size and composition, Pauw et al. (2006), Hoogeveen and Özler (2004) and van der Berg et al. (2006) used capita poverty lines because of its simplistic nature. Furthermore, results do not differ much between the two methods. Income equivalent measures can be described as a weighting procedure, where adults will have higher weights than children (Deaton and Muelbauer, 1980). Age is sometimes also considered, because a seven-month-old child can demand many resources as implied by Budlender (2003). Owing to the complex nature of per capita equivalent expenditure measures, this dissertation only adjusted for household size (i.e., per capita household expenditure).

4.4 Survey results by per capita expenditure measures

4.4.1 Description of households' per capita expenditure quintiles

In 2001, as shown in Table 4.4, 28,8 per cent of the population was found within the bottom 20 per cent of households. By definition, each quintile contains 20 per cent of households. In 2001 12,5 million people lived in households that fell within the bottom 20 per cent of the per capita expenditure group. This number increased in 2004 to 13,3 million, or 29,4 per cent of the population. In the bottom 40 per cent of households in 2001, there were 11,2 million people or 25,8 per cent.

Table 4.4 Description of household per capita expenditure quintiles

	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Share of population (%)	28,8	29,4	25,8	25,0	18,6	18,1	14,2	14,4	12,6	13,2	100	100
Number '000	12 519	13 311	11 244	11 313	8 101	8 190	6 171	6 505	5 469	5 995	43 503	45 311
Weighted n=102 467												

Sources: Stats SA, 2001; 2004a. Own calculations.

The share of people in the bottom 40 per cent of households overall decreased between 2001 and 2004, from 25,8 per cent to 25,0 per cent. Together the poor (20th percentile and 40th percentile) have decreased from 54,6 percent to 54,4 per cent of the total population. This could be due to the roll-out of extensive grant policies by government since 2001 (Woolard, 2003). What is worrying is that more than 50 per cent of the population falls within the bottom 40 per cent of poor households. The share of population situated in the top 40 per cent of households amounted to 26,8 per cent. This is most likely a reflection of South Africa's very unequal distribution of income. Table 4.5 shows that in 2001 the Gini coefficient for the whole of South Africa was 0,64 and for certain population groups within South Africa it has been worsening since then (SAIRR, 2006). According to Global Insight, the income inequality has worsened for

black Africans and coloureds since 2001, from 0,60 to 0,64 and 0,54 to 0,55 respectively in 2004 (SAIRR, 2005).

Table 4.5 Gini coefficients for population groups, 2001–2004

Source and year	Black African	Coloured	Indian	White	Total
Global Insight 2001	0,60	0,54	0,51	0,46	0,64
SARPN 2001	0,72	0,60	0,64	0,61	0,77
Global Insight 2002	0,61	0,54	0,51	0,46	0,64
Global Insight 2003	0,62	0,55	0,51	0,45	0,64
Global Insight 2004	0,64	0,55	0,51	0,45	0,64

Sources: SAIRR, 2004; 2005; HSRC, 2004.

The Southern African Regional Poverty Network (SARPN) reports Gini coefficients for the various racial groups that are much higher than reported by Global Insight (HSRC, 2004). According to the SARPN report, the Gini coefficient in 2001 for black Africans was 0,72, for coloureds 0,60, for Indians 0,64 and for whites 0,61. This is much higher than Global Insight's reported statistics, but it should be noted that the SARPN data set consists of a combination of sources. Inequality has also changed in that it has become an intra-group challenge, making it more difficult to address policy (Bhorat, van der Berg and van Aardt, 2003).

4.4.2 Demographic characteristics according to per capita expenditure quintiles

Because the South African population was segregated in the Apartheid years, the results clearly show what impact this had on the race dimension of poverty. These results, as emphasised by Pauw et al. (2006) can be seen in how the poor are distributed among regions and concentrated in certain areas. Furthermore, evidence shows that children and women are highly represented among the poor, and the unequal development of economic centres in different regions. Table 4.6 describes how racial groups are distributed within the per capita expenditure quintiles.

4.4.2.1 Population group

Table 4.6 Population group according to per capita expenditure quintiles

Population Group	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Black African	95,4	96,8	90,9	92,0	83,4	83,1	63,2	65,7	31,6	32,9	79,5	80,2
Coloured	4,3	3,0	8,0	7,1	12,0	13,3	14,1	15,3	9,0	10,3	8,7	8,6
Indian/Asian	0,1	0,1	0,5	0,5	2,0	1,4	7,3	6,0	7,1	8,0	2,4	2,3
White	0,1	0,1	0,5	0,4	2,6	2,2	15,4	13,1	52,3	48,8	9,4	8,8
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

Based on the results, of the total number of people situated in the 20th quintile, shown in Table 4.6, black Africans made up the largest share at 95,4 percent in 2001, increasing to 96,8 per cent in 2004, showing how severe black Africans are affected by poverty. What is interesting is that at the 100th quintile the share of black Africans increased from 31,6 per cent in 2001 to 32,9 per cent. The share of whites in the 100th quintile in contrast decreased from 52,3 per cent to 48,8 per cent.

4.4.2.2 Gender

Table 4.7 Gender and population group according to per capita expenditure quintiles 2004

Population Group	Quintile									
	20th Gender		40th Gender		60th Gender		80th Gender		100th Gender	
	Males	Females	Males	Females	Males	Females	Males	Females	Males	Females
Black African	44,7	50,8	42,3	48,6	42,2	41,2	34,5	28,6	19,4	12,3
Coloured	2,0	2,3	3,6	4,4	5,5	6,4	6,9	7,2	4,6	4,5
Indian	0,0	0,0	0,2	0,3	1,0	0,9	3,6	3,7	3,6	3,5
White	0,1	0,0	0,3	0,3	1,3	1,3	7,9	7,5	26,9	25,4
Total	46,8	53,2	46,5	53,5	50,1	49,9	52,9	47,1	54,4	45,6

Sources: Stats SA, 2001; 2004a. Own calculations.

As shown in Table 4.7, poverty also has a gender dimension, which is not only linked to race, but to history as a whole; where women were excluded from economic activity until the passing and adoption of the amended Constitution of South Africa in 1996, which included the Bill of Rights (Parliamentary Monitoring Group, 2007). Women in South Africa were under the control of their husbands or fathers and were not seen as being equal to men. Black African woman were even worse off. They were disadvantaged in both racial and gender terms. In 2004, as shown in Table 4.7, the largest share of the poor comprises females, and the largest share of representation in the 100th quintile is males. More specifically, of the poor the largest share was black African females, in contrast to those in the 100th quintile, where the largest share was white males, followed by white females. There is a small share of black African females (12,3 per cent) and black African males (19,4 per cent) who are in the 100th quintile. It is clear that disadvantages are still present today, where women, and particularly black African women, battle to find high-paid jobs. They are also struggling to enter into the economically active population, and when they do enter, they are mostly unemployed. Women in South Africa also experience a high rate of victimisation. Rape and domestic violence are often problems that authorities have to face.

4.4.2.3 Province

Table 4.8 Province according to per capita expenditure quintiles

Province	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Western Cape	4,3	3,1	6,9	6,8	10,6	12,7	13,5	14,0	18,2	19,6	9,2	9,5
Eastern Cape	21,1	21,8	19,0	16,6	13,2	12,2	8,7	10,7	7,6	8,1	15,6	15,4
Northern Cape	1,7	1,7	1,9	2,3	2,3	2,3	1,9	1,8	2,3	1,6	2,0	2,0
Free State	6,1	4,8	5,9	6,9	7,1	7,6	7,1	7,0	8,1	7,1	6,6	6,5
KwaZulu-Natal	18,2	22,0	24,4	21,0	22,6	19,5	23,7	21,4	17,8	17,3	21,4	20,6
North West	10,4	9,5	7,6	8,8	7,3	8,4	6,7	6,2	5,6	6,2	8,1	8,2
Gauteng	10,0	10,5	13,1	15,1	20,0	20,7	26,4	27,0	32,6	31,3	17,8	18,6
Mpumalanga	8,1	8,6	7,1	7,1	7,4	7,0	6,3	5,6	4,8	4,8	7,0	7,0
Limpopo	20,1	18,1	14,0	15,4	9,6	9,6	5,8	6,3	3,0	4,1	12,4	12,3
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

In 2001, as shown in Table 4.8, the largest share of the poor was in the Eastern Cape at 21,1 per cent. In 2004, the largest share of the poor was found in KwaZulu Natal, at 22,0 per cent. Of those in the 100th quintile, 30 per cent of people were situated in Gauteng, whereas Limpopo Province had the smallest share of people in this category. This is again evidence of regional segregation, and infrastructure backlogs due to uneven distribution of economic activity. It also seems that in provinces where poverty is high there is a correlation with low infrastructure development (Wildeman, 2002). The opposition political party, Democratic Alliance, argues that these provinces are highly impacted by corruption and mismanagement, and incorrect allocation of funds (Democratic Alliance, 2005).

4.4.2.4 Household structure

Table 4.9 Demographic distribution according to per capita expenditure quintiles

Household Distribution	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Children	44,9	44,0	40,9	39,8	32,7	31,3	25,1	24,0	21,9	21,6	35,9	34,8
Female adults	22,1	22,1	23,2	24,4	29,9	30,1	36,7	36,6	39,1	38,2	28,0	28,3
Male adults	29,2	29,9	30,5	30,6	32,0	32,6	32,9	32,4	33,6	33,7	31,1	31,4
Elderly	3,9	4,0	5,3	5,1	5,4	6,0	5,3	7,0	5,4	6,5	4,9	5,4
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
Average number of children per household	3,1	3,5	2,7	2,8	2,0	2,2	1,8	2,0	1,7	1,8	2,4	2,5

Sources: Stats SA, 2001; 2004a. Own calculations.

In the 20th quintile, as shown in Table 4.9, 44 per cent of the people were children, followed by male adults and then female adults. There is a slight increase in the share of elderly in this quintile. Moving towards the 100th quintile, the share of children decreases and the share of female adults and elderly increases. The average number of children per household also decreases when moving towards the 100th quintile. In the 20th quintile there is an increase in the average number of children per household from 3,1 to 3,5.

4.4.2.5 Age cohorts

Table 4.10 Age cohorts according to per capita expenditure quintiles

Age cohorts	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
15–25	44,4	44,8	41,6	40,9	35,8	34,0	29,4	28,1	22,3	21,4	36,2	35,4
26–35	22,1	22,3	23,7	24,0	24,9	27,5	29,0	31,2	29,1	28,8	25,2	26,2
36–45	14,9	14,4	15,9	15,3	19,0	17,0	20,1	19,9	21,7	21,8	17,8	17,2
46–55	11,0	10,9	10,6	11,3	11,8	12,7	12,9	12,8	16,8	16,8	12,2	12,6
56–65	7,7	7,6	8,2	8,5	8,5	8,9	8,6	8,0	10,2	11,2	8,5	8,7

Sources: Stats SA, 2001; 2004a. Own calculations.

Of those in the 20th quintile, around 44 per cent were between the ages of 15 and 25, whereas in the 100th quintile this age group was only represented by around 21 per cent (see Table 4.10). In the 100th quintile, however, around 28 per cent was found in the age group 26 to 35 years. Young adults and children are mostly affected by poverty. In the category 15–25 years, people are most likely to start looking for work, or enter into the economically active population. However, considering the high unemployment rate and low labour absorption rate, these people had a 3,9 out of 10 chance of getting a job at the end of 2004 (Stats SA, 2004a).

4.4.3 Labour market characteristics according to per capita expenditure quintiles

Incomes are usually generated by jobs in the economy. Other income, such as government transfers, subsidies and remittances, often mean the difference between survival and starvation. The poor who are able to find work are usually in low-paid positions in the informal economy (Frye, 2006). There are many poor people who are not economically active. Reasons for inactiveness could be because they are discouraged or because of other infrastructure disadvantages such as problems with transport or health. There is also a large proportion of the poor who are economically active, but cannot find work.

Table 4.11 Labour market characteristics according to per capita expenditure quintiles

Labour market Characteristics	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Not economically active (expanded)	39,1	38,7	38,6	38,3	32,9	32,1	26,2	25,7	22,4	26,7	33,1	32,8
Not economically active (official)	56,2	60,5	53,2	55,8	41,7	44,5	31,9	33,3	24,3	25,5	43,8	46,4
Economically active (expanded)	60,9	61,3	61,4	61,7	67,1	67,9	73,8	74,3	77,6	89,2	66,9	67,2
Economically active (official)	43,8	39,5	46,8	44,2	58,3	55,5	68,1	66,7	75,7	74,5	56,2	53,6

Labour market Characteristics	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Unemployed (expanded)	38,8	40,7	34,3	34,1	26,9	28,3	19,7	18,3	6,8	19,0	27,4	27,8
Unemployed (official)	21,7	18,8	19,6	16,6	18,1	15,9	14,0	10,7	4,9	4,3	16,7	14,1
Employed	22,1	20,7	27,2	27,6	40,2	39,6	54,1	56,0	70,8	70,2	39,5	39,4
Number of working age people '000	6 831	7 370	6 402	6 581	5 245	5 391	4 461	4 648	4 098	4 481	27 037	28 473
Share of working age people	25,3	25,9	23,7	23,1	19,4	18,9	16,5	16,3	15,2	15,7	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations. Values may not add up due to rounding

In total, based on the official definition of unemployment, 46,4 per cent of people in South Africa was not economically active in 2004, as shown in Table 4.11. This was an increase from 43,8 per cent in 2001. This went hand in hand with a decrease in the share of people who were economically active in South Africa, based on the official definition of unemployment. The number of employed, based on the analysis, decreased from 39,5 per cent in 2001 to 39,4 per cent in 2004. The number in the working age population increased from 27 million in 2001 to 28 million in 2004. Of the working age population, people in the 20th and 40th quintiles made out around 50 per cent. Of the adult population in the 20th percentile, 56,2 per cent was not economically active and this increased to 60,5 per cent in 2004. In 2004, 18,8 per cent of the adult population in the 20th percentile was unemployed, which was a decrease from 21,7 per cent in 2001. The share of people employed in this group also decreased from 22,1 per cent in 2001 to 20,7 per cent in 2004.

In the 40th percentile 53,2 per cent of the working age population was not economically active in 2001. This share increased to 55,8 per cent in 2004. Of those in this group there has been a decrease in the share of unemployed, from 19,6 per cent to 16,6 per cent.

There is an increase in employment for this category from 27,2 per cent to 27,6 per cent. From the 60th quintile there is an increase in the share of people who are employed.

Table 4.12 Highest education level according to per capita expenditure quintiles

Highest level of Education	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
No schooling	11,9	10,4	10,0	9,0	7,5	7,2	4,4	3,8	1,1	1,2	7,7	7,0
Grade R/0	0,3	0,4	0,3	0,2	0,1	0,2	0,1	0,1	0,1	0,1	0,2	0,2
Grade 1	1,0	0,8	0,8	0,7	0,6	0,6	0,3	0,2	0,2	0,1	0,7	0,5
Grade 2	2,1	1,6	1,8	1,4	1,1	0,8	0,8	0,5	0,1	0,3	1,3	1,0
Grade 3	3,9	3,1	3,3	2,6	2,3	2,2	1,3	1,1	0,6	0,4	2,5	2,1
Grade 4	5,1	4,4	4,4	3,7	3,5	3,1	2,2	1,8	0,9	0,8	3,5	3,0
Grade 5	6,3	5,0	5,5	4,4	4,4	3,5	2,8	2,5	0,7	0,8	4,3	3,5
Grade 6	8,7	7,7	6,9	7,2	6,0	5,7	3,9	3,5	1,5	1,1	5,9	5,5
Grade 7	12,0	10,1	10,6	9,2	10,1	8,7	6,4	6,1	2,6	2,0	9,0	7,7
Grade 8	11,1	11,8	12,1	11,6	11,2	11,2	9,0	8,4	4,1	3,6	10,0	9,8
Grade 9	9,8	11,0	9,7	12,2	9,4	10,2	7,9	8,7	4,6	4,1	8,6	9,7
Grade 10	9,3	11,8	10,7	11,7	11,2	12,6	12,9	13,5	9,7	8,6	10,6	11,7
Grade 11	8,3	9,5	8,9	9,8	10,1	10,9	9,3	10,3	5,7	5,3	8,6	9,3
Grade 12	8,8	10,9	12,1	14,1	17,5	19,0	26,8	29,1	35,0	39,0	18,2	20,6
NTC I	0,1	0,1	0,1	0,1	0,1	0,1	0,3	0,1	0,3	0,2	0,2	0,1
NTC II	0,0	0,1	0,0	0,1	0,1	0,1	0,1	0,1	0,3	0,3	0,1	0,1
NTCIII	0,1	0,1	0,1	0,1	0,3	0,3	0,6	0,6	1,5	1,6	0,4	0,5
Diploma/certificate with less than Grade 12	0,2	0,2	0,4	0,3	0,5	0,3	1,2	0,6	2,1	1,5	0,8	0,5
Diploma/certificate with Grade 12	0,7	0,8	1,6	1,3	2,9	2,6	6,4	6,5	14,3	14,1	4,4	4,3
Degree	0,1	0,1	0,3	0,3	0,7	0,4	2,0	1,9	8,3	10,3	1,8	2,1
Postgraduate degree	0,1	0,0	0,2	0,0	0,4	0,1	1,1	0,5	6,3	4,6	1,3	0,8
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

There has been an overall increase in the share of workers who had a Grade 12 and those who had a degree, as shown in Table 4.12. Of those in the 20th percentile, there has been a decrease in the share of people with no schooling. There has also been an increase from

8,8 to 10,9 per cent in the share of people who completed secondary education. The same holds true for those in the 40th quintile. Less than 1 per cent of people in the 20th, 40th and 60th quintile had a degree or a postgraduate degree. In 2004, in the 100th quintile, 10,3 per cent of workers had a degree, an increase from 2001 when the share was 8,3 per cent. Those in this quintile who had a post graduate degree decreased from 6,3 per cent in 2001 to 4,6 per cent in 2004. Having access to infrastructure plays an important role in education delivery, in that it provides a more conducive environment for children to be educated. In South Africa many schools, specifically previously classified township/rural schools are not performing as well as formerly seen white model C schools (Bloch, 2007). This is in terms of results as well as resources available. Schools in poor areas (usually rural), as reported by Wildeman (2002), often do not have textbooks, classrooms, telephones, water, electricity and toilets. This affects the quality of education and, in return, often leads to failing students, who fall into the vicious cycle of unemployment or low-level employment and ultimately poverty.

Table 4.13 Employment by sector according to per capita expenditure quintiles

	Quintile											
	20th		40th		60th		80th		100th		Total	
Sector	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Formal	39,3	40,9	52,7	54,1	66,6	66,2	79,8	79,0	91,6	93,5	70,3	71,4
Informal	45,1	46,1	34,0	33,9	23,5	23,7	14,0	13,3	6,9	5,4	21,5	20,8
Domestic workers	15,6	13,0	13,3	11,9	9,9	10,1	6,2	7,6	1,5	1,2	8,1	7,7
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

The informal economy plays an important role in South Africa. "Given the structural changes in the economy, growth of the informal sector cannot be ignored in evaluating the size of the labour market" (Bhorat et al., 2003). However, as reported by May (2000), those in the informal sector tend to remain in poverty, because of the subsistence nature of informal sector jobs. Table 4.13 shows that in the poorest quintile nearly half of the employed are in the informal sector. Around 40 per cent of the poor are in formal jobs. Moving towards the highest quintile, there is an increase in the share of people who are employed in the formal sector, where job pay is relatively higher.

Table 4.14 Employment by industry according to per capita expenditure quintiles

Industry	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Agriculture, hunting forestry and fishing	22,6	19,3	16,0	14,7	12,7	11,4	6,7	6,5	3,5	2,5	10,8	9,4
Mining and quarrying	1,9	1,6	2,6	2,0	4,0	2,1	6,8	4,0	7,6	6,0	5,1	3,6
Manufacturing	9,9	12,1	13,9	13,6	16,6	15,6	14,5	15,9	16,0	14,2	14,6	14,5
Electricity gas and water supply	0,3	0,5	0,8	0,5	0,7	0,4	1,0	1,2	1,1	1,3	0,8	0,9
Construction	8,0	10,4	7,9	9,0	6,4	8,9	5,4	6,6	3,2	3,8	5,8	7,1
Wholesale and retail trade	26,7	24,8	25,0	25,3	22,9	23,2	21,7	20,7	16,6	18,5	21,8	21,9
Transport, storage and communication	2,5	3,0	4,2	3,9	5,2	4,5	5,5	5,4	5,9	5,7	4,9	4,7
Financial intermediation, insurance, real estate and business services	2,4	3,1	3,7	4,4	5,5	7,0	10,1	9,9	17,5	17,8	9,0	9,8
Community social and personal services	6,8	8,9	11,4	11,8	14,5	14,4	21,0	21,1	26,9	28,7	17,8	18,8
Private households	18,8	16,4	14,6	14,8	11,5	12,3	7,3	8,8	1,7	1,4	9,4	9,4
Exterritorial organisations and foreign government	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

In South Africa the agricultural industry lost its importance in the job creation process between 2001 and 2004, as shown in Table 4.14. In 2001 22,6 per cent of the people in the 20th quintile relied on this industry for jobs, the share decreased to 19,3 per cent in 2004. Economically, the largest contributor to gross domestic product (GDP) is the

financial intermediation, insurance, real estate and business services industry. The industries where the share of employment of the poor increased were manufacturing and construction. Although wholesale and retail trade was the largest industry where the poor was employed, this share of employment decreased from 26,7 per cent in 2001 to 24,8 per cent in 2004. Manufacturing, wholesale and retail trade are two industries which are known as trade sectors. These two sectors are relient on external factors such as the stability of the rand and exports, which means that businesses can easily let go of their workforce in times of crises (Edwards and Stern, 2006). There has been an increase in the share of workers in the community, social and personal services industry, from 6,8 per cent to 8,9 per cent on the back of growing economic activity in this industry by the time of this survey. In the 40th and 60th quintile there has been a decrease in the share of workers employed in manufacturing from 13,9 per cent in 2001 to 13,6 per cent in 2004, and 15,6 in 2001 to 15,6 per cent in 2004. In the 100th quintile the community, social and personal services industry employed most of the workforce, and the share increased from 26,9 per cent in 2001 to 28,7 per cent in 2004. The second largest sector in this category was wholesale and retail trade, which employed 16,6 per cent in 2001 and 18,5 per cent in 2004. Another important sector that increased with the quintiles was the financial intermediation, insurance, real estate and business services.

Table 4.15 Employment by occupation according to per capita expenditure quintiles

Occupation	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Legislators, senior officials and managers	0,6	0,8	1,2	1,1	2,4	2,7	4,3	5,3	14,4	18,8	5,6	7,3
Professionals	0,2	0,2	0,6	0,4	1,4	0,7	3,0	2,6	11,8	10,6	4,3	3,8
Technical and associate professionals	2,6	3,0	4,7	4,3	6,7	6,3	12,4	9,8	18,6	18,8	10,3	9,8
Clerks	1,9	2,6	4,6	4,6	6,5	7,4	12,9	11,6	16,1	16,7	9,6	9,9
Service workers and shop and market sales workers	11,0	10,4	14,0	13,4	13,5	13,8	15,2	15,4	10,5	9,8	12,8	12,5
Skilled agricultural and fishery workers	9,9	8,3	5,9	5,6	5,3	2,3	2,8	1,0	2,4	0,7	4,7	2,9
Craft and related trades workers	14,6	16,7	14,8	14,6	14,4	14,0	14,3	14,3	11,9	9,9	13,8	13,4
Plant and machine operators and assemblers	8,2	7,0	11,3	10,7	13,4	11,1	11,8	12,4	7,3	7,4	10,3	9,7
Elementary occupation	35,6	38,1	29,9	33,4	26,7	31,6	17,2	20,1	5,5	6,2	20,5	23,0
Domestic workers	15,3	12,9	13,1	11,8	9,8	10,1	6,1	7,6	1,5	1,2	8,0	7,7
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

The largest share, as shown in Table 4.15, was in elementary occupations in 2004 at 23,0 per cent, which was an increase from 20,5 per cent in 2001. Within the 20th, 40th, 60th and 80th quintile the largest share of people was employed in this occupation. Elementary occupations and domestic workers are associated with low skills and low pay. More than 50 per cent of the poor and, according to Pauw et al. (2006), more than 60 per cent of the poor are employed in these low-skilled jobs. Less than 1 per cent of the poor are legislators, senior officials and managers or in professional occupations. Within the 100th quintile, in 2004 most people were employed as legislators, senior officials and managers (18,8 per cent), as technical and associate professionals (18,8 per cent) and clerks (16,7 per cent).

4.4.4 Unemployment characteristics according to per capita expenditure quintiles

Unemployment significantly impacts on poverty (May, 2000). Unemployment links on to various other factors such as education, health, access to information, transport and job related accessories, such as books, CVs and clothing. Unemployment is also more prominent among black Africans, and females, as well as the youth. Reasons for not working are usually related to the area where people live, their skills and their physical abilities, as well as the willingness to really find a job, and make the effort to keep it (Dias, 2002).

Table 4.16 Reasons for not working according to per capita expenditure quintiles

Reasons for not Working	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Has found a job, but is only starting at a definite date in future	0,6	0,4	0,8	0,4	0,8	0,9	2,0	1,2	4,3	3,5	1,1	0,7
Lack of skills or qualifications for available jobs	47,7	9,8	46,1	8,4	36,1	8,4	29,0	7,2	25,0	11,5	41,2	8,9
Scholar or student and prefers not to work	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Housewife/homemaker and prefers not to work	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Retired and prefers not to seek formal work	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Illness, invalid, disabled or unable to work (handicapped)	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Cannot find any work	0,0	85,3	0,0	84,7	0,0	83,0	0,0	82,4	0,0	63,1	0,0	83,3
Too young or too old to work	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Seasonal worker, e.g., fruit picker, wool-shearer	1,1	0,5	1,2	0,3	1,3	0,5	0,9	0,4	0,0	0,5	1,1	0,4
Cannot find suitable work (salary, location of work or conditions not satisfactory)	47,6	2,5	47,8	3,4	57,4	3,0	59,4	4,8	60,9	15,7	52,0	3,7
Contract worker, e.g., mine worker resting according to contract	0,4	0,1	0,4	0,1	0,7	0,2	1,3	0,6	0,4	0,3	0,6	0,2
Recently retrenched	2,6	1,6	3,7	2,7	3,8	4,0	7,4	3,4	9,4	5,4	4,1	2,8
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

In the 2001 questionnaire the question on reasons for not working did not include the category 'cannot find any work', whereas in 2004 this category was included as an option for reasons for not working. In 2001, as shown in Table 4.16, half of the respondents in the 20th quintile answered that they were not working due to lack of skills or qualifications for available jobs, and the other half responded that they 'cannot find suitable work (salary, location of work or conditions not satisfactory)'. In 2004, however, since the new category had been introduced, 85,3 per cent of respondents selected the easy answer by simply stating they 'cannot find any work'. When considering only 2001 along quintiles, the share of people in each quintile is more likely, as the quintiles increase, to select that they 'cannot find suitable work (salary, location of work or conditions not satisfactory)'. This suggests that the higher the quintile, the more room there is for respondents to be selective when considering employment, whereas the bottom quintile seems not to be able to find work largely due to lack of skills. The poor also lack the skills to network and look for work through connections. Therefore, it is important to analyse how the poor search for work, and what do they see as the reason for them being unemployed. It is expected that the method of job search should reflect the types of jobs for which the poor are looking.

Table 4.17 Activities carried out to find work according to per capita expenditure quintiles

Activities carried out to find work	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Waited/Registered at employment agency/trade union	4,9	5,1	6,7	4,6	7,7	5,9	8,9	7,5	10,0	7,9	6,8	5,6
Enquired at workplaces, farms, factories or called on other possible employers	63,2	58,0	59,1	61,0	59,0	59,1	52,3	50,5	41,7	46,9	58,7	57,6
Placed/Answered advertisement(s)	7,6	8,3	14,1	11,1	17,5	13,7	22,2	22,0	42,5	33,5	15,0	13,1
Sought assistance from relatives or friends	11,1	14,5	8,4	14,3	8,7	13,1	6,4	13,6	3,1	9,2	8,9	13,7
Looked for land, building, equipment or applied for permit to start own business or farming	0,4	0,8	0,3	0,6	0,3	0,5	0,8	1,4	0,4	1,1	0,4	0,8
Sought/Underwent training	0,2	0,0	0,1	0,0	0,0	0,0	0,1	0,0	0,4	0,0	0,1	0,0
Waited at the streetside where casual workers are found	12,6	13,3	11,2	8,5	6,8	7,7	9,3	5,1	1,8	1,3	10,1	9,2
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

Overall, most unemployed people enquired at workplaces, farms, factories or called on other possible employers for work. The share of unemployed who placed/answered advertisements decreased from 15,0 per cent in 2001 to 13,1 per cent in 2004, as shown in Table 4.17. There was also an increase in the share of unemployed persons who sought assistance from relatives or friends, from 8,9 per cent in 2001 to 13,7 per cent in 2004. The largest increase in this category was in the 80th quintile. Where the share of unemployed seeking assistance increased from 6,4 percent in 2001 to 13,6 percent in 2004. Of those in the 20th quintile, there was also a relatively large group who waited at the streetside where casual workers are found.

South Africa's unemployment is characterised by structural unemployment (Hirano, 2004). The longer people are unemployed the less likely they are to find a job and the more likely that they will become discouraged (Nattrass, 2002).

Table 4.18 Length of unemployment according to per capita expenditure quintiles

Length of unemployment	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Less than a month	6,1	8,1	6,2	8,2	6,3	6,7	6,5	7,6	7,2	8,9	6,3	7,8
One month to less than two months	4,5	5,3	4,2	4,6	5,7	4,6	3,8	3,9	2,9	7,3	4,5	4,9
Two months to less than three months	4,5	4,8	3,2	3,6	3,6	4,5	2,9	3,4	3,6	4,4	3,7	4,2
Three months to less than four months	3,2	4,0	3,5	3,3	3,9	2,8	3,1	3,0	2,5	4,5	3,4	3,5
Four months to less than six months	4,2	3,7	3,5	4,1	3,1	4,1	5,1	3,7	6,0	5,4	4,0	4,0
Six months to less than one year	11,8	13,2	11,8	10,0	10,5	11,9	13,7	12,2	15,1	10,7	11,9	11,8
One year to less than three years	26,5	21,3	26,8	22,7	27,4	25,7	30,3	29,6	34,2	29,4	27,6	24,0
Three years or more	39,3	39,6	40,9	43,5	39,5	39,8	34,8	36,5	28,5	29,4	38,7	39,8
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

The results in Table 4.18 show that of the people in South Africa that were unemployed 39,8 per cent had been trying to find work for three years or more. Around 39,6 per cent of people in the 20th quintile fell within this category and 21,3 per cent in the category: one year to less than three years. Of those in the 40th quintile, 43,5 per cent of the people had been trying to find work for three years or more, and 22,7 per cent one year to less than three years. Of the unemployed in the 100th quintile, 29,4 per cent had been trying to find work for three years or more, and the same share was trying to find work for one year to less than three years.

4.4.5 Household livelihood characteristics according to per capita expenditure quintiles

The environment and realities that the poor live in and face every day can broadly be considered as a livelihood (May, 2000). Households who struggle to provide food face this challenge as part of their reality. Similarly, income sources represent realities and means of living. For households where income depends on the sales of farm produce, their livelihood is affected should they not be able to sell those products, due to drought or other reasons.

Table 4.19 Share of people who struggled to satisfy food needs in the past 12 months according to per capita expenditure quintile

Struggle to satisfy food needs	**Quintile**											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Never	21,4	34,8	31,8	55,5	47,6	63,0	66,7	79,8	86,3	94,0	24,1	58,0
Seldom	8,5	10,1	11,3	1,2	10,5	10,3	10,6	6,9	5,5	2,6	5,3	8,8
Sometimes	42,6	37,0	39,8	33,3	32,1	21,5	17,8	11,5	6,5	2,9	17,6	24,3
Often	15,2	10,4	10,5	6,4	6,7	3,7	3,3	1,5	1,4	0,2	9,0	5,4
Always	12,3	7,7	6,5	3,6	3,1	1,5	1,7	0,3	0,2	0,3	3,4	3,4
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

What is very promising, as shown in Table 4.19, is that the share of people who were often experiencing problems satisfying their food needs decreased from 9,0 per cent in 2001 to 5,4 per cent in 2004. There is still, however, a large share of people within the bottom two quintiles who sometimes or often have problems satisfying their food needs.

Table 4.20 Receiving of welfare grants according to per capita expenditure quintile

Welfare Grants	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Old age pension	28,0	28,0	34,5	30,0	23,3	22,5	14,7	15,1	7,2	7,0	24,3	22,9
Disability grant	6,0	10,2	7,3	14,2	6,1	11,7	4,7	7,2	1,3	1,8	5,6	9,9
Child support grant	13,9	56,8	11,9	46,6	6,5	30,7	2,4	11,5	0,6	1,5	8,7	3,6
Care dependency grant	0,7	0,4	0,5	0,8	0,5	0,4	0,2	0,1	0,1	0,0	0,5	0,4
Foster care grant	0,5	0,5	0,3	0,7	0,2	0,5	0,3	0,5	0,3	0,0	0,3	0,5
Grant in aid	0,2	0,1	0,2	0,2	0,4	0,1	0,1	0,1	0,1	0,0	0,2	0,1
Social relief	0,1	0,2	0,2	0,2	0,2	0,1	0,1	0,1	0,1	0,0	0,1	0,1
Total	35,8	75,2	35,7	61,7	17,4	31,8	8,0	13,3	3,0	3,6	39,7	37,5

Sources: Stats SA, 2001; 2004a. Own calculations.

Based on the results shown in Table 4.20, of the total share of people who answered the question most people in the 20th quintile in 2004 received child support grants.

Table 4.21 Main income source according to per capita expenditure quintile

Main income Source	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Salaries and/or wages	34,4	28,4	42,6	40,6	63,8	59,5	80,8	79,3	87,0	86,5	55,1	52,1
Remittances	23,6	15,3	18,3	13,7	11,9	9,8	5,9	4,2	1,7	1,8	14,7	10,5
Pensions and grants	30,9	47,0	30,6	37,5	18,0	23,3	8,6	9,9	4,9	4,2	22,1	29,3
Sales of farm products	1,9	1,5	1,1	0,8	1,1	0,4	0,8	0,7	1,5	1,2	1,3	1,0
Other non-farm income	5,9	6,0	5,6	6,7	4,0	6,0	3,1	5,7	4,2	6,1	4,8	6,1
No income	3,4	1,8	1,7	0,7	1,2	1,0	0,9	0,3	0,7	0,2	1,9	0,9
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

In 2001 the main source of income for the poor was mostly salaries and wages, but in 2004 most of those in the 20th quintile were dependent on pensions and grants for their income, as shown in Table 4.21. This increase was from 30,9 per cent in 2001 to 47,0 per cent in 2004. The share of the poor who had no income decreased from 3,4 percent to 1,8 per cent in 2004. The poor is also highly dependent on remittances as a source of income, although the share of those in the 20th quintile decreased to 15,3 per cent in 2004. It is also shown that the poor are reliant on a combination of sources for income, namely salaries and wages, remittances, and pensions and grants (Woolard and Leibbrandt, 1999).

4.5 Conclusion

This chapter used information collected by the LFS to identify people living in poverty in South Africa. Monetary and non-monetary measures can be used to distinguish between the poor and the non-poor. This chapter focused on the monetary method of identifying the poor. To measure monetary poverty, the household expenditure of each household was used, adjusting for household size (i.e., per capita household expenditure). Households were then divided into quintiles, where as generally used, the lowest 20 and 40 per cent represent the poor (Leibbrandt and Woolard, 2006). The expenditure distribution for the poor and non-poor was compared, followed by an analysis of demographic, labour market characteristics, unemployment and household livelihoods.

Based on data from household expenditure quintiles, around 12,5 million people lived in households that fell within the bottom 20 per cent of the per capita expenditure group. This number increased in 2004 to 13,3 million. In the bottom 40 per cent the share of poor households decreased from 54,6 per cent to 54,4 per cent between 2001 and 2004, which could possibly be attributed to the roll-out of extensive grant policies by government since 2001. Furthermore, in terms of inequality, South Africa has a very uneven distribution of income in that over 50 per cent of the population fall in the lowest 40 per cent of poor households and 26,8 per cent of the population fall within the top 40 per cent. In 2004 South Africa's Gini coefficient was 0,64, unchanged from 2001.

However, evidence shows that within population groups inequality became worse between 2001 and 2004.

When demographic variables are analysed by household per capita expenditure quintiles, the results show that black African females make out the largest part of the poor 20th and 40th quintile. Among the provinces, Eastern Cape and KwaZulu-Natal are the provinces where the most people in the 20th and 40th quintile live, while Gauteng has the most people in the 100th quintile.

The average number of children in poor households is twice that of households in the 100th percentile. Of the poor, around 40 per cent were children in the 20th and 40th percentile, while children made out only 20 per cent of the households in the 100th percentile. Furthermore, poverty affects the youth more. Of the poor, most people were aged between the ages of 15 and 25 years.

Across household expenditure per capita quintiles, most of the poor people were employed in the informal sector. Within industries, most of the poor were employed in the categories wholesale and retail trade sales, followed by agriculture, hunting, forestry and fishing. Within the 100th percentile, the largest share of people was employed in community, social and personal services.

Of the poor, most had less than a matric as their highest qualification, whereas of the people in the 100th percentile, around 35–40 per cent had matric as their highest qualification. A high percentage of these also had some kind of postgraduate degree. Most of the not-economically active population falls within the bottom 20 per cent and 40 per cent, which indicates a link between unemployment and poverty. The poor are mostly unemployed for periods longer than three years, while persons in the 100th percentile are mostly unemployed for less than three years. In 2004, the poor were highly dependent on grants, especially the child support grant, as a means of income. The share of the poor who indicated salaries and wages as a main source of income declined between 2001 and 2004. Activities carried out to find work includes enquiries at workplaces, farms and

factories, or called on other possible employers, whereas a large share of the persons in the 100th percentile, apart from utilising the above method, also had the privilege of placing and answering advertisements.

Poverty has a multidimensional face, and it is clear that the poor can be characterised by selected characteristics such as unemployment, race, gender, age as well as education. It is clear that poverty in itself is a big challenge in South Africa. The next section will examine to what extent the poor have access to infrastructure, and whether there is a difference between the poor and non-poor when it comes to having access to basic infrastructure.

Chapter 5: Infrastructure development and the poor

5.1 Introduction

Social exclusion captures the essence of the multidimensional part of poverty, and includes exclusion from social networks and basic services, which in turn "deny people the opportunities available to others to increase their income and escape from poverty by their own efforts" (DFID, 2005). Without access to basic infrastructure, communities are excluded from a normal functioning society. The lack of service delivery is rapidly becoming as important a social challenge as poverty and unemployment in South Africa. In fact, the linkages between poverty, unemployment and service delivery are being highlighted more, after what can be seen as mass action taken by communities demanding what they proclaim as their right to better service delivery, as promised by government (Friedman, 2006). Apart from the pressures coming from civil society, the functioning of the economy as a whole is dependent on not only available, but reliable infrastructure and service delivery.

Government and municipalities in partnership with business are responsible for the organising, planning and managing of infrastructure delivery. South Africa is characterised by large inequalities and imbalances, largely as a result of apartheid and the rapid rising of informal settlements outside big towns (May, 2000). Owing to past land segregation and marginalisation, South Africa largely carries the scars of spatial inequality and pockets of economic activity, clustered unevenly across space. These pockets of economic activity result in uneven and lumpy development, drawing the poor closer to city centres where the hope of finding jobs is higher, and leaving the poor behind in mostly undeveloped areas where economic activity is low or even non-existent (urbanisation). Many of these poor job seekers end up in the large informal areas around the economic hubs, leading to high population density in certain provinces as well as additional pressure on service delivery. Large numbers of the poor are also found in provinces with a strong rural attribute (Heintz and Jardine, 1998).

Political agencies seems to come up with plans that look feasible on paper, but implementation, coordination, monitoring and evaluation seem to hamper the success of these plans (Goldman, Marumo and Toner, 2002). In many instances the money for projects are there, but the organisation of the projects fails due to the stalling of necessary documentation, lack of communication or due to corrupt agents. There is thus a necessity to revisit some policies and to identify gaps to establish new creative ways of addressing poverty in an ever-changing society.

This chapter will, firstly, focus on the problems of defining and measuring infrastructure development. As mentioned in Chapter 2, poverty can also be measured in non-monetary terms, which in this chapter relates to access to services and basic infrastructure. This relationship is not always so clear. Challenges arise when one has to determine what can be seen as access to "adequate" or non-poor infrastructure. The question that follows this is whether the poor necessarily benefit from infrastructure and service delivery, and then to what extent the poor have access to the following infrastructure: housing, water and sanitation, electricity, communication, financial and transport infrastructure. The provision of these services to the poor are then analysed by relying on information collected by the LFS. Lastly, the focus turns to the policy and financial environment, and how this environment impacts on the delivery of these services.

5.2 Problems with defining and measuring infrastructure development

'Infrastructure' can have different definitions; depending on the context in which it is investigated. Infrastructure is not only defined in terms of physical facilities, but also in terms of the services and potential benefit that can be gained from having access to these services (Masika and Baden, 1997). Physical infrastructure refers, among other things, to roads, airports, railways, water and sewage pipes, electricity cables, house structures, communication systems, and any other primary components. Services and potential benefits that accrue from utilising physical infrastructure can be associated with the secondary effect that results from these physical goods being utilised (Jerome and Ariyo, 2004).

Most literature on poverty and infrastructure, such as Bhorat et al. (2006), Aliber (2002), and Woolard and Leibbrandt (1999), tend to focus on the number or share of households with access to service delivery. This method, however, fails to consider to what extent infrastructure delivery affects the poor.

The different infrastructure categories that are considered in this dissertation are access to housing, electricity, water, roads, sanitation, financial and communication infrastructure. One of the challenges in assessing infrastructure delivery is how to differentiate between formal and informal infrastructure. As a guide, Statistics South Africa's general household survey methodology was used to determine what infrastructure can be classified as informal (Stats SA, 2004a).

Infrastructure delivery can also be measured in terms of a monetary value. The amount of finance that government invests in infrastructure can be seen as a measure for infrastructure development. This measure is, however, problematic because funds may be allocated through the infrastructure budget, but true expenditure may fall short. The quality of the infrastructure delivery is also not measured and therefore relying on the monetary value of infrastructure investment might be a misleading indicator to measure the level of infrastructure development in a country (Klitgaard, 2004). It is also challenging to determine whether infrastructure investment is indeed taking place on a local level to improve the lives of those who do not have access to basic infrastructure, or whether the infrastructure that is being delivered is on a macro-scale, only benefiting the big companies in terms of electricity and transport.

5.3 Does infrastructure development benefit the poor?

In the broader workings of everyday life, infrastructure is a channel through which development and economic growth are made possible in an economy. It is expected that the accrued potential benefits from this development and economic growth would trickle down to the poor in the long run. This, however, rarely happens, and in some instances,

infrastructure investment actually has a negative impact on the poor and on development in their communities (Masika and Baden, 1997).

Government has achieved numerous successes over the past decade in providing services. The number of houses delivered has increased and the number of households with sanitation and access to water has improved. The question that needs to be asked is more: who is the target to whom government is delivering these services? Has government successfully targeted the poorest communities across all spheres and provinces who are in dire need of services or are those who benefit from service delivery the ones who already live close to a city, who have a job, or who have an informal house with gas energy? It is therefore crucial to investigate how infrastructure development links to the poor and to determine whether the current infrastructure investment is adequate to alleviate poverty or whether infrastructure investment in its current form and rate is actually not benefiting the poor at all (Hemson and Owusu-Ampomah, 2005).

Infrastructure investment links to development and social enhancement as a result of the potential benefits. But do these potential benefits ever reach the poor? Some evidence shows that certain types of infrastructure provide the economic conditions needed for development, such as transport, electricity and communication, and that the poor are benefited through industrial growth and increased economic activity created by these economic conditions which, in turn, create employment and income-generating opportunities. However, these opportunities do not always trickle down to the poor, and therefore the poor in many instances gain little if nothing from improved infrastructure and infrastructure investment (Hemson and Owusu-Ampomah, 2005).

Goldblatt and Davies (2002) investigated linkages between the macroeconomy and water infrastructure investment as well as energy supply and the environment. They emphasise that macroeconomic reforms need to be compatible with sustainable development, and that there is "inadequate" information and comprehension of the impact of certain economic development choices. The research further showed that there are a number of unsustainable elements in the structure of the South African economy, and that policies

lack dynamic and long-term planning. The authors predict that the implication of not addressing these obstacles will be the underperforming of economic instruments. In 2005/06 the first signs of underperformance were widely experienced in the power industry, when Eskom failed to reach its transmission reliability target where national power failures left certain parts of provinces without electricity for a number of days (Newbery and Eberhard, 2007).

Set against the backdrop of current policy debates, such as appropriateness of infrastructure, financing costs, community participation, ownership and management, it becomes clear that infrastructure provision does not always translate into "benefiting the poor". For example, in Bangladesh over 80 per cent of public expenditure on infrastructure ends up in the pockets of the non-poor communities (Masika and Baden, 1997). Furthermore, in Brazil, an evaluation of the first phase of an electrification project called PRODEEM revealed how the lack of skills, funds for maintenance and lack of responsibility resulted in the project not benefiting the poor (Goldemberg, La Rovere, Coelho, 2004). Another way in which the poor do not benefit from infrastructure investment is when there is a lack of local community empowerment, as shown by Brook and Smith (2001). ICRW (2005) and Marc (2005) revealed that cultural and gender considerations played an important role in determining whether infrastructure delivery benefits poor communities. In South Africa, little research has been done on how infrastructure investment links to the poor and in which way, if any, they have benefited.

5.4 Access to infrastructure and the poor

The following section will provide information on the access to infrastructure, and on provision of services to the poor. The policy and finance environment in which these services are delivered will also be discussed. In this dissertation, 'infrastructure' refers to access to traditional infrastructure (Hassen, 2000):

5.4.1 Access to housing

Section 26 of the Constitution makes provision for every South African to have access to adequate housing (Republic of South Africa, 1996). More importantly, the Constitution mandates that government is to provide adequate housing progressively for all citizens. It states that:

> '(1) Everyone has the right to have access to adequate housing. (2) The state must take reasonable legislative and other measures, within its available resources, to achieve the progressive realisation of this right. (3) No one may be evicted from their home, or have their home demolished, without an order of court made after considering all the relevant circumstances. No legislation may permit arbitrary evictions' (Republic of South Africa, 1996).

The 3,3 million households who do not have access to formal housing in South Africa is largely blamed on the apartheid governments' refusal to provide housing in most black African urban communities (Cosatu, SACC and SANGOCO, 2006). The Reconstruction and Development Programme (RDP) aimed to change this by arguing that decent, affordable housing was a national priority. The expectations of the RDP were that the provision of infrastructure and housing to the poor will address asset poverty, giving families the basis for increased incomes as well as improving living conditions which, in turn, could be used for income-generating activities (Cosatu, SACC and SANGOCO, 2005). Before the shift to free basic services, poor households had to pass a means test to prove that their incomes were below R800 per month, which meant that many households could not get their housing subsidies. Even after the shift to free basic services, households received housing free of charge, but still had to pay for services. These services could cost anything from about R100 to R150 per month, resulting in many poor households having either to discontinue the use of the services or having to take on debt (Cosatu, SACC and SANGOCO, 2005).

Apart from these constraints at a national level, provincial housing delivery is just as complex. At a provincial level a major challenge is that resources are not equally distributed, and provinces run their municipalities and service delivery in different ways. Shortages in staff and skills are also obstacles impacting on service delivery. Demographic structures within provinces are also very different and some provinces have overcrowding while others have low density of population, making service delivery difficult due to remote access (ANC, 1997).

5.4.1.1 Housing provision and the poor

Even with the delivery of over 2,4 million houses over the past 12 years, current estimates put the housing backlog in South Africa still at 2,2 million households (Social Housing Foundation, 2007). It is also the first time that the backlog is less than the number of houses produced. It is estimated that subsidised housing delivery made in-roads into the backlog only in 1997, with the delivery of over 300 000 subsidised houses. Since 1994 rates of new construction have not kept pace with the growth in new families and the flow of foreigners into South Africa (Gardner, 2003). The housing backlog is still mainly located at the lower end of the income spectrum, although an important part of the backlog exists between the subsidy eligibility income cut-off (i.e., household incomes of above R3 500 per month) up to those households able to afford and access bond finance (i.e., household incomes of above R6 000 per month).

Housing infrastructure benefits the poor in that it provides a safe home for families, as well as other possible economic benefits of running a business from home (Social Housing Foundation, 2005). Housing further, in providing the poor with an asset of which they can take ownership, improves their self-worth. Households also have to start paying for services connected to these houses, such as electricity and water. Because of corruption, the benefits from housing delivery do not always reach the poor. Many of the new housing developments are on the outskirts of urban areas and far from employment opportunities. These houses are also small and sometimes poor in quality, reflecting deep cuts in the housing budget in the late 1990s (van der Westhuizen, 2004).

Having access to formal housing not only raises morale, but improves the standard of living of the poor. This, in turn, can lead to better health and safety (Cosatu, SACC and SANGOCO, 2005). Poor households' access to formal housing between 2001 and 2004 is shown in Table 5.1.

Table 5.1 Type of dwelling according to per capita expenditure quintiles

Type of dwelling unit	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Dwelling or brick structure on a separate stand or yard or on farm	55,6	54,1	59,4	59,4	63,6	63,6	68,8	66,0	77,8	75,9	62,7	61,5
Traditional dwelling/Hut/Structure made of traditional materials	28,5	29,2	23,9	20,6	12,6	10,9	4,3	0,5	1,0	0,5	17,5	16,3
Flat in a block of flats	0,7	0,5	1,2	0,8	2,7	2,9	6,9	8,7	9,3	9,3	3,1	3,2
Town/Cluster/Semi-detached house (simplex, duplex or triplex)	0,5	0,5	1,3	1,6	1,9	3,4	3,6	6,2	4,9	6,7	2,0	2,6
Unit in retirement village	0,0	0,0	0,0	0,0	0,1	0,0	0,1	0,3	0,2	0,3	0,1	0,1
Dwelling/Flat/Room in backyard	1,1	0,6	1,6	1,1	2,7	0,2	4,1	4,1	2,2	1,6	2,1	1,6
Informal dwelling/Shack in backyard	3,3	2,7	2,7	2,7	4,0	3,5	2,9	2,7	0,7	1,0	2,9	2,6
Informal dwelling/Shack not in backyard, e.g. in an informal/squatter settlement or on farm	9,6	11,4	9,0	12,1	10,4	12,7	6,4	8,2	1,3	1,9	8,1	10,1
Room/Flatlet	0,7	1,0	1,0	1,7	1,9	2,7	2,9	3,4	2,6	2,8	1,6	2,1
Caravan/Tent	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,1	0,0	0,0
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA 2001; 2004a. Own calculations.

The largest share of people stayed in dwellings or brick structures on a separate stand or yard or on a farm. The total share in this quintile decreased from 62,7 per cent in 2001 to 61,5 per cent in 2004. The highest share by quintile was the 100th quintile where 75,9 per cent of people stayed in brick structures on a separate stand or yard or on a farm. This was a decrease from 77,8 percent in 2001. There was an increase in households who lived

in an informal dwelling/shack not in a backyard, from 8,1 per cent in 2001 to 10,1 percent in 2004. Interestingly, in the 60th, 80th and 100th percentile there have been an increase in households living in an informal dwelling/shack not in a backyard.

5.4.1.2 Policy and Finance

South Africa's current housing policy is rooted in the *Housing White Paper*, published by government in December 1994, and has remained mainly unchanged ever since (Gardner, 2003). The Constitution as mentioned earlier was implemented in 1996 giving each citizen the right to housing. The Housing Act provides information on the roles and functions of the different spheres of government, together with government's functions in terms of facilitation, allocation and providing for the poor (Republic of South Africa, 1997a).

In 2000 the national housing code set out to fulfil the Constitution in meeting housing rights (Gardner, 2003). It clearly sets out the national housing policy in South Africa, which is formulated by the national Department of Housing. The National Housing Policy is implemented primarily by the three spheres of government namely (1) national government, (2) provincial government and (3) municipalities. Provincial government and departments have to approve housing subsidies and projects, and provide support to municipalities in terms of housing development. Local government is responsible for regulating local-level housing planning, development, funding and maintenance. During the 2007/08 Budget government allocated a further R2,7 billion over the medium term to housing, making the total allocation over the three-year period R32 billion (National Treasury, 2007).

5.4.2 Access to Water and Sanitation

Clean water is a significant resource for reducing poverty, disease and improving the life of the poor in South Africa (Mukheibir and Sparks, 2003). Until November 1997 the Department of Water Affairs and Forestry (DWAF) was the water services authority in South Africa, although water boards as well as some large cities themselves fulfilled the role of water service provider (Naraghi and Kebotlhale, 2004). However, these services were mostly limited to historically advantaged groups, and services in the rural areas, as was the case with housing, were largely underdeveloped or ignored.

A large percentage of those without access to clean water live in the historically disadvantaged rural areas, specifically in the previously demarcated homelands (Mukheibir and Sparks, 2004). In September 2000 national government developed a policy framework to bring about free basic services to the poor, which included 6 000 litres of free basic water per month (Naraghi and Kebotlhale, 2004).

Water is very critical for the development of the agricultural sector. This sector includes commercial, emerging and subsistence farming as well as other economic activities (National Treasury, 2003). As a water-poor country, South Africa has to manage its water resources carefully if it is to meet social and economic needs (National Treasury, 2003).

Sanitation also falls under the water services function of municipalities, and it has been under-prioritised over the last few years. Recent outbreaks of cholera have emphasised the need for prioritising sanitation (Hemson and Dube, 2004). This is particularly relevant in the context of the large number of unplanned settlements and overcrowding. Delivery of services is held up by factors such as lack of capacity and skills at the local government level, and the low priority that rural communities give to sanitation (DWAF, 2002).

5.4.2.1 Water and sanitation provision and the poor

There is clear evidence from the 1998 demographics and health survey that both water and sanitation have an acute effect on child mortality rates (Hemson and Dube, 2004). Hemson and Dube (2004) found that households that did not have piped water had a child mortality rate twice as high compared to those who had piped water, and those households that did not have flush toilets had a child mortality rate four times higher than those who did.

During 2005 and 2006 water was provided to approximately an additional 1 million people and sanitation services were provided to approximately 200 000 households (DWAF, 2006). Furthermore, a total of 5,06 million households lived in provinces where there was no refuse removal by local authorities. This was an increase from 2001, where 5 million households had no refuse removal by local authorities (Stats SA, 2001 and 2004).

Water and sanitation delivery face many challenges. In urban areas, municipalities choose to provide 6 kilolitres (kℓ) of free water per household per month because it was costing them more to bill low consumption clients than the consumers owed them (Grice, 2004). In smaller areas or communities where the number of rich and poor is not so well balanced the challenges are greater, because it is more expensive for municipalities to supply water (Grice, 2004).

Another challenge is the number of free services to be delivered. Groups, such as South African Municipal Workers Union (Samwu) have criticised government's 6 kℓ of free water. The World Health Organization (WHO) describes 'basic access' to water as one where consumers live at least 1 kilometre (km) or 30 minutes from the water supply, and therefore will not be likely to exceed around 20ℓ per capita per day. When comparing the kind of access that township residents have to the WHOs definition of intermediate

access (e.g., having at least a tap on site) the average water consumption would be around 50 kℓ per capita per month (Grice, 2004).

The absence of portable water and sanitation services makes people vulnerable to poor health, which reduces their quality of life and productive capacity, and burdens health care and social welfare services. Provision of dependable water supplies can have a strong positive effect on food security and income generation for rural women and livelihoods can be improved by releasing labour time spent on obtaining water. Water can also be used for small farming and other enterprises (National Treasury, 2003).

Within the past ten years, official records indicate that 9 million additional people have gained access to safe drinking water, more or less 3,7 million additional households were served between 1995 and 2003 (Hemson and Owusu-Ampomah, 2005). The proportion of people who have access to clean water has risen from 60 per cent in 1996 to 85 per cent in 2001. The number of households that have access to sanitation has also risen from 49 per cent in 1994 to 63 per cent in 2003. In 1993, 58,9 per cent of the national population had a piped water connection to their yards or houses; this had risen to 63,7 per cent in 2004. This is an improvement; however, there is still a great need for access to water (Hemson and Owusu-Ampomah, 2005).

Having access to water has many additional benefits. Having access to clean water not only improves health, but can also save time in not having to walk long distances to fetch water. Another benefit from having access to safe water is that people do not have to live close to rivers and dams, which prevents disasters from happening in the case of heavy rains and flooding (National Treasury, 2003).

Table 5.2 Water source according to per capita expenditure quintiles

	Quintile											
	20th		40th		60th		80th		100th		Total	
Water source	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Piped (Tap) water in dwelling	11,8	9,7	18,2	19,7	34,2	35,1	59,7	60,0	85,9	87,3	33,8	34,4
Piped (Tap) water on site or in yard	28,5	30,8	29,3	34,2	32,0	33,1	23,8	25,1	10,1	9,1	26,4	28,4
Neighbour's tap	4,0	1,0	3,5	1,5	2,7	1,5	1,6	1,5	0,2	0,8	2,8	1,2
Borehole on site	1,1	0,4	1,1	0,4	1,0	0,8	1,2	0,3	1,1	0,1	1,1	0,4
Rain water tank on site	0,5	3,2	0,5	2,9	0,5	2,9	0,3	0,9	0,2	0,2	0,4	2,3
Public tap	23,5	28,6	20,0	22,3	13,9	16,9	7,2	8,8	1,3	1,8	15,7	18,5
Water carrier/Tanker	2,3	1,8	1,4	1,3	1,3	1,3	0,6	0,4	0,1	0,1	1,4	1,1
Borehole off site/communal	4,7	4,6	4,7	3,1	3,7	1,8	1,4	0,8	0,3	0,0	3,5	2,6
Flowing water/Stream/River	11,0	12,4	12,1	8,4	6,6	4,4	2,2	1,3	0,2	0,2	7,8	6,7
Dam/Pool/Stagnant water	2,0	1,1	1,4	0,8	1,0	0,4	0,5	0,1	0,1	0,0	1,2	0,6
Well	5,0	2,5	3,3	2,2	1,7	0,7	0,8	0,2	0,2	0,1	2,7	1,5
Spring	5,7	3,9	4,5	3,2	1,5	1,3	0,8	0,5	0,2	0,2	3,2	2,3
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a.Own calculations.

As shown in Table 5.2, most of the people who are in the 20th percentile had access to piped water on site or in their yard. Around 28,6 per cent of people in the 20th quintile in 2004 made use of a public tap to access water. Overall, there has been an increase in the share of people in South Africa who have access to public taps and piped water on site or in a yard. There has also been an increase in the share of people who had piped water in their dwellings. There is still 6,7 per cent of people in South Africa who depend on a flowing stream or river as their source of water.

Most people without access to basic water supply and sanitation services live in rural areas, which is why the Community Water Supply and Sanitation Programme (CWSS) of the DWAF focuses on these areas. The CWSS aims to provide water supplies to 90 per

cent of the currently non-services population by 2004, and since its inception in 1994 provided basic water supply to 1,3 million people (May, 2000).

Having access to toilet facilities is an obvious necessity. It contributes to health and sanitation and improves people's self-respect. The share of the poor who have access to toilet facilities in dwellings or on site is shown in Table 5.3.

Table 5.3 Type of toilet facilities according to per capita expenditure quintiles

Type of toilet facility	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Toilet facility in dwelling	8,6	8,4	14,2	18,1	30,3	32,5	57,3	58,2	86,4	87,9	30,8	32,8
Toilet facility on site	63,7	68,2	66,3	66,6	57,1	56,9	37,4	37,2	12,6	11,1	53,0	53,8
Toilet facility off site	27,6	23,5	19,5	15,3	12,6	10,6	5,4	4,6	1,0	1,0	16,2	13,4

Sources: Stats SA, 2001; 2004a. Own calculations.

Of those in the 20th quintile, 68 per cent in 2004 had a toilet facility on site and a mere 8,4 per cent had a toilet facility in the dwelling. This share also decreased from 8,6 per cent in 2001. Of those in the 100th quintile 11 per cent had a toilet facility on site and 87 per cent had a toilet facility in the dwelling. In total, 53,8 per cent of the population had access to toilet facilities on site and 32,8 per cent to toilet facilities in the dwelling.

5.4.2.2 Policy and finance

The South African water sector is complex, with a wide range of organisations contributing to the development of water resources, and policy and legislation between water resources and water services. The responsibility for water services rests primarily with municipalities as established under the Water Services Act, 1997 (Republic of South Africa, 1997b). The actual act of service delivery is undertaken by the water services providers appointed by the water services authority (in case of South Africa, the municipalities). These service providers can be the municipalities themselves or other municipal entities such as water boards or contracted service providers. The DWAF

performs functions in both water resources, such as dams, rivers and water services, which are mainly bulk water services. There are also 15 water boards, that have been established in terms of the Water Services Act (National Treasury, 2003).

South Africa has a good policy framework for sanitation delivery; the problem lies in interpretation and implementation of these policies (WRC, 2003). In March 2006, 60 per cent of the backlog in water services and 40 per cent in sanitation had been eradicated. Since the basic services programme began in 1994; 10 million households had access to water services by the end of 2006. The National Water Resource Infrastructure Agency Limited Bill was passed at the start of 2008. Its purpose is to centralise the development and operation of the national water infrastructure to broaden access to water and sanitation services (Republic of South Africa, 2008).

Additional funding of R1,4 billion for bulk water services infrastructure over the Medium-Term Expenditure Framework (MTEF) period was proposed by government. This includes R300 million in 2007/08, R450 million in 2008/09 and R650 million in 2009/10 (PMG, 2007).

5.4.3 Access to Electricity

Access to electricity is important in the fight against poverty and a prerequisite for sustainable development. Providing electricity to rural and semi-urban areas in developing countries contributes to reducing poverty by enabling increased productivity, increased incomes and local economic growth (Nordström et al. 2004). Access to energy services also enables social development in that streets and neighbourhoods can be safer after dark, since there is adequate lighting. It also gives children the opportunity to do homework and learn in the evenings. Women and children also do not have to collect wood, which is time consuming, for cooking and heating (Nordström et al. 2004).

Emphasis on the productive use of energy services is important in helping people out of poverty. At a national level, energy drives economic development by serving as a launch

pad for industrial growth, and through transport and communications provides access to international markets and trade (Ashong, 2007) Given the state of technological advancement in an economy, capital and labour work together in converting, directing and amplifying energy to produce goods and services needed for growth and poverty reduction (Ashong, 2007). Reliable, efficient and cheap energy supplies also attract foreign companies for investment, which is important for economic growth. At a local level, energy facilitates economic development by improving productivity and enabling local income generation through improved agricultural development and through non-farm employment, including small businesses.

South Africa supplies two-thirds of Africa's electricity and is one of the four cheapest electricity producers in the world (SAIRR, 2006). Around 90 per cent of electricity in South Africa is generated by coal-fired power stations. Electricity generation is currently dominated by Eskom, the national wholly state-owned utility, which also owns and operates the national electricity grid. Eskom currently supplies about 95 per cent of South Africa's electricity (GCIS, 2007).

5.4.3.1 Electricity provision and the poor

Policy-makers and practitioners involved in the roll-outs of electricity services became aware of the difficulties in achieving sustainability (Hemson, 2004). Firstly, new connections to the electricity grid showed low usage, which translated into higher operating costs. Electrification of rural communities has been regarded as predominantly challenging due to the high cost of service payments and extensive distances covered by feed lines. Other challenges regarding the electrification programme is the lack of investment required to maintain and extend the networks. Although the average tariff applied across the country appears low by international comparisons there are significantly different tariff levels and service standards. Municipalities also levy a wide range of domestic tariffs, which bear no relationship to quality or cost (Hemson, 2004).

The following challenges further slowed down the roll-out of electricity services to the poor (Lloyd, Cowan, Mohlakoana, 2004):

- Households already possess appliances suited for cooking or heating with solid or liquid fuels, but not for cooking or heating with electricity
- Households cannot afford electrical appliances and believe they cannot afford the running cost of cooking or heating with electricity
- Households currently use solid or liquid fuel cookers or heaters for multiple purpose (e.g., simultaneous cooking and space-and-water heating), which are not readily achieved with electrical appliances
- Households are familiar with traditional fuels
- Households have access to cheap sources of energy, such as fuel wood or coal, which they employ in quantities that cannot be matched by electricity.

The real cost of using collected fuel wood is difficult to balance for households who have little money. It is important for low-income households to reduce their risks by managing their low income in a flexible way. They would thus rather choose a flexible hand-to-mouth approach, which allows for cutbacks when the need arises and a greater ability to give or receive exchanges from neighbours.

To alleviate this burden on the poor, government allocated 50 kiloWatt/hour (kWh) free electricity per month to poor households. According to the Department of Minerals and Energy (DME), this value does adequately reflects basic electricity needs (DME, 2003). However, electricity has transformed lives, especially that of women who used to cook using liquid or solid fuels (Lloyd, Cowan and Mohlakoana, 2004).

Energy and poverty reduction are linked in the following ways (Ashong, 2007):

- *Productivity:* Less time can be spent on fetching wood and more time on productive, income-generating work.
- *Income growth:* Energy consumption shows a strong correlation with national income and economic growth. Economic growth creates employment and raises incomes, which would not be possible without energy.
- *Health and education:* Electricity makes it possible to refrigerate vaccines, operate medical equipment and provide lighting for after-dark operations. In homes, electricity helps to improve children's education when they can study at night.
- *Gender dimension:* Modern cooking fuels free women and children from the burden of collecting and carrying large loads of firewood over long distances.
- *Social impact of energy extraction:* oil, gas and coal-mining operations should help reduce poverty by creating jobs and raising incomes.

Overall, there has been a decrease in the number of households with no electricity from 3,4 million in 2001 to 2,9 million in 2004. However, the total number of electrical connections made to households also decreased over the same period from 336 000 to 246 000 (The Presidency, 2007).

For those who have electricity, it is unimaginable to be without it. The poor who either cannot afford electricity or who do not have any access to it, have to make use of other means to generate energy, such as coal, gas or wood. These sources can be dangerous, especially when gas or candles are used in small shacks where a fire can easily be started. The percentage of the poor who relies on these sources of electricity for cooking, heating and lighting is shown in Table 5.4. Access to electricity for lighting is usually used for measuring poverty.

Table 5.4 Electricity source for cooking according to per capita expenditure quintiles

Electricity source used for cooking	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Electricity from mains	22,37	25,69	36,13	40,38	55,70	59,95	78,69	81,21	95,12	95,84	49,18	52,75
Electricity from generator	0,01	0,03	0,05	0,03	0,01	0,11	0,04	0,00	0,03	0,04	0,03	0,04
Gas	1,31	1,03	1,72	1,83	2,65	1,93	2,80	1,95	1,54	1,30	1,90	1,56
Paraffin	19,65	21,37	18,80	21,82	18,15	19,13	10,26	10,35	2,31	1,76	15,66	16,92
Wood	50,86	45,89	37,58	31,02	19,25	16,06	6,28	4,90	0,71	0,77	28,97	24,96
Coal	4,89	4,74	5,05	4,17	3,97	2,50	1,85	1,42	0,24	0,28	3,75	3,13
Candles	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Animal dung	0,91	1,23	0,67	0,76	0,27	0,31	0,03	0,17	0,00	0,00	0,49	0,63
Solar energy	0,00	0,02	0,00	0,00	0,00	0,01	0,05	0,00	0,05	0,01	0,01	0,01
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Sources: Stats SA, 2001; 2004a. Own calculations.

For cooking it seems that most of the people in the 20th quintile used wood; 45,9 per cent in 2004. Whereas 25,7 per cent used electricity from mains. This was an increase from the 22,4 per cent in 2001. In the 100th quintile more than 95 per cent of people used electricity from mains to cook. Overall there has been an increase in the share of people in South Africa who use electricity from mains to cook; from 49,2 per cent in 2001 to 52,7 per cent in 2004.

Table 5.5 Electricity source for heating according to per capita expenditure quintiles

Electricity source used for heating	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Electricity from mains	20,2	20,1	31,6	33,4	52,0	53,3	77,4	78,4	93,9	92,8	47,0	47,7
Electricity from generator	0,0	0,0	0,1	0,1	0,0	0,2	0,0	0,0	0,0	0,0	0,0	0,1
Gas	0,3	0,2	0,4	0,6	0,8	0,5	0,8	0,9	0,9	1,7	0,6	0,7
Paraffin	12,3	14,3	12,7	16,6	14,4	17,0	8,9	9,3	2,5	2,7	11,0	13,0
Wood	57,1	56,2	45,4	40,9	24,5	22,6	9,0	7,7	1,7	1,8	33,8	32,0
Coal	8,7	7,7	8,8	7,6	7,9	5,8	3,7	3,5	0,9	0,8	6,8	5,8

Electricity source used for heating	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Candles	0,2	0,1	0,1	0,1	0,1	0,2	0,1	0,0	0,0	0,0	0,1	0,1
Animal dung	1,1	1,3	0,8	0,7	0,3	0,4	0,1	0,1	0,0	0,0	0,6	0,7
Solar energy	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Source: Stats SA, 2001; 2004a. Own calculations.

As shown in Table 5.5, in 2004 56,2 per cent of people in the 20th quintile used wood for heating and 20 per cent used electricity from mains. Compared with the 100th quintile where more than 90 per cent used electricity from mains for heating. Overall, the total share of people using electricity from mains for heating remained unchanged from 2001 to 2004.

Table 5.6 Electricity source for lighting according to per capita expenditure quintiles

Electricity source used for lighting	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Electricity from mains	54,0	64,6	63,2	72,5	75,6	80,7	88,3	89,7	97,6	98,2	70,7	77,6
Electricity from generator	0,1	0,1	0,1	0,0	0,0	0,2	0,1	0,0	0,0	0,1	0,1	0,1
Gas	0,3	0,1	0,4	0,3	0,4	0,1	0,0	0,2	0,0	0,0	0,3	0,2
Paraffin	10,1	6,6	8,3	4,9	5,8	4,1	2,5	2,0	0,6	0,4	6,6	4,2
Wood	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Coal	0,0	0,3	0,0	0,1	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,1
Candles	35,4	28,1	27,9	22,1	18,0	14,8	8,9	8,0	1,7	1,2	22,2	17,7
Animal dung	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Solar energy	0,1	0,2	0,1	0,2	0,1	0,1	0,1	0,0	0,1	0,1	0,1	0,1
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Source: Stats SA, 2001; 2004a. Own calculations.

The share of people in the 20th quintile who used electricity from mains for lighting has increased from 54 per cent in 2001 to 64,6 per cent in 2004, as shown in Table 5.6. There

has been an increase in the share of people using electricity from mains for lighting across all quintiles. In total 77,6 per cent of households in South Africa used electricity from mains for lighting in 2004.

5.4.3.2 Policy and Finance

Electricity, as a key strategic sector in the economy, underpins growth and development. The DME has several policies in place to ensure adequate supply of electricity-generation capacity (GCIS, 2007). The Integrated National Electrification Programme (INEP) provides a socio-economic support system to ensure that previously unconnected households have access to electricity in South Africa (DME, 2007). While Eskom does not have exclusive generation rights, it practically has a monopoly on bulk electricity in South Africa (DME, 1998). In 1998 government released the *Energy White Paper*, which aims to restructure the electricity supply industry in an attempt to alleviate poverty, and better economic growth and redistribution of assets.

Since the inception of the INEP, Eskom has electrified around 3 million homes. In July 2003 the National Electricity Basic Services Support Tariff Policy was gazetted. This policy aims, through government intervention, to bring relief to low-income households and to address the electrification backlog by 2012. Basic energy provision consists of liquid fuels and electricity. In terms of liquid fuel, government approved a zero-value added tax (VAT) on paraffin, which was implemented in April 2001. In terms of electricity, customers who qualify (i.e., those who earn less than the national minimum wage and who are connected to the national electricity grid) are eligible for 50 kWh of free electricity per month. Eskom determined that 56 per cent of households in South Africa who are connected to the national grid (in Eskom licence areas) consume on average less than 50 kWh of electricity per month (DME, 1998). This adds up to the same number of people who fall in the last two quintiles of the population and who are classified as poor. Not all households connected to the national electricity grid have electricity.

Approximately 30 per cent of existing households in South Africa currently do not have access to grid electricity (DME, 1998). Although a National Electrification Programme (NEP) is rolling out electrification connections, it will take at least ten years before the existing backlog will be eliminated; perhaps even longer given the current rate of growth of households (Knight, 2004). Apart from grid-connected electrification, non-grid electrification is usually found in remote rural areas, where households use solar home systems (SHSs), which is funded through the National Electrification Fund (NEF). This service however, will cost households around R58 per month, of which a free basic subsidy of R48 per SHS will be provided. These are 2003/04 figures, and are revised from time to time (GCIS, 2007).

In South Africa municipalities are responsible by law for providing infrastructure and access to public services such as water, sanitation, telecommunications and electricity. While few municipal distributors generate the electricity they sell, the majority purchases it at wholesale prices from Eskom (Philpott and Clarke, 2002).

Through the MTEF, the Department of Minerals and Energy plans to accelerate electrification and the rehabilitation of electricity infrastructure leading to 2010. The department intend to electrify 150 000 grid-connected households and 12 000 non-grid households. Additional funds of R45 million, R90 million and R150 million over the MTEF will be allocated for reducing backlogs and for the provision of electricity to schools and clinics. Government will further allocate an additional R102 million in 2007/08, R230 million in 2008/09 and R668,8 million in 2009/10 to the INEP for the rehabilitation of infrastructure and the expansion of the electrification programme (National Treasury, 2007).

5.4.4 Communication

The importance of communication development in the infrastructure debate is often neglected. Communication – more importantly, reliable and easily accessible communication – provide access to information, employment opportunities, education

and health facilities; which, in turn, can improve productivity and social networks for the poor (May, 2000). Furthermore, the South African postal service is aiming at a long-term scenario in which each postal outlet will offer a full range of services.

5.4.4.1 Communication provision and the poor

Research has shown that information and communications technologies (ICTs) do have an impact on economic growth and social development (Moshapo and Hanrahan, 2003). Telephone services are probably the most important service that are provided to, and utilised by, the poor. The reason for this is that it is easy to use and it saves people from travelling long distances. It also eases the burden of communication for those who are not literate. In cases where these facilities are operated by members of the community, it serves as an income-generating process (Moshapo and Hanrahan, 2003).

Other ICT provision such as electronic-mail, facsimile and Internet browsing are currently available in poor communities; however, they are more readily available in areas where economic activity is relatively high. Because these services depend on expensive hardware such as computers, which needs to be upgraded frequently, they add extra cost to the poor. For many in poor communities, computer literacy is even more challenging than written literacy. It can also not be clearly determined whether the benefit of a computer or cellular telephone is equally shared between household members (HSRC, 2003).

It is not clear to what extent the importance of ICT in developing communities impacts on education, health and business services. As the world moves towards a global information village, it is sure to become an important question. ICT can help learners with their studies, giving them access to knowledge and business opportunities. In the health sector ICT can save lives by providing equipment and knowledge to assist in the early identification of diseases (Moshapo and Hanrahan, 2003). Access to telecommunications improves the linkages between people and communities, which enhance networking and can ultimately lead to the creation of opportunities.

Table 5.7 Telecommunication facilities according to per capita expenditure quintiles

Telecommunication facility	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Have a fixed telephone in the dwelling (percentage who answered yes to the question)	6,7	4,5	13,5	9,1	22,4	15,8	38,4	33,8	64,4	60,7	23,1	19,3
Own a cellular telephone (percentage who answered yes to the question)	12,3	40,8	17,8	47,1	27,6	48,6	43,6	62,9	67,2	83,7	27,9	52,6

Sources: Stats SA, 2001; 2004a. Own calculations.

As shown in Table 5.7, the share of people in South Africa who had a fixed telephone line in their dwelling decreased from 23,1 per cent in 2001 to 19,3 per cent in 2004 in total. Whereas those who owned a cellular telephone increased from 27,9 per cent in 2001 to 52,6 per cent in 2004. The largest increase in people who owned a cellular telephone took place within the bottom two quintiles.

5.4.4.2 Policy and finance

The cost of communication is identified as an impediment to growth and investment in South Africa. The Electronic Communications Act, 2005 and the Independent Communications Authority of South Africa Amendment Act, 2006 have been promulgated and should ensure reliable and cost-effective access to communication systems (Republic of South Africa, 2006). Expenditure on digitising infrastructure is expected to amount to R100 million in 2006/07, increasing to R125 million in 2007/08 and decreasing to R75 million in 2009/10 as projects come into completion (National Treasury, 2008).

5.4.5 Financial infrastructure

The strength and weaknesses of South Africa's financial sector reflect the dualistic nature of the economy. An advanced financial industry exists that serves the needs of large and medium enterprises, government and upper-income households, while the same industry fails to assist small businesses and poor households.

5.4.5.1 Financial infrastructure provision and the poor

When consulting literature on the South African banking sector, it is clear that the banking sector is one of the more advanced sectors in the world, providing a vast range of financial instruments from savings to credit lines, to advanced secondary financial instruments and markets. It is unfortunate that a large part of the South African population does not share in this advanced financial system. In 2003 there were 17 million unbanked individuals in South Africa (Porteous, 2003).

Banks have not had much success in providing the poor with access to basic banking services. In many instances the poor cannot afford the high cost associated with opening and maintaining a bank account. Owing to the lack of banking facilities in remote areas, people often have to travel long distances to make cash withdrawals (Nigrini, 2007). Many of the poor also do not have the knowledge or training to make informed decisions about their finances, and also lack training in operating infrastructure such as automated teller machines (ATMs) (Gush, Cambridge and Smith, 2004).

Apart from the four dominating commercial banks, there are alternative methods of banking available. Poor communities make use of informal financial structures, such as credit and savings associations (Coetzee and Cross, 2002). As financial instruments become more refined and the world move towards a moneyless economy where wealth is stored on cards, it becomes more necessary for individuals to have access to these services. Not only will it ensure that the poor are not left behind in the era of

modernisation, but it is also safer and more profitable to put money in to a savings account.

The National Credit Act (NCA) came into effect in June 2007. The NCA was introduced in an attempt to control credit extended to uninformed consumers who do not have the money to repay debt (Republic of South Africa, 2005). It is also difficult for banks to extend large loans, such as home loans to poor people, because in many cases these loans cannot be repaid. Already in 2004, South Africa introduced *Mzansi*, a low-cost national bank account which was specifically aimed at making banking more available to the poor (SAinfo, 2006).

The percentage of South Africans holding such an account increased from 2 per cent in 2005 to 6 per cent in 2006. The percentage of banked population using *Mzansi* grew from 3 per cent to 12 per cent. The Banking Association of South Africa reported that 3,3 million *Mzansi* accounts had been opened by June 2006 (SAinfo, 2006). Of the people who hold *Mzansi* accounts, 60 per cent claim this to be their first bank account. There is still, however, a large divide between access to services when it came to the rich and to the poor. To create a culture of investing and saving is a challenge in South Africa, mostly due to high unemployment and the large number of poor in the country (SAinfo, 2006).

Where people do borrow, they do so mainly to buy food, pay for funerals, school fees or medical expenses. Over 35 per cent of those who know about *Mzansi* still consider the account to be too expensive (SAinfo, 2006).

Having access to financial infrastructure empowers people to make economic decisions and to widen their choices. People who have access to finance can start their own business and the poor can put their money in the bank where it is safer. For the wealthy, financial infrastructure provides for risk diversification and the ability to accumulate interest income (Ardington and Leibbrandt, 2004).

Table 5.8 Financial assets according to per capita expenditure quintile

	Quintile											
	20th		40th		60th		80th		100th		Total	
Type of financial assets	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Money in a savings account at a bank	16,2	22,4	27,7	36,9	45,0	50,8	62,8	71,8	77,4	87,7	38,9	46,9
Savings in a *stokvel*	6,4	8,0	7,9	10,5	9,8	10,2	9,1	12,3	5,8	7,1	7,7	9,5
Savings in pension plan or retirement annuity	1,2	3,2	4,2	6,8	9,4	14,9	21,9	31,1	49,0	61,1	12,5	17,9
Unit trust, stocks or shares	0,2	0,1	0,6	0,3	1,6	1,0	5,3	4,0	23,2	17,9	4,2	3,2
Cash loans that are expected to be repaid	0,6	1,4	1,4	1,5	2,5	2,1	4,9	4,0	8,1	6,0	2,7	2,5
Life insurance	8,1	11,1	14,5	16,6	23,3	27,5	42,1	42,9	69,0	71,6	25,1	28,0

Sources: Stats SA, 2001; 2004a. Own calculations.

There has been an overall increase, as shown in Table 5.8, in the share of people who answered yes in each category; those who had a savings account at a bank increased from 38,9 per cent in 2001 to 46,9 per cent in 2004. There has also been an increase in the share of people who had savings in pension plan or retirement annuity. In the 20th quintile there was an increase in both these categories; however, the share remains small.

The poor often battle to find financing because of their lack of assets, which can act as collateral. Table 5.9 provides an indication of other types of financing instruments available to the poor.

Table 5.9 Access to finance according to per capita expenditure quintile

	Quintile											
	20th		40th		60th		80th		100th		Total	
Type of finance	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Family member	13,5	15,9	10,5	14,8	10,5	12,2	8,3	9,1	7,0	6,3	10,6	12,7
Neighbour	12,9	16,2	11,0	14,0	8,7	10,4	4,5	6,3	2,3	1,5	9,1	11,2
Local dealer/shop	18,3	19,6	16,8	19,8	14,2	18,1	15,1	16,1	13,6	16,0	16,1	18,4
Co-operative	0,3	0,2	0,4	0,4	1,0	0,4	1,8	1,0	1,9	1,2	0,9	0,5
Commercial bank or building society (including credit card)	0,3	0,2	0,9	0,5	2,9	2,7	8,7	9,0	24,0	26,6	5,1	5,5
Land Bank	0,1	0,2	0,2	0,2	0,5	0,3	0,8	1,1	2,4	1,6	0,6	0,5
Other government agency	0,1	0,2	0,2	0,2	0,3	0,5	0,4	0,4	0,3	0,2	0,2	0,3
Stokvel	1,4	1,5	1,2	1,3	1,0	1,5	1,2	1,1	0,6	0,8	1,2	1,3
Non-governmental organisation (NGO)	0,3	0,1	0,4	0,1	1,0	0,2	1,8	0,5	1,9	0,7	0,9	0,3
Money lender/mashonisa	1,9	3,7	2,8	3,6	3,3	3,7	4,0	3,2	2,0	1,7	2,7	3,3
Commercial farmer	0,5	0,1	0,3	0,2	0,4	0,3	0,2	0,2	0,2	0,1	0,3	0,2

Sources: Stats SA, 2001; 2004a. Own calculations.

Most of the poor in the 20th quintile said that they borrowed money from a local dealer or shop, and this increased from 18,3 per cent in 2001 to 19,6 per cent in 2004. Of those in the 100th quintile, 24,0 per cent borrowed money from a commercial bank or building society (including credit card) in 2001, increasing to 26,6 per cent in 2004.

The extent to which the poor have access to assets such as vehicles, televisions and books is set out in Table 5.10. These assets give some indication of asset poverty in South Africa; however, this is not always a true reflection of poverty (Adato, Carter and May, 2004).

Table 5.10 Assets owned by at least one person in the household according to per capita expenditure quintiles

Type of assets owned	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Vehicle (e.g. car, truck, bakkie, van)	6,8	5,9	11,3	9,0	18,6	13,9	40,7	38,1	76,1	78,6	23,7	22,4
Motorcycle	0,3	0,2	0,2	0,2	0,8	0,5	2,1	2,0	5,1	5,4	1,2	1,2
Tractor	1,1	0,9	1,2	0,9	1,1	0,1	1,8	1,1	3,3	2,0	1,5	1,0
Plough	10,7	8,1	11,2	6,9	7,6	4,9	5,1	3,7	4,8	3,2	8,7	6,0
Television	40,2	40,5	53,1	53,2	63,1	64,1	74,6	77,6	88,5	90,9	58,8	59,9
Bicycle	11,9	11,5	15,5	12,8	17,9	15,6	24,8	21,4	39,2	39,4	19,2	17,7
Radio	70,7	70,2	80,3	79,1	83,5	8,2	87,5	89,1	94,4	95,5	80,9	80,8
Bed	94,9	95,2	96,4	96,9	96,4	97,4	97,4	97,2	97,6	98,7	96,3	96,8
Watch or clock	76,5	74,0	84,7	82,5	88,3	86,0	92,5	91,8	97,3	97,2	85,7	83,9
Books	44,9	57,7	50,6	64,5	57,5	66,2	67,3	76,3	84,0	90,3	56,8	67,9

Sources: Statistics South Africa 2001; 2004a. Own calculations.

Of the poor, 5,9 per cent had a vehicle in 2004. This was a decrease from 6,8 per cent in 2001. Of those in the 100th quintile 78,6 per cent had a vehicle in 2004. What is important is that there has been an increase in the share of the poor who owned books, from 44,9 per cent in 2001 to 57,7 per cent in 2004.

5.4.5.2 Policy and finance

South Africa's financial sector charter was signed in August 2002 and committed the country's financial institutions to extending first-order retail banking products to 80 per cent of South Africans in the lowest income bracket by 2008 (Financial Sector Charter Council, 2002). The Minister of Finance introduced a co-operative banks Bill by mid 2007, which should ensure access to services and upliftment to all South Africans (Republic of South Africa, 2007).

5.4.6 Transport infrastructure

The geography of apartheid has scattered communities far from work opportunities and, as a result transport is a major item of expenditure for poor households (Meth and Dias, 2003). There is a growing need for investment in road infrastructure. However, total expenditure on the road network has been declining. As a result, the estimated backlog in expenditure on roads is rising to about R3 billion in the national road network and R27 billion in total, with a further backlog of R3 billion for access roads (diverse transport networks). It is estimated that as many as 13 per cent of all South Africans do not have access to, or cannot afford, existing public transport services. The major allocation in the transport department's budget is for bus and rural subsidies (SAinfo, 2007).

5.4.6.1 Transport infrastructure provision and the poor

Being isolated from national roads, rail and other transport networks makes it difficult for the poor to participate in labour and other economic activities. It also limits the opportunities available to them. The lack of adequate transport infrastructure directly impacts on the poor's ability to generate income and accumulate assets (Hanmer, Lovel, Chapman and Slaymaker, 2000).

Poor transport infrastructure delivery mostly affects the poor who are situated in rural areas. The lack of adequate transport delivery leads to constraints in access to water, health care services and schools (Hanmer et al., 2000). Apart from already being marginalised, the poor are further deprived from decision-making processes and participation; living in isolation, far from towns and work opportunities with no transport (Hanmer et al., 2000). They are literally cut off from the outside world.

In urban areas the rapid growth in the minibus taxi mode of transport, and a decline in investment and support of urban bus and rail modes are adding pressure to the existing

road infrastructure. Around 70 per cent of all commuters make use of taxis (Hanmer et al., 2000).

The taxi industry is largely informal and operates outside legal spheres (Dugard, 2001). Competition between taxi owners often leads to violence and has a very destructive nature. A R7,7 billion programme is currently under way to recapitalise the minibus industry (Barradas, 2006). The main objective is to improve safety, and to reduce unhealthy competition and overtrading.

Around 15 per cent of commuters use trains as a means of transport (GCIS, 2007). Railway transport is still an important means of transport for the poor. However, there has been a significant decline in the use of railway transport over the past decade, due to poor urban integration with other public transport means, as well as safety issues resulting from poor rail service and the poor conditions of South African trains (City of Johannesburg, 2003). In 2004 government announced a five-year expenditure plan to revamp infrastructure.

When roads are not designed to accommodate pedestrian traffic, it has an adverse effect on the poor, because a large part of the poor rely on walking as their main means of transport. Furthermore, very often, maintenance on these roads is not done. Transport can also be too costly, especially for the poor living in remote areas (Commission on Poverty, 2006).

Table 5.11 Means of transport to various places according to per capita expenditure quintiles

Transport means	Quintile											
	20th		40th		60th		80th		100th		Total	
	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004	2001	2004
Walk on foot	31,44	32,53	28,08	27,38	19,71	19,43	13,45	13,37	7,31	7,29	100,0	100,0
Take taxi	32,44	33,23	29,31	29,02	20,69	20,50	12,74	13,11	4,82	4,14	100,0	100,0
Take the bus	34,65	39,39	30,06	31,40	18,50	17,14	11,26	8,28	5,53	3,80	100,0	100,0
Take the train	21,92	28,46	21,95	31,47	26,83	17,10	18,77	14,80	10,54	8,17	100,0	100,0
Own transport	3,32	3,13	5,96	5,34	10,80	7,95	25,67	25,04	54,25	58,53	100,0	100,0

Sources: Statistics South Africa 2001; 2004a. Own calculations.

Table 5.11 gives a breakdown of the means of transport used by people who travel to either one of the following places: food market, travel to a public transport depot, pre-primary or pre-school centre, primary school, secondary school, clinic, hospital, post office or post office agent and welfare office.

It is clear that most of the people who walk or take the bus or train are poor (in the 20th and 40th quintile). Of those who take the train, the share of the poor increased between 2001 and 2004. Of those who take their own transport more than 50 per cent were in the 100th quintile and 75 per cent in the 80th and 100th combined.

5.4.6.2 Policy and Finance

National government is responsible for the policy formulation, monitoring and strategic implementation of delivering public transport. In terms of the Constitution, legislative and executive powers in respect of public transport fall on the provinces (Republic of South Africa, 1996). In October 2006 the draft public transport strategy for public consultation was approved with the vision to ensure speedy implementation of public transport networks. A feasibility study for a public–private partnership for commuter rail transport and the decentralisation of this function to local government was planned and R20 million will come out of the 2007/08 Budget (National Treasury, 2008).

To achieve this, government increased its total expenditure on transport from R6,2 billion in 2003/04 to R21,5 billion in 2009/10, at an average annual rate of 22,9 per cent. Expenditure over the medium term is expected to rise from R13,7 billion in 2006/07 to R21,5 billion in 2009/10 (National Treasury, 2008). This includes additional allocations for national road infrastructure of R350 million, R550 million and R1,2 billion over the medium term. Additional allocations from the public transport infrastructure systems grant of R800 million, R1,7 billion and R3 billion has been given for priority infrastructure transport projects for the 2010 FIFA World Cup. Expenditure on the Gautrain Rapid Rail Link will be R12 billion from 2006/07 to 2009/10, which will be paid as a conditional grant to Gauteng province and includes additional allocations in the

2007 Budget of R878,4 million, R1,5 billion and R1,2 billion. Additional allocations in the 2007 Budget of R200 million, R250 million and R656 million are for passenger rail infrastructure (National Treasury, 2008).

Non-motorised transport promotion aims to increase mobility and accessibility in rural areas, which include donkey carts and the Shova Kalula Bicycle Programme that was launched in 2000. The Shova Kalula Project is an initiative that aims to deliver one million bicycles throughout South Africa by 2015 (GCIS, 2007). During the demonstration phase of the programme, the department made R10 million available to supply 19 411 bicycles to schoolchildren. By March 2006, 22 Shova Kalula microbusiness enterprises had been established in six provinces.

5.5 Conclusion

The first section in this chapter looked at the problems in defining and measuring infrastructure development. It is concluded that infrastructure delivery is a process that requires careful planning in terms of financial resources, recovery mechanisms and long-term feasibility as well as high-quality project management. This is essential because the linkages between infrastructure and poverty create the channels through which development and poverty alleviation take place. It was also determined that it is difficult to distinguish between what is "adequate" or non-poor infrastructure, and what is not. Furthermore, this chapter shows that infrastructure delivery can have both positive and negative impacts on poverty.

The second section of this chapter addressed the question of whether infrastructure development benefits the poor. Even though it is expected that accrued benefits from economic growth would trickle down to the poor in the long run, it is concluded that this is not always the case. In some instances the poor can even be negatively affected by infrastructure development.

The third section makes use of LFS data, where possible, to determine to what extent the poor have access to infrastructure in South Africa, that is, housing, water and sanitation, electricity, communication, financial infrastructure and transport infrastructure. This section also considered the policy and financial environment in which infrastructure delivery must take place.

It was found that in terms of housing delivery, around 30 per cent of those in the bottom 20 per cent of households lived in traditional dwellings/huts or structures made of traditional materials. When the access to water and sanitation delivery was analysed, it was determined that around 24 per cent of those in the bottom 20 per cent of households only had access to a toilet facility off site, and of the same 20 per cent of households 29 per cent had to make use of a public tap for water. Apart from serious challenges with electricity supply constraints, around 65 per cent of those in the bottom 20 per cent of households had access to electricity for lighting. Access to communication increased over the 2001 and 2004 period. Around 41 per cent of households in the bottom 20 per cent had access to a cellular telephone in 2004, this was up from 12,3 per cent in 2001. Furthermore, the poor have not seen much benefit from the financial sector in terms of affordable and easily accessible service delivery. Lastly, it was found that of those households in the 20th quintile, 5,9 per cent of those who answered the question had a vehicle in 2004. This was much less than households in the 100th quintile where 78,6 per cent of households who answered the question had a vehicle in 2004.

Government has to be acknowledged for progress that has been made in terms of service delivery, however slow and challenging the process has been. In the Constitution people's hopes and dreams of having access to services in a free and equal world are reflected. After the 1994 election, it was widely understood that change takes time. To implement these change financial resources, management and effective policies are also required. It has now been more than ten years after people had been promised free service delivery. For those who still do not have access, it is ten years too long, and they are taking to the streets, making their voice heard (Nemeroff, 2005).

Service delivery backlogs are present in almost all sectors, and the challenge is to catch up and at the same time improve service delivery. To achieve the goals set by government, there should be coherence between institutions, specifically government and the various policies governing service delivery. The more important issue is that while failing at delivering services, government is actually failing those that need it the most; the poor.

Chapter 6: Econometric models to simulate survey data

6.1 Introduction

> The science of model building consists of a set of quantitative tools which are used to construct and then test mathematical representations of the real world. —R. S. Pindyck and D. L. Rubinfeld Econometric models and economic forecasts (p.xiii).

Model building is a useful tool for testing the statistical relationship between variables and can be used to forecast future values. Although descriptive statistics, as shown in the previous chapters, provide useful information about the distribution of the data, econometric models provide information about the interaction between variables by making use of econometric equations. The most commonly used estimation method is linear regression, where the ordinary least square (OLS) method can be used to estimate the equation. Usually in such cases the dependent variable is a continuous variable.

This chapter will look at a qualitative- dependent variable model where the outcome involves two or more qualitative choices (Pindyck and Rubinfeld, 1998). The explanatory variables are a combination of qualitative (dummy) variables and quantitative (continuous) variables. Data for these types of models are usually collected through surveys. As discussed in Chapter 3, survey responses are usually categorical in nature; therefore data from an LFS is ideal for estimating qualitative models.

In this chapter the first section looks at binary response modelling approaches and their usefulness as a tool to model survey data. In this section, the various types of qualitative models will also be discussed.

The first modelling approach that will be considered is the linear probability model, because it can be estimated using the familiar OLS estimation method. This part will also look at the problems associated with using OLS estimation on a binary-dependent

variable. The second section will consider the alternatives to linear probability modelling approaches. These models are the logit and probit models. Lastly, this chapter will look at how the results from these models can be evaluated by making use of the count R^2, the Mcfadden, R^2 and the likelihood ratio index (Pindyck and Rubinfeld, 1998).

6.2 Binary-dependent variable modelling approaches

When modelling survey data, such as data from the LFS, the responses are usually categorical, that is, dummy-type variables. When the dependent variable is a binary/dummy-type variable, as is the case in the model used in this dissertation, binary response models are usually used to estimate such variables (Pindyck and Rubinfeld, 1998). Binary response models involve models where the dependent variable can have two or more options. In most cases where survey data is analysed, the dependent variable only has two options: for example either the respondent is female or male. A binary value is then assigned to each response. The category 'male' will be assigned a value of 1 and 'female' a value of 0. The binary-dependent variable is then regressed against one or more independent variables that could be binary or continuous in nature (Pindyck and Rubinfeld, 1998).

When modelling a binary-dependent variable, the OLS method of estimation can be used to estimate the coefficients. OLS estimation is when a line is fitted through a set of observation in such a way that the slope and intercept minimises the sum of the squared differences between the actual and fitted response values (Cook, et al., 2001). A challenge arises when using OLS estimation on a binary-dependent variable, because the appropriate model will have to be represented by a curved relationship, because the values of a binary-dependent variable are bound by 0 and 1.

The solution, is partly that, instead of using an OLS estimation, a maximum likelihood estimation should be applied (Gujarati, 2003). The slope and intercept of the logistic model are fitted via maximum likelihood estimation (MLE). The MLE of a parameter is that value that maximises the probability of the observed data. The likelihood that needs

to be maximised is determined by joint probability density function. Qualitative response regression models violate the assumption that Y is a continuous variable and therefore the error term will not be normally distributed. If OLS is applied to such data, the model will be a linear probability model. However, it becomes difficult then to interpret the results. These difficulties will be discussed in section 6.2.1.1.

6.2.1 Linear probability model

The most elementary of the three binary choice models is the linear probability model (LPM). Linear probability modelling is when a straight line can be fitted to intersections where two variables are plotted on a graph, one on the Y axis and one on the X axis. OLS regression models this linear function in the form of (Gujarati, 2003):

$$\mathrm{E}(Y|x) = \beta_1 + \beta_2 x \tag{1}$$

Equation 1 can also be rewritten as a regression model:

$$Y_i = \beta_1 + \beta_2 x_i + \varepsilon_i, \qquad i=1,\ldots,n \tag{2}$$

where Y_i is the dependent variable with two options, which can either be equal to 1, should the respondents option be the first option, or equal to 0, should the respondent select the alternative option. The residual, ε_i is the independently distributed random variable with a mean of 0. One can further build on this model and add multiple explanatory variables (x), transforming the regression into a multiple regression model. Equation 2 is referred to as a linear probability model, because Y_i is a binary or dichotomous regressand (Gujarati, 2003).

Equation 2 can be interpreted as taking the expected value of each observed observation of the dependent variable (Y_i) as in equation 1. Y_i therefore follows a Bernoulli probability distribution, which can be expressed as

$$E(Y_i) = 1(P_i) + 0(1 - P_i) = P_i \tag{3}$$

where p_i is the probability of $Y_i = 1$, and therefore represent a success; and $1 - P_i$ equals the probability that $Y_i = 0$, therefore representing a failure. X_i therefore follows a binomial distribution, restricting P_i to:

$$0 \leq P_i \leq 1 \tag{4}$$

However, problems arise as soon as a binary-dependent variable (Y) is used when running an OLS model. These problems will be discussed in more detail in the next section.

6.2.1.1 Problems with OLS

When using OLS regression procedures, it is assumed that the dependent variable is continuous. However, if the dependent variable is a binary variable taking on the values of either 1 or 0, the following problems arise (Brüderl, 2003 and Gujarati, 2003):

- The error term is not normally distributed

  Although it is not a requirement, for statistical purposes, using OLS estimation, one of the assumptions is that the error term is normally distributed. When using a linear probability model, however, the error term also follows a Bernoulli distribution:

$$\varepsilon_i = Y_i - \beta_1 - \beta_2 X_i \tag{5}$$

However, the estimates remain unbiased, even though the error term is not normally distributed. Furthermore, as the sample size increases indefinitely; based on the central limit theorem, the OLS estimators tend to conform to a normal distribution in general (Gujarati, 2003).

- Predictions are not bound between 0 and 1

The real problem with the OLS estimation is that there is no guarantee that the predicted values $\hat{Y}_i$ will lie in the (0,1) interval. By estimating a usual OLS model with a dichotomous dependent variable, the value of $\hat{Y}_i$ could be either less than 0 or more than 1. If $\hat{Y}_i$ is greater than 1, then $\hat{Y}_i$ is assumed to be 1 and if $\hat{Y}_i$ is smaller than 0, then $\hat{Y}_i$ is assumed to be 0 (Gujarati, 2003).

- Errors are highly heteroscedastic and difficult to correct

Based on statistical theory, the theoretical mean and variance for a Bernoulli distribution is respectively: P and $P(1-P)$, where P is the probability of a success (Gujarati, 2003). The variance is, however, a function of the mean and therefore the variance of the error changes over time, resulting in the errors being heteroscedastic. As regards to equation 2, $P_i = \beta_1 + \beta_2 X_i$ if the intercept (β_0) is larger than 0 and the slope ($\beta_0 X_1$) is smaller than 1. The values 1 and 0 can be substituted into the values of Y_i. The variance of the error term will then be:

$$\sigma_i^2 = E(\varepsilon_i^2) = E(Y_i)[1 - E(Y_i)] \tag{6}$$

or in other words (Gujarati, 2003):

$$\text{var}(\varepsilon_i) = P_i(1 - P_i) \tag{7}$$

However, the variance shows that the error term is heteroscedastic, in that when P_i approaches 0 or 1 the variance will be relatively low, however, as P_i approaches ½, it will have a relatively higher variance. In order to correct for heteroscedasticity, the variance for each of the values of Y_i is estimated, and then the weighted least-squares estimation is applied, by using a two-step procedure (Gujarati, 2003). Step one is to run an OLS regression and obtain $\hat{Y}_i$, followed by obtaining the weights that will be equal to $\hat{Y}_i(1-\hat{Y}_i)$. Step two will then use the weights obtained to transform the data and estimate the OLS with the weighted least squares (Gujarati, 2003).

- Dubious R^2 values as a measure for goodness of fit

The R^2 explains the goodness of fit of a model. When using dichotomous response models the R^2 is usually not of much value, because all of the Y values will either fall on the X axis or along the line corresponding to 1, resulting in the regression line not fitting the scatter well (Gujarati, 2003). The R^2 will then as a result be much lower. This low value may well show an upper bound of $\frac{1}{3}$, should the assumption be that the true probabilities are uniformly distributed across an interval (Pindyck and Rubinfeld, 1998). A low value means that a large part of the variance in the model is still unexplained. However, the model can still be used to explain correlations between the variables (Pindyck and Rubinfeld, 1998).

6.2.1.2 Alternatives to the linear probability model

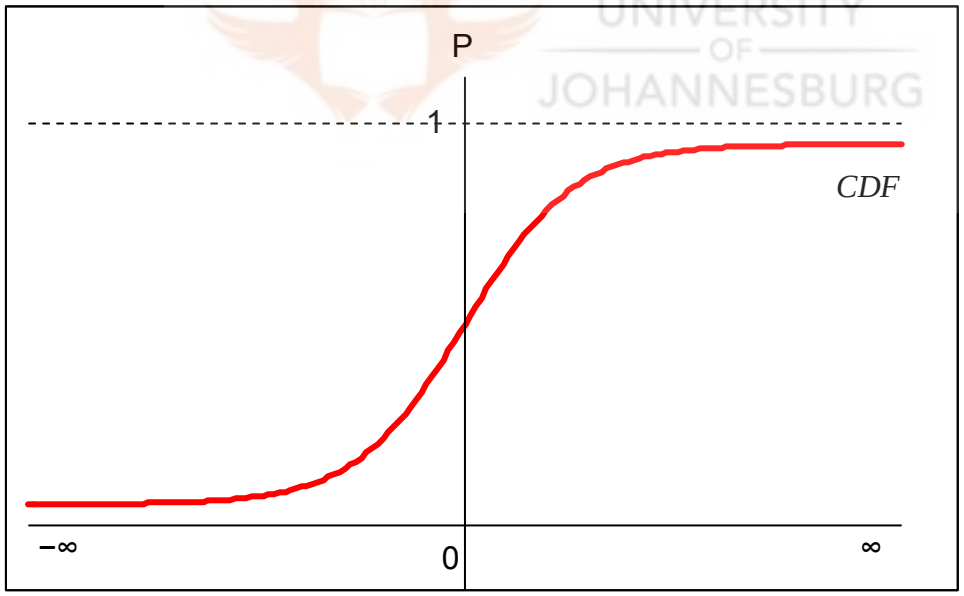
As discussed above, the linear probability model is not the most appropriate model for estimating binary-dependent variable models. These challenges can, however, be overcome by using the weighted least square method or to increase the sample size (Gujarati, 2003). These methods, however, do not address the fundamental problem, which is the assumption that the variables are linearly correlated. A more improved

approach will therefore be to use a probability model that incorporates the following two most important features (Kollamparambil, 2007) :

1. As X_i increases, $P_i = E(Y = 1|X)$ increases, but remains within the (0,1) interval and;
2. the relationship between P_i and X_i is non-linear.

A cumulative distribution function (CDF) best represents the geometry of the model, where the probabilities are limited between 0 and 1 and responds non-linearly with X, as represented in Figure 6.1. The cumulative distribution function ranges from $-\infty$ to $+\infty$, while P ranges between 0 and 1. The most common CDF models chosen for dichotomous models are the logistic CDF, as in the case of a logit model, and the normal CDF, as in the case of a probit model.

Figure 6.1 The cumulative distribution function



Source: Adapted from Gujarati (2003)

6.2.2 Logit model

A logit model uses a regression to obtain a *p* value that can be used to compute the predicted probability of a selected outcome of 0 or 1, based on the cumulative probability function which can be defined as (Nahlik, 2006 and Gujarati, 2003):

$$P_i\left[y = 1 \middle| x\right] = \frac{\exp(x\beta)}{1+\exp(x\beta)} = \beta_1 + \beta_2 X_i \tag{8}$$

However, P_i can also be expressed as a cumulative logistic probability function (Gujarati, 2003):

$$P_i = \frac{1}{1+e^{-z_i}} \text{ or } \frac{e^z}{1+e^z} \tag{9}$$

where $Z_i = \beta_1 + \beta_2 X_i$; e is the base of natural logarithm equal to 2.71828, and P_i is the probability that an individual will make a certain choice given a set of characteristics $Z_i(X)$ (Gujarati, 2003). Z_i ranges from $-\infty$ to $+\infty$ and P_i ranges from 0 to 1, and is non-linearly related to $Z_i(X)$. Instead of measuring proportions (probability), the odds are calculated. A probability is the chance of an event occurring written in proportional terms. The odds are equivalent to the ratio of the probability of an event occurring to the probability of an event not occurring (Sturgis, 2004). Per definition the following odds ratio can be defined (Kollamparambil, 2007):

$$e^{Z_i} = \frac{P_i}{1-P_i} \tag{10}$$

Then by taking the natural logarithm of both sides, the following linearised model or logit model is obtained (Pindyck and Rubinfeld, 1998):

$$Z_i = \log \frac{P_i}{1 - P_i} = \beta_1 + \beta_2 X_i \tag{11}$$

The Z_i is the odds that a particular choice will be made, and it is linear in X and linear in the parameters. The next part will be to estimate equation 11. Data can either be used on a micro or individual level or on a grouped or replicated level (Gujarati, 2003). Micro or individual level data are used when a person or household either owns a house or not. Group or replicated level data are used where information is aggregated, for example, if there are seventy families and forty families owned a house, while thirty families did not. In terms of micro or individual data the normal logit will be estimated, while for group or replicated level data the so-called glogit model can be applied (Kollamparambil, 2007). For group data, each group's probability will be calculated, after which a logit can be applied to each group. The results will be obtained by using the weighted least squares estimation method, after which confidence intervals can be established. Larger sample sizes tend to give more valid results (Gujarati, 2003). When using individual data, such as that used in this dissertation, the normal logit can be estimated, making use of the maximum likelihood estimation method.

As mentioned earlier in section 6.2, the MLE of a parameter is that value that maximises the probability of the observed data. If equation 9 is considered, and a random sample of n observations is drawn, allowing $f_i(Y_i)$ to denote the probability that Y_i equals either 1 or 0, the joint probability density function, or likelihood function, in other words, can be represented as (Gujarati, 2003):

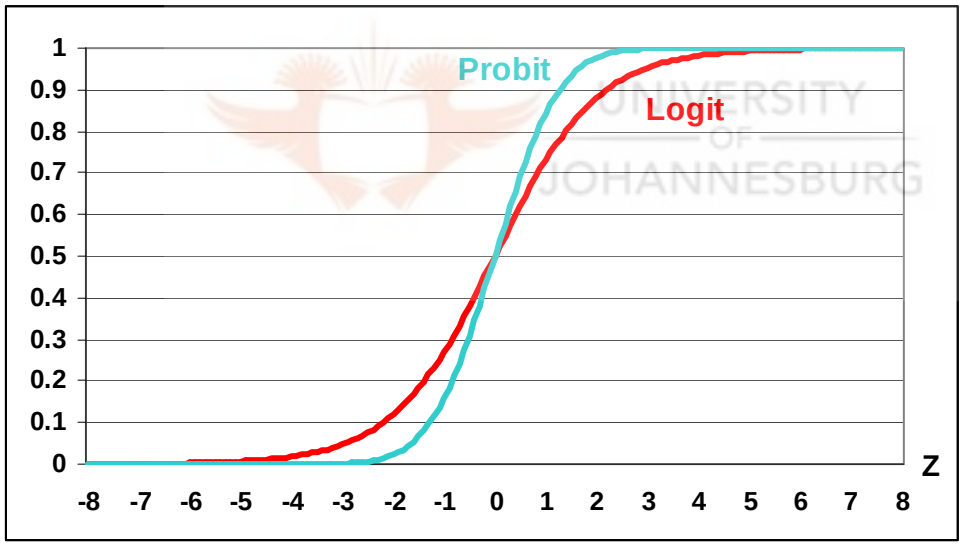
$$f(Y_1, Y_2, ...Y_n) = \prod_1^n f_i(Y_i) = \prod_1^n P_i^{Y_i} (1 - P_i)^{1 - Y_i} \tag{12}$$

The log-likelihood function can now be obtained by taking the natural logarithm of equation 12 (Gujarati, 2003):

$$\text{Ln } f(Y_1, Y_2, \ldots Y_n) = \sum_{1}^{n} Y_i(\beta_1 + \beta_2 X_i) - \sum_{1}^{n} \ln\left[1 + e^{(\beta_1 + \beta_2 X_i)}\right] \quad (13)$$

The final results can be calculated by maximising the log-likelihood function, and solving β_1 and β_2 by using methods of non-linear estimation. The maximum likelihood procedure for the logit and probit will be similar, except that the logit assumes a logistic cumulative distribution function, while the probit assumes a normal cumulative distribution function. The major difference between the two distributions is illustrated in figure 6.2, where the logistic cumulative distribution function has flatter tails than the normal cumulative distribution function (Vasisht, 2003). This means that the probit function approaches the axis faster than the logit curve.

Figure 6.2 Cumulative distribution functions of the logit and probit model



Source: Adapted from: Kollamparambil, 2007

6.2.3 Probit model

Because the two features discussed in section 6.2.1.2 needs to remain, a cumulative probability function is necessary for the probit model. As mentioned earlier, the logit model uses the logistic CDF, while the probit model makes use of the normal CDF. The

probit model is also sometimes known as the 'normit model' (Gujarati, 2003). The probit probability model can be represented as (Pindyck and Rubinfeld, 1998):

$$Z_i = \alpha + \beta X_i \tag{14}$$

where Z_i is a theoretical continuous index determined by variable X_i. Z_i represents the strength of feeling of individual i for one of two choices. Z_i can also be referred to as the 'unobservable utility index' or a 'latent variable' (Gujurati, 2003). The probit model provides a suitable means of estimation of the slope and intercept parameters, while at the same time providing information about the relationship between Z and X.

The probit model assumes that Z_i^* is a normally distributed random variable, so that the probability of $Z_i^* \leq Z_i$ can be calculated from the cumulative normal probability function, which in its standard form can be written as (Pindyck and Rubinfeld, 1998):

$$P_i = F(Z_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_i} e^{\frac{-s^2}{2}} ds \tag{15}$$

where s is a random variable that is normally distributed with a zero mean, and unit variance. P_i represents the probability that an event will occur by definition falling between the (0,1) interval. The higher the value of Z_i, the more likely an event is to occur. To obtain an estimate of the index Z_i, the inverse of the cumulative normal function is applied:

$$Z_i = F^{-1}(P_i) = \alpha + \beta X_i \tag{16}$$

the probability (P_i) is now an estimate of the conditional probability that an individual will choose one of the two options, given a certain set of characteristics (X). This is the same as the probability that a standard normal variable will be less than or equal to

$\alpha + \beta X_i$. The probit model uses the same maximum-likelihood approach to estimate the parameters as the logit model discussed earlier.

6.3 Interpreting LPM, probit and logit results

When a linear regression model is estimated, the slope coefficient measures the change in the average value of the dependent variable for a unit change in the value of the independent variable, while holding all other variables constant. Therefore, in a linear regression model, one change in X will result in a β change in Y, *ceteris paribus* (Golder, 2007).

The results from the logit and the probit model are not directly comparable. The reason for this is that even though the standard logistic (logit) and standard normal distribution (probit) both have a mean value of zero, the variances differ. The standard normal distribution has a variance of 1 and the logistic distribution has a variance of $\pi^2/3$ (Nahlik, 2006). By multiplying the logit coefficient of each explanatory variable by approximately 0,55 the coefficients can be made comparable to the probit coefficients. An alternative is to multiply the logit coefficient by 0,625 (Gujarati, 2003). The inverse is also true, thus by multiplying the probit coefficient by 1,81 and 1,6 respectively, a comparable logit coefficient can be obtained.

However, in the logit model, the interpretation of the estimated coefficients is different from the interpretation of the coefficient estimated in OLS. The marginal effect of X on P, that is, the probability that $Y = 1$, is given by the derivative (Gujarati, 2003):

$$\frac{dP_i}{dX_i} = \frac{dF(t)}{dt} \cdot \frac{dt}{dX} = f(\beta_1 + \beta_2 X_i)\beta_2 \qquad (17)$$

where $t = \beta_1 + \beta_2 X$ and $f(\beta_1 + \beta_2 X)$ is the logistic probability distribution function evaluated at $\beta_1 + \beta_2 X$. This process is also referred to as the 'chain rule of derivatives'. The result shows that the marginal effect of X on P depends on the slope of the logit function and the coefficient β_2 (Gujarati, 2003). The sign of dP/dX depends on the sign of β_2.

When using the probit model, the formula for calculating the marginal effect is the same as for the logit model, except for the probability distribution function used. The marginal effect of X on P, that is the probability that $Y = 1$ is given by the derivative (Gujarati, 2003):

$$\frac{dP}{dX} = \frac{dF(t)}{dt} \cdot \frac{dt}{dX} = f(\beta_1 + \beta_2 X_i)\beta_2 \qquad (18)$$

where $t = \beta_1 + \beta_2 X$ and $f(\beta_1 + \beta_2 X)$ is the standard normal probability distribution function evaluated at $\beta_1 + \beta_2 X$. This shows that the marginal effect of X on P depends on the slope of the probit function/normal cumulative distribution function, and the coefficient β_2. The sign of dP/dX depends on the sign of β_2 (Gujarati, 2003).

6.4 Evaluation of results

The best way to evaluate the LPM would be to evaluate how significantly the explanatory variables explain the dependent variable by looking at the partial slope coefficients. The zero hypothesis is that the coefficients are equal to each other and equal to 0. Before the slope coefficients can be evaluated, a confidence interval must be selected. If the p value of the slope coefficient is smaller than selected α the zero hypothesis will be rejected, and the slope coefficient is significantly different from each other and not equal to 0. The corresponding t-values can also be evaluated, and as a rule of thumb, by a t-value that is greater or equal to 2 (Gujarati, 2003). However, again, when estimating a binary-dependent variable model, the coefficients are not bounded to 0 and 1 and therefore

interpretation of these coefficients with OLS estimation should rather be avoided (Pindyck and Rubinfeld, 1998).

The R^2 provides information on how well the variation in the explanatory variables explains the variation in the dependent variable (Pindyck and Rubinfeld, 1998). However, as previously noted, when a binary-dependent variable is estimated using OLS estimation, the R^2 loses some of its value, because the dependent variable is restricted to a binary value of 0 or 1. The R^2 for such models will therefore most likely lie between 0,2 and 0,6 (Gujarati, 2003).

The adjusted R^2 adjusts for a loss in the degrees of freedom for each additional explanatory variable that is added (Gujarati, 2003). The adjusted R^2 seems to be preferred, because it provides more accurate information about the model. Pindyck and Rubinfeld (1998), however, suggest that the overall R^2, as well as the adjusted R^2, is not the best method of measuring the goodness of fit of a binary-dependent model, and other more relevant measures should be considered.

The F statistic tests the hypothesis that none of the explanatory variables helps to explain the variation of Y about its mean, considering $k-1$ and $N-k$ degrees of freedom (Pindyck and Rubinfeld, 1998). The hypothesis therefore states that the coefficients are all equal to one another and equal to 0. If the zero hypothesis holds the F statistic would be close to 0. If the F-statistic is large, the zero hypothesis is rejected, and the coefficients are significantly different from 0. Alternatively, if the probability of the F-statistic is very low, the zero hypothesis can be rejected, depending on the selected confidence interval. The F statistic therefore provides information about how well the variation in the explanatory variables jointly explains the overall model.

Because the maximum likelihood estimation method is used to estimate the logit model, it is assumed that a large sample is used, and therefore the estimated standard errors of the logit model are asymptotic (Gujarati, 2003). The standard normal Z statistic can therefore be used to evaluate the significance of the coefficients. Unlike in OLS models

where the R^2 provide information on the goodness of fit, in binary-dependent variables, there is no single measure that can measure this fit (Enders, 2003). A simulated R^2 measure called a 'pseudo R^2 measure' provides an alternative to the conventional R^2 measure. The pseudo R^2 measure consists of a range of measures that determines the goodness of fit for a binary-dependent model The pseudo R^2 measures consist of the McFadden R^2, the count R^2 and the likelihood ratio (LR) statistic (Gujarati, 2003).

- **McFadden R^2**

The McFadden R^2 value is defined as (Gujarati, 2003):

$$1 - \left(LLF_{ur} \Big/ LLF_{r} \right)$$

where LLF_{ur} is the unrestricted log likelihood function of the full model, and LLF_r is the restricted log likelihood function of the constant only model. The McFadden R^2 value ranges between 0 and 1. This formula is similar to the OLS R^2. Another interpretation of the McFadden R^2 is that it represents the ratio of the estimated information gain from using the current model compared to using a model where all possible explanatory variables are included (Shtatland, Kleinman and Cain, 2002).

- **Count R^2**

The count R^2 can be defined as:

$$\text{Count } R^2 = \frac{\text{number of correct predictions}}{\text{total number of observations}}$$

If the predicted probabilities are greater than 0,5, a value of 1 is assumed. However, if the predicted probabilities are less, a value of 0 will be assumed. The number of correct

predictions is then counted, and the R^2 is then calculated based on these correct predictions.

- **Likelihood Ratio (LR statistic)**

The LR statistic test is similar to the *F* test in linear regression analysis (Gujarati, 2003). The zero hypothesis states that all the slope coefficients are simultaneously equal to zero. The LR statistic follows an χ^2 distribution with the degrees of freedom equal to the number of explanatory variables. If the LR statistic is large and the probability of the LR statistic is small, the zero hypothesis is rejected. This will then mean that, together, all the repressors have a significant impact on the dependent variable (Gujarati, 2003).

6.5 Conclusion

This chapter provided a theoretical overview of the econometric modelling techniques that can be used to simulate survey data. It was established that data collected through surveys are usually in category format, and therefore it is likely that the variable of interest might be a binary variable (recall the example of male being assigned a value of 1 and female a value of 0). It was shown in the first section of this chapter that binary-dependent variables estimated using the OLS estimation method become difficult to interpret, because the linear model makes four very important assumptions about the underlying data. These assumptions are not met in that (1) the error term is not normally distributed; (2) predictions are not bound between 0 and 1, (3); errors are highly heteroscedastic and difficult to correct; and (4) the R^2 values, when used as a measure for the goodness of fit, give dubious results. There are, however, two ways in which these challenges can partly be overcome. One is to use the weighted least square estimation method and the other, by increasing the sample size, according to statistical theory, the OLS estimators will tend to conform to a normal distribution in general. However, interpretation will still be difficult. As an alternative, the maximum likelihood estimation method was introduced. The first part under this section looked at the LPM, together with the problems that arises when applying OLS estimation to a binary-dependent variable.

The second section concentrated on the Logit and Probit models as alternatives to the LPM, using maximum likelihood estimation. This method fitted a curve line through the data points, binding the slope coefficients to between 0 and 1. It was further shown that the cumulative distribution function for the logit model is a logistic cumulative distribution, while the cumulative distribution function for the probit assumes a normal cumulative distribution function.

The third section looked at the calculation of the marginal effects of both the logit and the probit models, making interpretation and comparison possible. The last section distinguished between evaluating an LPM model using OLS estimation, and the logit and probit model using maximum likelihood estimation. It was concluded that the R^2 and the adjusted R^2 are not as useful in maximum likelihood estimation as it is in OLS estimation, because it is likely to lie between 0,2 and 0,6, due to the binary-dependent variable that is bound to 0 and 1. A series of other measures for measuring the goodness of fit was discussed, called the pseudo R^2. These were the McFadden R^2, the count R^2 and the LR statistic.

In conclusion, for a binary-dependent variable either the logit or probit model is more appropriate. In practice, however, the logit model is more preferred because it is more simplistic (Gujarati, 2003).

Chapter 7: Modelling poverty in South Africa using Labour force survey data

7.1 Introduction

> Econometrics may be defined as the social science in which the tools of economic theory, mathematics and statistical inference are applied to the analysis of economic phenomena. — A.S. Goldberger. Econometric theory in Gujarati (p.1)

Econometric analysis is important because it provides evidence that economic relationships are not just theoretical. To establish whether infrastructure, as discussed in the previous chapters, does practically have an impact on poverty, an econometric model can be used to estimate the interaction between the poverty and a range of household characteristic variables, from data collected from households. The poor are usually associated with a lack of infrastructure (WWAP, 2006).

In chapter 4 it was established that over 50 per cent of the population in South Africa was situated in the bottom 40 per cent of households, based on their per capita household expenditure. It was further established that of all South Africans, the poor mostly lacked access to infrastructure. The question that arises is whether the provision of infrastructure to those in poverty will provide them with the opportunity to escape poverty. To test this hypothesis, an econometric model will be used to determine whether access to infrastructure, together with other characteristics, can decrease the likelihood that a household can be poor.

In this chapter the first part will identify the dependent and explanatory variables for estimating a poverty model. The explanatory variable will, firstly, be divided into demographic and status variables and, secondly, into infrastructure variables. The next part will specify the three different models. First, the LPM model will be specified, followed by the probit and logit models estimation.

The third part will estimate slope coefficients for the explanatory variables, using OLS estimation for the LPM and maximum likelihood estimation for the probit and logit models. Under the OLS model, the results as well as the evaluation of the goodness of fit will be discussed. Under the logit model the odds ratios will be analysed together with the marginal effects of the coefficients. Lastly, the goodness of fit of the logit will be evaluated.

7.2 Identifying the variables for estimating a poverty model

The poor are usually classified as being excluded from having freedom of economic choice. Poverty is usually associated with low standards of living and directly impacts on the level of development in a country (Burns, 2004). It is difficult to comprehend fully the characteristics of poverty and the factors that determine whether one person would be more likely than another to be poor. Besides infrastructure, it is also important to note other factors that contribute to poverty. To determine the variables that impact on poverty, the following dependent and explanatory variables are identified:

7.2.1 The Dependent Variable (POOR)

The dependent variable is the variable of interest in this dissertation. The econometric model aims to determine how various explanatory variables have an impact on poverty. As mentioned in Chapter 2, poverty can be determined in various ways. This dissertation makes use of the per capita household expenditure method. After dividing households into quintiles based on their per capita household expenditure, the bottom 40 per cent are identified as those living in poverty, as implemented by Woolard and Leibbrandt (1999), Büdlender (2003), and Edwards and Stern (2006). The dependent variable, POOR, is a binary variable, indicating whether a person is within the bottom 20th and 40th quintile as measured by their per capita household expenditure. This group of people were allocated a 0. Those who were in the 60th, 80th and 100th quintile were assigned a value of 1.

7.2.2 Explanatory variables

The explanatory variables identified in this model are divided into two categories. First, the impact of demographic and status variables are identified. Second, the research question is addressed by including infrastructure variables as explanatory variables in the model.

7.2.2.1 Demographic and status variables

- Gender (GEN)

The variable representing gender was coded 'GEN'. Gender is considered a very important factor in determining whether a person in South Africa would be considered poor or not. Owing to the historic exclusion of females in key economic decision-making processes, females are more likely to be poor than males. Many females may be in non-poor households but due to the inequalities in intra-household distribution of resources, they should be counted as poor (Woolard, 2002).

At present the share of females that are poor makes up 53 per cent of the total. The gender variable is a dummy variable, where males are coded 0 and females are coded 1. The reference group is males.

- Population group (POP)

The variable 'POP' represents the population group of a person. Living standards are closely correlated with race in South Africa (Woolard, 2002). Poverty is not race-specific, but more than 90 per cent of the poor in South Africa, based on the finding in Chapter 4 of this dissertation, per capita household expenditure analysis are black African. Whites by contrast constitute less than 1 per cent of the poor.

The race variable is a categorical variable, where black African-headed households are coded 0, households headed by a coloured person are coded 1, Indian-headed households are coded 2 and white-headed households are coded 3. The reference category is black African.

- Age (AGE)

The 'AGE' variable represents the age of the household head in years. According to du Toit (2003), the inequality and rates of unemployment across age groups is a matter of concern. Based on the results discussed in Chapter 4, the poor are more likely to be found in the age category 15–25 years. However, based on research done by Lam et al. (2004), older households are larger in size and consist of a large share of dependents and unemployed persons, making them more likely to be poor. The age variable is a continuous variable ranging from 15 upwards.

- Main income source of the household head (INC)

'INC' is the name for the variable that classifies the household heads' main source of income. The categories in this variable are salaries and/or wages coded as 0, remittances coded as 1, pensions and grants coded as 2, sales of farm products coded as 3, other non-farm income coded as 4 and lastly no income coded as 5. The reference category is salaries and/or wages.

The reason for including this variable is that the poor are classified by a lack of wage income, either as a result of unemployment or of low-paid jobs (Woolard, 2002). The poor are also heavily reliant on income from remittances and state transfers.

- Economic status of the household head (STATUS)

The status of the households head was represented by the variable 'STATUS'. The economic status of a person is the main determinant whether a person will have an

income or not. In South Africa poverty, unemployment and economically inactiveness are correlated (Woolard, 2002). Of the poor, more than 50 per cent are not economically active, based on the official definition. Only around 24 per cent of the working age of the poor are employed according to Woolard (2002). This compares with this study's analyses where in the 20th quintile 20,7 per cent of the poor are employed, and within the 40th quintile, 27,6 percent are employed in 2004. The coding for the status variable is 0 for employed and 1 for unemployed. The benchmark category is employed.

- Highest level of education of the household head (EDUC)

A household heads highest level of education was represented by the variable 'EDUC'. The highest level of education The poor are usually characterised by having little or no education. As shown in Chapter 4, most of the poor do not have more than a Grade 12. This has a direct impact on skills available and unemployment of the poor. There is, however, a slight improvement in the share of the poor finishing matric from 2001 to 2004. For non-poor people, 39 per cent finished matric, a further 10 per cent completed a degree, and 5 per cent a postgraduate degree. For poor people fewer than 0,5 per cent finished a degree or a post graduate degree, based on results from Chapter 4. The returns to education are an important determinant in this poverty analysis and therefore the household heads' education is analysed. Education is coded into the number of years of schooling, that is, a household head that had completed Grade 1 as the highest level of schooling, had one year of education, Grade 2 had two years of education and so on. Those who completed Matric have twelve years of education, while those with a diploma and matric had on average fourteen years of education. For those with a degree, the average years of education was sixteen years, while those with an Honours degree had seventeen years of education.

7.2.2.2 Infrastructure variables

- Access to formal housing (*HOUSE*)

'HOUSE' is the variable that measures access to formal, traditional and informal housing. This variable aims to evaluate whether a household is more likely to be poor if they live in a traditional dwelling or in an informal dwelling versus living in a formal house. Households who live in a formal house are coded 0, those in a traditional house are coded 1, and those in an informal house are coded 2 (Stats SA, 2007a).

A formal house consists of a dwelling or brick structure on a separate stand or yard or on a farm; a flat or apartment in a block of flats; a town/cluster/semi-detached house; a unit in a retirement village; a dwelling/flat/room in a backyard and a room or flatlet. A traditional dwelling consists of a traditional dwelling/hut/structure made of traditional materials. Informal housing consists of an informal dwelling/shack in a backyard or in an informal squatter settlement or on a farm as well as those living in a caravan or tent.

- Access to safe water (*WATER*)

'WATER' is the variable that measure access to clean water. This variable aims to evaluate whether a household is more likely to be poor if the household has no access to clean water. Households with access to piped water inside the dwelling is coded 0 and households with access to water outside the dwelling is coded 1, households with access to informal water such as water tanks, rivers, dams and public taps are coded 2 (Stats SA, 2007a). The reference category is households with piped water inside the dwelling.

- Access to electricity for lighting (*LIGHT*)

'LIGHT' is the variable that measure a household's access to electricity. This variable aims to evaluate whether a household is more likely to be poor if the households do not have access to electricity for lighting. Households with access to formal electricity from mains or from solar energy are coded 0 and households with access to traditional energy sources is coded 1 and households with access to other sources such as wood, coal, candles and animal dung is coded 2 (Stats SA, 2007a). Electricity from mains represents the benchmark category.

- Access to sanitation (*TOIL*)

'TOIL' is the variable that measures household's access to formal sanitation facilities. Households with access to formal toilet facilities are coded 0, that is, flush toilet facilities within the dwelling (Stats SA, 2007a). Households with access to toilet facilities on site are coded 1 and households with access to toilet facilities off site are coded 2.

- Access to rubbish removal (*RUB*)

Access to rubbish removal 'RUB' measures access to rubbish removal by local authorities. Households with access to rubbish removal by local authorities are coded 0 and those who have access to communal rubbish removal are coded 1. Those who have access to a communal dump or container were coded 2 and those who do not have access to rubbish removal are coded 3 (Stats SA, 2007a).

- Access to cellular telephony (*CELL*)

Access cellular telephony 'CELL' can contribute to enhancing the capabilities of the poor by assisting in job search methods, and increasing access to education, health and

entertainment. Access to telecommunications is measured by having access to a cellular telephone. Having access is coded 0 and no access is coded 1.

- Access to telecommunications (*TEL*)

Access to telecommunications 'TEL' plays an important role in alleviating poverty in South Africa (Telkom, 2007). Similar to cellular telephones, access to telephones provides households with the opportunity to find work, increase their education and enhance their capabilities. Access to telecommunications is measured by having access to a telephone in the dwelling. Having access is coded 0 and no access is coded 1.

7.3 Model specification

This analysis simply aims to see the impact of the demographic and infrastructural variables on poverty and how these variables contribute to the probability of being poor. The data used is from Statistics South Africa's LFS 2001 and 2004.

The sample of Enumeration Areas for the 2001 and 2004 LFS is drawn from the census 1996 and 2001, respectively. After omitting non-responses in the relevant variables, the total number of observations for 2001 was 16 593 and for 2004 it was 16 137. Demographic and labour market information is available on an individual level, while social characteristics are available on a household level. Household variables are used at a household level, and age, education and other individual characteristics of the household head are used. The next section will discuss the dependent variable and explanatory variables selected as regressors.

7.3.1 Specification: OLS model

The linear regression is specified in the form:

$$Y = \alpha_i + \beta_i X_i + \mu_i$$

and if the X_i is substituted with the explanatory variables, the equation will be as follow:

$$P_i = \beta_1 + \beta_2 HOUSE + \beta_3 AGE + \beta_4 POP + \beta_5 GEN + \beta_6 EDUC + \beta_7 INC + \beta_8 CELL + \beta_9 TEL + \beta_{10} RUB +$$

$$\beta_{11} TOI + \beta_{12} LIGH + \beta_{13} WATER + \varepsilon_i$$

Where P_i = Binary-dependent variable, 0 for poor and 1 for not poor;

β_1 = The intercept

X_i = List of explanatory characteristics as discussed in 8.2

ε_i = error term

As mentioned in Chapter 6, OLS estimation is not ideal for a binary-dependent variable model, mainly because the slope coefficients are not bounded to lie between 0 and 1. OLS estimation will aim to fit a linear line though the observations, even though the observations do not have a linear relationship. The alternative is to use either the logit or the probit model.

7.3.2 Specification: Logit

The logit function can be specified as:

$$P_i = \frac{1}{1+e^{-z_i}} \text{ or } \frac{e^z}{1+e^z} \text{ and}$$

$$Z_i = \beta_1 + \beta_2 X_1 + \beta_3 X_2 + \beta_4 X_3 + \beta_5 X_4 + \beta_6 X_5 + \beta_7 X_6 + \beta_8 X_7 + \beta_9 X_8 + \beta_{10} X_9 +$$

$$\beta_{11} X_{10} + \beta_{12} X_{11} + \beta_{13} X_{12} + \varepsilon_i$$

where

P = the likelihood of not being poor (where Y=1)

X_i = explanatory variables

β_i = parameters associated with X_i

Z can be rewritten as:

$$Z_i = \beta_1 + \beta_2 HOUSE + \beta_3 AGE + \beta_4 POP + \beta_5 GEN + \beta_6 EDUC + \beta_7 INC + \beta_8 CELL + \beta_9 TEL + \beta_{10} RUB +$$

$$\beta_{11} TOI + \beta_{12} LIGH + \beta_{13} WATER + \varepsilon_i$$

Because the logit function is non-linear and OLS estimation cannot be used, maximum likelihood estimation is employed. The MLE of a parameter is that value that maximises the probability of the observed data. The marginal effect of X on P, that is, the probability that $Y =$ becoming less poor, is given by the derivative (Gujarati, 2003):

$$\frac{dP_i}{dX_i} = \frac{dF(t)}{dt} \cdot \frac{dt}{dX} = f(\beta_1 + \beta_2 X_i)\beta_2$$

where $t = \beta_1 + \beta_2 X$ and $f(\beta_1 + \beta_2 X)$ is the logistic probability distribution function evaluated at $\beta_1 + \beta_2 X$ (Gujarati, 2003).

As mentioned in Chapter 6, the logit model, amongst others, is preferred to the LPM because in the logit model, the probability is bound to lie between 0 and 1. Even more so, the logit model is preferred overall because of its simplicity (Pindyck and Rubinfeld, 1998).

7.3.3 Specification: Probit

The probit probability model is represented as (Pindyck and Rubinfeld, 1998):

$$Z_i = \alpha + \beta X_i$$

$$Z_i = \beta_1 + \beta_2 X_1 + \beta_3 X_2 + \beta_4 X_3 + \beta_5 X_4 + \beta_6 X_5 + \beta_7 X_6 + \beta_8 X_7 + \beta_9 X_8 + \beta_{10} X_9 +$$

$$\beta_{11} X_{10} + \beta_{12} X_{11} + \beta_{13} X_{12} + \varepsilon_i$$

where

P = the likelihood of not being poor (where Y = 1)

X_i = explanatory variables

β_i = parameters associated with X_i

Z can be rewritten as:

$$Z_i = \beta_1 + \beta_2 HOUSE + \beta_3 AGE + \beta_4 POP + \beta_5 GEN + \beta_6 EDUC + \beta_7 INC + \beta_8 CELL + \beta_9 TEL + \beta_{10} RUB +$$

$$\beta_{11} TOI + \beta_{12} LIGH + \beta_{13} WATER + \varepsilon_i$$

The probit model assumes that Z_i^* is a normally distributed random variable, so that the probability of $Z_i^* \leq Z_i$ can be calculated from the cumulative normal probability function, which in its standard form can be written as (Pindyck and Rubinfeld, 1998):

$$P_i = F(Z_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_i} e^{\frac{-s^2}{2}} ds$$

where s is a random variable that is normally distributed with a zero mean, and unit variance. P_i represents the probability that $Y =$ becoming less poor. The higher the value of Z_i, the more likely an event is to occur. To obtain an estimate of the index Z_i, the inverse of the cumulative normal function is applied:

$$Z_i = F^{-1}(P_i) = \alpha + \beta X_i$$

The probit model uses the same maximum likelihood approach to estimate the parameters as the logit model.

Although the assumptions around the error term and the formula used for prediction are different for Logit and Probit models, after conversion the two models give similar results (Baum, 2006). However, the analysis for this dissertation is based on the marginal effects or the change in the probability of an event occurring as a result of a unit change in the value of the regressor, *ceteris paribus*.

7.4 Results and discussion

7.4.1 Ordinary least squares (OLS) regression

Estimating a binary-dependent variable model is not impossible, even though it is not preferable, because of the difficulty in interpreting the coefficients. The difficulties with using OLS to estimate a binary-dependent variable has been discussed in Chapter 6. The most significant of these is that OLS estimation on a binary-dependent variable fits a linear line through the observations even though the probabilities of the slope coefficient are not restricted to lie between 0 and 1. Despite these problems, OLS can be used as a sound method for screening variables to determine the significance as well as the direction of the effect that the independent variables have on the dependent variable. Table 7.1 includes the results from both 2001 and 2004.

Table 7.1 OLS regression results 2001 and 2004

Explanatory variable	2001		2004	
	Coefficient	t-ratio	Coefficient	t-ratio
Constant	0,8038	34,27	0,884	31,43
Demographic variables:				
Status				
Employed				
Unemployed	-0,139	-10,73	-0,179	-13,54
Age	-0,001	-4,75	-0,001	-2,52
Population				
Black African				
Coloured	0,025	-2,26	-0,005	-0,51
Indian	0,080	3,57	0,094	4,09
White	0,069	5,33	0,102	7,73
Gender				
Male				
Female	-0,070	-9,74	-0,051	-7,17
Education	0,019	19,86	0,021	21,62
Income				
Salaries and Wages				
Remittances	-0,117	-7,49	-0,091	-5,78
Pensions and grants	-0,129	-6,83	-0,239	-16,19
Sales of farm products	-0,122	-4,46	-0,117	-4,24
Other non-farm income	-0,145	-10,07	-0,135	-9,88
No income	-0,168	-6,23	-0,100	-3,02
Infrastructure variables:				
House				
Formal house				
Traditional dwelling	-0,040	-2,81	-0,085	-6,52
Informal house	0,005	0,51	-0,018	-1,63
Cellular telephony				
Access				
No access	-0,051	-6,60	-0,012	-1,71
Telecommunications				
Access				
No access	-0,046	-4,87	-0,055	-5,42
Rubbish removal				
Removal by local authorities				
Removal by community members	-0,034	-1,03	0,013	0,41
Communal dump	-0,052	-5,75	-0,019	-2,14
None	-0,049	-3,28	-0,051	-3,06
Toilet				
Access to formal toilet				
Access to toilet not in a dwelling	-0,087	-9,24	-0,081	-8,53
No toilet	-0,122	-7,95	-0,072	-4,74
Electricity				
Access to formal electricity from mains or solar power				
Traditional sources	-0,058	-3,67	0,015	0,87
Other sources	-0,039	-3,80	-0,015	-1,40
Water				
Access to piped water inside the dwelling				
Access to piped water outside the dwelling	-0,038	-3,61	-0,073	-6,92
Access to informal water	-0,091	-5,96	-0,109	-7,07
Number of observations	16593		16137	
F (25, 16567)	203,68		220,38	
Prob > F	0,0000		0,0000	
R-squared	0,2351		0,2548	
Adjusted R squared	0,2339		0,2537	
Root MSE	0,41344		0,41365	

Source: Stats SA 2001; 2004a. Stata output.

7.4.1.1 Discussion of results

As mentioned, the coefficients obtained from a binary-dependent variable using OLS estimation (LPM) is difficult to interpret. Results from the OLS model show that all the slope coefficients have the correct signs (see section 7.4.5.1 and 7.4.5.2). The coefficients are interpreted as, all other things being equal, a one percentage point increase in the explanatory variable will lead to a decrease or increase in the probability that the dependent variable will be equal to 1 (i.e., less poor), with the value of the coefficient (β). As an example, households where the household head is employed, keeping all other things constant, a one percentage point increase for the number of household heads that are employed, will lower the probability of a household being poor by 0,14.

7.4.1.2 Evaluation of the model

Table 7.1 reports the results obtained from using OLS regression procedure with 13 independent variables and one dependent variable. The adjusted R-squared value in 2001 was 0,23, which indicates that around 23 per cent of the variation in the dependent variable is explained by the explanatory variables. As mentioned in Chapter 6, R^2 measures do not provide very useful information in binary-dependent variable models. Overall, the probability of rejecting the F value is smaller than 0,05, which means that the model is overall significant. Except for two variables, informal housing and rubbish removal by community members, all explanatory variables were significant at the 95 per cent confidence level. The two variables that were found to be not significantly different from the base do not have a large impact on the R square when excluded from the regression. In 2004 the adjusted R square is 0,25, and the probability of rejecting the F value is smaller than 0,05. In 2004 four variables were not significant at the 95 per cent confidence level. These were traditional and other sources of energy, rubbish removal by community members, coloured population group. Informal housing and access to cellular telephones were close enough to the critical values to be accepted. The low R-square

value is to be expected in the case where the dependent variable is binary variable (see Pindyck and Rubinfield, 1998).

7.4.2 Results from the probit model

An alternative to OLS estimation is using maximum likelihood estimation to estimate the poverty model. With linear probability regression, the marginal effects of the exogenous variables are constant. Empirically, one cannot tell which model fits the data best (Golder, 2007). As discussed in the previous chapter, probit models assume a normal distribution function, while logit models assume a logistic distribution. The estimation results of the probit (see Appendix 1) and logit are so much the same, that only the results from the logit will be analysed. In binary-dependent variable models the R^2 is not likely to be close to 1. However, measures similar to R^2 called pseudo R^2 can be used to measure the goodness of fit, even though this measure is still of secondary importance. The results from the probit model are presented in Appendix 1 as well as the marginal effects of the slope coefficients on the probability that a household will be less poor.

7.4.3 Results from the logit model

The logit model applies a curved line to the observations by assuming a logistic cumulative distribution function. The results of the slope coefficients are now bounded to lie between 0 and 1. Owing to the simplicity of this model, the logit model is often preferred for estimating a binary-dependent variable. The results of the logit model are presented in Table 7.2.

Table 7.2 Logit regression results 2001 and 2004

	2001				2004			
Explanatory variable	**Coefficient**	**z-ratio**	**Odds ratio**	**Comparable probit coefficients**	**Coefficient**	**z-ratio**	**Odds ratio**	**Comparable probit coefficients**
Constant	1,787	12,82			2,146	12,95		
Demographic variables								
Status								
Employed								
Unemployed	-0,735	-10,45	0,48	-0,459	-0,926	-12,52	0,396	-0,579
Age	-0,009	-4,62	0,991	-0,006	-0,005	-2,38	0,995	-0,003
Population								
Black African								
Coloured	-0,161	-2,54	0,851	-0,101	-0,055	-0,9	0,946	-0,034
Indian	1,331	5,05	3,785	0,832	1,143	4,89	3,136	0,714
White	2,665	10,02	14,368	1,666	2,552	11,21	12,833	1,595
Gender								
Male								
Female	-0,389	-9,52	0,678	-0,243	-0,287	-6,97	0,751	-0,179
Education	0,099	17,93	1,104	0,062	0,108	19,3	1,114	0,068
Income								
Salaries and wages								
Remittances	-0,514	-6,17	0,598	-0,321	-0,371	-4,39	0,69	-0,232
Pensions and grants	-0,596	-5,73	0,551	-0,373	-1,197	-13,88	0,302	-0,748
Sales of farm products	-0,871	-5,06	0,419	-0,544	-0,875	-4,72	0,417	-0,547
Other non-farm income	-0,753	-9,48	0,471	-0,471	-0,71	-9,33	0,492	-0,444
No income	-0,8	-5,46	0,449	-0,500	-0,401	-2,22	0,67	-0,251
Infrastructure variables:								
House								
Formal house								
Traditional dwelling	-0,163	-2,13	0,85	-0,102	-0,398	-5,62	0,672	-0,249
Informal house	0,009	0,16	1,009	0,006	-0,102	-1,75	0,903	-0,064
Cellular telephony								
Access								
No access	-0,384	-7,87	0,681	-0,240	-0,086	-2,1	0,918	-0,054
Telecommunications								
Access								
No access	-0,385	-6,07	0,68	-0,241	-0,494	-6,83	0,61	-0,309
Rubbish removal								
Removal by local authorities								
Removal by community members	-0,218	-1,16	0,804	-0,136	0,054	0,27	1,055	0,034
Communal dump	-0,274	-5,46	0,76	-0,171	-0,119	-2,3	0,888	-0,074
None	-0,245	-3,08	0,783	-0,153	-0,263	-2,92	0,769	-0,164
Toilet								
Access to formal toilet								
Access to toilet not in a dwelling	-0,472	-8,7	0,624	-0,295	-0,418	-7,73	0,658	-0,261
No toilet	-0,61	-7,32	0,543	-0,381	-0,361	-4,31	0,697	-0,226
Electricity								
Access to formal electricity from mains or solar power								
Traditional sources	-0,25	-2,99	0,779	-0,156	0,083	0,9	1,087	0,052
Other sources	-0,147	-2,73	0,863	-0,092	-0,039	-0,69	0,962	-0,024
Water								
Access to piped water inside the dwelling								
Access to piped water outside the dwelling	-0,162	-2,87	0,85	-0,101	-0,319	-5,65	0,727	-0,199
Access to informal water	-0,401	-4,95	0,67	-0,251	-0,498	-6,01	0,608	-0,311

Sources: Stats SA 2001; 2004a. Stata output.

7.4.4 Discussion of the coefficient and the odds ratios

The coefficients of the logit model are difficult to interpret. However, they can be explained as follows: a one-unit increase in the explanatory variable will lead to a log odds increase or decrease in the dependent variable. To compare the coefficients from the logit with probit, as discussed in Chapter 6, the logit coefficient is multiplied by 0,625 to obtain a comparable probit coefficient (see Appendix 1).

By taking the antilog of the coefficients, that is, e^{β_i}, the odds ratio can be calculated, as shown in Table 7.2. These ratios provide more interpretable information. Because there is more than one regressor in the model, 1 is subtracted from the odds ratio and then multiplied by 100 to get the percentage change in the odds (Gujarati, 2003). The results are discussed below and divided into demographic and infrastructure variables.

7.4.4.1 Demographic variables

In households where the household head was unemployed, the odds of becoming less poor decreased by 0,52 or 52 per cent in 2001. In 2004 these odds decreased by 60,4 per cent.

As household heads become older, the odds of becoming less poor decrease by 0,9 per cent in 2001. These odds decreased by 0,5 per cent in 2004.

In households where the household head was coloured, compared to black Africans, the odds of becoming less poor decreased by 14,1 per cent in 2001, and improved in 2004 by only decreasing the odds by 5,4 per cent. In households where the household head was Indian, compared to black Africans, the odds of becoming less poor increased by 278,5 per cent in 2001, and in 2004 the odds increased by 213,6 per cent. In households where the household head was white, compared to black Africans, the odds of becoming less poor increased 1 336,8 per cent. In 2004, the odds only increased by 1 183,3 percent.

In households where the household head was female, compared to males, the odds of becoming less poor decreased by 32,2 per cent in 2001. In 2004 these odds decreased only by 24,9 per cent.

In households where the household heads had one additional year of education, the odds of becoming less poor increased by 10,4 per cent in 2001. These odds increased by 11,4 per cent in 2004.

In households where the household head's main income was remittances, compared to households whose main income was salaries and wages, the odds of becoming less poor decreased by 40,2 per cent in 2001, and decreased by 31 per cent in 2004. In households where the household head's main income was pension and grants, compared to salaries and wages, the odds of becoming less poor decreased by 44,9 per cent in 2001, and decreased by 69,8 per cent in 2004.

7.4.4.2 Infrastructure Variables

The odds that a household will become less poor when they live in a traditional dwelling decreased by 0,15 or about 15 per cent in 2001. In 2004 the odds decreased by 33 percent. The odds that a household who lived in an informal house will be less poor in 2001, increased by 0,09 per cent. These odds decreased by 9,7 per cent in 2004.

In households who had no access to cellular telephones, compared to households who did have access to cell phones, the odds of becoming less poor decreased by 32 per cent in 2001. In 2004 these odds decreased by 8,2 per cent.

The odds of becoming less poor decreased by 19,6 per cent in 2001 for households who had rubbish removal by community members, compared to households who had rubbish removed by local authorities. In 2004 these odds increased by 5,5 per cent. In households with a communal dump, the odds of becoming less poor decreased by 24 per cent in 2001, compared to households who had rubbish removal by local authorities. In 2004 the

odds decreased by 11,2 per cent. The odds for households with no refuse removal decreased by 21,7 per cent in 2001, compared to 23,1 per cent in 2004.

For households who had access to toilets not in a dwelling, compared to those who did have access, the odds of becoming less poor decreased by 22,1 per cent in 2001 and 34,2 per cent in 2004. In households with no toilet facilities, compared to households with toilet facilities, the odds of becoming less poor decreased by 45,7 percent in 2001, and 30,3 per cent in 2004.

In 2001 the odds for becoming less poor decreased by 22,1 per cent for households who made use of traditional household fuel sources instead of electricity, compared to households using electricity. These odds increased by 8,7 per cent in 2004.

In households who had access to piped water outside the dwelling, compared to households who had access to piped water inside the dwelling, the odds of becoming less poor decreased by 15 per cent in 2001, and 27,3 per cent in 2004. For households who had access to informal water, the odds of becoming less poor decreased by 33 per cent in 2001 and decreased by 39,2 per cent in 2004.

Although these odds ratios provide useful information, it is still a challenge to interpret the coefficients. Instead, the predicted probabilities or marginal effects can be calculated, which is much easier to interpret. The marginal effects of the logit and the probit are very similar (see Appendix 1). Table 7.3 presents the marginal effects for the logit model.

Table 7.3 Marginal effects of the logit model

	2001		2004	
Explanatory variable	**dy/dx**	**z ratio**	**dy/dx**	**z ratio**
Constant				
Demographic variables:				
Status				
Employed				
Unemployed	-0,155	-9,54	-0,187	-12,39
Age	-0,002	-4,61	-0,001	-2,37
Population				
Black African				
Coloured	-0,031	-2,47	-0,011	-0,89
Indian	0,172	8,34	0,172	7,23
White	0,276	30,12	0,301	30,29
Gender				
Male				
Female	-0,076	-9,16	-0,059	-6,81
Education	0,019	17,40	0,022	18,90
Income				
Salaries and wages				
Remittances	-0,106	-5,68	-0,080	-4,15
Pensions and grants	-0,126	-5,21	-0,279	-13,16
Sales of farm products	-0,193	-4,56	-0,202	-4,38
Other non-farm income	-0,162	-8,56	-0,160	-8,66
No income	-0,175	-4,91	-0,087	-2,08
Infrastructure variables:				
House				
Formal house				
Traditional dwelling	-0,032	-2,06	-0,085	-5,31
Informal house	0,002	0,16	-0,021	-1,72
Cellular telephony				
Access				
No access	-0,069	-8,18	-0,017	-2,10
Telecommunications				
Access				
No access	-0,068	-6,43	-0,093	-7,43
Rubbish removal				
Removal by local authorities				
Removal by community members	-0,043	-1,11	0,011	0,28
Communal dump	-0,053	-5,32	-0,024	-2,28
None	-0,048	-2,94	-0,056	-2,80
Toilet				
Access to formal toilet				
Access to toilet not in a dwelling	-0,088	-8,70	-0,084	-7,74
No toilet	-0,127	-6,71	-0,077	-4,10
Electricity				
Access to formal electricity from mains or solar power				
Traditional sources	-0,049	-2,85	0,017	0,92
Other sources	-0,028	-2,66	-0,008	-0,69
Water				
Access to piped water inside the dwelling				
Access to piped water outside the dwelling	-0,031	-2,80	-0,067	-5,44
Access to informal water	-0,081	-4,62	-0,109	-5,63
Number of observations	16593		16137	
LR chi2 (25)	4538,19		4739,42	
Prob > chi2	0,0000		0,0000	
Pseudo R2	0,2142		0,2256	
Log-likelihood	-8323,5002		-8136,3416	
Obs. Probability	0,75014527		0,71987406	

Sources: Stats SA 2001; 2004a. Stata output.

7.4.5 Discussion of results

The results from the logit, as well as the probit model are very similar, apart from the main difference which are that the logit model uses the logistic cumulative distribution function and the probit model uses the normal cumulative distribution function (Gujarati, 2003). Both the probit and the logit model provide better results than the LPM, because it makes provision for a non-linear relationship to exist between variables. In 2001 the continues variables, age, had a negative sign and, education, a positive sign, indicating that the higher the age of the head of the household, the probability of not being poor increased at a decreasing rate with age. The positive sign for education means that the likelihood of not being poor increased as the households head's education increased. The same is true for 2004.

7.4.5.1 Demographic variables

The coefficients for gender is negative, which means that being a female household head increases the probability of being poor by 0,07 in 2001 and 0,05 in 2004. The benchmark category for the population group is black African. The white population had a positive and significant coefficient in 2001 of 0,27 and 0,28 in 2004. This means that being white increases the probability of not being poor by 28 per cent when comparing the category with the benchmark category (*ceteris paribus*), black African-headed households. The coefficient for coloured-headed households is negative for both years, meaning that compared to the benchmark category, coloured-headed households were 0,03 more likely to be poor in 2001 and 0,01 in 2004.

The coefficients for the household's main source of income are all negative and significant at the 95 per cent confidence interval. Households where the main source of income was remittances in 2001 were more likely with a coefficient of 0,11 to poor than a household where the head's main income was salaries and wages. The same was true in 2004, with a coefficient of 0,08. The same went for households who depended on pensions and grants for income, as well as sales of farm products and those who had

other income, and those who claimed they had no income. The coefficient for pensions and grants increased from 0,12 in 2001 to 0,27 in 2004, indicating that households depending on pensions and grants as a main source of income were even more likely to be poor in 2001 than in 2004.

The sign for unemployment is negative for both years; meaning that a household head who is unemployed is more likely to be poor compared to one who is working. This makes sense, because working individuals have a better chance of having an income. The coefficient in 2001 was 0,16 and in 2004, 0,20.

7.4.5.2. Infrastructure variables

The sign of the coefficient of households who lived in a traditional dwelling was negative for both 2001 and 2004, and was significant at a 90 per cent confidence level. The coefficient for households who lived in informal houses was not significant in 2001, but was in 2004. In 2004 this coefficient was negative 0,02, whereas the coefficient for traditional dwelling was 0,03 in 2001 and 0,08 in 2004, which means that it is more likely for households who live in traditional structures to be poor than those who live in formal housing.

The coefficient for electricity from lights was significant in 2001, but not in 2004, which could mean that access to electricity as a measure of poverty could be unclear. Nevertheless, the coefficient for 2001 was negative for both traditional energy sources and other sources of energy such as wood, coal, candles and animal dung. The coefficient for traditional energy sources (generator, gas, paraffin) was 0,05 and for other sources, 0,03.

The coefficient for access to sanitation facilities was significant in both years and the sign was negative in both years. The coefficient in 2001 for a toilet on site was 0,09 and for a toilet off site or none was 0,12. In 2004 the coefficient for an on-site toilet was 0,08 and for off site it was 0,07. This could show that households who did not have access to

formal toilet facilities were less likely to become less poor, compared to households who did have access to formal toilet facilities, even though the coefficients for 2004 was lower than that for 2001.

Water had a negative significant coefficient in both years. In 2001 this coefficient for having a tap in the dwelling or on site was 0,03, and for other water sources such as public taps, dams and rivers it was 0,08. In 2004 these coefficients were 0,06 and 0,10. Water thus did have an impact on poverty, even more so in 2004. Households who only have access to alternative water sources are more likely to be poor than households who have a tap in the dwelling or on site.

Rubbish removal by community members were found to be not significant in both years, but those who had their rubbish removed by the community, had a community dump or had no rubbish removal was significant and negative. The coefficient in 2001 for those with their own refuse dump or communal container was 0,05 and households with no refuse removal was also 0,05. In 2004 these probabilities were 0,02 and 0,05 respectively. The probability of households being less poor, when they had their rubbish removed by community members decreased by 0,04 in 2001. However it increased by 0,01 in 2004.

Households who did not have access to a landline and/or a cellular telephone had a negative significant sign and coefficients for both years. In 2001 the coefficient for a landline and a cellular telephone was 0,07 and in 2004 a landline was 0,09 and a cellular telephone 0,01. This means that having no form of telecommunication increases the probability of being poor. This probability was much higher in 2004 than in 2001 for a landline.

7.5 Evaluation of the model

As mentioned in chapter 6, the traditional measures of the goodness of fit, that is. the R^2, does not provide useful information in binary-dependent variable models. A range of similar measures called 'pseudo R^2' can be used to evaluate models such as the probit

and the logit model. The first of these measures is the McFadden R^2, which ranges between 0 and 1. The pseudo R^2 for the logit model is 0,23, which means that a large part of the variation is not explained by the model. It should be noted, however, that in binary-dependent variable model, the goodness of fit is not as important as the signs and the size of the coefficients (Gujarati, 2003). The count R^2 is 0,74, which means that a large number of correct predictions were made in relation to the number of observations. The LR statistic for the logit model is large and the probability is small, meaning that, together, the demographics as well as the infrastructure variables significantly explain the changes in poverty.

7.6 Conclusion

This chapter started off by identifying the dependent variable as a binary variable, where households living in poverty based on their per capita household expenditure were coded 0 and those who were not poor were coded 1. A range of independent variables were then identified as explanatory variables of poverty. These were divided into demographic and status variables, such as age, gender, education, employment status, population group and infrastructure variables. The aim was to determine whether infrastructure can contribute to alleviate poverty. LFS data for South Africa were used.

Firstly, the OLS estimation was used to estimate an LPM to determine the slope coefficients of the explanatory variables. The results showed that around 23 per cent of the change in poverty is explained by the model. The F statistic was also significant, showing that the model is significant overall. However, owing to the problem of linearity, the estimated slope coefficients are not bounded by 0 and 1, and therefore results can be difficult to interpret. After determining that a linear model was not the best approach to estimate a binary-dependent variable model, a probit and logit model was estimated using maximum likelihood estimation. The logit coefficients were adjusted by a value of 0,625 to make the coefficients comparable to the probit coefficients. It was found that the logit and the probit coefficients were so similar, and therefore the most simplistic model, and the logit model was discussed in further detail. The odds ratios of the logit model were

then calculated. However, it was found that it is difficult to interpret the odds ratios and therefore the marginal effects were calculated for each coefficient. The results showed that, overall, the logit model was a good model to use for estimating poverty, because a large F statistic can be interpreted as a model that is overall significant. Furthermore, most of the explanatory variable coefficients were significant at the 95 per cent confidence level.

The results from the marginal effects or predicted probabilities showed that demographic variables such as population group, age, gender, education and employment status in general, have a significant effect on whether a household will be poor or not. However, for the purpose of this dissertation, attention is shifted towards the impact that the infrastructure variables have on poverty. The results showed that infrastructure variables significantly contributed to the probability that the household will become less poor, should they have access to these services. Households who did not have access to formal housing were less likely to become less poor, compared to households who did have formal housing. It was also found that households who had no access to communication infrastructure were also less likely to become less poor compared to those who had access to communication infrastructure. Rubbish removal by community members was found to be insignificant in both years. The coefficients for households who had their rubbish removed by the community, and/or who had a communal dump was significant and negative, meaning that households who did not have access to formal refuse removal was less likely to become less poor, compared to households who had their rubbish removed by local authorities. The coefficient for access to sanitation facilities was significant in both years and the sign was negative in both years. This could show that households who did not have access to formal toilet facilities were less likely to become less poor, compared to households who did have access to formal toilet facilities. The coefficient for electricity from lights was significant in 2001, but not in 2004, which could mean that access to electricity as a measure of poverty could be unclear. Nevertheless, the coefficient for 2001 was negative for both traditional energy sources and other sources of energy such as wood, coal, candles and animal dung. Water had a negative significant coefficient in both years. Water thus did have an impact on poverty, even more so in

2004. Households who only have access to alternative water sources are more likely to be poor than households who have a tap in the dwelling or on site.

After estimating the probabilities, it can be concluded that infrastructure does play an important roll in significantly lowering the probability of poverty. Of the demographic variables, population group and income had a greater impact on predicting the probability of poverty, compared to other variables. Most of the infrastructure variables were significant, and the predicted probabilities successfully showed that infrastructure plays an important role in alleviating poverty in South Africa.

Chapter 8: Conclusion and Recommendations

8.1 Conclusion

The purpose of this study was to determine whether the current infrastructure delivery rate is adequate to alleviate poverty in South Africa. To achieve this goal, an empirical study of the impact of infrastructure on poverty was conducted by using a logit model to estimate the expected probabilities that a person will become less poor given a set of explanatory variables.

In an aim to alleviate poverty, government allocates large parts of its budget towards providing infrastructure, especially to poor households in South Africa. It was therefore important to establish to what extent, if any, infrastructure delivery contributed to lessening poverty in South Africa. To determine this impact, it was important to firstly distinguish between the poor and the non-poor.

A review of the literature on how to measure poverty proved useful in setting a platform for selecting a poverty measure that best represented poor households in South Africa. Chapter 2 looked at various poverty measures, implemented domestically and internationally. It was shown that in a global context poverty is often measured in broad terms such as the GNI per capita, HDI and HPI. One of the major drawbacks of these measures is that they do not provide any information on the distribution of wealth in a country. Countries therefore started to employ more comprehensive measures of poverty, where information is collected from household surveys.

The next measurement challenge that was dealt with was whether a monetary or non-monetary poverty measure should be used. When a monetary measure, such as an income or expenditure measure is used, it was found that the process of selecting a monetary cut-off point, which best represents the poor, could be subjective. Instead of using average monetary expenditure measures, expenditure can be adjusted to take household size and

composition into account, that is, an expenditure-per-adult-equivalent measure, which adjusts for the amount of adults and children in the household. The literature shows, however, that these adjustments do not necessarily provide better results. Chapter 2 also looked at how to establish an official poverty line for South Africa, looking at both monetary and non-monetary poverty measures.

After establishing how the poor would be identified, Chapter 3 provided an overview of how information such as income and expenditure and other characteristics could be collected from households in a population. To collect accurate and useful information, a clear understanding of the various data collection methods and tools was required. To achieve this goal, Chapter 3 started off by looking at the importance of establishing a well-formulated research question, it was shown that before a question can be formulated, a researcher should be well acquainted with the facts required to meet the purpose of the study. Following the process of defining a question, a sample had to be designed from where the information would be collected. Sampling selection was divided into 5 steps, of which the first step was to define the target population. The second step was to choose a sampling frame, employing either probability sampling or non-probability sampling selection. The third step discussed the planning procedure for sampling unit selection. Fourthly it was discussed how to determine the sample size, which mainly depended on the sampling error that a researcher was willing to make. The last step subdivided the population into primary sampling units from which the sample was selected.

The following main data collection methods, together with their strengths and weaknesses, were distinguished: focus groups, interviews, monitoring, surveys and follow-up surveys/interviews. The questionnaire design and testing were then divided into three phases. Phase 1 looked at the type of data that were required to answer the research question. Phase 2 looked at how a questionnaire should be designed, while phase 3 discussed the testing of the instrument, that is, the pilot study.

The next section looked at how to train interviewers to use the survey instrument optimally and how to conduct fieldwork. The LFS, as conducted by Stats SA, uses a

collection of the methods discussed above to collect data from households. The last part of this chapter looked at the LFS design and use as a data collection tool in South Africa.

Chapter 4 implemented some of the poverty measures discussed in Chapter 2 by making use of information collected by the LFS to identify the poor in South Africa. Chapter 4 started off by identifying household expenditure as a monetary indicator to measure poverty. Households' expenditure was then adjusted for household size, assuming that expenditure was shared equally between household members. Based on this measure, households were then divided into quintiles. Individuals in the bottom 20th and 40th quintile was then classified as poor, and those in the 60th, 80th and 100th quintile was classified as non-poor. The results showed that around 12,5 million people lived in households that fell within the bottom 20 per cent of households based on their per capita household expenditure. In 2004 this number increased to 13,3 million. In the bottom 40th quintile the share of poor people decreased from 54,6 to 54,4 per cent.

This chapter also briefly looked at how wealth was distributed in South Africa. It was shown that the Gini coefficient in 2004 was 0,64, unchanged from 2001. Over this period however, inequality between population groups worsened. Returning to those households classified as poor and non-poor, chapter 4 analysed various demographic, labour market characteristics, unemployment and household livelihoods for each of the two groups.

The results obtained from the comparison between the demographic information on poor and the non-poor showed that black African females made up the largest part of those in the 20th and 40th quintiles. Furthermore, the people in the bottom 20th and 40th quintiles were mostly located in the Eastern Cape and KwaZulu-Natal Provinces, while those in the 100th quintile were mostly located in Gauteng Province. Most of the poor (those in the 20th and 40th quintile) were aged between 15 and 25 years.

Labour market information on the poor showed that most of the poor were employed in the informal sector, while industry information showed that, most of the poor were employed in the categories wholesale and retail trade sales, and agriculture, hunting,

forestry and fishing. Of the poor, most had less than a Matric as their highest qualification, while those in the 100th quintile mostly had a Matric as their highest qualification. Furthermore, the poor were mostly unemployed for periods of three years or longer. The poor were also very dependent on grants, especially the child support grant, as a means of income. The poor also mostly relied on enquiring at workplaces, farms and factories or called on other possible employers to find work.

Chapter 5 investigated how infrastructure development affected the poor. This chapter first looked at problems in defining infrastructure delivery. It was shown that infrastructure delivery was a process that involved careful planning in terms of financial resources, recovery mechanisms and long-term feasibility and high-quality project management. Development and poverty alleviation could only take place once a link between infrastructure and poverty was established. It becomes difficult to determine which infrastructure can be classified as adequate/non-poor infrastructure and which can-not. Infrastructure delivery can have both positive and negative effects on poverty. It was therefore important to look at to what extent infrastructure development benefited the poor. LFS data were used as a tool to determine the level of access to infrastructure for the poor. The following infrastructure deliverables were investigated: housing, water and sanitation, electricity, communication, financial infrastructure, and transport infrastructure. Consideration was also given to the policy and financial environment in which infrastructure delivery should take place.

The results for access to formal housing showed that of those households in the bottom 20th quintile approximately 30 per cent lived in traditional dwellings/huts or structures made of traditional materials. Only around 24 per cent of those in the bottom 20 per cent of households had access to a toilet facility off site, and of the same 20 per cent of households, 29 per cent had to make use of a public tap for water. The challenges with electricity supply were also highlighted in this chapter. Around 65 per cent of those in the bottom 20 per cent of households had access to electricity for lighting. Furthermore, an overview of the financial sector in terms of service delivery to the poor was also provided.

Lastly, this chapter stressed the importance of government in infrastructure delivery, even though the process had been slow and challenging. Improvement in delivery was evident across the various infrastructure services; however, there was still a large part of the South African population who did not have access to adequate services. As proved in this chapter, data from the LFS showed that many of those who are poor are also those who lack basic infrastructure. Should government succeed in providing infrastructure to these poor households, the question arose as to what extent this service delivery would contribute to alleviating poverty. To answer this question, given the LFS data available, econometric modelling became useful.

Econometric models can be very useful in simulating survey data, such as LFS data discussed in Chapter 5. Chapter 6 provided a theoretical overview of models that could be used when modelling survey data. Data collected through surveys were usually in categorical format and therefore the modelling technique of a binary-dependent variable was discussed. It was established that for various problems, such as that the error term was not normally distributed, predictions were not bounded to 0 and 1, errors were highly heteroscedastic and dubious R^2 values, the LPM model using OLS estimation did not provide the best results for a binary-dependent variable model. Two alternative models, the logit and probit, were then introduced. These models make use of maximum likelihood estimation and solve many of the problems associated with OLS estimation.

The difference between the logit and the probit model was shown to lie in their cumulative distribution function. Logit models assumed a logistic cumulative distribution function, while probit models assumed a normal cumulative distribution function. The calculation of the marginal effects was also discussed for both the logit and the probit models. In order to determine how well the data fitted the model, various evaluation measures were discussed. The R^2 and adjusted R^2 were not very useful as a measure of fit in maximum-likelihood estimation because in using binary-dependent variable models, the R^2 was likely to lie between 0,2 and 0,6. More suitable measures, referred to as 'pseudo R^2', which consist of the McFadden R^2, the count R^2 and the LR statistic, were

also examined. It was concluded that the logit model was the more preferred model to use in practice because of its simplicity.

Econometric modelling can be useful in simulating survey data. Chapter 7 made use of a micro-econometric model, using LFS data, to determine to what extent service delivery would contribute to alleviating poverty in South Africa. Based on the work done in chapter 4, the dependent variable was identified as a binary variable, where households living in poverty, were coded 0, while the non-poor were coded 1. A range of demographic and status variables were identified as explanatory variables. These were age, gender, education, employment status and population group. A group of infrastructure variables was then identified as explanatory variables. The aim of the model was to determine whether infrastructure could contribute towards alleviating poverty in South Africa.

Firstly, OLS estimation was tested by estimating an LPM. The results showed that around 23 per cent of the change in poverty was explained by the model. The overall significance of the model was good as indicated by the large F-statistic. However, due to the linearity problem, as discussed in chapter 6, the estimated slope coefficients were not bound by 0 and 1, making interpretation difficult. A different estimation method called maximum likelihood estimation was then applied, using both a logit and probit model for estimating the coefficients of the explanatory variables. To make the logit results comparable to the probit results, the logit models coefficients were adjusted by a value of 0,625. Because the only difference between the logit and the probit model was their cumulative distribution the results from the logit and the probit were more or less the same. The focus was then turned to the logit model, and the probit models' results were provided in appendix 1.

The logit model was then discussed in more detail, and the odds ratio were calculated. This calculation provided the percentage change in the odds of a person becoming less poor, given the set of explanatory variables. To answer the research question it was necessary to calculate the marginal effects, or in other words, the predicted probabilities

for each explanatory variable. The results showed that demographic explanatory variables such as population group, age, gender, education and employment status in general have a significant effect on whether a household would be less poor. The results from analysing the predicted probabilities of the infrastructure variables provided evidence that infrastructure plays an important role in alleviating poverty. Households who did not have access to formal housing were less likely to become less poor, compared to households who did have access to formal housing. It was also found that households who had no access to communication infrastructure were more likely to be poor compared to those who had access to communication infrastructure. Households who did not have access to formal refuse removal were less likely to become less poor, compared to households who had their rubbish removed by local authorities. Households who did not have access to formal toilet facilities were less likely to become less poor, compared to households who did have access to formal toilet facilities. Access to electricity as a measure of poverty was unclear due to the positive coefficient estimated in 2004 for traditional sources of electricity. In 2001 households who had access to both traditional energy sources and other sources of energy such as wood, coal, candles and animal dung were less likely to be less poor. Households who only had access to alternative water sources are more likely to be poor than households who have a tap in the dwelling or on site.

8.2 Recommendations

The data used for estimating the logit model was for 2001 and 2004. If more current data can be used, the logit model can be estimated to provide more updated results.

Different poverty measures can also be applied to estimate the logit model instead of the per capita household expenditure measures. These measures could include adjusting expenditure not only for household size, but also for household composition. Furthermore, income as a monetary poverty measure can be used in the place of expenditure. A more stringent poverty line can also be used to test how infrastructure impact on the very poor, for example only households located in the 20th quintile can be

classified as poor. This study can also be done for the different provinces in South Africa, considering each province's unique infrastructure situation and demographic demarcation.

References

Adato, M., Carter, M. and May, J. (2004). "Sense in sociability? Social exclusion and persistent poverty in South Africa." South African Regional Poverty Network. Online: http://www.sarpn.org.za/documents/d0001127/P1246-sense_in_sociability_Dec2004.pdf (accessed 01 October 2008).

African National Congress. (1997). *ANC 50th National Congress Report.* African National Congress. Online: http://www.anc.org.za/show.php?doc=ancdocs/history/conf/conference50/resolutions3.html#Infrastructure (accessed 18 October 2008).

Aliber, M. (2002). *Overview of the incidence of poverty in South Africa for the 10 year review*. Online: http://www.sarpn.org.za/documents/d0000875/docs/OverviewOfTheIncidenceOfPovertyInSAforThe10yearReview.pdf (accessed 03 April 2007).

ANC *see* African National Congress.

Ardington, C., Lam, O., Leibbrandt, M. and Welch, M. (2005). *The sensitivity of estimates of post-apartheid changes in South African poverty and inequality to key data imputations*. CSSR Working Paper No. 106. Cape Town: SALDRU. Online: http://www.sarpn.org.za/documents/d0001148/P1265-South_African_poverty_Febr2005.pdf (accessed 21 October 2008).

Ardington, C., Lam, D., Leibbrandt, M. and Welch, M. (2006). *The sensitivity to key data imputations of recent estimates of income poverty and inequality in South Africa*. Online: http://www.tips.org.za/files/forum/2006/papers/TheSensitivitytoKeyDataLeibbrandtLamArdingtonWelch.pdf (accessed 21 October 2008).

Ardington, C. and Leibbrandt, M. (2004). *Financial Services and the Informal Economy.* African development and poverty reduction: the macro-micro linkages. Cape Town: Development Policy Research Unit. Online: http://www.commerce.uct.ac.za/

Research_Units/DPRU/DPRU-Conference2004/Papers/ Financial Services Ardington.pdf (accessed 02 October 2008).

Ashong, N. (2007). "Energy, economic growth, and poverty reduction: the linkages". Background paper for F4D Conference, Accra, May 2007. Centre for Policy Analysis (CEPA).

Asraf, R. M. and Brewer, J. K. (2004). "Conducting tests of hypotheses: the need for an adequate sample size". The Australian Educational Researcher, Volume 31, No. 1, April. Australia: Australian Association for Research in Education.

Audience Dialogue. (2006). "Know your audience". Online: http://www.audiencedialogue.org/ (accessed 08 October 2007).

Auerbach, A. J. and Feldstein, M. (1987). *Handbook of public economics*. Volumes. I–II, Amsterdam: Elsevier.

Barradas, S. (2006). "Taxi recapitalisation programme", *Engineering News*. 08 December. Online: http://www.engineeringnews.co.za/article.php?a_id=40972 (accessed 22 September 2008).

Baum, C. F. (2007). *Limited dependent variables and sampling selection.* Online: http:// fmwww.bc.edu/ec-c/s2007/327/EC327.s2007.nn6.pdf (accessed 23 July 2008).

Bhorat, H., Naidoo, P. and van der Westhuizen, C. (2006). *Shifts in non-income welfare in South Africa: 1993–2004*. Cape Town: Development Policy Research Unit.

Bhorat, H., van der Berg, S. and van Aardt, C. (2003). "Breaking the grip of poverty and inequality in South Africa 2004–2014: Current trends, issues and future policy options". Coordinated by Landman, J. P. Online: http://www.sarpn.org.za/documents/d0000649/ P661-Povertyreport3b.pdf (accessed 08 October 2008).

Bloch, G. (2007). "Building education beyond crisis: development today". Development Bank of Southern Africa. Online: http://www.dbsa.org/Research/Documents/Building%20Education%20Beyond%20Crisis.doc (accessed 08 October 2008).

Brack, C. and Wood, G. (1998). "Introduction to forest sampling, measurement of the DBH, height, and the equipments used in forest inventory, sampling techniques and errors in forest inventory". Forest Mensuration Resources. Department of Forestry, School of Resource and Environmental Management, Australian National University, Canberra. Australia. Online: http://sres.anu.edu.au/associated/mensuration/BrackandWood1998/INDEX.HTM (accessed 08 October 2007).

Brook, P. and Smith, W. (2001). *Improving access to infrastructure services by the poor: institutional and policy responses.* World Bank: Washington DC.

Brüderl, J. (2003). "Applied Regression Analysis using STATA". Mannheim: University of Mannheim, Germany. Online: http://www.sowi.unimannheim.de/lehrstuehle/lessm/veranst/AppliedRegression.pdf (accessed 08 October 2007).

Budlender, D. (2003). "The social and human development context". Chapter 7 in *Human Research Development Review 2003*. Cape Town: Human Science Research Council.

Burns, C. (2004). *A review of the literature describing the link between poverty, food insecurity and obesity with specific reference to Australia.* Victoria: VicHealth. Available: http://www.vichealth.vic.gov.au/assets/contentFiles/Literature%20Review%20 Poverty_Obesity_Food%20Insecurity.pdf (accessed 21 October 2008).

Caldwell, J. G. (2007). "Vistas approach to sample survey design". Online: http://www.foundationwebsite.org (accessed 08 October 2007).

Carter, M. R. and May, J. (1999). *Getting ahead or falling behind? The dynamics of poverty in post-apartheid South Africa*. Madison: Department of Agricultural and Applied Economics. University of Wisconsin.

Casale, D. and Desmond, C. (2007). The *economic well-being of the family: Household's access to resources in South Africa, 1995–2003*. Cape Town: HSRC Press. Online: http://www.hsrc.ac.za/Document-2279.phtml (accessed 08 October 2008).

Chromy, J. R. and Abeyasekera, S. (2005). *Household surveys in developing and transition countries: design, implementation and analysis*. North Carolina: Research Triangle Park and Reading: The University of Reading.

City of Johannesburg. (2003). *Regional Spatial Development Framework.* Online: http://www.joburg-archive.co.za/2003/budget/idp/annex42.pdf (accessed 22 September 2008).

Cochran, C. (2007). "Data collection (2): quantitative questionnaire design". Aberdeen: Health Services Research Unit. University of Aberdeen. Online: aberdeen.ac.uk/public_health/course-materials/documents/ClaireLectureV4.ppt (accessed 07 October 2008).

Coetzee, G. and Cross, C. (2002). *The role of community banks in South Africa: can it contribute to improve access to financial services for the poor?* Pretoria: Department of Agricultural Economics. Extension and Rural Development. University of Pretoria.

Commission on Poverty. (2006). W*ork incentives for working poor – Transport support for those living in remote areas.* Hong Kong: Commission on Poverty. http://www.cop.gov.hk/eng/pdf/CoP%20Paper%205-2006(e)%20.pdf (accessed 21 October 2008).

Congress of South African Trade Unions, South African Council of Churches and South African Non-Governmental Organisation Coalition. (2005). *Peoples budget 2005/2006*. Available: http://www.sacc-ct.org.za/PB2005.pdf (accessed 10 August 2007).

Congress of South African Trade Unions, South African Council of Churches and South African Non-Governmental Organisation Coalition. (2006). *Peoples budget 2006/2007***.** Available: http://www.sacc-ct.org.za/PB2006.pdf (accessed 12 October 2008).

Cook, D., Dixon, P., Duckworth, W. M., Kaiser, M. S., Koehler, K., Meeker, W. Q. and Stephenson, R. W. (2001). "Binary response and logistic regression analysis". Chapter 3 in *Beyond traditional statistical methods*. Iowa: Iowa State University.

Cooper, D. R. and Schindler, P. S. (2003). *Business research methods***.** New York: McGraw-Hill.

Cosatu, SACC and SANGOCO *see* Congress of South African Trade Unions, South African Council of Churches and South African Non-Governmental Organisation Coalition.

Deaton, A. and Muellbauer, J. (1980). *Economics and consumer behaviour*. New York, Cambridge University Press.

Democratic Alliance. (2005). *The rot in ANC municipalities: five case studies of cronyism, corruption and ineptitude***.** Online: http://www.da.org.za/da/Site/Eng/ campaigns/DOCS/SA_Awards4.doc (accessed 10 October 2007).

Department for International Development. (2005). *Reducing poverty by tackling social exclusion*. DFID policy paper. Online: http://www.dfid.gov.uk/pubs/files/social-exclusion.pdf (accessed 08 October 2008).

Department of Minerals and Energy. (1998). *White Paper on the Energy Policy of the Republic of South Africa.* Online: http://www.dme.gov.za/pdfs/energy/planning/wp_energy_policy_1998.pdf (accessed 12 October 2008).

Department of Minerals and Energy. (2003). *Electricity Basic Services Support Tariff (Free Basic Electricity) Policy.* Online: http://www.dme.gov.za/pdfs/energy/electricity/fbe_policy.pdf (accessed 02 October 2008).

Department of Minerals and Energy. (2007). *Integrated National Electrification Programme (INEP)*. Department of minerals and energy. Online: http://www.dme.gov.za/energy/elect_inep.stm (accessed 10 August 2007).

Department of Water Affairs and Forestry. (2002). *Sanitation for a healthy nation. framework for a national sanitation strategy. bringing sanitation up to speed.* Online: http://www.dwaf.gov.za/dir_ws/content/lids/PDF/Strategy.pdf
(accessed 01 October 2008).

Department of Water Affairs and Forestry. (2006). *Annual report 2005/06*. Online: http://www.dwaf.gov.za/Documents/AnnualReports/2006.asp
(accessed 09 October 2007).

DFID *see* Department for International Development.

Dias, R. (2002). "Exploring the nature of unemployment in South Africa: Insights from the Labour Force Survey 2000". University of Natal. Online: http://www.commerce.uct.ac.za/Research_Units/DPRU/Conf2003PDF/Conf2002pdf/Rosa_Dias.pdf
(accessed 28 October 2008).

Dieden, S. (2002). "Households' main income sources in the geography of poverty in South Africa". Paper presented at the Second Annual Conference on Labour Markets and Poverty in South Africa. Johannesburg 22–24 October 2002.

di Martino, V. (2002). "Workplace violence in the health sector: country case studies: Brazil, Bulgaria, Lebanon, Portugal, South Africa, Thailand and an additional Australian study". *Synthesis report: IOL*. Online: http://www.ilo.org/public/english/dialogue/sector/papers/health/violence-ccs.pdf (accessed 08 October 2008).

DME s*ee* Department of Minerals and Energy.

Dugard, J. (2001). *From low intensity war to mafia war: Taxi violence in South Africa (1987 – 2000).* Centre for the study of violence and reconciliation. Online: http://csvr.org.za/wits/papers/papvt4.htm (accessed 21 September 2008).

Durrance, J. C., Fisher, K. E. and Hinton, M. B. (2005). *How libraries and librarians help: a guide to identifying user outcomes.* Chicago: American Library Association.

du Toit, R. (2003). "Unemployed youth in South Africa: the distressed generation?" Paper presented at the Minnesota International Counseling Institute (MICI), 27 July – 1 August 2003.

DWAF s*ee* Department of Water Affairs and Forestry.

Edwards, L. and Stern, M. (2006). *Trade and poverty research project in South Africa: Lessons and policy recommendations.* Johannesburg: DPRU. Online: http://www.commerce.uct.ac.za/Research_Units/DPRU/DPRUConference2006/Papers/Trade_and_Poverty%20ResearchLawrence%20Edwards_Matthew%20Stern.pdf (accessed 07 October 2008).

Enders, P. (2003). "Applied categorical and nonnormal data analysis. Model fit". Online: http://www.gseis.ucla.edu/courses/ed231c/notes3/fit.html (accessed 08 October 2008).

Financial Sector Charter Council. (2002). *Financial Sector Charter*. Online: http://www.fscharter.co.za (accessed 01 October 2008).

Friedman, S. (2006). "Putting the service back into service delivery". *IDASA Municipal Talk news letter*. June 2006. Online: http://www.idasa.org.za/index.asp?page=output_details.asp%3FRID%3D1004%26OTID%3D4%26PID%3D25 (accessed 08 October 2008).

Frye, I. (2005). "Constructing and adopting an official poverty line for South Africa: some issues for consideration". Synthesised document by Frye and Magasela. Johannesburg: National Labour and Economic Development Institute.

Gardner, D. (2003). *Getting South Africans under shelter: an overview of the South African housing sector*. Housing Finance Resource Programme. Online: http://web.wits.ac.za/NR/rdonlyres/E46C878A-D071-434B-A401-F8EFAC6540A5/0/HousingFinanceOverview.pdf (accessed 22 October 2008).

GCIS *see* Government Communication and Information System.

Goldberger, A. S. (1964). *Econometric theory.* John Wiley and Sons: New York in Gujarati, D. N. (2003). *Basic econometrics*. 4th ed. New York: McGraw-Hill.

Goldblatt, M. and Davies, G. (2002). *Water energy and sustainable development in South Africa*. Development Southern Africa 19(3): 370–387.

Goldemberg, J., La Rovere, E. L. and Coelho, S. T. (2004). "Expanding access to electricity in Brazil". *Energy for Sustainable Development* 8(4): 86–95. Online: www.ieiglobal.org/ESDVol8No4/brazil.pdf (accessed 21 September 2008).

Golder, M. (2007). *Methods IV: advanced quantitative analysis; binary response models*. Lecture handouts. Online: http://homepages.nyu.edu/~mrg217/binaryresponse.pdf (accessed 10 October 2007).

Goldman, I., Marumo, J. M. and Toner, A. (2002). *Goodbye to projects? The institutional impacts of a livelihood approach on development interventions. An appraisal of the use of livelihoods approaches in South Africa.* Research project no. R7908. Working paper series: paper no 2. London: Department for International Development.

Government Communication and Information System. (2007). *South African Year book, 2007*. Online: http://www.gcis.gov.za/docs/publications/yearbook.htm (accessed 14 February 2008).

Grice, J. (2004). "Special focus: What's happened to free basic services?" Online: www.numsa.org.za/article.php?cat=200403-bulletin&id=86 (accessed 10 February 2007).

Gujarati, D. N. (2003). *Basic econometrics*. 4th ed. New York: McGraw-Hill.

Gush, K., Cambridge, G. and Smith, R. (2004). *The digital doorway – minimally invasive education in Africa.* CSIR. Available: http://www.digitaldoorway.org.za/publications/ICT%20in%20Education%20conference%20paper%20The%20Digital%20Doorway_March_2004.pdf (accessed 12 October 2008).

Gyekye, A. B. and Akinboade, O. A. (2001). "Analysis of poverty in the Northern province of South Africa: Implications for empowerment policy". Paper presented at the 75th Anniversary Conference of the economic society of South Africa, Glenburg Lodge, Johannesburg.

Hanmer, L., Lovell, E., Chapman, R. and Slaymaker, T. (2000). *Poverty and transport: toolkit.* A report prepared for the World Bank in collaboration with DFID June 2000. London: Oversees development institute.

Haq, M. U. (1995). *Reflections on human development*. New York: Oxford University Press.

Hassen, E. (2000). "Infrastructure delivery and employment in South Africa". A paper presented at the Forum 2000 hosted by the Trade and Industrial Policy Secretariat at Glenburn Lodge Johannesburg 18–21 September 2000. National Labour and Economic Development Institute. Online: http://www.tips.org.za/files/400.pdf (accessed 08 October 2008).

Heintz, J. and Jardine, C. (1998). "Poverty and economics in South Africa. A briefing paper". National Labour and Economic Development Institute *(NALEDI).* Online: http://www.sangoco.org.za/site/images/stories/Pdf_Documents/economics.pdf (accessed 03 September 2008).

Hemson, D. (2004). *Beating the backlog: meeting targets and providing free basic service.* Online: http://www.sarpn.org.za (accessed 22 April 2007).

Hemson, D. and Dube, B. (2004). "Water Services and Public Health: the 2000–01 cholera outbreak in KwaZulu-Natal, South Africa". Presented at the 8th World Congress on Environmental Health. 22–27 February 2004. Durban: Human Sciences Research Council. Online: http://www.hsrc.ac.za/research/output/outputDocuments/2696_Hemson_Waterservicesandpublichealth.pdf (accessed 01 October 2008).

Hemson, D. and Owusu-Ampomah, K. (2005). *A better life for all? Service delivery and poverty alleviation*, in Daniel, J., Southall, R. and Lutchman, J. (eds) *State of the nation: South Africa 2005*. Cape Town: HSRC Press.

Hirano, K. (2004). "Mass unemployment in South Africa: a comparative study with East Asia". *Institute of Developing Economies, Japan External Trade Organisation (JETRO).*

Discussion paper no. 6. Online: http://www.ide.go.jp/English/Publish/Dp/pdf/006_hirano.pdf (accessed 22 October 2008).

Hoogeveen, J. G. and Özler, B. (2004). *Not separate, not equal: poverty and inequality in post-apartheid South Africa*. Washington DC: World Bank.

HSRC *see* Human Sciences Research Council.

Human Sciences Research Council. (2003). *Human resources development review 2003*. Education, Employment and Skills in South Africa. Pretoria: HSRC Press.

Human Sciences Research Council. (2004) *Fact sheet: poverty in South Africa*. Human Sciences Research Council, 26 July 2004. Posted with the permission of Craig Schwabe, GIS Centre, Human Sciences Research Council. Online: http://www.sarpn.org.za/documents/d0000990/index.php (accessed 06 June 2007).

ICRW *see* International Center for Research on Woman.

ILO s*ee* International Labour Organisation.

International Center for Research on Women. (2005). *Infrastructure shortfalls cost poor women time and opportunity: toward achieving the third millennium development goal to promote gender equality and empower women.* Online: http://www.icrw.org/docs/2005_brief_mdg-infrastructure.pdf (accessed 21 September 2008).

International Labour Organisation. (2003). *Resolution regarding household income and expenditure statistics.* International Labour Organisation. Online: http://www.ilo.org/wcmsp5/groups/public/---dgreports/---integration/---stat/documents/normativeinstrument/wcms_087503.pdf (accessed 07 October 2008).

Jerome, A. and Ariyo, A. (2004). *Infrastructure reform and poverty reduction in Africa*. Online: http://www.commerce.uct.ac.za/Research_Units/DPRU/DPRU-Conference2004/Papers /Infrastructure_Reform_Jerome.pdf (accessed 17 May 2007).

Keller, G. and Warrack, B. (2000). *Statistics for management and economics*. 5th ed. California, USA: Duxbury Thompson Learning.

KIDS *see* KwaZulu-Natal Income Dynamics Study.

Klasen, S. and Woolard, I. (2000). "Surviving unemployment without state support: unemployment and household formation in South Africa". IZA Discussion Papers 237. Germany: Institute for the Study of Labour (IZA).

Klitgaard, R. (2004). *On infrastructure and developm*ent. Online: http://www.adb.org/Documents/Events/2004/Infrastructure_Development/First-Workshop/paper-Klitgaard.pdf (accessed 27 August 2007).

Knight, R. (2004). *A decade of democracy: housing, services and land in South Africa.* South Africa Delegation Briefing Paper, 22 April – 02 May. Availalable: http://richardknight.homestead.com/files/sihousing2004.htm (accessed 12 October 2008).

Kollamparambil, U. (2007). "*Advanced Econometrics"*. Johannesburg: Wits University presentation.

KwaZulu-Natal Income Dynamics Study. (2003). *KwaZulu-Natal Income Dynamics Study*. KwaZulu-Natal: University of KwaZulu-Natal.

Lam, D., Seekings, J. and Sparks, M. (2006). *The Cape Area Panel Study: Overview and Technical Documentation for Waves 1-2-3.* Cape Town: University of Cape Town and University of Michigan.

Lehtonen, R. and Pahkinen, E. J. (1995). *Practical methods for design and analysis of complex surveys*. New York, NY: Wiley.

Leibbrandt, M., Levinsohn, J. and McCrary, J. (2005) "Incomes in South Africa since the fall of apartheid". Working Paper No. 11384. Cambridge: National Bureau of Economic Research.

Lloyd, P., Cowan, B. and Mohlakoana, N. (2004). *Improving access to electricity and stimulation of economic growth and social upliftment*. Contribution to the Conference Improving Access to Modern Energy Services through CDM and Technology Transfer. Eskom Conference Centre, 27–29 July. Energy Research Centre, Rondebosch: University of Cape Town.

Macro International Inc. (ORC Macro). (2005). M*alaria indicator survey: guidelines for sampling for the malaria indicator survey*. Calverton, Maryland April 2005. Online: http://www.who.int/malaria/docs/mis/cc8.pdf (accessed 14 September 2007).

Madow, W. G., Nisselson, J. and Olkin, I. (eds). (1983). *Incomplete data in sample surveys, volume 1: report and case studies*. New York: Academic Press.

Marc, A. (2005). *"Cultural diversity and service delivery: where do we stand?"*. World Bank. Arusha Conference, December 12-15, 2005. New Frontiers of Social Policy. Online: http://siteresources.worldbank.org/INTRANETSOCIALDEVELOPMENT/ Resources/Marcpaper.rev.pdf (accessed 21 September 2008).

Masika, R. and Baden, S. (1997). "Infrastructure and poverty: a gender analysis". *Report no. 51. Bridge (development-gender) infrastructure of development studies.* Sussex: University of Sussex. Online: http://www.bridge.ids.ac.uk/reports/re51.pdf (accessed 08 October 2008).

May, J. (2000). *Poverty and inequality in South Africa: meeting the challenge*. Cape Town: David Philip publishers, and London and New York: Zed Books.

Meth, C. and Dias, R. (2003). *Increases in poverty in South Africa, 1999–2002*. Cape Town: Development Policy Research Unit School of Economics, University of Cape Town.

Meyers, R. A. (2006). *Designing systematic inquiry: research methods and evidence gathering.* Lecture notes. Milwaukee: University of Wisconsin.

Moshapo, S. and Hanrahan, H. E. (2003). "ICT services for poor rural community development: what is missing in current implementations?" *Southern African Telecommunications and Applications Conference (SATNAC) 2003. Proceedings* (CD-ROM), ISBN 0-620-32632-8, 6–8 September 2004. Cambridge, MA, USA: MIT Press.

Mukheibir, P. and Sparks, D. (2003). "Water resource management and climate change in South Africa: visions, driving factors and sustainable development indicators." Cape Town: Energy and Development Research Centre, University of Cape Town.

Nahlik, A. (2006). *Econometrics notes for 3-9-06*. Online: http://www.clarejones.com/econ/2ndsemester/Econometrics/Notes/Econometrics%203.9.06.pdf (accessed 08 October 2008).

NAO *see* National Audit Office.

Naraghi, S. and Kebotlhale, T. (2004). "Water as a key to socio-economic development and poverty eradication in North-West Province, as projected for the entire South Africa". This paper was originally presented at the 2004 Water Institute of South Africa (WISA) Biennial Conference, Cape Town, South Africa, 2–6 May 2004.

Nassirpour, M. (2004). “Overview of Sample Surveys” (Presentation). Applied Sampling Workshop at the Annual TRB Conference, Washington, D.C., 2004.

National Audit Office. (2000). *A Practical guide to sampling*. National Audit Office of the United Kingdom: London. Online: http://www.nao.org.uk/publications/ samplingguide.pdf (accessed 13 September 2008).

National Treasury. (2003). *Intergovernmental fiscal review 2003*. Pretoria: National Treasury. Online: http://www.treasury.gov.za/documents/ifr/2003/default.htm (accessed 08 August 2007).

———. (2007). *Estimates of national expenditure*. Pretoria: National Treasury. Online: http://www.treasury.gov.za/ (accessed 08 August 2007).

———. (2008). *National budget 2008.* Available: http://treasury.gov.za/documents/ nationalbudget/2008/ene/33transport.pdf (accessed 12 October 2008).

Nattrass, N. (2002). *Should youth unemployment be targeted as part of a comprehensive welfare policy in South Africa? Online*: http://www.commerce.uct.ac.za/research_units/ dpru/ Conf2003PDF/Conf2002pdf/NNattrass.pdf (accessed 08 October 2008).

Nemeroff, T. (2005). *Probing the protests.* Institute for Democracy in South Africa. Online: http://www.idasa.org.za/index.asp?page=outputs.asp%3FPID%3D33%26OTID%3D4 (accessed 12 October 2008).

Newbery, D. and Eberhard, A. (2007). “Southern African network infrastructure review: Electricity”. Online: http://www.gsb.uct.ac.za/gsbwebb/mir/documents/SAElectricityPaper08.pdf (accessed 12 October 2008).

Nigrini, M. (2007). *Empowering poor rural villages through the provision of financial services by means of financial service cooperatives: a preliminary investigation into financial service cooperatives in South Africa.* Online: http://www.essa.org.za/download/papers/007.pdf (accessed 12 October 2008).

Nordström, A., Korneval, C., Segerberg, F. and De Grijp, M. (2004). *Access to electricity – an approach to sustainable reduction of rural poverty.* 19th World energy Congress Australia: Sydney. Online: http://87.224.35.42:8190/wecgeis/congress/papers/nordstroma0904.pdf (accessed 09 October 2008).

OHRO *see* Oral History Research Office.

Oral History Research Office. (2005). *Oral history project*. Oral History Research Office Collection of the Columbia University Libraries (OHRO/CUL). Online: http://www.columbia.edu/cu/lweb/digital/collections/oral_hist/carnegie/special-features/ (accessed 14 September 2007).

Owens, L. K. (2002). "*Introduction to survey research design"*. Survey Research Laboratory (SRL) Fall 2002. Online: http://www.srl.uic.edu (accessed 14 September 2007).

Palmquist, R. A. (1997). *Survey methods*. Austin: University of Texas. Online: http://www.ischool.utexas.edu/~palmquis/courses/survey.html (accessed 07 October 2008).

Parliamentary Monitoring Group. (2007). "General overview of the central tenants of government". Online: http://www.pmg.org.za/parlinfo/sectiona1.htm (accessed 9 October 2007).

Pauw, K. (2005). "*Creating an IES-LFS 2000 database in Stata"*. PROVIDE Technical Paper Series. Elsenburg: PROVIDE Project. Online: www.elsenburg.com/provide (accessed 13 September 2007).

Pauw, K., Leibbrandt, M., Edwards, L. and Dieden, S. (2006). *Trade and poverty in South Africa: traded sector employment and vulnerability*. Cape Town: Southern Africa Labour and Development Research Unit. University of Cape Town.

Philpott, J. and Clark, A. (2002). "South Africa: Electricity reform with a human face?" in Dubash, N. (ed.) *Power politics: equity and environment in electricity reform.* Washington DC: World Resources Institute: 139–156.

Pindyck, R. S. and Rubinfeld, D. (1998). *Econonometric models and economic forecasts*. 4th ed. New York: McGraw-Hill.

PMG *see* Parliamentary Monitoring Group.

Porteous, D. (2003). "The landscape of access to financial services in South Africa". Labour Market and Social Frontiers 3. Pretoria: South African Reserve Bank.

Poswell, L. (2003), *Comments on income and expenditure survey 2000*. Cape Town: Development Policy Research Unit School of Economics, University of Cape Town.

Project for statistics on living standards and development. (1994). *Project for statistics on living standards and development.* Cape Town: Southern African Labour Development Research Unit.

PSLSD *see* Project for statistics on living standards and development.

Ravallion, M. (1992). "Poverty comparisons: a guide to concepts and methods. Living Standards Measurement Study. "*Working Paper no. 88.* Online: http://siteresources.

worldbank.org/PGLP/ Resources/Ravallion1992PovertyComparisons.pdf (accessed 08 October 2008).

Ravallion, M. (2004). "*Pro-poor growth: a primer*". *World Bank Policy Research Working Paper.* Washington: World Bank.

Ravallion, M. and Chen, S. (2003). *Measuring Pro-poor Growth. Economic Letters 78(1).93-99.* UK: Elsevier.

Republic of South Africa. (1996). *Constitution of the Republic of South Africa, 1996. (No. 108 of 1996)*. Online: http://www.info.gov.za/documents/constitution/index.htm (accessed 08 October 2007).

———. (1997a). *Housing Act, 1997 (No. 107 of 1997)*. Online: http://www.info.gov.za/ view/DownloadFileAction? id=70765 (accessed 08 October 2008).

———. (1997b) *Water Services Act, 1997 (No. 108 of 1997)*. Department of Water affairs and forestry. South Africa.

———. (2005). *National Credit Act 2005. (No. 34 of 2005)*. Online: http:/ /www.ncr.org.za/ pdfs/NATIONAL_CREDIT_ACT.pdf (accessed 10 October 2008).

———. (2006). *Independent Communications Authority of South Africa Amendment Act, 2006. (No. 3 of 2006).* Online: http://www.info.gov.za/view/DownloadFileAction?id =67854 (accessed 02 October 2008).

———. (2007). *Co-operative Banks Bill, 2007. (Bill 13 of 2007)*. Online: http:// www.info.gov.za/gazette/bills/2007/bills07_bill13-07.pdf (accessed 01 October 2008).

———. (2008). *The South African National Water Resources Infrastructure Agency Limited Bill, 2008. (Bill 36* of 2008*).* Online: http://www.parliament.gov.za/live/commonrepository/Processed/20081006/76026_1.pdf (accessed 09 October 2008).

SABG *see* South African Business Guide.

SAinfo *see* South Africa Information.

SAIRR *see* South African Institute of Race Relations.

South Africa Information. (2006). *More South Africans Banking.* Online: http://www.southafrica.info/business/economy/development/finscope2006.htm (accessed 22 September 2008).

———. (2007). *Diverse transport network*. Online: http://www.southafrica.info (accessed 06 July 2007).

South African Institute of Race Relations. (2004). *The South African survey 2003/2004.* Johannesburg: SAIRR.

———. (2005). The *South African survey 2004/2005*. South African Institute of Race Relations. Johannesburg: SAIRR.

———. (2006). *The South African survey 2005/2006*. South African Institute of Race Relations. Johannesburg: SAIRR.

Sandene, B., Horkay, N., Benette, R., Allen, N., Braswell, J., Kaplan B. and Oranje, A. (2005). "Online assessment in mathematics and writing report from the NAPE technology based assessment project". Research and Development Series. Online: http://nces.ed.gov/nationsreprotcard/pdf/studies/2005457.pdf (accessed 17 August 2007).

Scott, K. (2008). *Generating relevant household level data: multi-topic household surveys.* Washington: World Bank. Available: http://go.worldbank.org/LGWDCZ87J0 (accessed 07 September 2008).

Shtatland, E. S., Kleinman, K. and Cain, E. M. (2002). "*One more time on R2 measures of fit in logistic regression". NESUG 15 Proceedings.* Online: http://www.lrz-muenchen.de/~wlm/ ST004.pdf (accessed 08 October 2008).

Simkins, C. (2003). "A critical assessment of the 1995 and 2000 income and expenditure surveys as sources of information on incomes". Johannesburg: University of the Witwatersrand. (Unpublished work) in Pauw, K., Leibbrandt, M., Edwards, L. and Dieden, S. (2006). *Trade and poverty in South Africa: traded sector employment and vulnerability*. Cape Town: Southern Africa Labour and Development Research Unit. University of Cape Town.

Simkins, C. E. W. (2004). *What happened to the distribution of income in South Africa between 1995 and 2001?* Johannesburg: University of the Witwatersrand, 22 November.

Social housing Foundation. (2005). "Housing Entrepreneurs". A research project sponsored by the Gauteng Department of Housing, the Social Housing Foundation, Nedbank, the National Department of Housing and the FinMark Trust. Online: http://66.102.9.104/search?q=cache:Zp_yfKmpii0J:ftp.shf.org.za/landlords_HE1.pdf+Housing+Entrepreneurs,+2005&hl=en&ct=clnk&cd=2&gl=za (accessed 25 July 2007).

Social housing foundation. (2007). *The Housing backlog apex is broken*. Housing in Southern Africa. Available: http://ftp.shf.org.za/hisa_071.pdf (accessed 25 July 2008).

Statistics South Africa. (1999). *October Household Survey.* Various years 1995 – 1999. Pretoria: Statistics South Africa.

———. (2000a). *Labour Force Survey 2000*. Pretoria: Statistics South Africa.

———. (2000b). *Income and Expenditure Survey 1995 – 2000*. Pretoria: Statistics South Africa.

———. (2001). *Labour Force Survey 2001*. Pretoria: Statistics South Africa.

———. (2003). *Labour Force Survey 2003*. Pretoria: Statistics South Africa.

———. (2004a). *Labour Force Survey 2004*. Pretoria: Statistics South Africa.

———. (2004b). *Comparative guide to Stats SA censuses and surveys*. (2004). Pretoria: Statistics South Africa. Online: http://www.statssa.gov.za/CRS/pdfdocs/guidedraft.pdf (accessed on 10 August 2007).

———. (2004c). *Labour Force Survey: Metadata 2004 (Revised)*. Statistics South Africa.

———. (2007a). *General Household Survey*. 2002 – 2007. Statistics South Africa.

———. (2007b). *Labour Force Survey*. 2000 – 2007. Statistics South Africa.

Statpac inc. (2007). *Sampling methods*. Online: http://www.statpac.com/surveys/sampling.htm (accessed 12 September 2007).

Stats SA *see* Statistics South Africa.

Stats SA/Treasury. (2007). *A national poverty line for SA*. Statistics South Africa and National Treasury. Online: http://www.treasury.gov.za/publications/other/povertyline/Treasury%20StatsSA%20poverty%20line%20discussion%20paper.pdf (accessed 24 February 2008).

Sturgis, P. (2004). "Analysing complex survey data: clustering, stratification and weights". Social Research Update, 43: United Kingdom: University of Surrey. Published journal papers from the Department of Sociology. Online: Paper 2. http://epubs.surrey.ac.uk/publsoc2/2 (accessed 10 August 2007).

Taylor-Powell, E. (1998). *Questionnaire design: asking questions with a purpose*. Wisconsin: University of Wisconsin. Online: http://learningstore.uwex.edu/pdf/G3658-2.pdf (accessed 08 September 2008).

Telkom. (2007). "International ICT Companies". Online: http://www.telkom.co.za/common/wholesale/MobileCelullarOperator/InternationalOperators.html (accessed 04 October 2008).

The Presidency. (2007). *Development Indicators Mid-Term Review*. The Presidency of the Republic of South Africa. Pretoria: The Presidency. Online: http://www.info.gov.za/otherdocs/2007/developmentindicator/development_indicator.pdf (accessed 18 October 2008).

Thirlwall, A. P. (1999). *Growth and development: with special reference to developing economies.* 6th ed. Basingstoke: Macmillan.

Todaro, M. and Smith, C. (2003). *Economic development*. 8th ed. Edinburgh Gate, Harlow, Essex: Pearson Education Limited.

Townsend, P. (1979). *Poverty in the United Kingdom.* Harmondsworth: Penguin Books.

Trochim, W. M. K. (2006). "*Probability sampling"*. Online: http://www.socialresearchmethods.net/kb/smpprob.php (accessed 09 October 2007).

van der Berg, S., Burger, R., Burger, R., Louw, M. and Yu, D. (2006). *Trends in poverty and inequality since the political transition*. Stellenbosch: Development Policy Research Unit. University of Stellenbosch.

van der Berg, S. and Louw, M. (2004). "C*hanging patterns of South African income distribution: towards time series estimates of distribution and poverty"*. South African Journal of Economics 72(3): 546–572.

van der Westhuizen, C. (2004). "*Comparative provincial housing brief 2004"*. Budget Brief no. 145, IDASA. Online: http://www.idasa.org.za/index.asp?page=outputs.asp%3FTID%3D10%26OTID%3D15 (accessed 08 October 2008).

van Teijlingen, E. and Hundley, V. (2001). *The importance of studies.* Surrey: University of Surrey. Online: http://sru.soc.surrey.ac.uk/SRU35.pdf (accessed 08 September 2008).

Vasisht, A. K. (2003). *Logit and probit analysis*. New Delhi: I. A. S. R. I. Online: http://www.iasri.res.in/ebook/EBADAT/6-Other%20Useful%20Techniques/5-Logit%20and%20 Probit%20Analysis%20Lecture.pdf (accessed 08 October 2008).

Vermaak, C. (2005). "*Trends in income distribution, inequality and poverty in South Africa, 1995–2003"*. Paper presented at the Economic Society of South Africa Conference, Durban 7–9 September.

Water Research Commission. (2003). "*Sanitation research strategy"*. Online: http://www.wrc.org.za/downloads/special%20publications/SanitationResearchStrategyFinal1.pdf (accessed 06 June 2007).

Wildeman, R. A. (2002). *Infrastructure provisioning in schools and the right to basic education.* Education Rights Project, South Africa. Online: http://www.erp.org.za/htm/issue2.htm (accessed 08 October 2008).

Woolard, I. (2002). “An overview of poverty and inequality in South Africa”. Working paper prepared for the Department for International Development (DFID) South Africa.

———. (2003). *Social assistance grants, poverty and economic growth in South Africa*. Cape Town: Development Policy Research Unit School of Economics, University of Cape Town.

Woolard, I. and Leibbrandt, M. (1999). *Measuring poverty in South Africa*. Cape Town: Development Policy Research Unit: University of Cape Town.

Woolard, I. and Leibbrandt, M. (2006). *Towards a poverty line for South Africa: background note*. Cape Town: Southern Africa Labour and Development Research Unit. University of Cape Town.

World Bank. (2007). *World development indicators 2007*. Washington: World Bank.

———. (2008). *World Bank updates poverty estimates for the developing world*. Online: http://go.worldbank.org/C9GR27WRJ0 (accessed 05 October 2008).

World Water Assessment Programme. (2006). “*Water, a shared responsibility”*. *The United Nations World Water Development Report 2.* World Water Assessment Programme. Paris: UNESCO, and New York: Berghahn Books.

WRC *see* Water Research Commission.

WWAP *see* World Water Assessment Programme.

APPENDIX 1

Table 1.1 Probit results

Explanatory variable	2001		2004	
	Coefficient	**Z**	**Coefficient**	**Z**
Constant	1,051	12,81	1,272	13,02
Demographic variables:				
Status				
Employed				
Unemployed	-0,438	-10,32	-0,563	-12,83
Age	-0,005	-4,50	-0,002	-2,14
Population				
African				
Coloured	-0,085	-2,27	-0,027	-0,75
Indian	0,635	5,18	0,601	5,18
White	1,177	11,16	1,173	12,44
Gender				
Male				
Female	-0,230	-9,45	-0,166	-6,79
Education	0,060	18,23	0,065	19,63
Income				
Salaries and wages				
Remittances	-0,314	-6,25	-0,218	-4,30
Pensions and grants	-0,356	-5,78	-0,699	-14,07
Sales of farm products	-0,499	-4,96	-0,506	-4,77
Other non-farm income	-0,447	-9,38	-0,426	-9,34
No income	-0,484	-5,54	-0,234	-2,19
Infrastructure variables:				
House				
Formal house				
Traditional dwelling	-0,099	-2,17	-0,234	-5,51
Informal house	0,005	0,17	-0,060	-1,72
Cellular telephony				
Access				
No access	-0,228	-8,03	-0,056	-2,34
Telecommunications				
Access				
No access	-0,224	-6,19	-0,281	-6,89
Rubbish removal				
Removal by local authorities				
Removal by community members	-0,129	-1,15	0,039	0,33
Communal dump	-0,164	-5,45	-0,073	-2,36
None	-0,147	-3,07	-0,159	-2,95
Toilet				
Access to formal toilet				
Access to toilet not in a dwelling	-0,276	-8,67	-0,252	-7,88
No toilet	-0,359	-7,19	-0,216	-4,33
Electricity				
Access to formal electricity from mains or solar power				
Traditional sources	-0,155	-3,09	0,049	0,89
Other sources	-0,093	-2,89	-0,022	-0,66
Water				
Access to piped water inside the dwelling				
Access to piped water outside the dwelling	-0,098	-2,88	-0,194	-5,70
Access to informal water	-0,242	-4,95	-0,298	-6,01

Sources: Stats SA 2001; 2004. Stata output.

Table 1.2 Marginal effects of the Probit model

Explanatory variable	2001		2004	
	Coefficient	Z	Coefficient	Z
Constant				
Demographic variables:				
Status				
Employed				
Unemployed	-0,156	-9,74	-0,195	-12,79
Age	-0,001	-4,50	-0,001	-2,14
Population				
African				
Coloured	-0,028	-2,23	-0,009	-0,75
Indian	0,165	7,17	0,170	6,83
White	0,264	22,93	0,285	23,94
Gender				
Male				
Female	-0,078	-9,23	-0,058	-6,69
Education	0,019	18,10	0,022	19,55
Income				
Salaries and wages				
Remittances	-0,111	-5,91	-0,078	-4,14
Pensions and grants	-0,127	-5,42	-0,266	-13,65
Sales of farm products	-0,184	-4,64	-0,191	-4,55
Other non-farm income	-0,162	-8,78	-0,159	-8,89
No income	-0,178	-5,17	-0,085	-2,10
Infrastructure variables:				
House				
Formal house				
Traditional dwelling	-0,033	-2,12	-0,084	-5,31
Informal house	0,001	0,17	-0,021	-1,71
Cellular telephony				
Access				
No access	-0,073	-8,31	-0,019	-2,34
Telecommunications				
Access				
No access	-0,071	-6,48	0,092	-7,34
Rubbish removal				
Removal by local authorities				
Removal by community members	-0,044	-1,11	0,013	0,34
Communal dump	-0,055	-5,36	-0,025	-2,35
None	-0,050	-2,97	-0,057	-2,86
Toilet				
Access to formal toilet				
Access to toilet not in a dwelling	-0,091	-8,70	-0,087	-7,91
No toilet	-0,127	-6,79	-0,078	-4,18
Electricity				
Access to formal electricity from mains or solar power				
Traditional sources	-0,053	-2,98	0,016	0,90
Other sources	-0,031	-2,84	-0,007	-0,66
Water				
Access to piped water inside the dwelling				
Access to piped water outside the dwelling	-0,033	-2,83	-0,069	-5,55
Access to informal water	-0,084	-4,72	-0,109	-5,75
Number of observations	16593		16137	
LR chi2 (25)	4552,88		4751,46	
Prob > chi2	0,0000		0,0000	
Pseudo R2	0,2149		0,2261	
Log-likelihood	-8316,155		-8130,3241	
Obs. Probability	0,73077719		0,70322999	

Sources: Stats SA 2001; 2004. Stata output.